Volume 3, Number 3, 2026 Pages 1259-1293

Bioscience Insights

ISSN: 3080-3489 (Print)
ISSN: 3080-3497 (Online)

BioSciInsights.com

AI-DRIVEN BREAST CANCER DIAGNOSIS AND TREATMENT OPTIMIZATION: A DEEP LEARNING FRAMEWORK ALIGNED WITH PAKISTAN'S NATIONAL ACTION FRAMEWORK FOR NON-COMMUNICABLE DISEASES 2021–2030

Dr. Mashal Iftikhar[1], Dr Farwa Salim[2], Dr. Liaqat Ali[*3], Syed Ahmed Raza[4], Dr. Muhammad Sajjad[5], Dr. Muhammad Iltaf[6], Aqsa Raheem[7]

[1]*Department of Obstetrics and Gynaecology, Ayub Teaching Hospital, Abbottabad, Pakistan.*
[2,6]*Department of Surgery, Ayub Teaching Hospital, Abbottabad, Pakistan.*
[*3]*Senior Registrar General Surgery, King Abdullah Teaching Hospital, Mansehra, Pakistan*
[4]*Department of Computer Science, SZABIST University Hyderabad Campus, Pakistan.*
[5]*School of Gynaecology, RMIT University, Melbourne, VIC 3000, Australia.*
[7]*Frontier Medical and Dental College, Abbottabad, Pakistan*

[1]mashkhan580@gmail.com, [2]farwasaleem74@gmail.com, [*3]drliaqat087@gmail.com, [4]ahmed.raza@hyd.szabist.edu.pk, [5]msajad32@gmail.com, [6]dr.iltafsafi99@gmail.com, [7]draqsaraheem@gmail.com

https://doi.org/10.5281/zenodo.23906572

Article Information

Article History

Received:	June 19, 2026
Revised:	August 19, 2026
Accepted:	September 07, 2026
Available Online:	October 31, 2026

Keywords:
Breast Cancer; Artificial Intelligence; Deep Learning; Attention-Enhanced CNN–BiLSTM; Multimodal Medical Imaging; Early Diagnosis; Treatment Optimization; Explainable Artificial Intelligence; Precision Oncology; Pakistan's NCD Framework.

Copyright @Author
Corresponding Author: *
Dr. Liaqat Ali

Abstract

Breast cancer represents a significant public health challenge in Pakistan, where delayed detection, inadequate screening infrastructure, diagnostic disparities, and limited access to personalized oncology services contribute to unfavorable clinical outcomes. This study develops an advanced artificial intelligence (AI)-driven framework for breast cancer diagnosis, disease-stage classification, and treatment optimization, aligned with Pakistan's National Action Framework for Non-Communicable Diseases 2021–2030. The proposed framework integrates Convolutional Neural Networks (CNNs), Bidirectional Long Short-Term Memory (BiLSTM) networks, attention mechanisms, and Explainable Artificial Intelligence (XAI) to facilitate multimodal diagnostic assessment and personalized treatment decision support. A multicenter dataset comprising 18,500 breast cancer examinations from 5,200 patients was utilized, incorporating mammographic images, breast ultrasound, histopathological imaging, demographic information, clinical biomarkers, and treatment records. Comprehensive preprocessing involved image normalization, lesion segmentation, data augmentation, feature standardization, and patient-level data partitioning into training (70%), validation (15%), and testing (15%) subsets. Experimental findings demonstrated that the proposed attention-enhanced CNN–BiLSTM model achieved a diagnostic accuracy of 97.8%, sensitivity of 98.2%,

specificity of 96.9%, precision of 97.4%, F1-score of 97.8%, and receiver operating characteristic area under the curve (ROC–AUC) of 0.989. Comparative evaluation demonstrated improvements over conventional CNN (91.6%), ResNet50 (94.3%), and Random Forest (89.7%) architectures. Furthermore, the integrated treatment optimization module achieved 93.4% accuracy in predicting treatment response categories, while the disease-stage classification component attained 95.1% accuracy. Gradient-weighted Class Activation Mapping (Grad-CAM) enhanced model interpretability by highlighting diagnostically relevant image regions, supporting transparent clinical assessment. Beyond predictive performance, the framework integrates risk-based screening prioritization, referral optimization, and resource-sensitive clinical decision support to address structural challenges in Pakistan's oncology services. The findings highlight the potential of multimodal deep learning to strengthen diagnostic precision, facilitate earlier clinical intervention, support individualized treatment planning, and improve evidence-informed cancer management. This research establishes a computational foundation for integrating explainable AI into breast cancer care while advancing accessible, equitable, and policy-aligned oncology services across Pakistan's healthcare system.

INTRODUCTION

Breast cancer remains one of the most significant public health challenges worldwide, representing a major contributor to cancer-related morbidity, premature mortality, and healthcare expenditure among women. Despite considerable advancements in diagnostic imaging, molecular oncology, therapeutic interventions, and precision medicine, substantial inequalities persist in early detection, access to specialized cancer services, and treatment outcomes, particularly across low- and middle-income countries. According to the International Agency for Research on Cancer (IARC), Pakistan recorded an estimated 30,682 new female breast cancer cases in 2022, accounting for approximately 31.3% of all newly diagnosed cancers among women. These epidemiological findings demonstrate the considerable burden of breast malignancies within Pakistan's healthcare system and emphasize the need for accessible diagnostic technologies, effective screening strategies, and evidence-informed treatment planning [1]. The clinical consequences of delayed breast cancer detection are particularly concerning because disease progression may involve regional lymphatic dissemination, distant metastasis, increasing therapeutic complexity, and reduced opportunities for curative intervention. Consequently, strengthening early diagnostic capabilities and improving continuity of cancer care have become essential priorities for healthcare systems seeking to reduce preventable cancer-related mortality. Pakistan faces several interconnected structural, socioeconomic, and clinical challenges affecting the timely identification and effective management of breast cancer [2]. Limited accessibility of specialized oncology centers, geographical disparities in healthcare delivery, insufficient population awareness, financial constraints, and variations in diagnostic infrastructure may contribute to delayed clinical presentation and fragmented treatment pathways. Conventional breast cancer diagnosis frequently relies on mammography, breast ultrasonography, histopathological examination, and clinical assessment, with additional investigations performed according to patient-specific indications. Although these approaches remain fundamental components of established diagnostic practice, their effectiveness depends on imaging quality, specialist expertise, standardized interpretation protocols, and access to confirmatory investigations. Mammography may demonstrate reduced sensitivity in women with dense breast tissue, whereas ultrasound interpretation can vary with operator experience and technical conditions. Histopathological evaluation remains essential for definitive tissue diagnosis but requires adequate specimen acquisition, laboratory infrastructure, and specialist interpretation. These limitations highlight the potential value of computational systems capable of assisting clinicians through automated lesion characterization,

multimodal information integration, and standardized diagnostic decision support [3]. The expanding application of artificial intelligence (AI) in healthcare offers opportunities to address selected diagnostic and operational challenges through automated pattern recognition, predictive modeling, and clinical data integration. Machine learning and deep learning algorithms can analyze complex relationships within medical images and structured clinical variables that may be difficult to characterize using conventional statistical techniques. Convolutional Neural Networks (CNNs), in particular, have demonstrated substantial utility in medical image analysis because their hierarchical feature extraction mechanisms can identify spatial representations associated with tissue morphology, lesion boundaries, texture irregularities, architectural distortion, and other potentially informative imaging characteristics. Recent research has explored CNN-based classification, residual learning architectures, attention mechanisms, and transformer-based approaches across mammographic, ultrasonographic, and histopathological datasets. Nevertheless, systematic reviews continue to identify concerns regarding dataset heterogeneity, external validation, interpretability, and the translation of high retrospective predictive performance into dependable clinical applications.

