INTERNATIONAL JOURNAL OF ADVANCED INNOVATIVE TECHNOLOGY IN ENGINEERING

Published by Insightive Research Pvt. Ltd. Home page: www.ijaite.co.in

Machine Learning and Deep Learning Approaches for Breast Cancer Recurrence Prediction: A Systematic Literature Review

[1]Prachi Damodhar Shahare, [2]Abha Mahalwar

[1,2]Department of Computer Science and Engineering, ISBM University, Chhattisgarh, India

[1]prachi_shahare@yahoo.com, [2]tamrakar.abha@gmail.com

Article History

Received on: 05 Dec 2023

Revised on: 28 Dec 2023

Accepted on: 15 Jan 2024

Keywords: Breast Cancer Recurrence; Machine Learning; Deep Learning; Hybrid and Ensemble Models

e-ISSN: 2455-6491

DOI:10.65809/IJAITE/24/v09i01/001

Production and hosted by
www.insightiveinc.org

ABSTRACT

Breast cancer recurrence remains a major clinical challenge despite advances in diagnosis and treatment, significantly affecting long-term survival and patient management. Accurate recurrence prediction is essential for personalized therapy planning and optimized follow-up strategies. In recent years, artificial intelligence–driven approaches, including machine learning, deep learning, hybrid, and ensemble models, have been widely explored to address limitations of traditional prognostic methods. This study presents a systematic literature review of computational models developed for breast cancer recurrence prediction, following PRISMA guidelines. Peer-reviewed studies published between 2017 and 2025 were systematically identified, screened, and categorized based on learning paradigm, data modality, prediction objective, and evaluation strategy. The review provides a structured synthesis of methodological trends, comparative performance outcomes, and clinical applicability across machine learning, deep learning, hybrid, and ensemble frameworks. Key challenges related to data heterogeneity, class imbalance, interpretability, longitudinal modeling, and external validation are critically analyzed. The findings highlight hybrid and ensemble approaches as the most promising solutions for robust recurrence prediction. Future research directions emphasize explainable, survival-aware, and clinically deployable AI models for precision oncology.

1. INTRODUCTION

Breast cancer is the most frequently diagnosed malignancy among women worldwide and remains a leading cause of cancer-related mortality despite continuous advances in screening, diagnosis, and therapeutic interventions. Improvements in early detection, surgical techniques, radiotherapy, chemotherapy, hormonal therapy, and targeted treatments have significantly increased survival rates [1]. However, a substantial proportion of patients experience disease recurrence following initial treatment, which remains a major contributor to long-term morbidity and mortality. Recurrence may manifest as local, regional, or distant metastatic disease, often associated with aggressive progression, limited treatment options, and reduced survival outcomes. Consequently, accurate and early prediction of breast cancer recurrence is a critical requirement for personalized treatment planning, optimized follow-up strategies, and improved patient prognosis [2].

Traditionally, breast cancer recurrence risk has been assessed using clinic pathological indicators such as tumour size, lymph node involvement, histological grade, hormone receptor status, HER2 expression, and molecular subtypes [3]. These variables form the basis of widely used prognostic tools and staging systems. While clinically valuable, such approaches rely on simplified assumptions and linear relationships that fail to capture the underlying biological complexity of cancer progression. Moreover, traditional statistical models particularly Cox proportional hazards regression assume proportionality and independence among predictors, which often do not hold in heterogeneous patient populations [4].

In parallel, the healthcare ecosystem has witnessed an unprecedented growth in medical data availability, driven by the widespread adoption of electronic health records (EHR), high-resolution histopathology and radiological imaging, genomic and transcriptomic profiling, and long-term survival databases [5]. These high-dimensional, multimodal datasets contain rich latent information that cannot be effectively exploited using conventional analytical techniques. As a result, artificial intelligence (AI) driven approaches, particularly machine learning (ML) and deep learning (DL), have emerged as powerful alternatives for modeling complex, nonlinear relationships inherent in breast cancer recurrence dynamics [6].

ML algorithms such as support vector machines, random forests, gradient boosting, and survival-aware models have demonstrated improved predictive performance over traditional methods. More recently, DL architectures including convolutional neural networks, recurrent neural networks, and attention-based frameworks have enabled automated feature learning from unstructured data such as histopathology whole-slide images and radiological scans. Hybrid learning and ensemble strategies further enhance predictive robustness by integrating complementary modeling paradigms and multimodal data sources [7].

The rapid proliferation of AI-based models for breast cancer recurrence prediction, several fundamental challenges persist. Existing studies are often fragmented across learning paradigms, focusing either on machine learning, deep learning, hybrid approaches, or ensemble models in isolation [8]. Many reviews emphasize specific data modalities such as imaging or clinical records without offering a comprehensive methodological synthesis across computational strategies.

Furthermore, reported performance gains are difficult to compare due to heterogeneous datasets, inconsistent evaluation metrics, varying prediction horizons, and limited external validation. Issues such as class imbalance, lack of interpretability, overfitting, and poor generalizability are frequently acknowledged but rarely analyzed systematically. As a result, clinicians and researchers lack a unified understanding of which computational approaches are most effective, under what conditions, and with what limitations [9].

Therefore, there exists a clear need for a systematic, structured, and comparative review that consolidates current knowledge on AI-driven breast cancer recurrence prediction, identifies methodological trends, and highlights unresolved research gaps.

This study addresses the aforementioned challenges by conducting a Systematic Literature Review (SLR) of breast cancer recurrence prediction models developed using ML, DL, hybrid learning, and ensemble learning approaches. Unlike prior reviews that focus narrowly on specific techniques or datasets, this work provides a holistic synthesis across computational paradigms and data modalities.

The reviewed studies are systematically categorized based on learning strategy, input data characteristics, prediction objectives, and evaluation metrics. Comparative analyses are performed to identify performance trends, interpretability considerations, and clinical applicability. By integrating findings across diverse methodologies, this SLR enables an evidence-based understanding of the current state of recurrence prediction research and clarifies the relative strengths and weaknesses of competing approaches.

The motivation for this SLR stems from three key considerations. First, breast cancer recurrence remains a life-threatening clinical problem, and inaccurate risk stratification can lead to suboptimal treatment decisions and follow-up planning. Second, the growing body of AI-driven recurrence prediction studies necessitates a critical synthesis to distinguish genuine methodological advances from dataset-specific or overfitted solutions. Third, the absence of standardized evaluation practices and interpretability frameworks hampers clinical adoption, underscoring the need for consolidated insights to guide future research and deployment.

By systematically analyzing recent advancements, this review aims to bridge the gap between computational innovation and clinical relevance, supporting the development of reliable, explainable, and deployable recurrence prediction systems.

The major contributions of this SLR are summarized as follows:

- Comprehensive categorization of breast cancer recurrence prediction studies into machine learning–based, deep learning–based, hybrid learning, and ensemble learning frameworks.

- Comparative analysis of predictive performance, data characteristics, and evaluation metrics across diverse computational approaches.

- Critical synthesis of methodological strengths and limitations, including interpretability, generalizability, and clinical applicability.
- Identification of open research challenges such as class imbalance, lack of external validation, limited longitudinal modeling, and deployment barriers.
- Future research directions emphasizing explainable AI, multimodal data fusion, federated learning, and prospective clinical validation.

The remainder of this paper is organized as follows.

Section 2 describes the background concepts and methodological foundations relevant to breast cancer recurrence prediction. Section 3 outlines the systematic review methodology, including data sources, inclusion criteria, and study categorization. Sections 4 through 7 present detailed analyses of machine learning–based, deep learning–based, hybrid learning, and ensemble learning approaches, respectively. Section 8 discusses the key challenges in breast cancer recurrence prediction. Section 9 outlines the limitations of the reviewed studies. Section 10 provides a critical discussion synthesizing insights across methodologies. Finally, Section 11 concludes the paper and highlights future research directions.

2. BACKGROUND DETAILS

Breast cancer recurrence prediction is inherently complex due to tumor heterogeneity, patient-specific biological variability, and treatment-related factors. Recurrence is influenced not only by primary tumor characteristics but also by microenvironmental interactions, immune response, and molecular signaling pathways. From a computational perspective, recurrence prediction can be framed as a classification, regression, or survival analysis problem, depending on the clinical endpoint of interest.

Machine learning approaches typically operate on structured clinical data, extracting discriminative patterns using algorithms such as logistic regression, support vector machines, decision trees, random forests, and gradient boosting techniques. These models are computationally efficient and relatively interpretable, making them attractive for early clinical adoption.

Deep learning techniques extend this capability by enabling automated feature learning from unstructured data such as histopathology images, mammograms, MRI scans, and genomic sequences. Convolutional neural networks (CNNs) dominate image-based recurrence prediction, while recurrent neural networks (RNNs) and long short-term memory (LSTM) models are increasingly used for temporal and survival modeling.

