A Systematic Literature Review on Artificial Intelligence Applications for Breast Cancer Classification

Nursakinah Abdullah[1], Qi Wei Oung[2], Chee Chin Lim[3],
Chiew Chea Lau[4], Vrshni Menaka R Siva Nathan[5], Hui Wen Tiu[6]
Faculty of Electronic Engineering & Technology (FKTEN), Universiti Malaysia Perlis, Arau, 02600, Perlis, Malaysia [1, 2, 3, 6]
Sport Engineering Research Centre (SERC), Universiti Malaysia Perlis, Arau, 02600, Perlis, Malaysia[3]
Advanced Communication Engineering (ACE), Universiti Malaysia Perlis, Arau, 02600, Perlis, Malaysia[2]
School of Medical Sciences-Health Campus, Universiti Sains Malaysia,
Kubang Kerian, Kota Bharu, 16150, Kelantan, Malaysia[4, 5]

***Abstract*—Breast cancer remains one of the most prevalent and life-threatening diseases worldwide, needing to be diagnosed early and properly classified for effective treatment. Advancements in artificial intelligence (AI), deep learning, and machine learning techniques have shown great potential in automating breast cancer diagnosis and molecular subtyping using medical imaging. This systematic literature review explores the application of AI in breast cancer classification, focusing on mammographic imaging and its application in distinguishing molecular subtypes. The study follows the PRISMA guideline, investigating studies from multiple digital libraries published between 2020 and November 2024. Findings show that while deep learning models have significantly improved breast cancer detection, challenges remain in optimizing classification models for molecular subtypes, balancing accuracy and interpretability, and integrating AI-based tools into clinical practice workflows. Besides, heterogeneity in preprocessing pipeline algorithms and dataset limitations highlights the importance of conducting additional research to develop robust and generalized classification models. This review underscores the importance of AI-driven solutions in advancing breast cancer diagnosis and treatment planning while providing insights into future research directions.**

Keywords*—Artificial intelligence; breast cancer; classification; Convolutional Neural Network; deep learning; machine learning; mammography; medical imaging; molecular subtypes; Vision Transformer

I. Introduction

Breast cancer occurs when abnormal cells in the breast grow uncontrollably, forming tumors. This often involves changes in hormone receptors, such as estrogen (ER), progesterone (PR), and human epidermal growth factor receptor 2 (HER2). According to the World Health Organization, breast cancer affected 2.3 million women worldwide in 2022, causing around 670,000 deaths, making it the second most common cancer globally [1][2]. In Asia, particularly in Malaysia, the incidence continues to rise, with breast cancer being the most frequently diagnosed cancer, accounting for 16.2% of all cases [3].

Classifying breast cancer into molecular subtypes is crucial for determining appropriate treatment strategies. Based on gene expression profiling, the main subtypes include Luminal A, Luminal B, HER2-enriched, Basal-like, and Claudin-low [4][5][6][7]. Each subtype exhibits distinct biological characteristics and responds differently to therapeutic interventions, making accurate classification critical for personalized care [8]. For instance, luminal tumors are typically treated with endocrine therapy, HER2-enriched tumors respond well to targeted antibody therapy, while Basal-like or triple-negative tumors, although aggressive, may show responsiveness to chemotherapy.

Currently, molecular subtyping is mainly done through immunohistochemistry (IHC), which requires invasive tissue biopsies. However, this approach presents several limitations, including patient discomfort, procedural risks, and diagnostic inaccuracies. Studies have reported that up to 30% of biopsies may yield false-negative results due to sampling errors, potentially delaying appropriate treatment. Although imaging modalities such as mammography, magnetic resonance imaging (MRI), and ultrasound provide valuable tumor-related information, they are not routinely utilized for molecular subtype classification. Among these, digital mammography stands out as a cost-effective and widely accessible screening tool, particularly in resource-limited settings. It also offers practical advantages in terms of acquisition time, typically requiring only 15 to 20 minutes compared to longer durations for ultrasound and MRI. Among these, digital mammography is particularly notable for its cost-effective and widespread accessibility, especially in resource-limited settings. It also offers practical advantages over other imaging techniques, with screening procedures typically taking just 15 to 20 minutes, compared to over 30 minutes for ultrasound and between 30 and 90 minutes for MRI scans [9]. Furthermore, mammography enables early detection of breast cancer by identifying abnormalities two to three years before they become clinically palpable [10].

Despite such advances, a substantial research problem remains, namely the lack of accurate, dependable, and non-

invasive techniques for the categorization of breast cancer genetic subtypes utilizing imaging data. Current diagnostic methods predominantly depend on invasive techniques, whereas contemporary imaging modalities primarily emphasize tumor identification rather than subtype classification. Furthermore, the incorporation of sophisticated artificial intelligence (AI) methodologies, especially deep learning architectures like Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs), has not been thoroughly examined regarding non-invasive subtype classification. This deficiency constrains the advancement of effective computer-aided diagnostic systems that can assist in clinical decision-making. Non-invasive AI-based approaches have the potential to reduce unnecessary biopsies and improve diagnostic efficiency.

Therefore, the primary objective of this study is to systematically review and analyze the application of artificial intelligence techniques in breast cancer molecular subtype classification using imaging data. Specifically, this study aims to: 1) categorize existing approaches into CNN-based, Vision Transformer-based, and hybrid CNN–ViT models; 2) evaluate their performance, strengths, and limitations; and 3) identify current research gaps and future directions for improving non-invasive diagnostic methods.

The significance of this study lies in its contribution to advancing non-invasive breast cancer diagnosis by providing a comprehensive synthesis of state-of-the-art AI techniques. By critically analyzing existing methods, this review offers valuable insights into the effectiveness of different deep learning architectures and highlights opportunities for integrating imaging, radiomic features, and advanced models to enhance classification accuracy. Ultimately, this work supports the development of reliable, AI-driven diagnostic tools that can reduce dependence on invasive biopsies, improve early detection, and facilitate personalized treatment planning.

In this study, a systematic literature review based on the PRISMA 2020 framework is conducted to evaluate recent advancements in AI-driven breast cancer molecular subtype classification. The selected studies are analyzed and categorized according to their methodological approaches, including CNN-based models, Vision Transformer-based models, and hybrid architectures. Through this structured analysis, the study aims to provide a clear understanding of current trends, challenges, and future research opportunities in this domain.

This review bridges AI-based imaging analysis with clinically relevant molecular subtype classification, directly linking computational advancements to personalized treatment strategies and decision-making in breast cancer care.

II. Research Methodology

This systematic literature review adopts a focused (scoping) approach to explore AI's clinical applications in mammography imaging for classifying breast cancer molecular subtypes. The study specifically emphasizes recent and highly relevant works that address non-invasive subtype classification using imaging data, ensuring a targeted and in-depth analysis within a clearly defined research scope. Additionally, it examines other diagnostic methods, highlighting their current limitations and discussing future directions for AI in breast cancer diagnosis and treatment planning. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guiding principles were employed for the systematic review. This approach ensures a consistent and replicable process for identifying, selecting, and evaluating the relevant studies [11]. Fig. 1 presents the PRISMA flow diagram, illustrating the inclusion and exclusion process for articles considered in this literature review.