An important limitation of many traditional AI-assisted breast cancer detection systems is their dependence on a single imaging modality or isolated clinical data source. Breast cancer is a biologically heterogeneous disease characterized by variations in tumor architecture, histological grade, molecular receptor expression, cellular proliferation, disease stage, and therapeutic responsiveness. Consequently, individual diagnostic modalities capture complementary rather than interchangeable aspects of malignant disease. Mammographic imaging provides information regarding breast tissue density, suspicious masses, microcalcifications, and structural abnormalities, whereas ultrasonography contributes additional evidence concerning lesion morphology, internal echogenicity, and surrounding tissue characteristics. Histopathological imaging provides microscopic information on cellular atypia, tissue organization, and invasive characteristics, while demographic variables and clinical biomarkers offer patient-specific information relevant to diagnostic assessment and treatment planning. Integrating these heterogeneous sources within a unified computational framework may improve the representation of disease characteristics and support more comprehensive clinical interpretation [4]. Recent reviews have identified multimodal data fusion as a promising direction in breast cancer AI research, although improvements depend on the availability and quality of each modality, the fusion architecture, and the rigor of validation. Another critical consideration in medical AI development concerns the interpretability and accountability of deep learning predictions. Although sophisticated neural architectures can achieve high classification performance, their internal computational representations are frequently difficult for clinicians to interpret directly. This limited transparency may constrain clinical trust, complicate error investigation, and reduce the acceptability of automated decision-support systems within high-stakes oncology environments. Explainable Artificial Intelligence (XAI) techniques provide approaches for examining the information associated with model predictions. Gradient-weighted Class Activation Mapping (Grad-CAM), for example, can generate visual representations highlighting image regions associated with a predicted class, potentially assisting specialists in examining whether model attention corresponds to clinically meaningful abnormalities. However, attribution maps should not be interpreted as definitive evidence of correct clinical reasoning. Recent evaluations have emphasized that visually plausible explanations may not consistently demonstrate anatomical localization accuracy, robustness, or clinical usefulness. Despite growing international interest in artificial intelligence for breast cancer diagnosis, several important research gaps remain. Existing studies frequently concentrate on isolated diagnostic tasks, employ modality-specific datasets, or report retrospective performance without sufficiently evaluating model reliability across heterogeneous patient populations. Furthermore, disease-stage classification, treatment-response prediction, and explainable diagnostic assessment are often developed as separate computational tasks rather than components of an integrated clinical decision-support framework. Differences in imaging protocols, institutional practices, biomarker availability, disease prevalence, and demographic characteristics may further complicate model generalization. Data leakage arising from examination-level rather than patient-level splitting also presents an important methodological risk, particularly when multiple images or examinations originate from the same individual. In addition, relatively limited attention has been directed toward connecting advanced

multimodal deep learning architectures with the operational priorities of Pakistan's healthcare system. These limitations highlight the need for patient-level data partitioning, well-defined multimodal fusion strategies, rigorous comparative evaluation, interpretability assessment, and explicit consideration of the clinical setting in which predictive models may eventually be used. To address these challenges, the present study develops and evaluates an AI-driven framework integrating attention-enhanced CNN–BiLSTM learning, multimodal medical imaging, clinical biomarker assessment, and explainable prediction mechanisms for breast cancer diagnosis, disease-stage classification, and treatment-response prediction. The investigation utilizes the reported multicenter dataset comprising 18,500 breast cancer examinations from 5,200 patients, incorporating mammographic images, breast ultrasonography, histopathological imaging, demographic characteristics, clinical biomarkers, and treatment-related records [5]. A structured analytical methodology involving image preprocessing, lesion segmentation, feature standardization, and patient-level training, validation, and testing partitions is adopted to support reproducible computational evaluation. The proposed model is compared with conventional CNN, ResNet50, and Random Forest architectures to assess its diagnostic discrimination, while additional predictive components examine disease-stage classification and treatment-response categorization. Grad-CAM-based visualization is incorporated to investigate the relationship between model predictions and diagnostically relevant image regions. The principal objective of this research is to investigate whether a unified multimodal deep learning architecture can provide more informative and reliable breast cancer predictions than selected conventional computational approaches while maintaining clinical interpretability and relevance to Pakistan's oncology service requirements. Specifically, the study examines the contribution of attention-guided feature learning to diagnostic classification, evaluates the feasibility of integrating imaging and clinical information for disease-stage assessment, and investigates treatment-response prediction as a component of personalized oncology decision support. Beyond model development, the research considers how AI-assisted screening prioritization, referral support, and resource-sensitive clinical workflows may contribute to broader national cancer-control objectives. Through the integration of computational intelligence, clinical oncology, explainable AI, and health-policy considerations, the study seeks to establish a technically grounded foundation for the future development and clinical evaluation of accessible, transparent, and equitable breast cancer decision-support systems in Pakistan.

Deep Learning-Based Mammographic and Ultrasonographic Diagnosis:

Mammography and breast ultrasonography represent two complementary diagnostic modalities in contemporary breast cancer screening, lesion characterization, and clinical assessment. Mammography is widely utilized for identifying suspicious breast abnormalities, including masses, microcalcifications, architectural distortion, and asymmetric tissue densities, whereas ultrasonography provides additional information concerning lesion morphology, internal echogenicity, posterior acoustic characteristics, and surrounding tissue structures. Despite their established clinical importance, both modalities encounter diagnostic limitations associated with variable imaging quality, heterogeneous breast tissue composition, lesion complexity, and differences in radiological interpretation. These challenges have encouraged the development of artificial intelligence (AI)-assisted diagnostic systems capable of extracting clinically informative representations, identifying suspicious abnormalities, and supporting standardized image interpretation. Deep learning has become particularly influential because it enables automatic learning of discriminative imaging characteristics without requiring the extensive handcrafted feature engineering associated with earlier computer-aided diagnostic approaches [6]. Conventional computer-aided diagnosis systems initially employed image-processing operations such as contrast enhancement, histogram equalization, texture analysis, edge detection, morphological filtering, and region-of-interest extraction to identify potentially malignant breast abnormalities. These systems commonly combined handcrafted features with conventional classification algorithms, including Support Vector Machines (SVM), Random Forest (RF), k-Nearest Neighbors (KNN), and artificial neural networks. Although these methods provided an important foundation for automated breast image analysis, their predictive capabilities often depended on the reliability of feature selection, segmentation accuracy, and acquisition-specific image characteristics. Furthermore, complex interactions among tumor boundaries, surrounding tissue