Hybrid learning approaches combine ML and DL paradigms, often integrating feature selection, optimization algorithms, or multimodal fusion strategies. Ensemble learning further enhances predictive stability by aggregating multiple base learners, reducing variance and mitigating overfitting—an essential requirement in high-stakes medical applications.

Understanding these methodological foundations is crucial for interpreting the comparative performance and limitations of recurrence prediction systems, which this review systematically evaluates.

3. METHODOLOGY

This study adopts a Systematic Literature Review (SLR) methodology to comprehensively analyze and synthesize existing research on breast cancer recurrence prediction using machine learning, deep learning, hybrid learning, and ensemble learning approaches. The review protocol is designed to ensure transparency, reproducibility, and methodological rigor, following the principles recommended by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

A. Review Protocol and Research Scope

The primary objective of this SLR is to systematically identify, categorize, and critically analyze computational models developed for breast cancer recurrence prediction. The review focuses on studies published between 2017 and 2025, a period that captures the rapid evolution of AI-driven techniques in medical prognosis. Both journal articles and peer-reviewed conference papers were considered to ensure comprehensive coverage of methodological advancements. The scope of the review encompasses:

- Machine learning–based recurrence prediction models
- Deep learning–based recurrence prediction models
- Hybrid learning frameworks combining ML and DL
- Ensemble learning strategies for recurrence risk estimation

B. Data Sources and Search Strategy

A comprehensive literature search was conducted across multiple reputable scientific databases to minimize publication bias and ensure broad

coverage. The following digital libraries were queried:

- IEEE Xplore
- SpringerLink
- ScienceDirect (Elsevier)
- PubMed
- Scopus-indexed journals
- arXiv (for recent high-impact preprints)

The search strategy employed a combination of controlled vocabulary and free-text keywords. Core search terms included:

- "Breast Cancer Recurrence Prediction",
- "Machine Learning",
- "Deep Learning",
- "Hybrid Learning",
- "Ensemble Learning",
- "Survival Analysis",
- "Artificial Intelligence".

Boolean operators (AND/OR) were used to refine the queries and retrieve relevant studies. Reference lists of selected articles were also manually screened to identify additional relevant publications.

C. Inclusion and Exclusion Criteria

To ensure relevance and quality, explicit inclusion and exclusion criteria were defined prior to study selection.

Inclusion Criteria:

Studies were included if they:

- Explicitly addressed breast cancer recurrence, relapse, or survival-related prediction.
- Employed machine learning, deep learning, hybrid, or ensemble-based computational techniques.
- Reported quantitative performance metrics such as accuracy, AUC, sensitivity, specificity, F1-score, or concordance index.
- Were published in peer-reviewed journals or reputable conference proceedings.
- Were written in English.
- Exclusion Criteria
- Studies were excluded if they:
- Focused solely on breast cancer diagnosis or detection without recurrence or prognosis analysis.
- Were review articles, editorials, or opinion papers.
- Lacked sufficient methodological details or evaluation results.
- Were non-English publications.
- Included retracted or withdrawn articles.

D. Study Selection Process

The study selection process followed a multi-stage screening procedure. Initially, all retrieved articles were screened based on titles and abstracts to remove irrelevant or duplicate records. Subsequently, full-text screening was performed to assess eligibility according to the predefined inclusion and exclusion criteria.

The selection process involved:

- Identification of potentially relevant studies through database search.
- Screening of titles and abstracts.
- Eligibility assessment through full-text review.
- Final inclusion of studies for qualitative synthesis.

This process ensures that only methodologically sound and relevant studies contribute to the final analysis. A PRISMA flow diagram illustrating the study selection procedure is recommended to be included in the final manuscript.

E. Data Extraction and Synthesis

For each selected study, relevant information was systematically extracted using a predefined data extraction template. The extracted attributes included:

- Author(s) and year of publication
- Dataset characteristics (size, modality, source)
- Computational approach (ML, DL, hybrid, ensemble)
- Model architecture and feature selection strategies
- Prediction task (binary recurrence, multi-class recurrence, survival prediction)
- Evaluation metrics and reported performance
- Key findings and limitations

The extracted data were synthesized qualitatively and summarized in comparative tables to facilitate cross-study analysis. Studies were grouped according to learning paradigm to identify performance trends and methodological differences.

F. Quality Assessment of Selected Studies

To assess the methodological quality of the included studies, a lightweight qualitative quality assessment was conducted. Each study was evaluated based on the following criteria:

- Clarity of problem formulation and clinical relevance
- Adequacy of dataset description and preprocessing
- Appropriateness of modeling techniques
- Use of validation strategies (cross-validation, independent test set)
- Reporting of evaluation metrics and interpretability analysis

Although a formal meta-analysis was not performed due to dataset heterogeneity, this quality assessment ensured that only credible and well-reported studies contributed to the synthesis.

G. Categorization of Studies

Based on the extracted information, the selected studies were categorized into four major groups:

- Machine learning–based recurrence prediction
- Deep learning–based recurrence prediction
- Hybrid learning–based recurrence prediction
- Ensemble learning–based recurrence prediction

This categorization enabled structured comparison and facilitated identification of paradigm-specific strengths, limitations, and research gaps.

H. Threats to Validity

Several potential threats to validity were identified in this SLR. First, the review is limited to published and publicly accessible studies, which may introduce publication bias. Second, variations in datasets, evaluation protocols, and prediction horizons limit direct quantitative comparison across studies. Third, the absence of standardized reporting practices may affect the consistency of extracted information.

Despite these limitations, adherence to a structured review protocol and systematic screening process mitigates bias and enhances the reliability of the synthesized findings.

A. Recurrence breast cancer prediction using Machine Learning:

Machine learning-based approaches have laid the foundation for computational breast cancer recurrence prediction. Studies consistently demonstrate that tree-based and boosting algorithms outperform linear classifiers due to their ability to model nonlinear feature interactions. Large-scale cohort studies utilizing survival analysis combined with ML further enhance prognostic reliability. Interpretability techniques such as SHAP and Bayesian networks have improved clinical trust by aligning model outputs with known prognostic factors.

Breast cancer recurrence prediction has been extensively explored using machine learning (ML) and artificial intelligence (AI) techniques, with the objective of improving prognostic accuracy and supporting clinical decision-making. Recent studies demonstrate that advanced ML models, ensemble techniques, and deep learning architectures consistently outperform traditional statistical approaches.

Noman et al. [10] developed predictive models for breast cancer recurrence and metastasis using recurrence-free survival analysis combined with ML techniques. By integrating multiple data sources, a large-scale dataset comprising 190,789 samples and 23 features was constructed. Survival analysis identified key prognostic factors, enhancing recurrence risk interpretation. Among the evaluated models, LightGBM achieved the highest performance with an AUC of 0.92, while XGBoost and Random Forest effectively distinguished recurrence types with accuracies reaching 86%.

Vadthe et al. [11] demonstrated that neural network models outperform conventional ML algorithms such as decision trees, K-nearest neighbors (KNN), and logistic regression for long-term recurrence prediction. The proposed neural network achieved a classification accuracy of 94% and an AUC of 0.98, highlighting its strong discriminative capability.

Jiang et al. [12] proposed an interpretable ML pipeline for predicting distant recurrence-free survival at 5-, 10-, and 15-year intervals. The best-performing models achieved AUC values of 0.79, 0.83, and 0.89, respectively, significantly surpassing traditional methods. The Markov Blanket and Interactive Risk Factor Learner (MBIL) reduced feature dimensionality by more than 80% without compromising accuracy. Moreover, features identified by MBIL aligned with SHAP-based importance rankings, improving model interpretability. Grid search optimization further enhanced performance by 25.3%–60%.

Lauritzen et al. [13] addressed the challenge of identifying recurrent breast cancer cases in Denmark, where 10%–30% of patients experience recurrence despite adequate treatment. Using national registry data from 79,483 patients, the proposed ML-based identification system achieved an AUC-ROC of 0.93 in the development cohort and 0.86 in the validation cohort, demonstrating robust generalization.

Li [14] evaluated several ML classifiers and reported that logistic regression (LR) and support vector machines (SVM) outperformed other methods, achieving AUC values of 0.9977 and 0.9974, respectively. Both models attained an accuracy of 97.37%, precision of 97.62%, recall of 95.35%, and an F1-score of 96.47, along with high Cohen's Kappa scores, indicating strong agreement between predicted and actual outcomes.

Mengad et al. [15] compared multiple ML algorithms and found that artificial neural networks achieved the highest recurrence prediction accuracy of 91%, followed by decision trees (90.10%) and KNN (88.20%). Logistic regression showed comparatively lower performance. The study emphasized the inclusion

of psychological and behavioral variables to enhance risk stratification.

See et al. [16] applied an XGBoost-based approach to predict recurrence in patients undergoing breast-conserving surgery. Using data from 1,518 patients, the model achieved an accuracy of 0.947 and a precision of 0.897. Surgical margin status emerged as a critical prognostic factor, while age and race were also influential predictors.