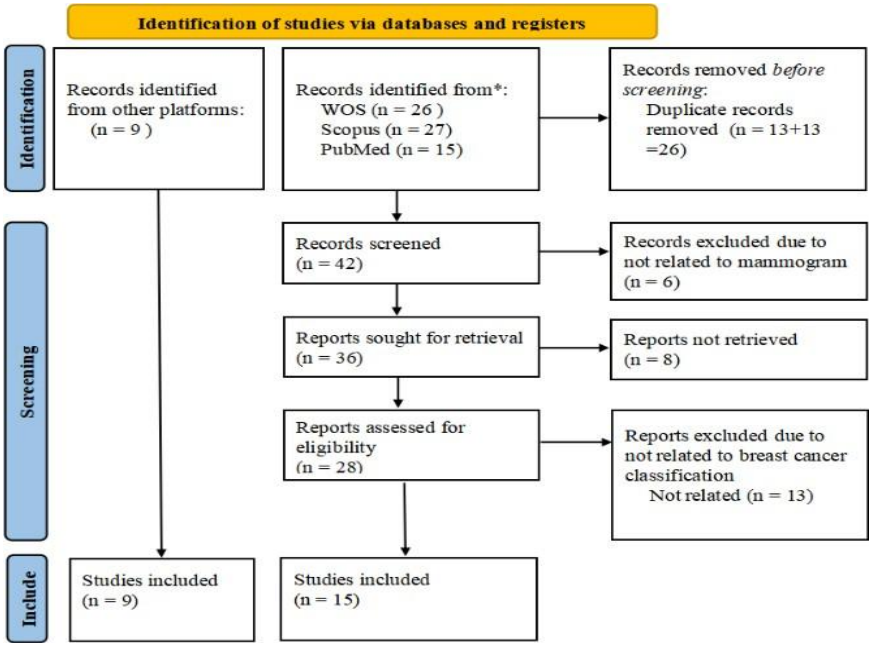

Fig. 1. PRISMA flow diagram.

A. Research Question

The research questions for this study are outlined in Table I, defining the overall purpose and expected outcomes. Each question is presented along with its motivation.

TABLE I. RESEARCH QUESTIONS AND MOTIVATIONS

Research Question	Main Motivation
What is the most effective image processing technique applied to extract tumor features from mammographic images?	To develop image processing technique to extract radiomic features, including shape and texture, relevant to breast cancer molecular subtypes from digital mammographic images.
What is the most effective model architecture for breast cancer molecular subtypes classification in terms of performance?	To construct classification models to analyze radiomic features and classify breast cancer into molecular subtypes.
How can a GUI be designed to support seamless interaction and decision-making for breast radiologists for identifying molecular subtypes of breast cancer?	To develop a Graphical User Interface (GUI) that allows users to automatically identify breast cancer molecular subtypes from digital mammographic images.

B. Search Scheme

The primary digital libraries used for this search were Web of Science (WOS), Scopus, and PubMed. These databases were selected for their comprehensive collections of peer- reviewed studies and their capability to offer access to the latest advancements in biomedical engineering and artificial intelligence applications. Additionally, several breast cancer classification journals were identified on platforms outside these three databases, including Google Search, Google Scholar, ScienceDirect, and IEEE Xplore Library.

C. Systemic Review Process

The systematic review followed the PRISMA framework, consisting of three key phases, which are identification, screening, and analysis. The literature search was conducted from 2020 to November 6, 2024, using the search string: "breast" AND ("cancer" OR "tumour") AND "subtype*" AND "deep learning" AND "mammogra*".

During the identification stage, a total of 68 studies were obtained from Web of Science (WOS), Scopus, and PubMed. However, 26 studies were removed after filtering for duplicate papers. Besides, an additional 9 studies were identified from other platforms. In the screening stage, the 27 papers were further excluded, as 6 studies were unrelated to mammograms, 8 studies could not be retrieved, and 13 studies were not focused on breast cancer classification. Finally, a total of 24 research articles were selected for review after performing an identification from other platforms. Furthermore, some studies that explored other modalities for breast cancer classification, such as Magnetic Resonance Imaging (MRI), ultrasound, and histopathology techniques, were considered as references. While some of these studies were included in the final set of 24 research articles, others were excluded but were still referenced in several sections of the literature review. Fig. 2 displays the distribution of the final set of selected research articles.

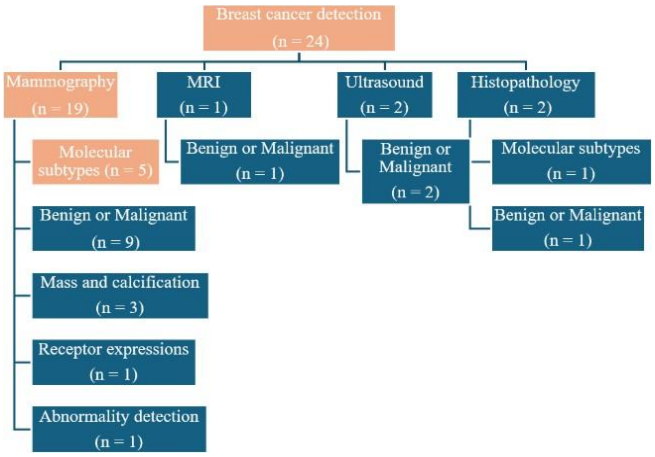

Fig. 2. Distribution of the final set of selected research articles.

D. Inclusion and Exclusion Principles

Table II lists the inclusion and exclusion principles for the chosen research. Stringent inclusion and exclusion criteria were applied to ensure the relevance and quality.

TABLE II. RESEARCH INCLUSION AND EXCLUSION CRITERIA

Inclusion criteria	Exclusion criteria
IC1 – Search will be conducted across all fields.	EC1 – Studies not related to breast cancer.
IC2 – Limited to open access and open archive materials.	EC2 – Studies lack focus on tumor radiomics feature extraction and classification.
IC3 – Articles and research articles.	EC3 – Review articles, books, chapters in book series, news articles, conference proceedings, case reports, and abstracts without full papers.
IC4 – The publication date range is set from 2020 to January 2026.	EC4 – Articles with inaccessible full texts.
IC5 – Studies published in English.	EC5 – Non-English publications.

III. RESULTS AND DISCUSSION

Table IV presents all the included studies, detailing their sample size, sample source, processing methods, architecture, and SWOT analysis (which examines strengths, weaknesses, opportunities, and threats).

Recent studies have explored the application of artificial intelligence techniques, particularly Convolutional Neural Networks (CNNs) and emerging Vision Transformer (ViT) architectures, for breast cancer classification using medical imaging. CNN-based approaches have demonstrated strong performance in tumor detection and classification tasks due to their ability to capture local spatial features. Meanwhile, Vision Transformer-based models have shown potential in modeling global contextual relationships within medical images. Hybrid CNN–Transformer architectures have also been introduced to leverage the strengths of both approaches.

However, most existing studies primarily focus on binary classification tasks, such as distinguishing between benign and malignant tumors, rather than performing detailed molecular subtype classification. Furthermore, many models rely on limited datasets or single imaging modalities, which may affect

generalizability and robustness. These limitations highlight the need for a comprehensive analysis of current AI-based methods specifically targeting non-invasive breast cancer molecular subtype classification. Therefore, this study aims to bridge this gap by systematically reviewing and comparing existing approaches while identifying key challenges and future research directions.

The overall workflow of AI-based breast cancer analysis adopted in this study is illustrated in Fig. 3, which presents a systematic framework encompassing data acquisition, preprocessing, feature extraction, model development, and classification.