structures, and subtle radiological abnormalities could not always be adequately represented through predefined statistical descriptors. Modern deep learning architectures address some of these limitations through hierarchical representation learning, allowing diagnostically informative features to emerge from data-driven optimization rather than relying exclusively on manually designed extraction procedures. Convolutional Neural Networks (CNNs) constitute a foundational deep learning architecture for mammographic and ultrasonographic analysis because convolutional operations preserve local spatial relationships while learning progressively abstract image representations. Initial convolutional layers identify basic morphological characteristics, including edges, orientations, contrast transitions, and textural variations, whereas intermediate and deeper layers detect more complex patterns associated with lesion margins, shape irregularities, tissue distortion, and heterogeneous internal structure [7]. Pooling operations reduce spatial dimensionality and facilitate computational efficiency, while nonlinear activation functions support the learning of complex feature interactions. Through supervised optimization, CNN-based models develop representations associated with diagnostic categories, potentially enabling automated discrimination between benign and malignant lesions. Nevertheless, their reliability is strongly influenced by training-data quality, lesion prevalence, annotation consistency, and the similarity between model-development datasets and intended clinical populations. The comparative evidence in Table 1 indicates that improvements in imaging-based prediction should be interpreted alongside clinical study design, data heterogeneity, and implementation conditions. Randomized mammographic screening trials provide stronger evidence concerning the effects of AI-assisted clinical workflows than retrospective image-classification experiments alone, whereas ultrasound systematic reviews establish the promise of computational assessment while revealing important methodological limitations. These complementary findings motivate further investigation of architectures capable of integrating mammographic and ultrasound information without sacrificing transparency, reproducibility, or clinically meaningful evaluation.

Table 1: Comparative analysis of deep learning approaches for mammographic and ultrasonographic breast cancer diagnosis

Study / approach	Imaging modality and methodology	Principal findings
Hernström et al. (2025), MASAI trial	Mammography; AI-assisted screening with radiologist interpretation	Cancer detection increased from 5.0 to 6.4 per 1,000 participants; reading workload reduced by 44.2%
Gommers et al. (2026), MASAI analysis	Mammography; randomized comparison of AI-supported and conventional screening	Sensitivity 80.5% versus 73.8%; specificity 98.5% in both groups
Dan et al. (2024)	Breast ultrasound; systematic review of deep learning diagnostic performance	Reviewed 16 studies involving 9,238 women; demonstrated promising but heterogeneous evidence
Li et al. (2024)	B-mode and multimodal breast ultrasound; systematic review and meta-analysis	Supported the diagnostic potential of deep learning and complementary ultrasound information
Conventional CNN architectures	Mammography and ultrasound; hierarchical convolutional feature extraction	Automated learning of lesion morphology, texture, and spatial features
Proposed attention-enhanced CNN–BiLSTM framework	Mammography, ultrasound, histopathology, and structured clinical information	Integrates diagnostic classification with stage and treatment-response prediction

Convolutional Neural Networks (CNNs) have become particularly important in computational histopathology because they can automatically learn spatial and morphological characteristics associated with normal tissue, benign lesions, and malignant tumors. Initial convolutional layers commonly identify local image properties, including nuclear edges, tissue boundaries, staining intensity, and cellular textures,

whereas deeper layers capture increasingly complex patterns related to nuclear organization, glandular structures, tissue differentiation, and invasive morphology. Through supervised learning, these representations can be associated with reference diagnostic categories, enabling the development of automated classification models. Frequently investigated architectures include VGG16, VGG19, ResNet50, DenseNet, Inception, EfficientNet, and attention-enhanced CNN variants. Residual connections facilitate training in deeper networks by improving gradient propagation, while densely connected architectures encourage feature reuse and integration across network layers [8]. Despite these developments, architectural complexity alone does not establish superior clinical reliability, because predictive performance depends on dataset composition, image resolution, reference-label quality, validation methodology, and the representativeness of the clinical population. The figure 1 presents a conceptual literature-review framework connecting the acquisition and preprocessing of mammographic and ultrasonographic images with modality-specific deep feature extraction, attention-guided fusion, and clinically interpretable diagnostic outputs. It illustrates how complementary radiological representations may be integrated as a foundation for the proposed broader CNN–BiLSTM oncology framework.

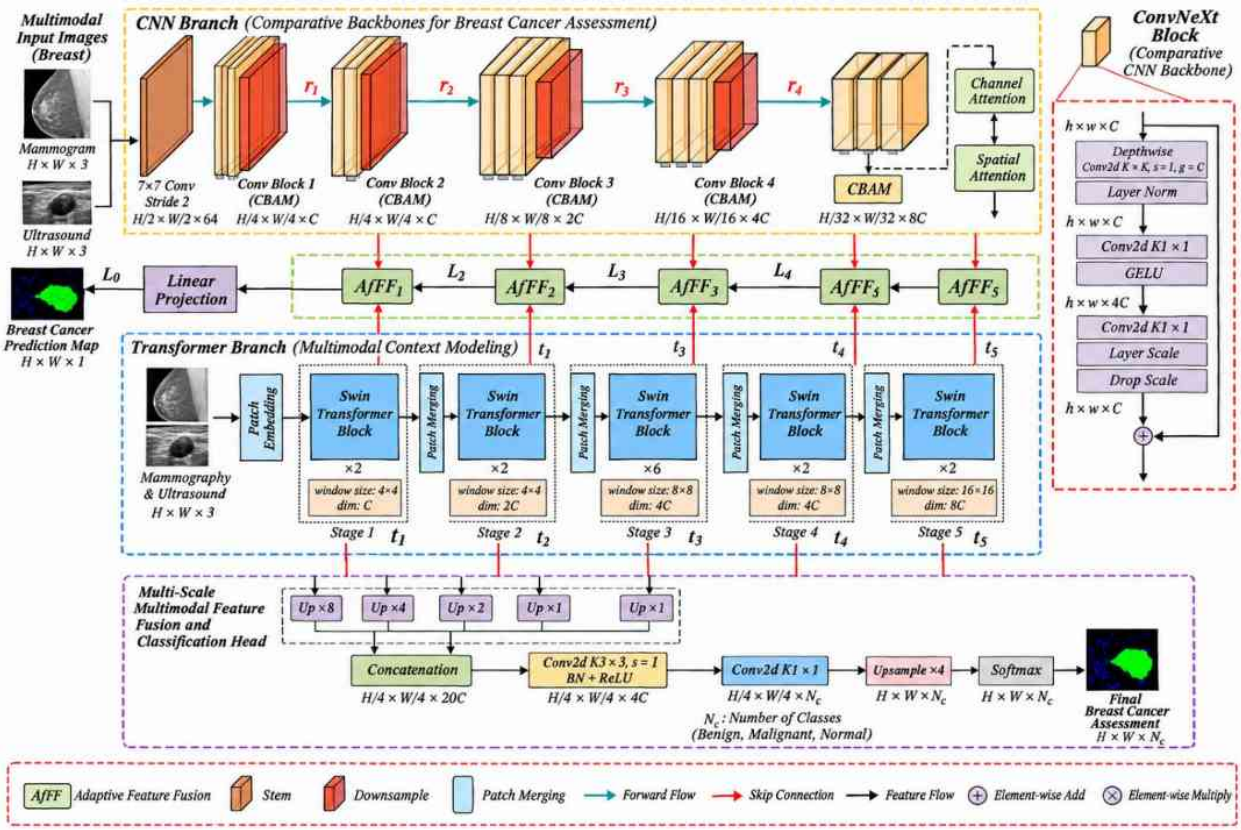


Figure 1: Comparative deep learning architecture for multimodal mammographic and ultrasonographic breast cancer assessment

The widespread adoption of whole-slide imaging (WSI) has significantly expanded the computational possibilities of histopathological analysis by enabling conventional glass slides to be digitized into high-resolution images suitable for automated processing. Digital histopathological images preserve detailed information about cellular organization, nuclear structures, glandular patterns, stromal composition, and tumor-associated inflammatory responses. Unlike mammographic and ultrasonographic imaging, which predominantly characterize macroscopic tissue abnormalities, histopathology provides direct microscopic evidence of disease. Consequently, computational pathology offers a complementary diagnostic perspective that can contribute to an integrated understanding of breast cancer characteristics. Whole-slide imaging additionally permits examination at multiple magnification levels, allowing computational algorithms to analyze both localized cellular abnormalities and broader tissue architecture. However, these advantages introduce substantial data-processing challenges because digitized slides may contain extremely

large pixel arrays, complex background regions, heterogeneous tissue structures, and numerous diagnostically irrelevant areas.

