ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jul1461

A Comprehensive Literature Survey on CNN-Based Breast Cancer Prognosis Using Multi-Modal Data

Jayachitra L.[1]; Dr. S. Subatra Devi[2]

[1] Research Scholar, Department of Computer Applications, Dr. M.G.R. Educational and Research Institute, Maduravoyal, Chennai.

[2] Professor, Department of Computer Applications, Dr. M.G.R. Educational and Research Institute, Maduravoyal, Chennai.

Publication Date: 2025/08/01

Abstract: Breast cancer remains one of the most prevalent diseases affecting women worldwide. Accurate prognosis plays a vital role in guiding treatment decisions and improving survival rates. In recent years, Convolutional Neural Networks (CNNs) have gained significant attention for their ability to automate diagnostic and prognostic tasks. This paper reviews recent CNN-based models developed for breast cancer prognosis, particularly those integrating multi-modal data such as clinical, imaging, and molecular profiles. We explore key trends in model design, data fusion strategies, and common datasets used in research. Although CNNs show promising results, challenges such as limited interpretability and poor generalization remain. To address these, we suggest future research directions involving attention-based data fusion and explainable CNN architectures, with the goal of enhancing clinical adoption and reliability.

Keywords: *Breast Cancer Prognosis, Convolutional Neural Networks (CNN), Multi-Modal Learning, Deep Learning, Medical Imaging, Attention-Based Fusion, Interpretability.*

How to Cite: Jayachitra L.; Dr. S. Subatra Devi (2025) A Comprehensive Literature Survey on CNN-Based Breast Cancer Prognosis Using Multi-Modal Data. *International Journal of Innovative Science and Research Technology*, 10(7), 2448-2452. https://doi.org/10.38124/ijisrt/25jul1461

I. INTRODUCTION

Globally, breast cancer continues to affect women at alarming rates and stands among the top contributors to cancer-related fatalities, highlighting its critical public health impact. According to the World Health Organization, approximately 2.3 million new cases and nearly 685,000 deaths were reported globally in 2020 alone [1]. Timely detection and individualized prognosis are essential for enhancing patient survival and maximizing the effectiveness of treatment strategies. Traditionally, clinicians have relied on manual analysis of histopathology slides, radiographic images, and genetic biomarkers—procedures that are inherently time-intensive and prone to subjective interpretation.

In response to these limitations, Artificial Intelligence (AI) and deep learning models have shown considerable promise. Among these, Convolutional Neural Networks (CNNs) have demonstrated robust performance in various medical domains by learning data-driven features directly from raw inputs [2], [3]. Their hierarchical structure makes them particularly effective at detecting patterns in complex data, including mammograms, histopathological images, and genomic sequences.

Recent developments have shifted toward multi-modal learning, where CNNs are used to process and integrate diverse biomedical data types. For breast cancer prognosis, this includes clinical parameters (e.g., age, tumor grade, hormone receptor status), imaging data, and molecular profiles such as gene expression and copy number alterations. Several studies have proposed multi-branch CNN architectures to accommodate these inputs, resulting in improved prognostic accuracy. However, these models often fall short in critical aspects: they rarely incorporate interpretability mechanisms, overlook modality-specific feature contributions, and are rarely validated across multiple datasets [4]– [6].

This literature survey systematically reviews CNN-based approaches for breast cancer prognosis with a focus on multi-modal data integration. Our goals are:

- To synthesize CNN methodologies applied to multi-modal breast cancer prognosis.
- To compare widely used datasets and data fusion strategies;

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jul1461

- To highlight unresolved issues such as interpretability gaps and dataset-specific bias.
- And to propose future directions involving explainable and attention-guided CNN architectures.

II. BACKGROUND: CONVOLUTIONAL NEURAL NETWORKS IN MEDICAL PROGNOSIS

Convolutional Neural Networks (CNNs) have become foundational in deep learning, particularly for tasks involving pattern recognition in visual and structured data. Initially developed for image classification and digit recognition [7], since then, CNNs have been extensively utilized in medical imaging and diagnostics because of their capability to automatically extract hierarchical features directly from raw input data. [8], [9].