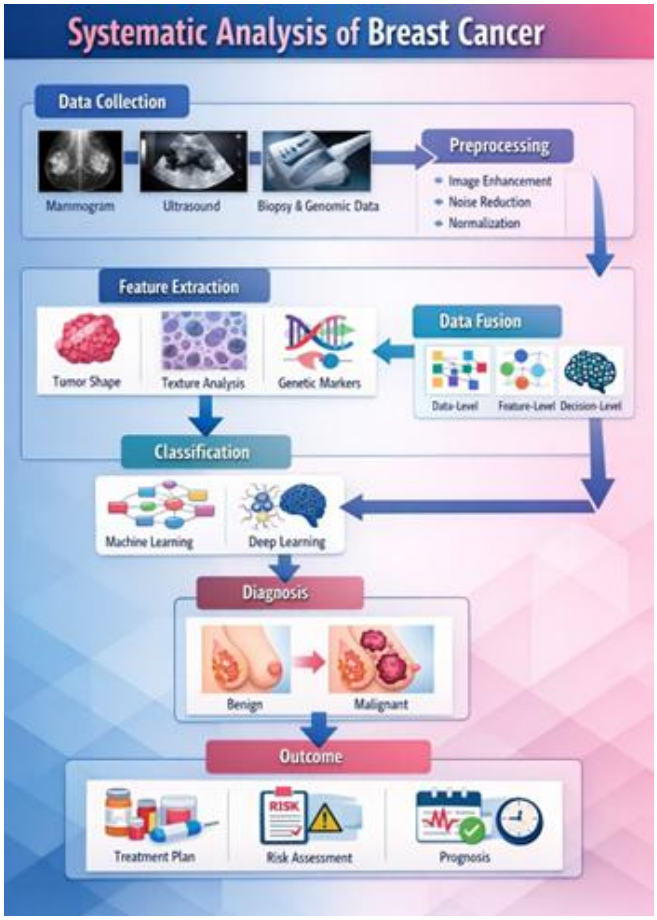

Fig. 3. Systematic framework for breast cancer analysis.

A. Machine Learning Approaches

Studies such as Khalid et al. [12] and Ma et al. [8] have employed machine learning approaches to classify mammogram images. Khalid et al. [12] implemented six classification models to distinguish between benign and malignant breast tumors, concluding that the Random Forest algorithm achieved the highest accuracy of 96.49%. On the other hand, Ma et al. [8] proposed a Naïve Bayes-based binary classification model to identify various molecular subtypes of breast cancer. These approaches, however, rely on manual feature selection and engineering, which necessitates human input and expertise. For example, the former utilized three modules, the low-variance features removal, univariate feature selection, and recursive feature elimination, to refine their models. Similarly, the latter combined manual segmentation with automated feature selection and extraction using LASSO (Least Absolute Shrinkage and Selection Operator) method.

B. Deep Learning Approaches

In contrast, deep learning approaches offer a significant advantage through their fully automated methodologies, from segmentation to classification. This is particularly beneficial in domains where the relevant features are not predefined or easily identifiable. For example, Carvalho et al. [13] introduced an end-to-end classifier leveraging SegNet and UNet for segmentation, thereby processing raw images without manual intervention. While deep learning typically requires large datasets for training, it is better suited for complex tasks such as image recognition, surpasses the traditional machine learning approaches.

C. Histopathology

Histopathology, which involves the diagnosis and study of diseases of the tissues level, plays a critical role in breast cancer analysis. It typically involves examining tissues or cells under a microscope and often requires a biopsy [14]. Biopsy is a procedure to remove cells, tissue or fluid from the targeted organ for examination. As the reference standard for breast cancer diagnosis, histopathology enables classification into molecular subtypes [15], which is essential for specialist to determine the type of breast cancer and devise an appropriate treatment plan.

Murtaza et al. [16] introduces a two-level hierarchical classifier employing various types of deep neural networks to classify breast cancer using histopathology images. The framework identifies eight distinct output classes, which are classifier distinguishes images as benign and malignant at the first level, while the second level further categorizes subtypes within each category. Among the architectures evaluated, the k-Nearest Neighbors (kNN) algorithm demonstrated the best performance, achieving 95.48% accuracy for first-level classification, 94.62% for benign subtypes, and 92.45% for malignant subtypes. Similarly, Rafiq et al. [17] utilizes histopathology images from the Kaggle dataset to evaluate the classification performance of a fusion-based convolutional neural network (CNN) framework. The authors proposed three CNN-based architectures for benign and malignant cancer classification, which are a straightforward CNN model (1-CNN), a fusion CNN model (2-CNN) and three CNN model (3-CNN). Among these, the 3-CNN architecture recorded the highest accuracy of 97.90%.

D. MRI and Ultrasound Imaging

Although histopathology remains a critical step in diagnosing breast cancer, non-invasive screening methods such as mammography, MRI and ultrasound have become increasingly favored as feasible options for large-scale screening. These methods offer advantages in terms of accessibility and reduced patient discomfort, often outweighing the invasive nature of histopathology in routine screening contexts. For instance, Carvalho et al. [13] utilized 20 dynamic contrast-enhancing magnetic resonance imaging (DCE-MRI) images from the public Quantitative Imaging

Network (QIN) database, recorded from 10 patients, to evaluate the accuracy of classifying tumor malignancy using the three-time-point (3TP) method, and the model achieved an impressive 100% classification accuracy. However, the small dataset used in the study raises concerns about potential overfitting, which could limit the model's generalizability. Beyond MRI, ultrasound imaging has also demonstrated utility in breast cancer classification. For example, Alshowarah [18] analyzed 780 ultrasound images from the Breast Ultrasound Images (BUSI) dataset to compare the performance of various classification architectures. Among these, the Random Forest algorithm outperformed others, achieving an accuracy of 88.63%. Similarly, T. Zhang et al. [19] and Jiménez-Gaona et al. [20] employed ResNet-50 and ResNet-18 architectures, respectively, to classify benign and malignant breast cancer using ultrasound and mammogram images, demonstrating the growing potential of advanced deep learning techniques in non-invasive diagnostics.

E. Mammography Imaging

Digital mammography is a cost-effective and widely available diagnostic tool, particularly in less-developed countries. It offers significant advantages over other imaging methods, as screening mammograms typically take only 15 to 20 minutes, compared to the 30 minutes or more required for ultrasound and the 30 to 90 minutes needed for MRI scans [9]. Additionally, mammography is particularly suited for early breast cancer detection, capable of identifying lumps two or three years before they can be felt by the patient [10]. Therefore, regular breast cancer screenings using mammography play a vital role in reducing modality risks and facilitating earlier and more effective treatment. Numerous studies have leveraged mammogram images for breast cancer classification. For example, Bobowicz et al. [21] used mammogram images from Medical University of Gdansk (MUG) and Chinese Mammography Database (CMMD) to train and test the models for classifying benign and malignant breast cancers. In their study, ResNet-34 demonstrated the best performance, achieving an accuracy of 82.2%. Mammography has also been employed in other applications, such as the classification of receptor expressions [22], feature extraction and early-stage breast abnormality detection [23], and the classification of breast masses and calcification [24]. These advancements highlight the versatility and effectiveness of mammography in breast cancer diagnosis and research.

F. Breast Cancer Molecular Subtypes Classification

As previously highlighted, breast cancer subtyping has critical therapeutic implications for clinical management of the disease, thus several studies have focused on developing classifiers for breast cancer molecular subtypes, including luminal A, luminal B, human epidermal growth factor receptor 2 (HER2)-enriched, and triple-negative types.

G. Convolutional Neural Network and Machine Learning

Molecular subtyping remains challenging due to subtle radiographic differences between subtypes, heterogeneity in tumor morphology, and frequent class imbalance in datasets (e.g., Luminal A dominance). Binary classifiers often fail to capture subtype-specific biomarkers, leading to accuracy <80% (Table IV). For instance, binary classification model which used to classify luminal A and non-luminal A subtypes is proposed by Liu et al. [25] with ResNet-50, achieving an accuracy of 53.3%. Similarly, Ma et al. [8] employed a Naive Bayes-based binary classification model to differentiate molecular subtypes recording accuracies of 0.796 for triple-negative, 0.748 for HER2-enriched and 0.788 for luminal subtypes. In contrast to these binary approaches, studies such as Mota et al. [26], Nissar et al. [15] and T. Zhang et al. [19] developed models capable of distinguishing multiple molecular subtypes. For example, Mota et al. [26] achieved an area under the curve (AUC) of 60.62% using a ResNet-101 deep CNN, Nissar et al. [15] reported an accuracy of 90% using MobileNet-V3 —highlighting that high accuracy may mask poor sensitivity in minority classes (e.g., HER2-enriched); and T. Zhang et al. [19] achieved an accuracy of 88.5% with ResNet-50. These multi-class classification model demonstrate the potential for more comprehensive diagnostic tools, advancing precision medicine in breast cancer treatment.