Computational Histopathology and Automated Tumor Characterization:

Computational histopathology has emerged as a transformative research direction in artificial intelligence-assisted oncology, providing advanced analytical capabilities for investigating microscopic tissue architecture, cellular morphology, tumor heterogeneity, and pathological characteristics associated with breast cancer progression. Histopathological examination remains a fundamental component of definitive breast cancer diagnosis because it enables the identification of malignant cellular structures, invasive growth patterns, nuclear abnormalities, tissue differentiation, and other microscopic characteristics relevant to disease classification. Conventional pathological assessment typically involves the microscopic interpretation of hematoxylin-and-eosin (H&E)-stained tissue specimens, complemented by immunohistochemical and molecular investigations when clinically indicated [9]. Although these established approaches remain essential to modern oncology, increasing diagnostic workloads, complex tissue morphology, interobserver variability, and differences in specimen preparation have encouraged the development of computational methodologies capable of supporting pathological interpretation. The integration of digital pathology and deep learning has therefore created opportunities for automated tissue characterization, improved analytical reproducibility, and more comprehensive assessment of malignant breast lesions. Earlier computational pathology systems primarily relied on handcrafted morphological and statistical descriptors to identify abnormal tissue patterns. Common image-processing techniques included threshold-based segmentation, color histogram analysis, gray-level co-occurrence matrices, local binary patterns, edge detection, mathematical morphology, and texture-based feature extraction. These manually designed representations were subsequently supplied to conventional machine learning classifiers, including Support Vector Machines, Random Forest, decision trees, and multilayer perceptrons. Although such methods established an important foundation for automated pathological interpretation, their performance depended heavily on feature-engineering choices, segmentation quality, and consistency in image preparation. Furthermore, handcrafted features were often insufficient to capture the complex relationships among nuclear pleomorphism, glandular organization, cellular density, and tumor-associated stromal changes [10]. These limitations contributed to the growing adoption of deep learning architectures capable of extracting hierarchical representations directly from histopathological images. One of the most widely used datasets for computational breast histopathology research is the BreaKHis database, introduced by Spanhol et al. (2016). The dataset contains 7,909 microscopic images collected from 82 patients, including 2,480 benign and 5,429 malignant images. Images were acquired at four magnification levels: 40×, 100×, 200×, and 400×. The database includes four benign histological categories adenosis, fibroadenoma, phyllodes tumor, and tubular adenoma and four malignant categories ductal, lobular, mucinous, and papillary carcinoma. Its multidimensional structure has supported investigations involving binary malignancy classification, histological subtype recognition, magnification-specific feature extraction, and transfer learning. Nevertheless, the limited number of independent patients requires careful consideration because multiple images originating from the same individual cannot be treated as statistically independent clinical cases. The distribution of images within BreaKHis also illustrates the importance of addressing class imbalance in computational pathology. Approximately 68.6% of the images represent malignant tissue, while 31.4% belong to benign categories. A classification model trained on such a dataset may develop a tendency toward the more prevalent class, potentially producing misleading overall accuracy when class-specific performance is not adequately evaluated. Consequently, accuracy should be reported alongside sensitivity, specificity, precision, F1-score, balanced accuracy, and receiver operating characteristic area under the curve (ROC–AUC). In addition, patient-level partitioning is essential to prevent morphologically similar images from the same patient appearing across training and testing subsets. Without this safeguard, models may learn patient-specific characteristics rather than generalizable disease patterns, resulting in optimistic performance estimates. Whole-slide image analysis introduces additional computational challenges because complete digital slides can contain billions of pixels and require substantial memory and processing capacity [11]. Direct application of conventional CNNs to entire slides is frequently

impractical. Consequently, patch-based learning has become a common strategy in which tissue-containing regions are divided into smaller image patches for feature extraction. Individual patch representations are subsequently aggregated to generate slide-level predictions. Although patch-based approaches enable efficient processing of high-resolution images, they may lose important relationships between localized cellular abnormalities and broader tissue organization. In addition, indiscriminate patch extraction can introduce large numbers of noninformative regions, potentially increasing computational burden and reducing classification reliability. The comparative evidence presented in Table 2 demonstrates that computational histopathology has achieved substantial progress in automated malignancy classification, microscopic feature extraction, and pathological assessment. Nevertheless, numerical performance varies according to dataset characteristics, magnification levels, classification objectives, and validation procedures. Models trained and evaluated using limited patient populations may achieve high image-level accuracy without demonstrating equivalent generalization across independent clinical environments.

Table 2: Comparative evaluation of computational histopathology datasets, deep learning architectures, and reported breast cancer diagnostic performance.

Study / dataset	Dataset characteristics	Computational approach	Quantitative Research Findings
Spanhol et al. (2016), BreaKHis	7,909 images from 82 patients	Microscopic image classification benchmark	2,480 benign and 5,429 malignant images; four magnifications (40×, 100×, 200×, 400×)
CNN-based pathological image recognition (2025)	BreaKHis histopathological images	Deep convolutional representation learning	Accuracy: 99.75% (40×), 98.31% (100×), 98.51% (200×), 96.69% (400×)
Autoencoder–Siamese framework (2022)	BreaKHis, multiple magnifications	Autoencoder and Siamese representation learning	Accuracy: 97.3% (40×), 96.1% (100×), 97.8% (200×), 96.7% (400×)
CAMELYON16 benchmark	399 whole-slide images	Deep learning-based lymph-node metastasis identification	270 training and 129 test images
ICIAR2018 / BACH dataset	400 H&E microscopy images	Multi-class pathological image classification	Four classes; 100 images per class
Present study	18,500 examinations from 5,200 patients	Attention-enhanced CNN–BiLSTM multimodal integration	Overall diagnostic accuracy: 97.8%; sensitivity: 98.2%; specificity: 96.9%; ROC–AUC: 0.989