In breast cancer prognosis, CNNs are used to analyze diverse biomedical inputs, including histopathological slides, mammograms, and multi-omics data. A standard CNN architecture includes convolutional layers to detect spatial patterns, pooling layers to reduce feature dimensions, and fully connected layers to perform classification or regression. This architecture enables the extraction of discriminative features, particularly those that may not be readily apparent through conventional analysis [10].

With the expansion of accessible biomedical data, CNNs are increasingly adapted for multi-modal integration, combining imaging data with structured clinical and genomic inputs. This integration, while promising, presents challenges related to data alignment, modality fusion, and model interpretability [11].

III. BENCHMARK DATASETS FOR MULTI-MODAL CNN MODELS

Selection of datasets plays a crucial role in CNN model performance, particularly when addressing issues like class imbalance, modality diversity, and generalization. This section outlines widely used datasets relevant to breast cancer prognosis. Table I summarizes key datasets used in CNN-based multi-modal prognosis.

➢ *MIAS and DDSM*

The Mammographic Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM) datasets contain grayscale mammographic images labeled by expert radiologists. While commonly used for CNN-based image classification, these datasets are limited by relatively low resolution and significant class imbalance between benign and malignant cases [12].

➢ *BreakHis*

The BreakHis dataset includes 7,909 microscopic breast tissue images captured at four magnification levels (40×, 100×, 200×, and 400×). It enables multi-scale learning of tumor morphology and has been widely employed in CNN-based histopathology analysis [13].

➢ *METABRIC*

The METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) dataset contains detailed clinical profiles, gene expression data, and copy number alterations from over 2,000 patients [14]. It is the most widely used multi-modal dataset for survival and prognosis prediction using CNNs.

➢ *The Cancer Genome Atlas (TCGA-BRCA)*

The TCGA-BRCA repository provides extensive molecular and clinical information, including RNA-Seq, DNA methylation, and protein expression data [15]. TCGA's granularity allows for modeling complex genomic patterns, but its multi-omics nature necessitates careful preprocessing and integration.

While these datasets offer powerful resources for model training, they are often dataset-specific and differ in format, patient demographics, and data modalities. This makes cross-dataset validation challenging and contributes to limited generalizability in many CNN-based studies.

IV. CNN-BASED METHODOLOGIES IN BREAST CANCER PROGNOSIS

This section highlights prominent CNN architectures designed for multi-modal breast cancer prognosis, as presented in three key studies from the literature. Each model reflects a different strategy for processing and integrating diverse data modalities.

➢ *Multi-Input CNN with Parallel Modality Streams*

Azhir et al. [4] developed a deep learning framework designed to handle multiple data sources—specifically, clinical data, gene expression, and copy number alterations—using individual CNN branches for each modality. These parallel streams allow the model to learn distinct feature representations before merging them through concatenation. Their results showed improved prognostic performance over single-modality models, highlighting the benefit of modality-specific processing.

However, the model lacks mechanisms for transparency or interpretability. While effective in terms of classification, it does not reveal which modalities or features most significantly contribute to its decisions—an important drawback in clinical environments.

➢ *Stacked Ensemble CNN with Modality-Wise Training*

The stacked ensemble approach by Zhang et al. [5] incorporates modular CNNs, each trained independently on a specific modality. The final prognosis prediction is achieved by combining their outputs using a support vector machine (SVM) classifier. This framework emphasizes modularity and strong generalization performance, achieving a higher AUC than non-ensemble baselines.

Although the model architecture shows strong generalization, it lacks attention layers or interpretability modules to explain how each input contributes to the final decision. Like many ensemble-based models, its complexity

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jul1461

can obscure transparency, making it difficult to understand or visualize the learning process.

➢ *MMDCNet: A Deep Multi-Modal Detection Model*

Wang et al. [6] introduced MMDCNet, a hybrid CNN architecture that integrates mammographic imaging features with structured clinical data. The model uses early fusion to embed both modalities into a joint representation space prior to classification. The inclusion of image and tabular features led to improvements in classification performance.

However, the authors did not explore feature-wise or modality-wise interpretability. There was also no detailed analysis on how different input types influenced the final decisions—reducing the model's potential for explainable clinical deployment.