H. Vision Transformers

Despite the success of convolutional neural networks (CNNs) in medical imaging, their limited local receptive field restricts their ability to capture global context information [27]. A recent innovation, Vision Transformers (ViTs) has demonstrated a great potential to be an alternative to CNNs in image analysis due to their capacity to process large datasets and learn complex patterns. Gheflati & Rivaz [27] highlights the application of Vision Trans- formers (ViTs) in classifying breast ultrasound images, showcasing their ability to capture global dependencies within images through self-attention mechanisms by achieving 86.7% of accuracy, outperforming several state-of-the-art CNN architectures, including ResNet-50, VGG16, InceptionV3 and NASNetLarge. This study underscores ViTs' adaptability to small datasets using transfer learning, al- though computational requirements remain a challenge. Zeng et al. [28] also applied a Vision Transformer Multi-head Attention (ViT-MHA) architecture for breast mass diagnosis, leveraging bilateral mammography images to enhance diagnostic accuracy to 92.7%.

I. Hybrid CNN-ViTs

The hybridization of ViTs with CNNs has further improved image classification tasks. For example, Boudouh & Bouakkaz [29] combines ViTs for contextual feature extraction with CNNs for local detail recognition using the CBIS-DDSM dataset. ViTs outperform CNNs in capturing global tumor context but require 3–5× more training data and GPU resources. Hybrid ViT-CNN models (Boudouh & Bouakkaz, 2024 [29]) mitigate this at the cost of interpretability. Among the proposed models, the fusion of ViT++ and VGG16 achieved the highest accuracy of 99.22% underscoring the effectiveness of this hybrid approach. These studies highlight the potential of Vision Transformers in medical imaging while challenges such as computational demands and data diversity remain areas for future research and optimization.

J. Graphical User Interface

A Graphical User Interface (GUI) is a critical feature for transitioning a proposed classifier into clinical use. Zhang et al. [30] developed a web server using the open-source Dash framework to host their trained BSNet model. This web application enables data scientists to analyze mammogram images by uploading them to the server, which outputs probabilities for three breast disease classifications along with extracted features. Besides, Jiménez-Gaona et al. [20] created an open-source mobile app named "BraNet" for the segmentation and classification of benign and malignant breast cancer using mammography and ultrasound breast images. This app employs ResNet-18 deep learning algorithms, offering a practical tool for healthcare applications. Despite promising accuracy, GUI tools face regulatory (e.g., FDA/CE certification), interoperability (DICOM/HL7 integration), and physician-trust barriers.

K. Recommendations and Limitations

The literature review reveals that only a limited number of studies have specifically focused on developing classification models for breast cancer molecular subtypes using mammographic images. In particular, only five studies have directly addressed this problem, as summarized in Table III. This highlights a significant research gap in non-invasive subtype classification using imaging data.

To provide a broader perspective and enable cross-study comparison, a comprehensive summary of relevant studies, including model architectures, datasets, and performance metrics, is presented in Table IV.

Furthermore, to facilitate systematic comparison and evidence synthesis across studies, a standardized summary of included works is presented in Table V, highlighting model architectures, datasets, sample sizes, imaging modalities, and performance metrics.

TABLE III. SUMMARY OF CNN TECHNIQUES USING MAMMOGRAPHIC IMAGES TO CLASSIFY MOLECULAR SUBTYPES IN PAST STUDIES.

Author	Method	Accuracy	Remark
Mota et al. (2024) [26]	ResNet-101	60.62%	Low accuracy
Nissar et al. (2024) [15]	MobileNet-V3 with Convolutional Block Attention Module	90%	High complexity
T. Zhang et al. (2023) [19]	ResNet 50 with intra-modality and inter-modality	88.5%	High complexity
Ma et al. (2019) [8]	Naïve Bayes	75% to 80%	Low accuracy
Liu et al. (2023) [25]	ReNet 50	53.3%	Low accuracy

TABLE IV. SUMMARY OF THE PAST RESEARCH

Author	No. of sample and sample source	Processing methods	Architecture	Result	SWOT analysis (S = Strengths, W = Weaknesses, O = Opportunities, T = Treats)	Remark
Mammogram						
Mota et al., 2024 [26]	> 660 patients, 1397 images > Mammography images from OPTIMAM imaging database	> Cropping > Sampling for imbalanced data (imbalanced data, oversampling and under sampling) > Data augmentation > 3-fold cross-validation	ResNet-101 deep CNN	Binary classification: Best performance = Oversampling and data augmentation Average accuracy = 79.02% Average AUC = 64.69% multi-class classification: Best performance = Oversampling and data augmentation AUC = 60.62%	S = Access to a larger dataset of mammography images. Implementation of robust pre-processing and data augmentation techniques, enhancing model generalizability. W = Dependence on labelled mammographic data. Lack of data variability, potentially affecting the model's performance and generalizability. O = Potential to address class imbalance issues for more balanced model training. T = Exclusion of Ki-67 data may reduce the model's ability to differentiate between subtypes. Challenges in distinguishing subtle differences among molecular subtypes.	Classification molecular subtype
Nissar et al., 2024 [15]	> 2358 images before augmentation, 4987 images after augmentation > Mammography	> Image cropping and filtering > Image enhancement and augmentation > Image normalization	MobileNet-V3 architecture with a Convolutional Block Attention Module (MOB-CBAM)	Best performance = Median filter Accuracy = 0.90 Precision = 0.91 Recall = 0.90 F1-score = 0.89 Matthew's correlation	S = High accuracy. Capable of identifying various types of breast cancer, including masses, calcifications and benign or malignant tumour, as well as predicting molecular subtypes. W = Small sample size for certain classification tasks. High complexity of deep learning models.	Classification of benign or malignant tumour, as well as molecular subtypes