González-Castro et al. [17] demonstrated the effectiveness of ML techniques for 5-year recurrence prediction using electronic health records. Among five evaluated algorithms, XGBoost achieved the best results with a precision of 0.900, recall of 0.907, F1-score of 0.897, and AUROC of 0.807.

Shah [18] developed an AI-based model that achieved an accuracy of 92.94% in distinguishing recurrent from non-recurrent breast cancer cases. The study identified positive ovarian status, negative HER2 status, and negative estrogen receptor status as significant predictors of recurrence.

Cartron et al. [19] focused on 5-year recurrence prediction using ML models including LR, SVM, decision tree, and random forest. The random forest model achieved the best performance with an accuracy of 0.69, precision of 0.75, recall of 0.66, F1-score of 0.70, and a C-index of 0.71. SHAP analysis identified tumor size, lymph node involvement, age, tumor stage, and HER2 status as key contributors.

Zeng et al. [20] proposed AI-based models, including LSTM, XGBoost, and SVM, for post-surgical recurrence risk prediction. The LSTM model outperformed others with an accuracy of 0.89 and an AUC of 0.98, validated on a cohort of 1,841 patients.

Shankar et al. [21] introduced an ensemble-based approach for classifying recurrent and non-recurrent breast cancer. The proposed model achieved an accuracy of 0.97, recall of 0.97, F1-score of 0.969, and Cohen's Kappa of 0.9655. A soft voting classifier further improved classification accuracy to 98.24% across multiple datasets.

Fanizzi et al. [22] presented a novel CNN-based framework that transformed structured clinical data into image representations. The approach successfully predicted invasive disease events at 5- and 10-year follow-ups, achieving AUC values of 92.07% and 92.84%, respectively.

Sahoo et al. [23] proposed an ensemble learning framework combining deep neural networks, artificial neural networks, and classical ML techniques. The model achieved AUC values of 0.987 and 0.978 on the UMCIO and WPBC datasets, respectively, demonstrating superior performance in relapse prediction.

Kim et al. [24] developed ML-based models to predict high- and low-risk Oncotype DX recurrence scores. The proposed models achieved AUC values of 0.917 for high-risk and 0.744 for low-risk groups, indicating the potential of ML techniques in genomic recurrence risk stratification.

Table 1: Comparative Analysis of Machine Learning-Based Studies

Author	Dataset Size	ML Techniques	Prediction Task	Findings
Noman et al. (2025) [10]	190,789	LightGBM, XGBoost, RF	Recurrence & metastasis	AUC = 0.92 (LightGBM)
Vadthe et al. (2024) [11]	–	Neural Network	Long-term recurrence	Acc = 94%, AUC = 0.98
Jiang et al. (2025) [12]	–	Interpretable ML, MBIL	DRFS (5–15 yrs)	AUC = 0.79–0.89
Lauritzen et al. (2023) [13]	79,483	ML-based identification	Recurrence detection	AUC = 0.93 (dev), 0.86 (val)
Li, J. (2024) [14]	–	LR, SVM	Recurrence prediction	AUC ≈ 0.998
Mengad et al. (2023) [15]	–	ANN, DT, KNN	Recurrence prediction	Acc = 91% (ANN)
See et al. (2023) [16]	1,518	XGBoost	Post-BCS recurrence	Acc = 0.947
González-Castro et al. (2023) [17]	–	XGBoost	5-year recurrence	AUROC = 0.807
Shah, A. (2025) [18]	–	AI-based model	Recurrence classification	Acc = 92.94%
Cartron, M. (2022) [19]	–	RF, SVM, LR	5-year recurrence	Acc = 0.69 (RF)
Zeng et al. (2023) [20]	1,841	LightGBM, XGBoost, RF	Recurrence & metastasis	AUC = 0.92 (LightGBM)
Shankar et al. (2023) [21]	–	CNN-based AI	IDE prediction	AUC ≈ 0.93
Fanizzi et al. (2024) [22]	–	Ensemble DNN/ANN	Relapse prediction	AUC = 0.987
Sahoo et al. (2024) [23]	76 (val)	Decision Jungle, NN	Oncotype DX RS	AUC = 0.917

B. Recurrence breast cancer prediction using Deep Learning

Deep learning models have shown superior performance, particularly in image-driven recurrence prediction. CNN-based frameworks applied to histopathological WSIs and mammographic images demonstrate strong generalization across independent cohorts. Hybrid CNN–RNN architectures and survival-based deep

networks such as DeepSurv further enhance temporal modeling of recurrence risk.

However, DL models require large annotated datasets and are often criticized for limited interpretability. Recent efforts incorporating attention mechanisms, weak supervision, and explainability tools partially address these concerns, positioning DL as a powerful yet evolving solution.

Deep learning (DL) techniques have gained significant attention for breast cancer recurrence prediction due to their ability to automatically learn complex, high-dimensional representations from heterogeneous clinical, imaging, and pathological data. Recent studies demonstrate that convolutional neural networks (CNNs), recurrent neural networks (RNNs), hybrid architectures, and ensemble deep learning frameworks consistently outperform traditional machine learning models in recurrence risk stratification.

Su et al. [25] introduced Deep-BCR-Auto, a deep learning-based computational pathology framework for predicting breast cancer recurrence using hematoxylin and eosin (H&E)-stained whole slide images (WSIs). The model achieved AUROC values of 0.827 on the TCGA-BRCA dataset and 0.832 on an independent Ohio State University (OSU) cohort, demonstrating strong generalization. On the OSU dataset, the framework achieved an accuracy of 82.0%, specificity of 85.0%, and sensitivity of 67.7%, effectively stratifying patients into low- and high-risk recurrence groups.

Azman et al. [26] developed a deep neural network (DNN) model to predict distant breast cancer recurrence. Using dimensionality reduction with principal component analysis, the proposed approach achieved an accuracy of 0.80 when utilizing three principal components. The study identified critical recurrence-related factors, including patient age, tumor size, surgical procedure, molecular subtype, hormone receptor status, chemotherapy administration, and lymph node involvement.

Su et al. [27] proposed BCR-Net, a deep learning framework for recurrence risk prediction using WSIs stained with H&E and Ki-67 biomarkers. The model achieved AUC values of 0.775 and 0.811 on H&E and Ki-67 WSIs, respectively. BCR-Net demonstrated superior performance compared to existing state-of-the-art WSI classifiers while maintaining low computational complexity, making it suitable for deployment in resource-constrained clinical environments.

Fagbuagun et al. [28] explored the application of deep learning techniques for breast cancer analysis by proposing a convolutional neural network (CNN)-based framework for automated diagnosis using mammographic images. The study employed a dataset comprising 569 mammograms representing both benign and malignant breast cancer cases. Experimental results demonstrated that the proposed CNN model achieved a classification accuracy of 98.25% and a sensitivity of 99.5% after 80 training iterations, indicating its strong capability in distinguishing malignant from benign cases.

Phan et al. [29] developed a weakly supervised deep learning approach to predict recurrence risk directly from pathological WSIs without requiring region-of-interest annotations. Using 233 WSIs from 139 patients, the Xception-based model achieved an overall accuracy of 0.87 at the patch level and patient-wise accuracies of 0.90 and 1.00 for high-risk and low-risk groups, respectively.

Shi et al. [30] investigated the use of deep learning for early recurrence prediction using H&E-stained tumor images. The proposed model achieved a cross-validation accuracy of 62.4%, comparable to conventional prognostic markers such as tumor grade and estrogen receptor status. Notably, the model identified 70% of early recurrent cases among low- to intermediate-grade tumors, indicating its complementary prognostic value.

Chandran et al. [31] proposed a deep convolutional neural network (DCNN) for recurrence prediction, achieving an accuracy of 97.63%, precision of 98.57%, recall of 96.84%, and an F1-score of 97.89% on the Wisconsin Breast Cancer dataset. The results highlight the capability of deep CNN models to support clinical decision-making by accurately identifying high-risk patients.

Sankar et al. [32] employed artificial neural networks and survival analysis models for breast cancer relapse prediction. The ANN achieved an accuracy of 0.95, while the DeepSurv model significantly outperformed the traditional Cox proportional hazards model, improving the concordance index from 0.41 to 0.71.

Srivastava et al. [33] proposed a CNN-based deep learning model for automatic breast cancer prognosis, achieving an accuracy of 96.37%, sensitivity of 96.38%, and specificity of 96.35%. The model demonstrated substantial improvements over traditional ML techniques.

Comes et al. [34] utilized transfer learning on dynamic contrast-enhanced MRI (DCE-MRI) images for early recurrence prediction following neoadjuvant chemotherapy. The proposed CNN-based framework achieved accuracies of 91.7% on fine-tuning data and 85.2% on an independent test set. The inclusion of clinical variables further enhanced predictive performance.

Kalafi et al. [35] compared ML and DL models for breast cancer survival prediction using 4,902 patient records. The multilayer perceptron (MLP) achieved the highest accuracy of 88.2%, outperforming random forest, decision tree, and support vector machine models. Tumor size was identified as the most influential prognostic factor.