Collectively, these studies demonstrate the emerging trend of multi-modal CNNs for breast cancer prognosis. Yet, across all three implementations, common limitations persist: reliance on basic fusion techniques, minimal use of interpretability frameworks, and lack of cross-dataset validation.

Table 1: Common Datasets for Breast Cancer Prognosis

Dataset	Modalities	Description
MIAS/DDSM	Mammograms	Early image datasets, grayscale, imbalanced
BreakHis	Histopathology images	7,909 images at 4 magnifications
METABRIC	Clinical, Gene Expression	>2,000 samples with survival labels
TCGA-BRCA	RNA-Seq, DNA, Proteomics	Comprehensive molecular profiling

V. PERFORMANCE METRICS AND EVALUATION PRACTICES

Evaluating CNN models for breast cancer prognosis requires careful consideration of both performance and clinical relevance. The following metrics are commonly used:

- Accuracy: Measures the overall proportion of correct predictions.
- Sensitivity (Recall): Reflects the model's ability to correctly identify true positives (e.g., high-risk patients).
- Specificity: Assesses its ability to correctly classify true negatives (e.g., non-cancer or low-risk cases).
- Precision and F1-Score: Particularly important in imbalanced datasets to evaluate the balance between false positives and false negatives.
- AUC-ROC: The area under the receiver operating characteristic curve captures performance independent of classification threshold.

Azhir et al. [4] reported AUC improvements from 0.81 to 0.87 with their multi-input CNN model. Zhang et al. [5] demonstrated stable performance across different metrics by leveraging ensemble diversity. Wang et al. [6] reported a 9.4% increase in accuracy when combining mammographic and clinical data over unimodal models.

Despite these reported gains, metric reporting varies widely across studies. Some use train-test splits while others apply k-fold cross-validation. Moreover, very few evaluate models on external or independent datasets, leading to concerns about reproducibility and model robustness across diverse patient populations [16].

VI. IDENTIFIED RESEARCH GAPS

Despite notable progress in applying CNNs to breast cancer prognosis, a number of unresolved challenges continue to limit clinical deployment and scientific generalization. This section outlines the most pressing limitations observed across reviewed studies.

➢ *Interpretability Deficiencies*

A significant limitation is the absence of interpretability mechanisms in most CNN-based models [17], [18]. Given their black-box nature, CNNs typically offer no insight into how individual inputs—especially across modalities—contribute to final outcomes. Without visual or analytical explanations, these models lack transparency, which can reduce clinician trust and hinder practical integration into diagnostic workflows.

➢ *Inadequate Fusion Strategies*

Many models implement data fusion by simply concatenating features from various input branches or relying on ensemble classifiers without learning dynamic modality interactions [5], [6]. These static fusion methods fail to capture per-patient relevance of clinical, genomic, or imaging data. As a result, the contribution of each modality remains unquantified and inconsistent, leaving room for improvements in personalized feature weighting.

➢ *Limited Generalizability*

The majority of CNN models are trained on single, static datasets such as METABRIC or BreakHis. While these datasets are valuable, their lack of demographic diversity and domain heterogeneity raises concerns about model overfitting and external validity [14], [19]. Very few studies conduct multi-cohort or cross-dataset evaluations, which are essential for demonstrating robustness in real-world applications.

➢ *Minimal Clinical Utility Assessment*

Although several models report high accuracy and AUC values, they rarely assess how predictions could assist clinical tasks like risk stratification, treatment planning, or follow-up scheduling. There is often a disconnect between statistical performance and actual decision-making needs of oncologists [20].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jul1461

➢ *Underuse of Emerging Modalities*

While datasets now offer molecular, proteomic, and longitudinal data, these additional modalities are seldom incorporated into CNN pipelines. The potential synergy from incorporating radiomics or time-series follow-up data is largely unexplored [21].

VII. PROPOSED RESEARCH DIRECTION

In light of these challenges, we propose designing a CNN architecture that incorporates both interpretability and attention mechanisms to enhance multi-modal breast cancer prognosis. The model would consist of distinct branches for each input type—clinical, genomic, and imaging—allowing tailored feature extraction for each data stream.