				coefficient = 0.85	O= Potential to integrate model-agnostic explainability methods, which would improve interpretability and assist in error and bias detection. T = Risks of errors or biases in predictions.	
T. Zhang et al., 2023 [19]	> 3360 paired cases > Mammography and ultrasound	> Feature extraction > Feature map > Intra-Modality Attention > Refined feature > Inter-Modality Attention	ResNet 50 with intra- and inter-modality attention modules	Accuracy = 88.5% MCC = 0.837	S = Combines multiple image views to improve classification of molecular subtypes. Employs multi-modal image analysis by integrating mammography and ultrasound. Uses intra- and inter-modality attention modules to effectively integrate features from different modalities, further boosting accuracy. W = Retrospective study design may lead to missing values. Small sample size within certain subgroups. Cases are sourced from a single centre, which may restrict the study's applicability across diverse populations. O = Potential to incorporate MRI into the multimodal models, which could further enhance the performance. T = Risks of model bias and errors.	Classification of molecular subtypes
Ma et al., 2019 [8]	> 331 patients, 662 images > Mammography	> Radiomic feature extraction > Normalization > LASSO feature selection > Oversampling technique > 10-fold cross validation	Naïve Bayes machine learning	Best performance = CC and MLO Accuracy: Triple-negative vs non-triple-negative = 0.796 HER2-enriched vs non-HER2-enriched = 0.748 Luminal vs non-luminal = 0.788	S = Combines multiple mammogram views (MLO and CC) to improve accuracy. Utilizes Naive Bayes and LASSO feature selection to enhance the prediction accuracy. W = Focuses on binary classification instead of multi-class classification. Limited dataset diversity may reduce the model's generalizability. Does not classify between luminal A and luminal B subtypes. Manual lesion segmentation. O = Develops multi-class classifier for molecular subtypes prediction. Incorporates larger and more diverse datasets. T = Less reliability due to data imbalance.	Classification molecular subtype
Liu et al., 2023 [25]	> 158 patients, 170 lesions > Mammography, MRI	> ROI labelling > ROI segmentation > Normalization	ResNet 50	Mammography: Accuracy = 53.3% Sensitivity = 66.7% Specificity = 44.4% AUC value = 0.593 (95%CI, 0.436-0.737) T2WI models = optimal	S = Multi-modal imaging approaches provide more comprehensive insights. W = Reliance on manual labelling of ROI. Only focus to Luminal A and non-Luminal A breast cancer. O = Opportunity to leverage ResNet model to classify all molecular subtypes. T = Small sample size could lead to model inaccuracies.	Classification Luminal A and non-Luminal A
Saber et al., 2021 [31]	> 322 images from MIAS > Mammography	> Noise removal > Histogram equalization > Morphological analysis > Threshold-based segmentation > Image resizing	> Inception V3 > ResNet 50 > VGG 19 > VGG 16 > Inception-V2 ResNet (Transfer learning)	Best performance = VGG 16 Accuracy = 98.96% Sensitivity = 97.83% Specificity = 99.13% Precision = 97.35%	S = Achieves high accuracy. Reduces training time and improves classification performance by automatic extracting of affected patch using segmentation. Compares many pre-trained CNN. W = The method focuses solely on benign and malignant cases, without addressing molecular subtypes critical for treatment decisions. Limited dataset diversity. O =	Classification between normal, benign and malignant

		> Data splitting > Data augmentation		F-score = 97.66% AUC = 0.995	Extends the function to classify molecular subtypes. T = Potential biases in classification outcomes as dataset lacks diversity.	
Khalid et al., 2023 [12]	> 1501 individuals, 3002 images (CC and MLO) > Mammography, kaggle.com	> Normalization and standardization > Data balancing > Feature extraction > Creating variants of photos > Noise reduction	Machine learning > Random Forest (RF) > Decision tree (DT) > k-nearest neighbors (KNN) > Logistic regression (LR) > Support vector classifier (SVC) > Linear support vector classifier (linear SVC)	Accuracy: RF = 96.49% DT = 93.86% LR = 92.98% KNN = 92.11% Linear SVC = 89.47% SVC 87.72%	S = High accuracy achieved for the majority of models. W = Classification limited to malignant and benign cases, rather than including molecular subtypes. Manual feature selection and engineering. O = Exploration of the RF model to classify molecular subtypes. T = Risk of false positives post-processing. Loss ability to discover complex patterns due to the manual engineered features.	Classification between benign and malignant
Bobowicz et al., 2023 [21]	> 789 patients, 1968 images from MUG > 1775 patients, 3744 images from CMMD > Mammography	> Data augmentation > Feature extraction > Patch features > 5-fold cross validation	> ImageNet (ResNet-18, ResNet-34, ResNet-50 and EfficientNet-B0) (Feature extraction) >Multiple instance learning-based	Best performance = ResNet-34 AUC-ROC = 0.892±0.023(-) F1-Score = 82.2±2.6(%) Accuracy = 82.2±2.6(%) Precision = 82.4±2.4(%) Recall = 82.2±2.6(%)	S = Access to two large datasets, which enhances accuracy and robustness. Integration of both CC and MLO views, leading to improved performance. Utilization of weakly supervised learning methods reduces reliance on fully annotated datasets. W = Computational complexity may increase deployment costs. O = Potential to expand classification capabilities to include molecular subtypes. T = The presence of pectoral muscle regions within image patches could introduce high-intensity pixel disturbances. Model bias due to data imbalances.	Classification between benign and malignant
M. Zhang et al., 2023) [30]	> 2321 images > Mammography cases	> Feature extraction > Data augmentation > 5-fold cross validation > Balanced sampling > Cross-entropy	Weakly supervised deep learning framework BSNet	Average AUC: Test set = 0.89 External validation set = 0.92	S = Compare BSNet model with others to prove its best performance. Weak supervision reduces dependency on extensive manual pixel-level annotations. Accessible via a web server. W = Limited comparison scope, as the model's fine-tuning performance is evaluated only on pre-trained models based on the ImageNet classification task. Detection capability is limited to breast cancer with HR status and benign tumours. O = Potential to expand classification capabilities to include molecular subtypes. Enhance model training and validation by collecting more extensive datasets. T = Robustness may be low. Lack of reliable benchmark for comparative analysis.	Detect HR status and benign
Mishra et al., 2023) [32]	> No indicated the number of images > Digital database for screening mammography (DDMS)	> Cleaning the dataset > Data exploratory analysis > Data distribution analysis	> Logistic regression > Neural network architecture with 4 hidden layers	Best performance = Neural network architecture with 4 hidden layers Train accuracy	S = Achieves high accuracy in detecting breast cancer. W = The model does not classify molecular subtypes, which are crucial for identification of treatment. Unbalanced dataset.	Classification between benign and malignant

	> Mammography	> Data encoding > Dataset splitting > Feature scaling		= 98.99% Test accuracy = 98.83%	O = Potential to extend the proposed method for classifying molecular subtypes. Incorporating more diverse and reliable datasets to improve sensitivity and overall robustness. T = Risk of bias due to dataset imbalances.	
Harish et al., 2024 [33]	> 30 different features, 570 data > CT mammogram from Kaggle	> Data preprocessing (resizing images and normalizing pixel values) > Feature localization and extraction	Convolutional neural network (CNN) equipped with a region proposal network, RPN	Accuracy = 89%	S = A faster, sequential CNN model is proposed for breast cancer classification. W = Limited dataset used, with only 50 malignant and 50 benign cases for training. Focuses only on classification of normal, malignant and benign cases. O = Potential to enhance the accuracy, precision and recall by integrating with other deep learning model without increasing time complexity. Opportunity to expand the method to classify molecular subtypes of breast cancer. T = The small dataset increases the risk of overfitting, which may affect the performance on unseen data. Potential for reduced accuracy in real-world applications.	Classification of normal, malignant and benign
Agnes et al., 2020 [34]	> 322 images from MIAS database > 4500 images after data augmentation > Mammography	> Noise and background removal > Data normalization > Data augmentation > Feature reduction	Multiscale All Convolutional Neural Network (MA-CNN)	> Accuracy = 96.47% > Sensitivity = 96% > AUC = 0.99	S = Highest in accuracy compared with others. A larger stride convolution operation is used for dimension reduction, which preserves the neighborhood pixel relationship and eliminates the need for separate pooling layer. Implements multiple dilated convolution operation to extract multiscale features for classification. Operates as an end-to-end system, eliminating the manual feature engineering from the raw images. W = Limited dataset diversity as it relies on a single dataset only, which may affect the model's generalizability. O = Incorporating patch details to enhance the accuracy of the current classifier model. Expands the model to classify molecular subtypes. Training on larger and more diverse datasets to enhance reliability. T = May result in overfitting as reliance on a single dataset.	Classification between normal, benign and malignant
Jiménez-Gaona et al., 2024 [20]	> 5892 images > BUSI, collected by Rodrigues et al., UDIAT diagnostic centre, CBIS-DDSM, mini-MIAS, Inbreast, VinDr-Mammo > Mammography and ultrasound images	> Data normalization and automatic ROI annotation > Data augmentation > ROI segmentation	> SNGAN (generate synthetic images) > ResNet-18 (image classification) > SAM model (image segmentation)	Accuracy: US = 94.7% benign and 93.6% malignant DM (I) = 80.9% benign and 76.9% malignant DM (II) = 73.7% benign and 72.3% malignant	S = Large and diverse datasets. Offers an open-source API for mobile breast image processing and classification. W = Limited output provided, only benign and malignant classification is available. Less accurate for mammography images. Impacted by data quality. O = Integration with advanced models to improve accuracy and reliability. Expands the function to classify molecular subtypes. T = Risks of data privacy. Ethical challenges in AI.	Classification between benign and malignant