Prasad et al. [36] demonstrated that advanced deep learning techniques significantly enhance recurrence prediction accuracy. A hypercomplex-valued CNN achieved an accuracy of 98%, outperforming both conventional CNNs and SVM models, particularly after hyperparameter tuning.
Gupta et al. [37] evaluated multiple deep learning models for breast cancer survival prediction. The Restricted Boltzmann Machine achieved the highest accuracy of 0.97, followed by deep autoencoders (0.96) and CNNs (0.92), confirming the effectiveness of deep architectures in prognostic modeling.

Table 2: Comparative Analysis of Deep Learning-Based Studies

Author	Dataset Size	DL Techniques	Prediction Task	Findings
Su et al. (2024) [25]	TCGA-BRCA, OSU (WSI)	Deep-BCR-Auto (CNN)	Recurrence risk	AUROC = 0.832
Azman et al. (2022) [26]	Clinical data	DNN	Distant recurrence	Acc = 0.80
Su et al. (2023) [27]	H&E, Ki-67 WSIs	BCR-Net	Recurrence risk	AUC = 0.811
Fagbuagun et al. (2022) [28]	Mammogram images	CNN	569 images	Acc=98.25% ,
Phan et al. (2021) [29]	WSIs	Xception	Risk stratification	Acc = 0.90–1.00
Shi et al. (2023) [30]	H&E images	CNN	Early recurrence	Acc = 62.4%
Chandran et al. (2024) [31]	Clinical data	LR-CNN-LSTM	Recurrence prediction	Acc = 98.24%
Sankar et al. (2022) [32]	Survival data	ANN, DeepSurv	Relapse prediction	C-index = 0.71
Srivastava et al. (2024) [33]	Clinical data	CNN	Prognosis	Acc = 96.37%
Comes et al. (2021) [34]	DCE-MRI	Transfer CNN	3-year recurrence	Acc = 91.7%
Kalafi et al. (2019) [35]	Clinical data	MLP	Survival prediction	Acc = 88.2%
Prasad et al. (2023) [36]	Molecular data	Hypercomplex CNN	Relapse prediction	Acc = 98%

C. Recurrence breast cancer prediction using Hybrid Learning

Hybrid learning approaches leverage the complementary strengths of ML and DL, achieving consistently high accuracy across diverse datasets. Feature optimization techniques such as PSO, PCA, and ANOVA enhance discriminative learning, while multimodal fusion enables comprehensive risk modeling. These models demonstrate strong resilience to class imbalance and limited sample sizes, making them particularly suitable for real-world clinical datasets.
Hybrid learning approaches, which integrate machine learning (ML) and deep learning (DL) techniques or combine multiple learning paradigms, have emerged as effective solutions for breast cancer recurrence prediction. These methods leverage complementary strengths such as robust feature extraction, improved generalization, and enhanced predictive accuracy, particularly when dealing with heterogeneous clinical, imaging, and genomic data.
Chandran et al. [38] proposed hybrid architectures combining feature selection and advanced neural networks, including LR-CNN-LSTM and ANOVA-GRU models. The LR-CNN-LSTM model achieved mean accuracy, precision, recall, and F1-score values of 98.24%, 99.14%, 98.30%, and 98.14%, respectively, demonstrating superior performance over conventional approaches
Kumari et al. [39] introduced the BCR-HDL framework, a hybrid deep learning approach integrating MLP, VGG, ResNet, and Xception architectures with classical ML classifiers such as SVM, decision tree, random forest, and logistic regression. The framework generated 16 hybrid models and effectively addressed challenges related to limited data, class imbalance, and interpretability. The hybrid MLP+RF and Xception+RF models achieved a diagnostic accuracy of 97%.
Mohebian et al. [40] proposed a hybrid predictor of breast cancer recurrence (HPBCR) for 5-year recurrence estimation using clinicopathologic data from 579 patients. Statistical feature selection followed by Particle Swarm Optimization (PSO) was employed to identify discriminative attributes, which were then classified using a Bagged Decision Tree (BDT) ensemble. The selected features included age at diagnosis, tumor size, lymph node ratio, hormone receptor status, therapy type, and surgical procedure. The proposed model achieved a sensitivity of 77%, specificity of 93%, precision of 95%, and an overall accuracy of 85%, demonstrating the effectiveness of hybrid feature optimization and ensemble classification.
Edeh [41] developed a hybrid machine learning framework combining Random Tree, Logistic Regression, XGBoost, and Multilayer Perceptron classifiers for breast tumor growth prediction. The proposed system achieved an accuracy of 99.65%, significantly outperforming conventional diagnostic approaches and highlighting the potential of hybrid classifiers for early breast cancer prognosis

Othman et al. [42] introduced a multimodal hybrid deep learning model for breast cancer survival prediction using clinical, gene expression, and copy number alteration data from the METABRIC dataset. Convolutional neural networks were used for feature extraction, followed by sequence modeling using LSTM and GRU classifiers. The decision fusion of LSTM and GRU achieved the highest accuracy of 98.0%, outperforming single-modality and single-classifier models.

Rao et al. [43] proposed a hybrid recurrent neural network framework integrating convolutional layers with long short-term memory (LSTM) units for breast cancer prediction from gene expression data. Dimensionality reduction through normalization and feature selection was applied prior to learning. The proposed CNN–RNN architecture achieved an accuracy of 97.5%, surpassing traditional ML models such as SVM, Random Forest, and standalone deep neural networks.

Mishra et al. [44] presented HAXM, a novel hybrid ML model designed for breast cancer prediction. The proposed framework achieved an accuracy of 99.41% with a minimal error rate, demonstrating strong classification capability and improved patient outcome prediction.

Sajiv et al. [45] proposed a hybrid deep learning classifier for histopathological breast cancer image analysis. By integrating Multilayer Perceptron and LightGBM classifiers, the model achieved a classification accuracy of 98.28% on a dataset of 3,104 images, improving diagnostic efficiency and reducing clinical workload.

Pandey et al. [46] investigated a hybrid deep convolutional neural network combined with traditional classifiers for histopathological breast cancer classification. The integration of CNN-based feature extraction with SVM, Decision Tree, and KNN classifiers resulted in superior performance. The SVM with PCA-based features achieved the highest accuracy of 99.5%, while the Decision Tree achieved 99.4% accuracy without PCA.

Swathi [47] proposed a hybrid CNN–SVM framework for breast cancer classification using mammographic images from the CBIS-DDSM dataset. The approach addressed overfitting and class imbalance issues, achieving an accuracy of 91.7%, and demonstrated the effectiveness of combining deep feature extraction with robust ML classifiers.

Saini et al. [48] developed a hybrid deep learning framework integrating CNNs and Vision Transformers (ViTs) for breast cancer detection and classification. The model employed preprocessing, segmentation, and classification stages and achieved up to 100% accuracy on the MIAS dataset and over 99% accuracy on other benchmark datasets, while enhancing interpretability using Grad-CAM visualizations.

Raghuramaiah et al. [49] proposed BreastHybridNet, a hybrid CNN–BiLSTM architecture with spatial attention mechanisms for mammogram-based breast cancer diagnosis. The framework achieved an accuracy of 98.30%, outperforming existing hybrid models such as LMHistNet and BreastMultiNet.

Chakravarthy et al. [50] introduced a hybrid deep feature fusion approach combining VGG16, VGG19, ResNet50, and DenseNet121 for multi-label breast cancer classification. Evaluated on MIAS, CBIS-DDSM, and INbreast datasets, the proposed FHDF method achieved accuracies of 98.71%, 97.73%, and 98.83%, respectively.

Lakshminarayanan et al. [51] proposed a hybrid CNN–Random Forest framework for mammogram-based breast cancer detection and classification. The model achieved an overall accuracy of 98.6%, sensitivity of 96.5%, and specificity of 98%, demonstrating superior performance over state-of-the-art techniques.