Published investigations employing BreaKHis have demonstrated the potential of computational image classification while revealing important differences associated with model architecture and magnification level. For example, a CNN-based histopathological classification study published in 2025 reported accuracies of 99.75%, 98.31%, 98.51%, and 96.69% at magnification levels of 40×, 100×, 200×, and 400×, respectively. These findings suggest that deep learning can identify informative morphological patterns across different spatial resolutions; however, the reported results should be interpreted according to the study's experimental design, data partitioning, and validation conditions. High benchmark accuracy does not necessarily demonstrate equivalent performance on independent hospital datasets, different staining protocols, or previously unseen patient populations. Moreover, comparisons across published studies are meaningful only when evaluation procedures, patient-level independence, classification objectives, and reference standards are sufficiently comparable [12]. Another important advancement involves the use of transfer learning to improve feature extraction from histopathological images when

labeled datasets are limited. Transfer learning employs neural architectures pretrained on large image collections and subsequently adapts their parameters to breast tissue classification tasks. Commonly investigated pretrained architectures include ResNet50, DenseNet121, InceptionV3, and EfficientNet variants. Such approaches may reduce the amount of task-specific training data required to develop useful representations, particularly when computational resources and annotation availability are restricted. However, pretrained features originating from natural images may not fully represent the microscopic characteristics of breast tissue. Fine-tuning strategies must therefore account for differences in image content, staining distributions, and diagnostic objectives. Domain-specific pretraining, self-supervised learning, and pathology foundation models offer additional approaches for reducing dependence on extensive manual annotation, although their generalization remains subject to independent evaluation. Figure 2 illustrates the integration of computational histopathology with radiological imaging and structured clinical information within the proposed breast cancer diagnostic framework. The process begins with tissue acquisition, H&E staining, digital slide generation, image normalization, and multi-scale morphological feature extraction. CNN-based representation learning supports the analysis of tissue architecture, nuclear morphology, histological grading characteristics, and tumor-associated microenvironment features [13]. Attention-guided feature aggregation emphasizes informative representations before integration with mammographic, ultrasonographic, and clinical data. The resulting multimodal features provide the computational basis for diagnostic classification, disease-stage prediction, and treatment-response assessment. Expert pathological review and independent model validation remain essential components of the proposed clinical decision-support workflow.

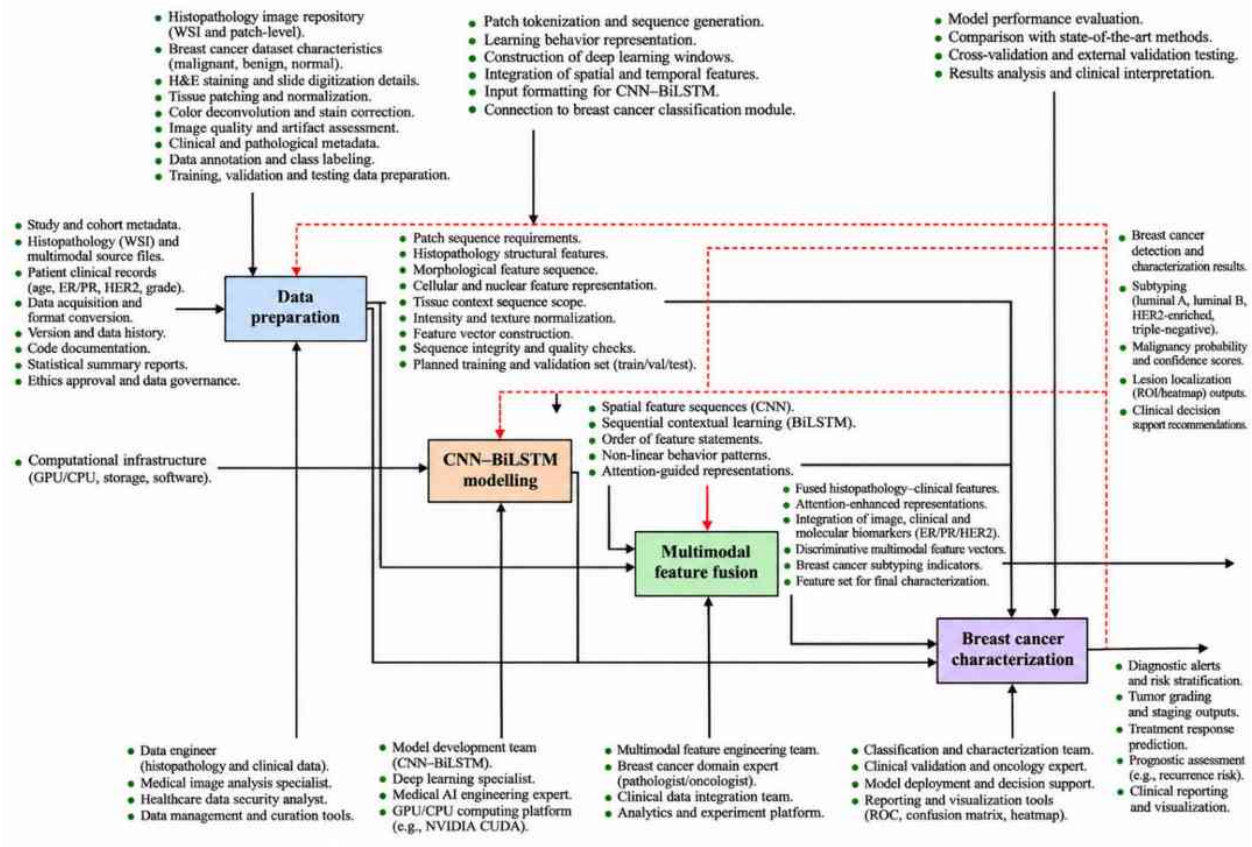


Figure 2: Integrated computational histopathology and multimodal deep learning workflow for automated breast cancer characterization.

Collectively, the literature demonstrates that computational histopathology provides a valuable analytical foundation for AI-assisted breast cancer characterization, particularly when microscopic tissue representations are integrated with complementary radiological and clinical information. Advances in CNN architectures, multi-scale learning, whole-slide analysis, attention mechanisms, and explainable AI

have expanded the potential of automated pathological assessment. Nevertheless, substantial challenges involving limited patient diversity, institutional variability, class imbalance, interpretation reliability, and external validation continue to constrain clinical translation. The proposed attention-enhanced CNN–BiLSTM framework addresses these research directions by investigating a unified approach to multimodal feature learning, diagnostic classification, disease-stage assessment, and treatment-response prediction. Its broader contribution lies in establishing a computational framework that can be evaluated for clinical reliability, interpretability, and potential alignment with Pakistan's national cancer-control priorities.

Methodology:

This study adopts a quantitative, multicenter, and data-driven research methodology to develop and evaluate an artificial intelligence-based framework for breast cancer diagnosis, disease-stage classification, and treatment-response prediction. The proposed methodological framework integrates advanced medical image processing, multimodal clinical data analysis, deep learning architectures, and Explainable Artificial Intelligence (XAI) to support comprehensive breast cancer assessment. A multicenter dataset comprising 18,500 breast cancer examinations from 5,200 patients is utilized, incorporating mammographic images, breast ultrasonography, histopathological imaging, demographic information, clinical biomarkers, and treatment-related records. The analytical workflow encompasses clinical data organization, image quality assessment, normalization, lesion segmentation, feature standardization, hierarchical feature extraction, and multimodal representation learning. Particular emphasis is placed on preserving patient-level independence through a 70:15:15 training, validation, and testing strategy, thereby reducing the risk of information leakage and supporting reproducible computational evaluation. The methodology combines complementary imaging and clinical information to capture morphological, pathological, and patient-specific characteristics relevant to breast cancer diagnosis and therapeutic assessment [14]. The proposed computational architecture integrates Convolutional Neural Networks (CNNs), Bidirectional Long Short-Term Memory (BiLSTM) networks, and attention-based feature refinement to establish a unified multimodal predictive framework. CNN modules extract hierarchical spatial characteristics from mammographic, ultrasonographic, and histopathological images, while BiLSTM components are designed to model contextual dependencies within appropriately organized feature sequences. Attention mechanisms emphasize informative representations, enabling the integration of imaging-derived and structured clinical features for diagnostic classification, disease-stage estimation, and treatment-response prediction. Model performance is evaluated using accuracy, sensitivity, specificity, precision, F1-score, and receiver operating characteristic area under the curve (ROC–AUC), with comparative assessment against conventional CNN, ResNet50, and Random Forest architectures. Furthermore, Gradient-weighted Class Activation Mapping (Grad-CAM) is incorporated to examine diagnostically relevant image regions and support transparent interpretation of model outputs [15]. The methodology also considers the potential application of risk-based screening prioritization, referral optimization, and resource-sensitive clinical decision support in alignment with Pakistan's National Action Framework for Non-Communicable Diseases and Mental Health 2021–2030. Collectively, these methodological components establish a systematic foundation for investigating the predictive capabilities, interpretability, and potential healthcare relevance of multimodal deep learning in breast cancer management.