Boudouh & Bouakkaz, 2024 [29]	> 1871 images from CBIS-DDSM > Mammography	> Balance collected data > Noise reduction and enhancement filters > Data augmentation > Feature extraction and classification	> Vision Transformer, ViT++ (Contextual features) > CNN based on transfer learning techniques includes Xception, VGG16, RegNetX002 (Visual features)	Best performance = ViT++ with VGG16 Accuracy = 99.22%	S = Combines ViT++ for capturing global context with CNNs for extracting local spatial details. High accuracy. Employs pre-processing techniques to reduce noise and enhance the quality of input data. Offers flexibility in model design and provides a robust feature representation. W = High computational requirements for the complex fusion approach. The feature extraction phase does not include a feature selection step. Limited datasets diversity. O = Incorporates other pre-training models in the CNN branch. Expands to diverse datasets. Extends the fusion approach proposed to classify molecular subtypes. T = High costs of implementation. Risk of redundancy and increased computational load due to irrelevant or overlapping features.	Classification breast calcifications to distinguish between benign and malignant cases
Boudouh & Bouakkaz, 2023 [24]	> No indicated the number of images > CBIS-SSSM collection > Mammography	> Noise removal > Image enhancement > Data augmentation > Feature extraction (transfer learning techniques)	> ResNet 50 V2 > Xception (Feature extractor)	Best performance = Xception Accuracy = 99.99% Recall = 100% Specificity = 100% AUC = 1.00	S = A high accuracy method proposed. Employs advanced pre-processing techniques such as Non-Local Means, Gaussian and Median filters to remove noise. W = Unbalanced dataset, which is lack of data on the Calc class. The model does not classify molecular subtypes, which are crucial for identification of treatment. O = Potential to improve accuracy by adding another dataset or hybridization of pre-trained models for feature extraction. Expand the proposed method to classify the molecular subtypes. T = The model may struggle to detect breast Calc.	Classification between breast Mass and breast Calc
Huang et al., 2022 [35]	> 322 images from MIAS > Mammography	> Image normalization and denoising > Image enhancement (sharpening) and histogram modification > Segmentation of ROI > Feature extraction	SVM (Support vector machine)	Best performance = Gaussian radial basis kernel function (RBF) Average accuracy = 78.89%	S = Employs kernel functions for optimal classification performance. The method extracts diverse features including grayscale, geometric and texture properties. W = Dependency on labelling information. The training time increases with larger datasets. Requires extensive experimentation to determine kernel function parameters. Limited dataset diversity as it relies on MIAS dataset only. O = Combines with other effective techniques can enhance classification performance. Includes larger and more diverse datasets to improve model's reliability. Opportunity to expand the method to classify molecular subtypes of breast cancer. T = Risk of biased results if critical features are mislabeled or omitted. Long training time may require.	Detection of breast calcification clusters and classification of masses
Zeng et al., 2024 [28]	> 1003 mass and 694 normal cases from DDSM	> Input embedding > multi-head attention > Addition and	ViT's MHA	Accuracy = 0.927	S = Integrates bilateral information from both the left and right breast images to improve the diagnostic accuracy of breast mass detection. Eliminates the need of manual intervention by	Classification between mass and normal tissue

	> Mammography	normalization > Feed-forward			automating the process of feature extraction and classification. Overcomes registration difficulties by adopted ViT's MHA architecture. W = Limited datasets diversity. Fails to fully exploit the symmetric region details of the images. High complexity in bilateral image integration. O = Expands to diverse datasets. Extends the method proposed to classify molecular subtypes. T = Low performance on unilateral images.	
Ueda et al., 2021 [22]	> 1448 images (data set) > 225 images (test data) > Mammography	> Ground truth labelling > Image cropping and resizing > Binary cross-entropy > Data augmentation > Five-fold cross-validation	> VGG 16 > Inception V3 > ResNet 52 >DenseNet 121 (Combination)	Best performance = Inception V3 + ResNet 52 and VGG 16 + Inception V3 AUC: ER+ or - = 0.67 (0.58-0.76) PgR+ or - =0.61 (0.53-0.68) HER2-enriched or -not enriched = 0.75 (0.68-0.82)	S = Trained model is available online with Apache License 2.0. Blending multiple models is used to improve the predictive performance. W = Insufficient number of mammograms. Used the results from the preoperative biopsy report rather than the postoperative tissue report for patients with NAC as NAC can alter receptor expressions. O = Adopted visualization technology such as gradient-weighted class activation mapping. T = Lower accuracy.	Classification of receptor expressions
Thangavel et al., 2024 [23]	> No indicated the number of images > Mammography	> Adaptive fuzzy based median filtering > Noise removal > Normalization > Image enhancement > Feature extraction	> ResNet (Hierarchical feature learning) > U-Net (Segmentation)	Accuracy = 99% Precision = 98.6% Recall = 99.01% Specificity = 98.9%	S = A high accuracy feature extraction and early-stage identification model by combining two deep learning models. W = The model not involve in the classification of molecular subtypes, which is important for identification of treatment. O = Potential to use the proposed method to classify the molecular subtypes. Integrating the model with real-time data streams. T = May face challenges related to interpretability, data dependency, computational complexity and ethical concerns regarding patient privacy and consent when it integrated into clinical practice.	Feature extraction and early-stage breast abnormality detection
Magnetic Resonance Imaging (MRI)						
Carvalho et al., 2024 [13]	> 20 DCE-MRI, 10 patients from QIN > MRI image	> Image segmentation > Image registration	> SegNet and Unet (segmentation) > Three-time-point, 3TP (classification)	Best performance = Unet (0.9799 IoU, 0.9332 Dice) Accuracy for classification = 100%	S = Fully automated tumor segmentation and classification methodology. Achieves high accuracy. Use three-time-point (3TP) approach to enhance dynamic contrast behavior. W = Small dataset may reduce generalizability. High number of parameters required for deep learning models increases complexity. Long training time. O = Incorporating larger datasets to enhance reliability. Use optimization methods to obtain the best parameters for the architecture. Expands to include other imaging modalities such as mammography. Expand classification to include molecular subtypes could enhance treatment planning.	Classification the malignancy of the tumors into malignant type III (WashOut) and malignant type II (Plateau)