Table 3: Comparative Analysis of Hybrid Learning-Based Studies

Author	Dataset Size	Hybrid Techniques	Prediction Task	Findings
Chandran et al. (2023) [38]	Wisconsin dataset	DCNN	Recurrence prediction	Acc = 97.63%
Kumari et al. (2025) [39]	Clinical + imaging	Hybrid DL + ML	Recurrence prediction	Acc = 97%
Mohebian et al. (2017) [40]	Clinical data (579 pts)	PSO + BDT Ensemble	5-year recurrence	Acc = 85%
Edeh, M. O. (2022) [41]	Clinical data	Multi-classifier Hybrid	Tumor growth	Acc = 99.65%
Othman et al. (2023) [42]	METABRIC (multi-omics)	CNN + LSTM/GRU Fusion	Survival prediction	Acc = 98.0%
Rao et al. (2024) [43]	Gene expression	CNN + LSTM	Cancer prediction	Acc = 97.5%
Mishra et al. (2023) [44]	Clinical data	HAXM Hybrid ML	Cancer prediction	Acc = 99.41%
Sajiv et al. (2024) [45]	Histopathology images	MLP + LightGBM	Diagnosis	Acc = 98.28%
Pandey et al. (2024) [46]	Histopathology images	CNN + SVM/DT/KNN	Classification	Acc = 99.5%
Swathi, K. (2025) [47]	Mammograms (CBIS-DDSM)	CNN + SVM	Classification	Acc = 91.7%
Saini et al. (2025) [48]	Mammograms (MIAS)	CNN + ViT	Detection & classification	Acc = 100%
Raghuramaiah et al. (2025) [49]	Mammograms	CNN + BiLSTM + Attention	Diagnosis	Acc = 98.30%

D. Recurrence breast cancer prediction using Ensemble learning

Ensemble learning techniques have been widely adopted for breast cancer recurrence prediction due to their ability to combine multiple base learners and mitigate the limitations of individual models. By aggregating predictions through bagging, boosting, or voting strategies, ensemble methods often achieve improved robustness, generalization, and sensitivity—critical factors in medical prognosis where misclassification of recurrent cases can have severe clinical consequences.

Almuhaidib et al. [52] conducted a comparative evaluation of machine learning models for breast cancer recurrence prediction, emphasizing the role of ensemble-based classifiers. Among the evaluated models, the Random Forest classifier demonstrated the best performance, achieving an accuracy of 0.6522, sensitivity of 0.6250, and specificity of 0.6593. Decision Tree models followed closely with an accuracy of 0.6261, while Naïve Bayes exhibited comparatively lower performance. The study highlighted that relying solely on accuracy may be misleading in recurrence prediction tasks and stressed the importance of sensitivity to avoid overlooking recurrent cases.

Sahoo [53] investigated ensemble learning strategies for breast cancer relapse prediction by integrating artificial neural networks (ANNs) and deep neural networks (DNNs) using weighted averaging, majority voting, and minority voting schemes. The proposed ensemble framework significantly improved predictive performance, achieving an accuracy of 96.21%, precision of 96.59%, sensitivity of 98.84%, specificity of 84.62%, and an F1-score of 97.41%. These results demonstrate the effectiveness of ensemble fusion in enhancing recall and overall diagnostic reliability.

Elshafey et al. [54] proposed a hybrid ensemble deep learning framework for breast cancer detection that combines feature extraction using a fine-tuned Xception model with temporal feature aggregation via a stacked LSTM-based regression module, followed by classification using a Support Vector Machine (SVM). Evaluated on the BreakHis dataset with extensive data augmentation, the ensemble approach improved accuracy and precision by 10.65% and 11.6%, respectively, with further gains of 3.43% and 5.22% attributed to the SVM classifier. Although primarily focused on detection, the methodology demonstrates the potential of ensemble deep learning in recurrence-related prognostic tasks.

Chandra et al. [55] introduced a hybrid ensemble learning framework leveraging a pre-trained ResNet50V2 model in combination with ensemble-based machine learning classifiers for histopathological breast cancer analysis. The proposed approach achieved an overall accuracy of 95% and showed notable improvements in precision, recall, and F1-score compared to state-of-the-art models. The ensemble architecture effectively captured complex feature representations, supporting its applicability to recurrence risk stratification.

Choudhury [56] proposed a hybrid ensemble model integrating Random Forest, Multilayer Perceptron, and Deep Belief Network classifiers. The individual models were trained independently, and their outputs were combined using a weighted averaging strategy. The ensemble achieved an accuracy of 96.5%, outperforming the standalone classifiers and demonstrating enhanced diagnostic performance. This study underscores the advantage of heterogeneous ensemble learning in improving reliability and robustness for breast cancer prognosis.

Table 4: Comparative Analysis of Ensemble Learning-Based Studies

Author	Dataset Size	ML Techniques	Prediction Task	Findings
Almuhaidib et al. (2018) [52]	Clinical data	Bagging	Random Forest, DT	Acc = 0.6522
Sahoo, G. (2023) [53]	Clinical data	Voting (Weighted /Majority/ Minority)	ANN, DNN	Acc = 96.21 %, Sens = 98.84 %
Elshafey et al. (2021) [54]	BreakHis (Histopathology)	Hybrid Ensemble	Xception + LSTM + SVM	Acc = 10.65 %
Chandra et al. (2023) [55]	Histopathology images	Hybrid Ensemble	ResNet50V2 + ML classifiers	Acc = 95%
Choudhury, Z. H. (2023) [56]	Clinical data	Weighted Averaging	RF, MLP, DBN	Acc = 96.5%

Challenges in breast cancer Recurrence prediction

Despite significant progress in applying machine learning, deep learning, hybrid, and ensemble-based approaches for breast cancer recurrence prediction, several critical challenges continue to limit their reliability, interpretability, and clinical deployment. These challenges arise from data-related constraints, methodological limitations, evaluation inconsistencies, and translational barriers between computational models and real-world oncology practice. Addressing these issues is essential to move recurrence prediction systems from experimental settings to routine clinical use.

Data Heterogeneity and Limited Standardization: One of the most fundamental challenges in breast cancer recurrence prediction is the heterogeneity of data sources. Clinical

datasets differ substantially across institutions in terms of patient demographics, diagnostic protocols, treatment regimens, follow-up durations, and outcome definitions. Variability in feature representation—such as tumor grading systems, biomarker measurement techniques, and imaging acquisition parameters—introduces distributional shifts that negatively impact model generalization.

Moreover, many studies rely on single-center or retrospective datasets, which may not capture population-level variability. The absence of standardized data collection protocols further complicates cross-study comparison and model reproducibility. As a result, models trained on one cohort often demonstrate degraded performance when evaluated on external datasets, raising concerns about their robustness in real-world clinical environments.

Class Imbalance and Rare Event Prediction: Breast cancer recurrence is a relatively rare event, particularly in early-stage or well-treated patient cohorts. This results in severe class imbalance, where non-recurrent cases significantly outnumber recurrent ones. Traditional learning algorithms tend to be biased toward majority classes, leading to inflated accuracy but poor sensitivity in identifying recurrent cases. Although several studies employ resampling techniques such as oversampling, under sampling, or synthetic data generation, these methods may introduce noise or artificial patterns that do not reflect true biological processes. In recurrence prediction, false negatives are clinically unacceptable, as missed high-risk patients may not receive intensified monitoring or adjuvant therapy. Balancing sensitivity and specificity under extreme class imbalance remains a persistent and unresolved challenge.

Limited Longitudinal and Survival-Aware Modeling: Recurrence is inherently a time-dependent clinical event, yet many studies treat it as a static binary classification problem. Such formulations fail to capture temporal dynamics, disease progression trajectories, and varying follow-up durations across patients. Survival analysis-based models and time-to-event deep learning frameworks offer more clinically meaningful outputs but are underutilized due to increased modeling complexity and data requirements.

Additionally, censored data—where recurrence status is unknown beyond a certain follow-up period—poses significant analytical challenges. Improper handling of censoring can bias recurrence risk estimation and reduce the reliability of prognostic predictions. Developing survival-aware, temporally robust models remains an open research problem.

Limited Dataset Size for Deep Learning Models: The deep learning models demonstrate strong performance in recurrence prediction, their effectiveness is highly dependent on the availability of large, well-annotated datasets. In practice, high-quality recurrence datasets with long-term follow-up are scarce, particularly for imaging and genomic modalities.

Small sample sizes increase the risk of overfitting, leading to overly optimistic performance estimates that do not translate to external cohorts. Although transfer learning, weak supervision, and data augmentation techniques partially mitigate this issue, they cannot fully compensate for the lack of diverse and representative recurrence-specific datasets.

Interpretability and Clinical Trust Deficit: One of the most frequently cited challenges in AI-driven recurrence prediction is the lack of model interpretability, especially in deep learning and ensemble frameworks. Clinicians require transparent explanations for predictions to support treatment decisions, yet many models function as “black boxes” with limited insight into feature importance or causal relationships.

Although explainable AI (XAI) techniques such as SHAP, attention mechanisms, and feature attribution maps have been introduced, their adoption remains inconsistent. Moreover, explanations are often post hoc and may not fully align with established clinical knowledge. The absence of standardized interpretability frameworks undermines clinician trust and limits regulatory acceptance.

Inconsistent Evaluation Metrics and Reporting Practices: Another major challenge is the lack of uniform evaluation standards across studies. While accuracy is frequently reported, it is often misleading in imbalanced recurrence datasets. Sensitivity, specificity, AUC, concordance index, and calibration metrics are inconsistently used, making objective comparison across models difficult.

Furthermore, many studies report only internal validation results without independent test sets or external cohort evaluation. The absence of confidence intervals, statistical significance testing, and robustness analysis further limits the credibility of reported improvements. Establishing standardized evaluation protocols is crucial for fair benchmarking and clinical relevance.