A central component of this approach would be an attention-driven fusion layer. This module would assign dynamic weights to each data modality depending on patient-specific characteristics, supporting personalized predictions and enhancing the interpretability of the model's outputs.

➢ *Integrated Explainability Tools*

Incorporation of methods such as SHAP, Grad-CAM, or attention heatmaps to visualize input importance and support clinical decision confidence.

➢ *Cross-Dataset Benchmarking*

Training and validating the model on multiple independent cohorts (e.g., METABRIC and TCGA-BRCA) to improve generalizability and reduce data bias.

By combining attention mechanisms with interpretability and external validation, this framework aims to produce reliable and clinically meaningful predictions. The proposed architecture directly builds on the structural innovations presented in prior models while addressing their critical gaps in transparency and real-world readiness.

VIII. CONCLUSION

CNN-based models have significantly advanced the field of breast cancer prognosis by enabling automated learning from complex data.

This survey reviewed how recent models incorporate multi-modal inputs and discussed their performance across key benchmarks. However, limitations around interpretability, simplistic data fusion, and dataset-specific training continue to restrict real-world adoption. Based on our analysis, future efforts should prioritize transparent model design, robust cross-dataset validation, and alignment with clinical workflows. By integrating explainability into model architecture and expanding the diversity of training data, we can move closer to developing AI tools that support meaningful decision-making in oncology.

REFERENCES

[1]. World Health Organization, "Breast cancer," *WHO*, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[2]. G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[3]. M. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[4]. S. U. Kandan, M. M. Alketbi, and Z. Al Aghbari, "Multi-input CNN: A deep learning-based approach for predicting breast cancer prognosis using multi-modal data," *Discover Data*, vol. 3, no. 2, Feb. 2025. [Online]. Available: https://link.springer.com/article/10.1007/s44248-025-00021-x

[5]. A. Maigari, Z. Zainol, and C. Xinying, "Multi-modal stacked ensemble model for breast cancer prognosis prediction," *Stat. Optim. Inf. Comput.*, vol. 13, no. 3, pp. 1013–1034, Oct. 2024. [Online]. Available: https://doi.org/10.19139/soic-2310-5070-2100

[6]. N. U. H. Shah *et al.*, "Deep multi-modal breast cancer detection network," *arXiv preprint*, arXiv:2504.16954, Apr. 2025. [Online]. Available: https://arxiv.org/abs/2504.16954

[7]. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[8]. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 3431–3440.

[9]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 1097–1105.

[10]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, arXiv:1409.1556, 2014.

[11]. S. Minaee *et al.*, "Deep learning-based mammography classification for breast cancer screening: A review," *Phys. Med.*, vol. 83, pp. 231–241, Jan. 2021.

[12]. M. Heath *et al.*, "The digital database for screening mammography," in *Proc. 5th Int. Workshop Digit. Mammogr.*, 2000, pp. 212–218.

[13]. F. Spanhol, L. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.

[14]. C. Curtis *et al.*, "The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups," *Nature*, vol. 486, no. 7403, pp. 346–352, 2012.

[15]. The Cancer Genome Atlas Network, "Comprehensive molecular portraits of human breast tumours," *Nature*, vol. 490, no. 7418, pp. 61–70, 2012.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jul1461

[16]. A. Haque, M. Neubert, and N. Demirci, “Benchmarking deep learning models for breast cancer prognosis prediction,” *IEEE Access*, vol. 9, pp. 103795–103805, 2021.
[17]. M. Tjoa and C. Guan, “A survey on explainable artificial intelligence (XAI): Toward medical XAI,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4793–4813, Nov. 2021.
[18]. R. R. Selvaraju *et al.*, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 618–626.
[19]. A. Momeni *et al.*, “Cross-dataset generalization in breast cancer classification using deep learning,” *Comput. Biol. Med.*, vol. 136, p. 104706, 2021.
[20]. A. Holzinger *et al.*, “What do we need to build explainable AI systems for the medical domain?” *arXiv preprint*, arXiv:1712.09923, 2017.
[21]. S. Saha *et al.*, “Breast cancer prognosis through the use of multi-modal classifiers: Current state of the art and the way forward,” *Brief. Funct. Genomics*, vol. 23, no. 1, pp. 1–15, 2024.