					T = Risks of overfitting. Less reliability.	
Ultrasound						
Alshowarah, 2023 [18]	> 780 images from Dataset of Breast Ultrasound Images (Dataset BUSI) > 5872 images after data augmentation >Ultrasound images	> Feature extraction > Data Augmentation, Statistical operations (Avg, Min, Max and fusion of fc6 and fc7)	> VGG-19 (features extraction) > KNN, Naïve Bayes, Random Forest, Decision Tree (classifiers)	Best performance = Random Forest Accuracy = 88.63% Precision = 0.88 Recall = 0.88 F-Measure = 0.88	S = Comprehensive comparison of multiple classifiers. Exploration of multiple scenarios in the training dataset. Application of various statistical operations to generate more datasets for better learning of CNN. W = Limited dataset diversity as it relies on a single dataset only, which may affect the model's generalizability. The method focuses solely on benign and malignant cases, without addressing molecular subtypes critical for treatment decisions. Insufficient feature exploration. O = Potential to improve the classification accuracy and reduce training time by conducting more investigation and using different algorithms. Extend the proposed method to include molecular subtypes classification. Incorporating different imaging modalities such as mammography. T = Risk of bias due to class imbalances in the dataset. Limited testing in diverse clinical environments.	Classification between benign and malignant
Gheflati & Rivaz, 2022 [27]	> 780 images from Dataset of Breast Ultrasound Images (Dataset BUSI) > 163 images from dataset B >Ultrasound images	> 5-fold cross-validation > SoftMax activation function (CNN TL) > Fine-tuning, weighted cross-entropy loss (ViT) > Data augmentation	> Vision Transformer (ViT) > State-of-the-art CNN includes ResNet-50, VGG16, InceptionV3, NASNetLarge	Best performance = ViT-B/32 (BUSI+B) Accuracy = 86.7% AUC = 0.95	S = ViT architecture excels in capturing global dependencies in images through its self-attention mechanism. Comprehensive comparison of ViT and multiple CNNs. ViT adapted well to small datasets. Diverse datasets used. W = High computational requirements for ViT. O = Combines ViT with other imaging modalities such as mammography. Expands the method proposed to classify molecular subtypes. T = High costs for implementation.	Classification of benign, malignant or normal
Histopathology						
Murtaza et al., 2020 [16]	> 58 patients, 1339 images from BreakHis > Histopathology image	> Data collection > Image augmentation > Feature extraction > Feature reduction	> AlexNet (Feature extraction) > KNN, SVM, NB, DT, LDA, LR	Best performance = k-Nearest Neighbors (kNN) First-level classifier = 95.48% Second-level classifier = 94.62% (Benign subtypes) 92.45% (Malignant subtypes)	S = The BMIC_Net model proposed is less complex, computationally efficient, reliable and more accurate. Eliminates the need for handcrafted feature extraction. Employs a two level hierarchical classifier, with eight distinct output classes are identified where first level classifies images as benign and malignant, second level further classifies the subtypes of benign or malignant. Feature reduction techniques is used to improve accuracy. W = Limited dataset diversity may reduce the model's generalizability. Dataset imbalance between benign and malignant subtypes. Invasive method as histopathological images	Classification between benign and malignant, and their subtypes

					are required. O = Expand to include imaging modalities such as mammography or MRI. Expand classification to include molecular subtypes could enhance treatment planning. T = Risk of bias. Limited testing in diverse clinical environments.	
Rafiq et al., 2023 [17]	> 220,025 images from Kaggle dataset > Histopathology image	> Data augmentation > Image segmentation	> Straightforward CNN model (1-CNN) > Fusion CNN model (2-CNN) > Three CNN model (3-CNN)	Best performance = 3-CNN Accuracy = 97.90% Recall = 96.80% Precision = 96.90% f1-score = 97.20%	S = Utilizes hybrid CNNs to improve accuracy and optimize training efficiency. Utilizes a larger dataset. W = High computational complexity of fusion approach. Limited dataset diversity may reduce the model's generalizability. Relies on invasive biopsy method to obtain histopathological images. O = Expands dataset with external source. Extends the function to classify molecular subtypes. Incorporate other imaging modalities such as mammography. T = Risk of overfitting. Potential biases in classification outcomes as dataset lacks diversity.	Classification between benign and malignant

TABLE V. SUMMARY OF AI-BASED BREAST CANCER MOLECULAR SUBTYPE CLASSIFICATION STUDIES

Study	Year	Modality	Dataset Size	Model	Category	Task	Performance	Key Insight
Mota et al. [26]	2024	Mammography	1397 images	ResNet-101	CNN	Multi-class subtype	Acc: 79.02%, AUC: 64.69%	Data balancing improves performance
Nissar et al. [15]	2024	Mammography	4987 images	MobileNet V3 + CBAM	CNN	Subtype	Acc: 90%	Attention improves feature learning
T. Zhang et al. [19]	2023	Mammo + US	3360 cases	ResNet50 + Attention	CNN	Subtype	Acc: 88.5%	Multi-modal enhances accuracy
Ma et al. [8]	2019	Mammography	662 images	Naïve Bayes	ML	Binary subtype	Acc: ~79%	Limited to binary classification
Liu et al. [25]	2023	Mammo + MRI	170 lesions	ResNet50	CNN	Partial subtype	Acc: 53.3%	Small dataset limits performance
Ueda et al. [22]	2021	Mammography	1448 images	CNN ensemble	CNN	Receptor classification	AUC: up to 0.75	Limited subtype accuracy

The results presented in Table III indicate significant variability in classification performance across different models. CNN-based architectures such as MobileNet-V3 and ResNet-based models demonstrate relatively higher accuracy, particularly when enhanced with attention mechanisms or multi-modal inputs. However, these approaches often involve increased computational complexity. In contrast, traditional machine learning methods such as Naïve Bayes exhibit lower performance, highlighting their limitations in capturing complex feature representations required for multi-class subtype classification.

Furthermore, several studies report relatively low accuracy, particularly in cases with limited datasets or partial subtype classification, suggesting that data availability and model design remain critical challenges. These findings emphasize the need for more advanced architectures, such as hybrid CNN–Transformer models, which can effectively capture both local and global features to improve classification performance.

These existing models face several limitations. First, most primarily classify tumors as benign or malignant, which is insufficient for guiding precise treatment decisions tailored to breast cancer subtypes. Besides, some proposed models are overly complex and lack interpretability, making them less intuitive and harder to understand. This hinders their usability and acceptance in clinical settings. Further research is needed to develop a highly accurate yet less complex classifier for classifying breast cancer molecular subtypes using mammographic images.

Moreover, different studies employed different image processing methods. This inconsistency highlights the need

for further research to assess the effectiveness of each technique in improving segmentation and feature extraction process.

Despite these advancements, many studies face challenges related to bias, reliability and overfitting due to the small or imbalanced datasets. To address these issues, larger and more diverse datasets are essential for effectively training and testing models. In addition, data augmentation techniques, which generate new data samples from preexisting data can significantly improve model optimization and generalizability, making AI-based tools more robust and applicable in clinical settings.

IV. Conclusion

This review highlights the growing impact of artificial intelligence, in breast cancer classification using mammographic imaging. While AI-based models have shown promise in improving diagnostic accuracy and distinguishing between molecular subtypes, existing approaches still face challenges related to model complexity, interpretability, and dataset diversity. A majority of studies primarily focus on benign versus malignant classification, limiting their clinical utility in precision treatment planning. In addition, heterogeneity in preprocessing pipelines techniques underscore the need for standardized methodologies in order to enhance segmentation and feature extraction. To fully leverage AI in breast cancer diagnostics, future research should focus on developing more interpretable, highly accurate classification models while ensuring their seamless integration into clinical workflows. The adoption of larger, more diverse datasets and advanced machine learning techniques, such as Vision Transformers and hybrid CNN approaches, may potentially result in further improved classification performance. Ultimately, AI-driven solutions hold great potential in revolutionizing breast cancer detection, subtyping, and personalized treatment strategies.

Acknowledgment

The author acknowledges the support of the Ministry of Higher Education (MoHE) Malaysia through the Fundamental Research Grant Scheme (FRGS) under grant number FRGS-EC/1/2024/ICT02/UNIMAP/02/8.

References

[1] World Health Organization, “Breast cancer.” Accessed: Oct. 22, 2024. [Online]. Avail- able: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[2] World Cancer Research Fund, “Breast cancer statistics.” Accessed: Jan. 09, 2025. [On- line]. Available: https://www.wcrf.org/cancer-trends/breast-cancer-statistics/

[3] Global Cancer Observatory, “Cancer Today.” Accessed: Oct. 22, 2024. [Online]. Available: https://gco.iarc.fr/today/en

[4] C. M. Perou et al., “Molecular portraits of human breast tumours,” Nature, vol. 406, no. 6797, pp. 747–752, Aug. 2000, doi: 10.1038/35021093.