Overfitting and Lack of External Validation: A substantial proportion of recurrence prediction studies evaluate models on the same dataset used

for training or rely on cross-validation alone. While such approaches are useful for preliminary assessment, they do not guarantee real-world generalization. External validation across geographically and demographically diverse cohorts is rarely performed due to data access limitations.

This lack of external validation raises concerns about model overfitting, particularly in complex hybrid and ensemble architectures. Without rigorous validation, high reported performance may reflect dataset-specific biases rather than true predictive capability.

Integration of Multimodal and High-Dimensional Data: Modern recurrence prediction increasingly relies on multimodal data, including clinical variables, histopathology images, radiological scans, and genomic profiles. While multimodal fusion improves predictive accuracy, it introduces challenges related to feature alignment, dimensionality imbalance, missing data, and computational scalability.

High-dimensional genomic data, in particular, require aggressive feature selection or dimensionality reduction, which may discard biologically relevant information. Designing models that effectively integrate heterogeneous data while preserving interpretability and computational efficiency remains a significant challenge.

Computational Complexity and Resource Constraints: Advanced deep learning and ensemble models often require substantial computational resources for training and inference. This limits their feasibility in low-resource clinical settings and increases barriers to deployment in routine hospital workflows. Real-time recurrence risk assessment, particularly in resource-constrained healthcare systems, demands lightweight and efficient models without compromising predictive accuracy.

Translational and Clinical Deployment Barriers: Despite promising experimental results, very few recurrence prediction models are integrated into clinical decision-support systems. Barriers include regulatory approval requirements, lack of prospective clinical trials, interoperability issues with hospital information systems, and ethical concerns related to algorithmic bias.

Additionally, clinicians may be reluctant to adopt AI-based tools without clear evidence of improved patient outcomes. Bridging the gap between algorithmic performance and real-world clinical impact remains one of the most pressing challenges in this domain.

Ethical, Privacy, and Bias Considerations: Recurrence prediction models trained on historical data may inadvertently encode biases related to ethnicity, socioeconomic status, or access to care. Such biases can exacerbate healthcare disparities if deployed without careful auditing. Privacy concerns further restrict data sharing, limiting large-scale, multi-center model development.

Federated learning and privacy-preserving AI techniques offer promising solutions but introduce new challenges related to communication overhead, model convergence, and interpretability.

Limitation of the Study

Although this systematic literature review provides a comprehensive and structured synthesis of recent advances in breast cancer recurrence prediction using machine learning, deep learning, hybrid learning, and ensemble-based approaches, several limitations should be acknowledged to ensure transparent interpretation of the findings.

First, this review is inherently limited by its dependence on published literature. Only peer-reviewed journal articles and reputable conference papers were considered, which may introduce publication bias, as studies reporting negative or inconclusive results are less likely to be published. Consequently, the synthesized performance trends may overrepresent successful models and underreport methodological failures or limitations encountered during model development.

Second, the heterogeneity of datasets across included studies restricts direct quantitative comparison. The reviewed works employ diverse data sources, including clinical records, histopathology images, radiological scans, and genomic profiles, with varying sample sizes, feature definitions, follow-up durations, and recurrence endpoints. Differences in cohort characteristics, treatment protocols, and outcome definitions limit the feasibility of meta-analysis and may affect the generalizability of aggregated conclusions.

Third, inconsistencies in evaluation metrics and validation strategies pose a significant limitation. While accuracy, AUC, sensitivity, and F1-score are commonly reported, their usage is not standardized, particularly in the presence of severe class imbalance. Many studies rely on internal validation or cross-validation without independent external cohorts, which may lead to optimistic performance estimates. The lack of confidence intervals, statistical significance testing, and calibration analysis further constrains objective comparison across models.

Fourth, although this review categorizes studies by learning paradigm, it does not perform a formal meta-analytic comparison of algorithmic performance due to methodological diversity and incomplete reporting. As a result, conclusions regarding the superiority of specific approaches—such as hybrid or ensemble models—are based on qualitative synthesis rather than pooled statistical evidence.

Fifth, the review is limited in its ability to assess clinical deployment readiness. Many studies focus primarily on predictive performance and provide limited discussion of real-world implementation aspects, such as integration with clinical workflows, computational efficiency, regulatory compliance, and clinician interpretability. Consequently, the translational feasibility of several high-performing models cannot be fully evaluated within the scope of this review.

Finally, the rapid evolution of AI methodologies implies that newly emerging techniques—particularly large-scale foundation models, federated learning frameworks, and self-supervised learning approaches—may not be fully captured, especially for studies published after the review cut-off period. This temporal limitation is inherent to systematic reviews and underscores the need for periodic updates.

Despite these limitations, the structured review protocol, clear inclusion criteria, and comprehensive synthesis adopted in this study provide a reliable overview of current research trends and challenges in breast cancer recurrence prediction. The identified limitations also highlight critical areas for future research and methodological standardization.

Discussion

This systematic literature review provides a consolidated perspective on recent advancements in breast cancer recurrence prediction using machine learning, deep learning, hybrid learning, and ensemble-based approaches. The synthesis of findings across diverse studies reveals clear methodological trends, performance hierarchies, and persistent research gaps that collectively shape the current state of the field.

A key observation emerging from this review is the progressive shift from traditional machine learning models toward hybrid and ensemble learning frameworks. Early ML-based approaches demonstrate strong baseline performance, particularly when applied to structured clinical datasets, owing to their computational efficiency and relative interpretability. Tree-based and boosting models, such as Random Forest and Gradient Boosting, consistently outperform linear classifiers by capturing nonlinear feature interactions. However, their predictive capacity is often constrained when confronted with high-dimensional or unstructured data, such as histopathology images or genomic profiles.

Deep learning models, particularly convolutional neural networks and survival-aware architectures, substantially improve recurrence prediction performance by enabling automated feature extraction from complex data modalities. Image-driven DL frameworks applied to histopathology whole-slide images and radiological scans demonstrate superior discriminative capability and improved generalization across independent cohorts. Nevertheless, these models remain heavily dependent on large annotated datasets and are frequently criticized for limited interpretability and insufficient external validation. As a result, despite high reported accuracy and AUC values, their standalone clinical deployment remains challenging.

The review highlights hybrid learning approaches as a pragmatic and increasingly dominant paradigm. By integrating machine learning classifiers with deep learning–based feature extractors, hybrid frameworks effectively balance predictive performance, data efficiency, and interpretability. Feature optimization techniques and multimodal fusion strategies further enhance robustness, particularly in scenarios involving limited sample sizes and heterogeneous data sources. This explains why hybrid models consistently outperform standalone ML or DL models across multiple studies.

Similarly, ensemble learning approaches emerge as particularly effective for recurrence prediction, especially in addressing class imbalance and improving sensitivity. Ensemble fusion techniques—such as weighted voting, bagging, and boosting—reduce model variance and mitigate the risk of false negatives, which is clinically critical in recurrence risk stratification. However, the increased computational complexity and reduced transparency of ensemble systems necessitate careful design choices to ensure clinical feasibility.

Despite these advances, the discussion underscores several unresolved challenges. The lack of standardized datasets, inconsistent evaluation metrics, and limited use of external validation restrict the generalizability of reported results. Many studies emphasize accuracy while underreporting calibration, robustness, and clinical interpretability. Moreover, recurrence is inherently a time-dependent event, yet survival-aware and longitudinal modeling remains underexplored relative to static classification approaches.

This review suggests that future progress in breast cancer recurrence prediction will depend less on isolated algorithmic innovation and more on methodological integration, interpretability, and

clinical alignment. Hybrid and ensemble models, supported by explainable AI techniques and validated across multi-institutional cohorts, represent the most promising pathway toward reliable and deployable recurrence prediction systems.

4. Conclusion And Future Scope

This systematic literature review comprehensively examined recent advancements in breast cancer recurrence prediction using machine learning, deep learning, hybrid learning, and ensemble-based approaches. The synthesized evidence indicates a clear methodological progression from conventional machine learning models toward hybrid and ensemble frameworks, which consistently demonstrate superior predictive performance, robustness to data heterogeneity, and improved sensitivity in recurrence risk stratification. While deep learning models excel in extracting complex patterns from high-dimensional and unstructured data, their limitations in interpretability and data dependency remain significant barriers to clinical adoption. Hybrid and ensemble strategies emerge as the most promising solutions by effectively balancing performance, interpretability, and generalizability. Despite notable progress, challenges such as class imbalance, limited longitudinal modeling, inconsistent evaluation protocols, and lack of external validation persist. Future research should prioritize explainable and survival-aware AI models, multimodal data fusion, federated and privacy-preserving learning frameworks, and large-scale prospective clinical validation to ensure reliable, transparent, and deployable recurrence prediction systems capable of supporting personalized oncology care.

Conclusion

It can be concluded from this paper that creating a database that is totally authentic is very difficult but making a semi authentic database is comparatively easy. Asking the subjects to watch certain video and then capturing their expressions creates the semi authentic databases. Semi supervised learning techniques are useful for labeling of data. And for a system to be more effective it should be able to detect micro expressions and deal with the different angles of head.