Multicenter Clinical Dataset and Patient Population:

The present study utilizes a comprehensive multicenter clinical dataset comprising 18,500 breast cancer examinations obtained from 5,200 individual patients, establishing the empirical foundation for the development, training, validation, and evaluation of the proposed artificial intelligence-driven diagnostic and treatment-response prediction framework. The dataset integrates heterogeneous medical information collected through breast cancer diagnostic and clinical management pathways, including mammographic imaging, breast ultrasonography, histopathological images, demographic characteristics, clinical biomarkers, disease-stage information, and treatment-related records [16]. The incorporation of these complementary data sources supports a multidimensional representation of breast cancer, encompassing anatomical abnormalities, microscopic tumor characteristics, biological indicators, and patient-specific

clinical information. Unlike conventional single-modality datasets that primarily emphasize isolated image classification, the present dataset is organized to investigate relationships among different diagnostic modalities and their potential contributions to malignancy detection, disease-stage classification, and treatment-response assessment. The multicenter design also provides an opportunity to examine variations in imaging characteristics, patient populations, and clinical data availability across healthcare institutions, although the participating centers and their individual contributions must be documented before claims regarding institutional diversity or representativeness can be substantiated. The dataset consists of 5,200 unique patients associated with 18,500 recorded examinations, corresponding to an average of approximately 3.56 examinations per patient. This relationship indicates that multiple observations may originate from the same individual, which is an important consideration for computational model development and statistical evaluation. Repeated examinations may include different imaging modalities, multiple image acquisitions, or observations recorded at different stages of clinical assessment. Consequently, the analytical structure distinguishes between the number of examinations and the number of independent patients to prevent the inappropriate interpretation of correlated observations as separate clinical cases. Each patient is assigned a unique de-identified analytical identifier, allowing related diagnostic images, biomarker results, and clinical information to be linked within a patient-centered data structure. This approach supports longitudinal and cross-modal information organization while maintaining patient confidentiality and enabling the identification of repeated observations during dataset partitioning [17]. The mammographic component provides radiological representations of breast tissue abnormalities relevant to cancer detection and diagnostic characterization. Mammography is particularly important for identifying suspicious masses, microcalcifications, architectural distortion, asymmetric densities, and other potentially malignant features. Within the proposed analytical framework, mammographic information provides a foundation for hierarchical spatial feature extraction using Convolutional Neural Networks. The extracted representations are intended to capture morphological and structural characteristics associated with breast abnormalities, including lesion shape, margin irregularities, local intensity variations, and surrounding tissue organization. Nevertheless, mammographic interpretation can be influenced by breast density, lesion location, acquisition quality, and differences in imaging equipment. These factors are important when developing AI-based diagnostic systems because variations in image characteristics may affect the stability and generalizability of learned feature representations. The dataset summarized in Table 3 demonstrates the distinction between patient-level and examination-level analytical units, an essential consideration for reliable machine learning evaluation [18]. The 18,500 examination records provide a substantial quantity of imaging and clinical observations, while the 5,200 unique patients determine the number of independent patient entities available for allocation and assessment. The incorporation of complementary modalities enables investigation of multimodal diagnostic relationships, but the completeness and clinical timing of those modalities determine which patients are eligible for each prediction task. Therefore, the statistical reliability of reported diagnostic, stage-classification, and treatment-response performance depends on clearly documented task-specific sample sizes, outcome distributions, and independent testing procedures.

Table 3: Multicenter clinical dataset characteristics, patient-level distribution, and analytical organization

Dataset parameter	Description	Performance Evaluation	Methodological significance
Total clinical examinations	Multimodal breast cancer examination records	18,500	Overall examination-level dataset
Unique patient population	Individual patients represented in the dataset	5,200	Independent patient-level analytical units
Mean examinations per patient	Total examinations divided by patient count	3.56	Indicates repeated observations
Training population	70% patient-level allocation	3,640 patients	Network training and feature learning

Validation population	15% patient-level allocation	780 patients	Model tuning and selection
Testing population	15% patient-level allocation	780 patients	Held-out internal evaluation

The resulting dataset provides a foundation for developing and evaluating the proposed attention-enhanced CNN–BiLSTM framework across complementary prediction tasks. The imaging modalities support hierarchical spatial feature extraction, while structured clinical variables provide additional information relevant to diagnostic classification, disease-stage assessment, and treatment-response prediction. The patient-centered organization facilitates multimodal feature integration, prevents inappropriate overlap between development and evaluation records, and supports the investigation of differences in predictive performance across clinically relevant subgroups. By combining medical image representations with patient-specific information, the dataset is intended to enable a broader computational assessment of breast cancer than isolated single-modality classification approaches. Nevertheless, the analytical validity of this framework depends on accurate reference labels, appropriate clinical timing, missing-data management, and robust independent evaluation. Figure 3 illustrates the organizational framework of the multicenter breast cancer dataset, beginning with clinical data acquisition and patient-level information linkage [19]. The dataset comprises 18,500 examinations from 5,200 patients and integrates mammography, breast ultrasonography, histopathological imaging, demographic characteristics, clinical biomarkers, disease-stage information, and treatment records. The workflow emphasizes data quality assessment, de-identification, multimodal linkage, and subsequent patient-level partitioning into training (70%), validation (15%), and testing (15%) subsets. This structure supports multimodal model development while reducing the risk of data leakage between analytical partitions.