[5] “Comprehensive molecular portraits of human breast tumours,” Nature, vol. 490, no. 7418, pp. 61–70, Oct. 2012, doi: 10.1038/nature11412.

[6] J. I. Herschkowitz et al., “Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumours,” Genome Biol, vol. 8, no. 5, p. R76, May 2007, doi: 10.1186/gb-2007-8-5-r76.

[7] S. Łukasiewicz et al., “Breast Cancer—Epidemiology, Risk Factors, Classification, Prog- nostic Markers, and Current Treatment Strategies—An Updated Review,” Cancers (Basel), vol. 13, no. 17, p. 4287, Aug. 2021, doi: 10.3390/cancers13174287.

[8] W. Ma et al., “Breast Cancer Molecular Subtype Prediction by Mammographic Radiomic Features,” Acad Radiol, vol. 26, no. 2, pp. 196–201, Feb. 2019, doi: 10.1016/j.acra.2018.01.023.

[9] C. Okonkwo, “Breast MRI vs Mammogram vs Ultrasound: Which Imaging of the Breast Should You Choose?,” scan.com. Accessed: Nov. 20, 2025. [Online]. Available: https://uk.scan.com/health-hub/breast-mri-vs-mammogram-vs-ultrasound-which-imaging-of-the-breast-should-you-choose

[10] NSBSP, “Benefits and Limitations of Mammograms.” Accessed: Nov. 20, 2025. [Online]. Available: https://nsbreastscreening.ca/benefits-and-limitations-mammograms

[11] M. Nasser and U. K. Yusof, “Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction,” Diagnostics, vol. 13, no. 1, p. 161, Jan. 2023, doi: 10.3390/diagnostics13010161.

[12] A. Khalid et al., “Breast Cancer Detection and Prevention Using Machine Learning,” Diagnostics, vol. 13, no. 19, p. 3113, Oct. 2023, doi: 10.3390/diagnostics13193113.

[13] E. D. Carvalho, O. P. da Silva Neto, and A. O. de Carvalho Filho, “Deep learning-based tumor segmentation and classification in breast MRI with 3TP method,” Biomed Signal Process Control, vol. 93, p. 106199, Jul. 2024, doi: 10.1016/j.bspc.2024.106199.

[14] Cleveland Clinic, “Biopsy.” Accessed: Nov. 15, 2024. [Online]. Available: https://my.clevelandclinic.org/health/diagnostics/15458-biopsy-overview

[15] I. Nissar et al., “MOB-CBAM: A dual-channel attention-based deep learning generalizable model for breast cancer molecular subtypes prediction using mammograms,” Comput Methods Programs Biomed, vol. 248, p. 108121, May 2024, doi: 10.1016/j.cmpb.2024.108121.

[16] G. Murtaza, L. Shuib, G. Mujtaba, and G. Raza, “Breast Cancer Multi-classification through Deep Neural Network and Hierarchical Classification Approach,” Multimed Tools Appl, vol. 79, no. 21–22, pp. 15481–15511, Jun. 2020, doi: 10.1007/s11042-019-7525-4.

[17] A. Rafiq et al., “Detection and Classification of Histopathological Breast Images Using a Fusion of CNN Frameworks,” Diagnostics, vol. 13, no. 10, p. 1700, May 2023, doi: 10.3390/diagnostics13101700.

[18] S. A. Alshowarah, “Breast Cancer Detection System using Deep Learning Based on Fusion Features and Statistical Operations,” 2023. [Online]. Available: www.ijacsa.thesai.org

[19] T. Zhang et al., “Predicting breast cancer types on and beyond molecular level in a multi-modal fashion,” NPJ Breast Cancer, vol. 9, no. 1, p. 16, Mar. 2023, doi: 10.1038/s41523-023-00517-2.

[20] Y. Jiménez-Gaona et al., “BraNet: a mobile application for breast image classifica- tion based on deep learning algorithms,” Med Biol Eng Comput, vol. 62, no. 9, pp. 2737–2756, Sep. 2024, doi: 10.1007/s11517-024-03084-1.

[21] M. Bobowicz et al., “Attention-Based Deep Learning System for Classification of Breast Lesions—Multimodal, Weakly Supervised Approach,” Cancers (Basel), vol. 15, no. 10, p. 2704, May 2023, doi: 10.3390/cancers15102704.

[22] D. Ueda et al., “Training, Validation, and Test of Deep Learning Models for Classification of Receptor Expressions in Breast Cancers From Mammograms,” JCO Precis Oncol, no. 5, pp. 543–551, Nov. 2021, doi: 10.1200/PO.20.00176.

[23] Y. Thangavel, H. Garg, M. Alagarsamy, and D. Pradeep, “Revolutionizing breast cancer diagnosis . . . digital mammogram-based feature extraction,” Biomed Signal Process Control, vol. 94, p. 106268, Aug. 2024, doi: 10.1016/j.bspc.2024.106268.

[24] S. S. Boudouh and M. Bouakkaz, “Breast Abnormalities Classification Using Preprocessing and Deep Transfer Learning Techniques,” IET Conference Proc., pp. 53–58, 2023, doi: 10.1049/icp.2024.0900.

[25] M. Liu et al., “Identification of Luminal A breast cancer by using deep learn- ing analysis based on multi-modal images,” Front Oncol, vol. 13, Nov. 2023, doi: 10.3389/fonc.2023.1243126.

[26] A. M. Mota, J. Mendes, and N. Matela, "Breast Cancer Molecular Subtype Prediction: A Mammography-Based AI Approach," Biomedicines, vol. 12, no. 6, p. 1371, Jun. 2024, doi: 10.3390/biomedicines12061371.

[27] B. Gheflati and H. Rivaz, "Vision Transformers for Classification of Breast Ultra- sound Images," in 44th Annual Int. Conf. IEEE EMBC, 2022, pp. 480–483, doi: 10.1109/EMBC48229.2022.9871809.

[28] T. Zeng et al., "Vision Transformer-Based Breast Mass Diagnosis in Mam- mography Using Bilateral Information," in IICAIET, 2024, pp. 147–152, doi: 10.1109/IICAIET62352.2024.10730097.

[29] S. S. Boudouh and M. Bouakkaz, "Advancing precision in breast cancer detection: a fusion of vision transformers and CNNs," Applied Intelligence, vol. 54, no. 17–18, pp. 8170–8183, Sep. 2024, doi: 10.1007/s10489-024-05619-3.

[30] M. Zhang et al., "Developing a weakly supervised deep learning framework for breast cancer diagnosis with HR status based on mammography images," Comput Struct Biotechnol J, vol. 22, pp. 17–26, Jan. 2023, doi: 10.1016/j.csbj.2023.08.012.

[31] A. Saber et al., "A Novel Deep-Learning Model for Automatic Detection and Classifica- tion of Breast Cancer Using Transfer Learning," IEEE Access, vol. 9, pp. 71194–71209, 2021, doi: 10.1109/ACCESS.2021.3079204.

[32] P. Mishra et al., "Classification of Breast Cancer using a Novel Neural Network-based Architecture," in ICCCNT, 2023, pp. 1–6, doi: 10.1109/ICCCNT56998.2023.10306841.

[33] M. N. Harish et al., "CNN-Based Breast Cancer detection using Noninvasive methods," in ICBSII, 2024, pp. 1–5, doi: 10.1109/ICBSII61384.2024.10564089.

[34] S. A. Agnes et al., "Classification of Mammogram Images Using Multiscale All Con- volutional Neural Network (MA-CNN)," J Med Syst, vol. 44, no. 1, Jan. 2020, doi: 10.1007/s10916-019-1494-z.

[35] Y. Huang et al., "Application of Support Vector Machines for Breast Calcification Clus- ter Detection and Mass Classification," J. Phys. Conf. Ser., 2022, doi: 10.1088/1742- 6596/2400/1/012003.