Conflict Of Interest

The authors declare that they have no conflict of interest.

Funding Support

The author declare that they have no funding support for this study.

References

[1] Monisha, J. A., Sarathi, S., Hemavathy, V., Pandian, B., & Nightingale, F. (2023). Management of Breast Cancer: an Overview. CARDIOMETRY, 26, 589–592. https://doi.org/10.18137/cardiometry.2023.26.589592

[2] S. Hugar and S. V, "Early Detection of Breast Cancer Recurrence using Machine Learning for Improved Patient Outcomes," 2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bangalore, India, 2025, pp. 1-6, doi: 10.1109/IITCEE64140.2025.10915526

[3] Acs, G., Esposito, N., Kiluk, J. V., Laronga, C., Lee, M. C., Loftus, L., Soliman, H., Boughey, J. C., Reynolds, C., Acs, P., & Gordan, L. (2009). Estimation of Risk of Recurrence of Early Stage Estrogen Receptor Positive Breast Carcinoma by Surgical and Medical Oncologists and Pathologists Compared to the Oncotype Dx® Recurrence Score. 69, 4061. https://doi.org/10.1158/0008-5472.SABCS-09-4061

[4] Rowley, M., Garmo, H., Van Hemelrijck, M., Wulaningsih, W., Grundmark, B., Zethelius, B., Hammar, N., Walldius, G., Inoue, M., Holmberg, L., & Coolen, A. C. C. (2015). Practical survival analysis tools for heterogeneous cohorts and informative censoring. arXiv: Applications. https://arxiv.org/abs/1511.06849

[5] Zhang, H., Hussin, H., Hoh, C.-C., Cheong, S. K., Lee, W.-K., & Yahaya, B. H. (2024). Big data in breast cancer: Towards precision treatment. Digital Health, 10. https://doi.org/10.1177/20552076241293695

[6] Singh, D., Sachdeva, D., & Singh, L. (2024). Advancing Breast Cancer Drug Delivery: The Transformative Potential of Bioinformatics and Artificial Intelligence. Current Cancer Therapy Reviews. https://doi.org/10.2174/0115733947287709240229104857

[7] Renukadevi, M., & Gomathi, S. (2025). Multi-Modal Hybrid Ensemble Framework for Breast Cancer Prognosis: Integrating Advanced Feature Engineering with Stacked Meta-Learning for Enhanced Clinical Decision Support. 1166–1170. https://doi.org/10.1109/icscds65426.2025.11166857

[8] Al-Duais, M., AL-KHULAIDI, A. A. G., Mohamad, F. S., Yousef, W., Al-Fuhaidi, B., Al-Taweel, S., Mohamad, M., Husen, M. N., & Yusoff, N. (2025). Comparative Analysis of Machine Learning and Deep learning Techniques for Early Prediction of Breast Cancer. Journal of Future Artificial Intelligence and Technologies, 2(2), 242–254. https://doi.org/10.62411/faith.3048-3719-68

[9] El Haji, H., Souadka, A., Patel, B. N., Sbihi, N., Ramasamy, G., Patel, B. K., Ghogho, M., & Banerjee, I. (2023). Evolution of Breast Cancer Recurrence Risk Prediction: A Systematic Review of Statistical and Machine Learning-Based Models. JCO Clinical Cancer Informatics, 7, e2300049. https://doi.org/10.1200/cci.23.00049

[10] Noman, S. M., Fadel, Y. M., Attia, N., Tarek, M., Essam, M., ElMasarawi, S. M., Fouad, F. A., GEltasawi, E., & Al-Atabany, W. (2025). Leveraging Survival Analysis and Machine Learning for Accurate Prediction of Breast Cancer Recurrence and Metastasis. https://doi.org/10.21203/rs.3.rs-5059228/v1

[11] Vadthe, N., Reddy, R. T., Kadiyala, R., Paritala, C., Shariff, V., & Rakesh, V. (2024). Assessing the Resilience of Machine Learning Models in Predicting Long-Term Breast Cancer Recurrence Results. 416–422. https://doi.org/10.1109/icisc62624.2024.00077

[12] Jiang, X., Zhou, Y., Wells, A., & Brufsky, A. (2025). Leveraging Deep Learning, Grid Search, and Bayesian Networks to Predict Distant Recurrence of Breast Cancer. Cancers, 17(15), 2515. https://doi.org/10.3390/cancers17152515

[13] Lauritzen, A. D., Jensen, M.-B., Lillholm, M., & Knoop, A. (2023). Identifying recurrent breast cancer patients in national health registries using machine learning. Acta Oncologica, 62, 350–357. https://doi.org/10.1080/0284186X.2023.2201687

[14] Li, J. (2024). Evaluative Comparison of Machine Learning Algorithms for Precision Diagnosis in Breast Cancer. Highlights in Science Engineering and Technology, 85, 354–362. https://doi.org/10.54097/40fmfw48

[15] Mengad, A., & Amali, S. M. J. (2023). A Comparative Study of Machine Learning Techniques to Predict Types of Breast Cancer Recurrence. International Journal of Advanced Computer Science and Applications, 14(5). https://doi.org/10.14569/ijacsa.2023.0140531

[16] See, M., Tan, Q., Lai, L., Ng, J., Abd Haleem, N., Yahya, A., Lim, P., Wu, Y.-S., Ling, P., & Pai, T. (2023). A Data-Driven Approach: Investigating Prognostic Factors for Overall Survival in Breast Conserving Surgery (BCS) using Machine Learning. https://doi.org/10.21203/rs.3.rs-3671157/v1

[17] González-Castro, L., Chavez, M., Duflot, P., Bleret, V., Martin, A., Zobel, M., Nateqi, J., Lin, S., Pazos-Arias, J. J., Del Fiol, G., & López-Nores, M. (2023). Machine Learning Algorithms to Predict Breast Cancer Recurrence Using Structured and Unstructured Sources from Electronic Health Records. Cancers, 15(10), 2741. https://doi.org/10.3390/cancers15102741

[18] Shah, A. (2025). Training an Artificial Intelligence Model to Predict Breast Cancer Recurrence. Journal of Student Research, 14(1). https://doi.org/10.47611/jsrhs.v14i1.8783

[19] Cartron, M. (2022). Prediction of Breast Cancer Recurrence in Five Years using Machine Learning Techniques and SHAP. Lecture Notes in Electrical Engineering, 441–453. https://doi.org/10.1007/978-981-19-0252-9_40

[20] Zeng, L., Liu, L., Chen, D., Xue, Y., Bi, H.-L., & Yang, W. (2023). The innovative model based on artificial intelligence algorithms to predict recurrence risk of patients with postoperative breast cancer. Frontiers in Oncology, 13. https://doi.org/10.3389/fonc.2023.1117420

[21] Shankar, R. S., Chigurupati, R. S., Voosala, P., & Pilli, N. (2023). An extensible framework for recurrent breast cancer prognosis using deep learning techniques. Indonesian Journal of Electrical Engineering and Computer Science, 29(2), 931. https://doi.org/10.11591/ijeecs.v29.i2.pp931-941

[22] Fanizzi, A., Bove, S., Comes, M. C., Di Benedetto, E. F., Latorre, A., Giotta, F., Nardone, A., Rizzo, A., Soranno, C., Zito, A., & Massafra, R. (2024). Prediction of breast cancer Invasive Disease Events using transfer learning on clinical data as image-form. PLOS ONE, 19(11), e0312036. https://doi.org/10.1371/journal.pone.0312036

[23] Sahoo, G., Nayak, A. K., Tripathy, P. K., Pati, A., Panigrahi, A., Rath, A., & Moharana, B. (2024). Breast cancer relapse disease prediction improvements with ensemble learning approaches. Indonesian Journal of Electrical Engineering and Computer Science. https://doi.org/10.11591/ijeecs.v35.i1.pp335-342

[24] Kim, I., Choi, H. J., Ryu, J. M., Lee, S. K., Yu, J. H., Kim, S. W., Nam, S. J., & Lee, J. E. (2019). A predictive model for high/low risk group according to oncotype DX recurrence score using machine learning. Ejso, 45(2), 134–140. https://doi.org/10.1016/J.EJSO.2018.09.011.