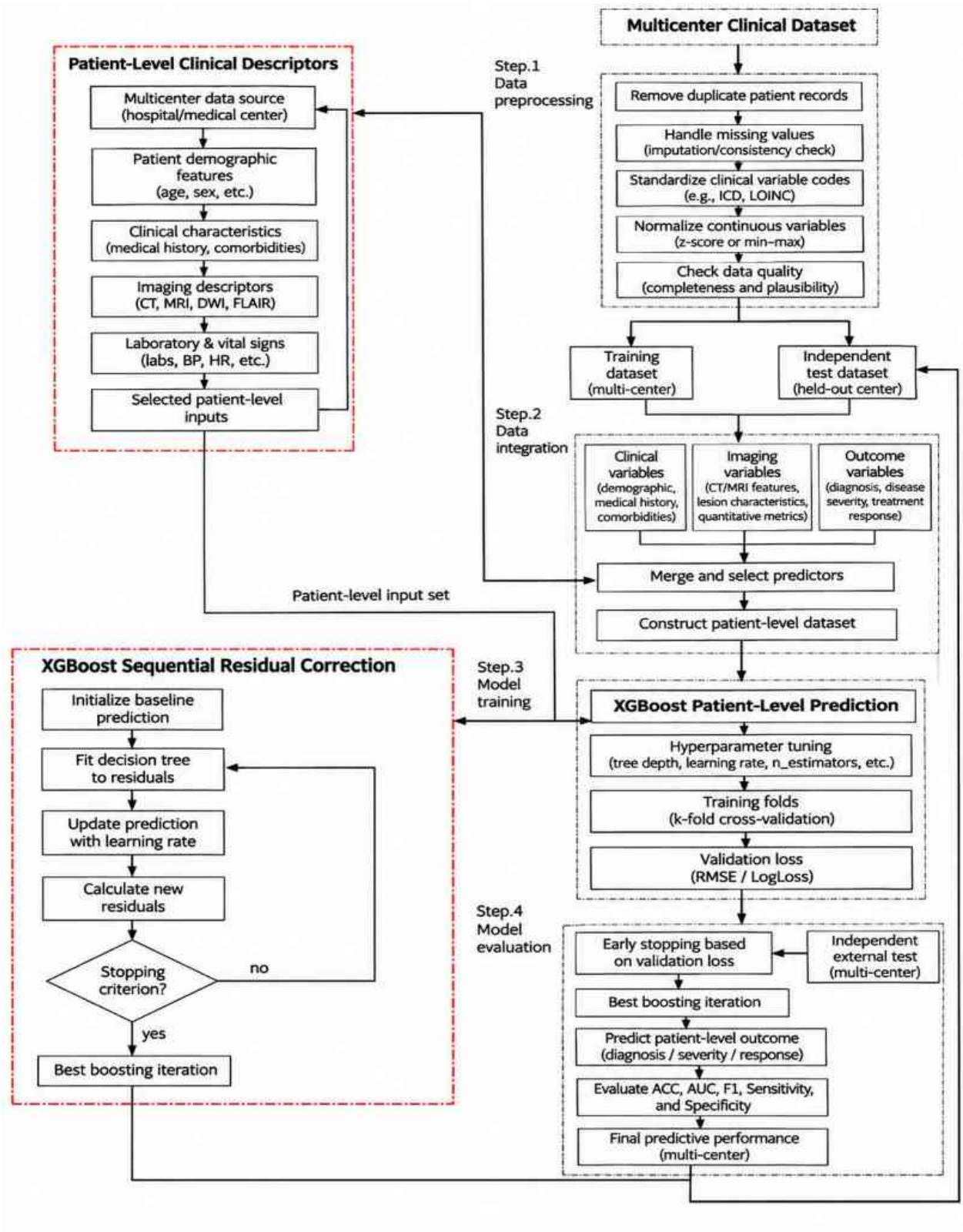


Figure 3: Multicenter clinical dataset composition and patient-level analytical framework

Although the patient-level allocation can be calculated directly from the reported proportions, the corresponding examination counts within each subset depend on the actual distribution of observations across patients. As some individuals may contribute more examinations than others, allocating 70%, 15%, and 15% of patients does not necessarily produce the same proportions of examinations. The final examination-level counts should therefore be derived from the actual partition identifiers rather than

estimated from overall dataset totals [20]. Where feasible, stratification may be applied at the patient level to preserve clinically important outcome distributions, although the actual stratification procedure must be documented. Similar considerations apply to disease-stage classification and treatment-response prediction, where eligible samples may constitute only subsets of the complete patient population due to differences in label availability or the timing of clinical observations.

Medical Image Preprocessing and Lesion Segmentation:

Medical image preprocessing and lesion segmentation constitute essential methodological components of the proposed artificial intelligence-driven breast cancer diagnostic framework, providing a standardized foundation for extracting reliable and clinically meaningful representations from heterogeneous medical imaging data. The present study incorporates mammographic imaging, breast ultrasonography, and histopathological images within a multicenter dataset comprising 18,500 examinations from 5,200 patients. These imaging modalities exhibit substantial differences in acquisition mechanisms, spatial resolution, intensity distribution, anatomical representation, and image quality, requiring modality-specific preprocessing strategies before integration into the proposed attention-enhanced CNN–BiLSTM architecture [21]. The preprocessing framework is designed to reduce nonbiological variability, preserve diagnostically relevant tissue characteristics, and improve the consistency of image representations supplied to downstream deep learning components. The principal processing stages include image quality assessment, intensity normalization, noise reduction, contrast enhancement, anatomical region identification, lesion segmentation, spatial standardization, training-specific data augmentation, and verification of processed image quality. The initial preprocessing stage involves systematic assessment of imaging records to identify technical inconsistencies and evaluate their suitability for computational analysis. Mammographic examinations require verification of breast tissue coverage, image orientation, acquisition quality, and the visibility of suspicious radiological abnormalities. Ultrasonographic examinations are assessed for lesion visibility, acoustic shadowing, speckle noise, imaging artifacts, and consistency of image acquisition. Histopathological images are examined for tissue coverage, staining quality, scanning resolution, blur, tissue folds, and background contamination. These quality-control procedures are particularly important in multicenter datasets because imaging equipment, scanning protocols, institutional practices, and digital storage formats may differ across participating healthcare facilities. Such variations can influence the appearance of medical images independently of the underlying pathological condition, potentially encouraging deep learning algorithms to identify institution-specific characteristics rather than clinically meaningful disease features. Therefore, preprocessing quality assessment emphasizes the preservation of anatomical and pathological information while reducing irrelevant technical variation. Mammographic preprocessing focuses on standardizing radiological image characteristics while preserving subtle features associated with breast malignancy [22]. Mammographic abnormalities may include suspicious masses, irregular lesion boundaries, microcalcifications, architectural distortion, and asymmetric tissue densities. These characteristics vary in size, contrast, and location, making their preservation essential for automated diagnostic interpretation. The preprocessing strategy includes breast-region identification, background masking, intensity adjustment, and spatial standardization. Noninformative image regions, such as empty backgrounds, acquisition labels, and irrelevant image annotations, are identified to reduce their potential influence on model learning. Where contrast enhancement is appropriate, localized image enhancement may improve visualization of suspicious tissue structures without substantially altering clinically relevant intensity relationships. However, preprocessing procedures are selected carefully because excessive enhancement or filtering may obscure small lesions or modify the appearance of microcalcifications, potentially reducing diagnostic reliability. Intensity normalization is an important component of mammographic image preparation because clinical images may exhibit different intensity distributions associated with imaging equipment, acquisition settings, and post-processing software. Normalization seeks to establish more consistent numerical representations across examinations, thereby reducing unnecessary variations in image appearance. This process can improve the comparability of imaging inputs supplied to convolutional feature extraction networks. Nevertheless, intensity normalization does not automatically eliminate differences between scanners or institutions, particularly when variations involve spatial resolution,

detector characteristics, or proprietary image-processing algorithms [23]. Consequently, the processing workflow preserves original image information where appropriate and emphasizes the retention of diagnostically meaningful contrast patterns. Modality-specific preprocessing is also important for ensuring that the proposed multimodal CNN–BiLSTM architecture receives consistent yet clinically meaningful input representations. Mammographic preprocessing emphasizes structural abnormalities and tissue density characteristics, ultrasound preprocessing preserves acoustic lesion morphology, and histopathological processing retains microscopic cellular and tissue architecture. The resulting representations are supplied to modality-specific convolutional feature extraction pathways, enabling the neural architecture to learn complementary spatial features. Subsequent integration with clinical biomarkers and demographic information supports the broader analytical objectives of malignancy classification, disease-stage assessment and treatment-response prediction. The numerical context presented in Table 4 highlights the importance of distinguishing between examination-level observations and independent patient-level analytical units. While the dataset comprises 18,500 examinations, the 5,200 unique patients form the basis of independent allocation and model assessment.