[25] Su, Z., Guo, Y., Wesolowski, R., Tozbikian, G., O'Connell, N. S., Khalid, M., Niazi, K., & Gurcan, M. N. (2024). Computational Pathology for Accurate Prediction of Breast Cancer Recurrence: Development and Validation of a Deep Learning-based Tool. arXiv.Org. https://doi.org/10.48550/arxiv.2409.15491

[26] Azman, B., Hussain, S. S., Azmi, N. H. M., Ghani, M., & Norlen, N. (2022). Prediction of Distant Recurrence in Breast Cancer using a Deep Neural Network. Revista Internacional De Metodos Numericos Para Calculo Y Diseno En Ingenieria, 38(1). https://doi.org/10.23967/j.rimni.2022.03.006

[27] Su, Z., Niazi, M. K. K., Tavolara, T. E., Niu, S., Tozbikian, G., Wesolowski, R., & Gurcan, M. N. (2023). BCR-Net: A deep learning framework to predict breast cancer recurrence from histopathology images. PLOS ONE, 18(4), e0283562. https://doi.org/10.1371/journal.pone.0283562

[28] Fagbuagun, O., Folorunsho, O., Adewole, L. B., & Akin-Olayemi, T. (2022). Breast Cancer Diagnosis in Women Using Neural Networks and Deep Learning. Journal of ICT Research and Applications, 16(2), 152–166. https://doi.org/10.5614/itbj.ict.res.appl.2022.16.2.4

[29] Phan, N. N., Hsu, C.-Y., Huang, C.-C., Tseng, L.-M., & Chuang, E. Y. (2021). Prediction of Breast Cancer Recurrence Using a Deep Convolutional Neural Network Without Region-of-Interest Labeling. Frontiers in Oncology, 11, 734015. https://doi.org/10.3389/FONC.2021.734015

[30] Shi, Y., Olsson, L. T., Hoadley, K. A., Calhoun, B. C., Marron, J. S., Geradts, J., Niethammer, M., & Troester, M. A. (2023). Predicting early breast cancer recurrence from histopathological images in the Carolina Breast Cancer Study. NPJ Breast Cancer, 9. https://doi.org/10.1038/s41523-023-00597-0

[31] Chandran, A., Mary, V., & Bai, A. (2024). Breast cancer recurrence prediction with deep neural network and feature optimization. Automatika, 65, 343–360. https://doi.org/10.1080/00051144.2023.2293280

[32] Sankar, P., & Bai, R. (2022). Survival Analysis on Relapse Predictions of Breast Cancer using Optimized ANN and Deepsurv Model. 1–6. https://doi.org/10.1109/IC3SIS54991.2022.9885556

[33] P. Srivastava, R. Devrani and M. Kukreja, "Automated Diagnosis of Breast Cancer Using Deep Learning," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-6, doi: 10.1109/ICCCNT61001.2024.10725541

[34] Comes, M. C., La Forgia, D., Didonna, V., Fanizzi, A., Giotta, F., Latorre, A., Martinelli, E., Mencattini, A., Paradiso, A., Tamborra, P., Terenzio, A., Zito, A., Lorusso, V., & Massafra, R. (2021). Early prediction of breast cancer recurrence for patients treated with neoadjuvant chemotherapy: A transfer learning approach on dce-mris. Cancers, 13(10), 2298. https://doi.org/10.3390/CANCERS13102298

[35] Kalafi, E. Y., Nor, N. A. M., Taib, N. A., Ganggayah, M. D., Town, C., & Dhillon, S. K. (2019). Machine Learning and Deep Learning Approaches in Breast Cancer Survival Prediction Using Clinical Data. Folia Biologica, 65, 212–220. https://pubmed.ncbi.nlm.nih.gov/32362304

[36] Prasad, V. vara, Venkataramana, L., Keerthana, S., & Subha, R. (2023). Optimization of Deep CNN Techniques to Classify Breast Cancer and Predict Relapse. Journal of Advanced Zoology, 44(4), 774–787. https://doi.org/10.17762/jaz.v44i4.2182

[37] Gupta, S., & Gupta, M. K. (2021). A Comparative Analysis of Deep Learning Approaches for Predicting Breast Cancer Survivability. Archives of Computational Methods in Engineering, 1–17. https://doi.org/10.1007/S11831-021-09679-3

[38] Chandran, A., & Bai, M. A. (2023). Optimized Deep Convolutional Neural Network for the Prediction of Breast Cancer Recurrence. Journal of Applied Engineering and Technological Science. https://doi.org/10.37385/jaets.v5i1.3384

[39] Kumari, D., Naidu, M. V. S. S., Panda, S., & Christopher, J. (2025). Predicting breast cancer recurrence using deep learning. Discover Applied Sciences, 7, 113. https://doi.org/10.1007/s42452-025-06512-5

[40] Mohebian, M. R., Marateb, H. R., Mansourian, M., Mañanas, M. A., & Mokarian, F. (2017). A Hybrid Computer-aided-diagnosis System for Prediction of Breast Cancer Recurrence (HPBCR) using optimized ensemble learning. Computational and Structural Biotechnology Journal, 15, 75–85. https://doi.org/10.1016/J.CSBJ.2016.11.004.

[41] Edeh, M. O. (2022). A hybrid machine learning model for timely prediction of breast cancer. International Journal of Modeling, Simulation, and Scientific Computing. https://doi.org/10.1142/s1793962323410234

[42] Othman, N. A.-H., Abdel-Fattah, M. A., & Ali, A. T. (2023). A Hybrid Deep Learning Framework with Decision-Level

Fusion for Breast Cancer Survival Prediction. Big Data and Cognitive Computing, 7(1), 50. https://doi.org/10.3390/bdcc7010050

[43] Rao, U., & Katariya, Y. (2024). Breast cancer prediction from gene expression data using recurrent neural networks. 6(1), 735–738. https://doi.org/10.21917/ijdsml.2024.0150

[44] Mishra, S., Angurala, M., & Sharma, Y. (2023). A Next-Gen Approach to Breast Cancer Prediction. 1, 1–7. https://doi.org/10.1109/icaiihi57871.2023.10489356

[45] G. Sajiv, G. Ramkumar, S. Shanthi, Arunachalam Chinnathambi, Sulaiman Ali Alharbi, RETRACTED: Predicting Breast Cancer Risk from Histopathology Images using Hybrid Deep Learning Classifier, Medical Engineering & Physics, 2024, 104149, ISSN 1350-4533, https://doi.org/10.1016/j.medengphy.2024.104149

[46] Pandey, S. K., Rathore, Y., Ojha, M. K., Janghel, R. R., Sinha, A., & Kumar, A. (2024). BCCHI-HCNN: Breast Cancer Classification from Histopathological Images Using Hybrid Deep CNN Models. Deleted Journal. https://doi.org/10.1007/s10278-024-01297-2

[47] Swathi, K. (2025). Breast Cancer Classification Using CNN and SVM: A Hybrid Approach. International Journal For Science Technology And Engineering, 13(8), 889–894. https://doi.org/10.22214/ijraset.2025.73626

[48] Saini, V., Khurana, M., & Challa, R. K. (2025). Hybrid CNN–ViT Model for Breast Cancer Classification in Mammograms: A Three-Phase Deep Learning Framework. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 7(4), 977–990. https://doi.org/10.35882/jeeemi.v7i4.920

[49] Raghuramaiah, B., & Chittineni, S. (2025). BreastHybridNet: A Hybrid Deep Learning Framework for Breast Cancer Diagnosis Using Mammogram Images. International Journal of Computational and Experimental Science and Engineering, 11(1). https://doi.org/10.22399/ijcesen.812

[50] Chakravarthy, S. R. S., Bharanidharan, N., Bhatia, S., Kumar, V. V., Mahesh, T. R., Almusharraf, A., & Albalawi, E. (2024). Multi-class Breast Cancer Classification Using CNN Features Hybridization. International Journal of Computational Intelligence Systems, 17(1). https://doi.org/10.1007/s44196-024-00593-7

[51] Lakshminarayanan, K., Robinson, Y. H., Vimal, S., & Kang, D. (2021). A Hybrid Deep Learning Approach for Detecting and Classifying Breast Cancer Using Mammogram Images (pp. 851–856). Springer, Cham. https://doi.org/10.1007/978-3-030-70296-0_64

[52] Almuhaidib, D. A., Shaiba, H., Alharbi, N. G., Alotaibi, S. M., Albusayyis, F. M., Alzaid, M. A., & Almadhi, R. M. (2018). Ensemble Learning Method for the Prediction of Breast Cancer Recurrence. https://doi.org/10.1109/CAIS.2018.8442017 32

[53] Sahoo, G. (2023). Enhanced breast Cancer Relapse Prediction Based on Ensemble Learning Approaches. International Journal on Recent and Innovation Trends in Computing and Communication. https://doi.org/10.17762/ijritcc.v11i10.8619

[54] Elshafey, M. A., & Ghoniemy, T. E. (2021). A hybrid ensemble deep learning approach for reliable breast cancer detection. International Journal of Advances in Intelligent Informatics, 7(2), 112–124. https://doi.org/10.26555/IJAIN.V7I2.615

[55] Chandra, P., Reddy, S., Nithyapriya, S., Sharma, N., Maheswari, S., Jayaram, B., & Katale, T. (2023). A Novel Ensemble Deep Learning Framework for Breast Cancer Prediction. 1–6. https://doi.org/10.1109/icses60034.2023.10465347

[56] Choudhury, Z. H. (2023). Integrating Random Forest, MLP and DBN in a Hybrid Ensemble Model for Accurate Breast Cancer Detection. https://doi.org/10.31219/osf.io/sdjqf