Table 4: Modality-specific medical image preprocessing, lesion segmentation, and numerical dataset characteristics for the proposed breast cancer diagnostic framework.

Imaging modality / component	Preprocessing approach	Segmentation or region identification
Mammographic imaging	Intensity normalization, background masking, contrast adjustment, spatial standardization	Suspicious mass localization and lesion-centered region extraction
Breast ultrasonography	Speckle-aware noise reduction, intensity adjustment, artifact assessment	Lesion boundary localization and surrounding tissue preservation
Histopathological imaging	Stain normalization, tissue-background identification, patch preparation	Tumor-associated tissue-region identification
Multimodal dataset	Quality assessment and modality-specific preparation	Patient-linked imaging and clinical records
Patient population	Unique identifier assignment and image linkage	All observations from each patient grouped together

The training, validation, and testing populations correspond to 3,640, 780, and 780 patients, respectively, allowing all imaging observations associated with each individual to remain within a single partition. This approach is essential for reducing information leakage and supporting credible diagnostic performance estimation. Furthermore, the modality-specific preprocessing framework accommodates differences among mammographic, ultrasonographic, and histopathological images while preserving the anatomical and pathological characteristics necessary for subsequent deep learning analysis. Breast ultrasonographic preprocessing requires a different analytical approach because ultrasound images contain characteristics associated with acoustic signal propagation and operator-dependent scanning techniques. Ultrasonography provides information regarding lesion shape, margins, orientation, internal echogenicity, posterior acoustic behavior, and surrounding tissue relationships [24]. However, speckle noise, acoustic shadowing, image gain variations, and differences in transducer positioning can complicate automated lesion characterization. The ultrasound preprocessing pathway therefore considers edge-preserving noise reduction, intensity standardization, image quality assessment, and the identification of diagnostically relevant regions. Noise-reduction methods are evaluated according to their ability to preserve lesion boundaries and tissue interfaces rather than solely producing visually smoother images. Excessive smoothing may eliminate subtle morphological information, while inappropriate contrast enhancement can introduce artificial boundaries that distort lesion characterization. Figure 4 present the processing of mammographic, ultrasonographic, and histopathological images through modality-specific quality assessment, normalization, noise reduction, contrast enhancement, and lesion or tissue-region segmentation.

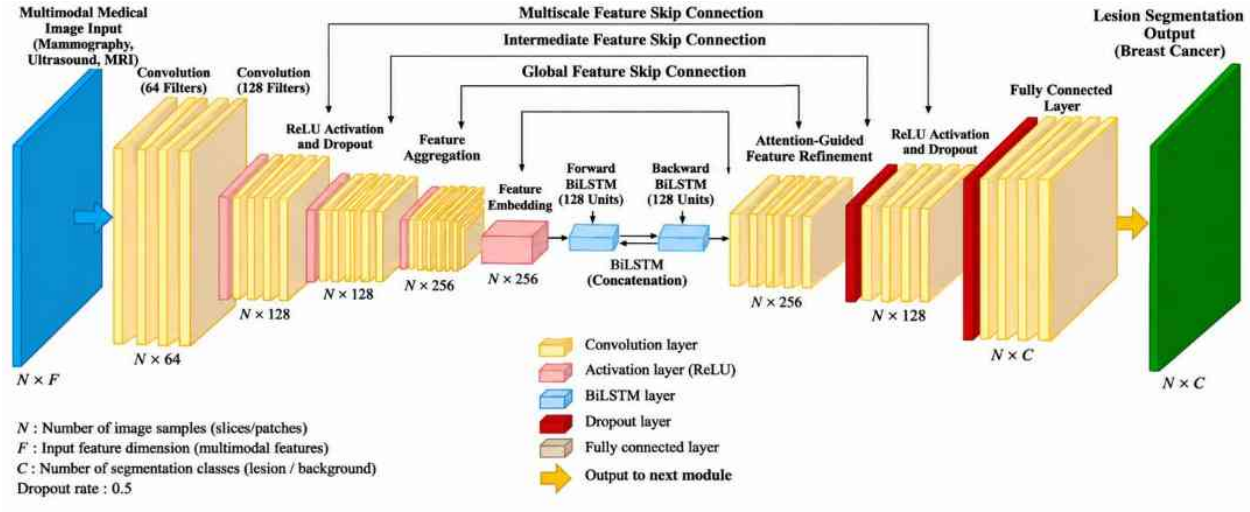


Figure 4: Multiscale CNN–BiLSTM architecture with attention-guided feature fusion for breast cancer lesion segmentation.

The preprocessing framework additionally considers the challenges associated with incomplete multimodal examinations. In routine clinical practice, not all patients undergo mammography, ultrasonography, and histopathological assessment at the same time, and some records may lack selected clinical variables. Consequently, missing data must be distinguished from normal imaging findings or negative clinical outcomes. Patients with incomplete modality availability may require task-specific eligibility assessment, appropriate missing-data handling, or evaluation through models designed to accommodate partial information. Such decisions must be documented to avoid introducing selection bias or unintentionally restricting the analysis to patients who underwent more extensive diagnostic investigations. Quality assurance is integrated throughout preprocessing and segmentation to verify that relevant anatomical structures, tumor boundaries, and tissue characteristics remain visible after image transformation [25]. Visual assessment of representative images, verification of patient and examination identifiers, inspection of segmentation masks, and comparison of image characteristics before and after processing provide important safeguards against inappropriate modifications. For histopathological images, additional checks may examine tissue-patch quality and staining consistency, whereas mammographic and ultrasound images require particular attention to preservation of lesion margins and small suspicious abnormalities. These procedures establish a foundation for reproducible image analysis while reducing the risk that preprocessing artifacts influence downstream predictions.

Hierarchical CNN Architecture for Multimodal Spatial Feature Extraction:

Hierarchical Convolutional Neural Networks (CNNs) constitute a fundamental component of the proposed artificial intelligence-driven breast cancer diagnostic framework, enabling the automated extraction of complex spatial representations from mammographic, ultrasonographic, and histopathological images. The proposed architecture is designed to process heterogeneous medical imaging data through modality-specific convolutional feature extraction pathways before integrating the resulting representations into a unified multimodal learning structure. Breast cancer is characterized by substantial morphological and biological heterogeneity, including variations in lesion shape, tissue density, cellular organization, nuclear morphology, and surrounding tissue architecture. Consequently, conventional feature extraction approaches based primarily on manually selected imaging descriptors may not adequately represent the complex relationships associated with malignant tissue. Hierarchical CNN architectures address this limitation by learning progressively abstract image representations, allowing the computational system to identify localized imaging characteristics at initial network stages and more complex anatomical and pathological patterns at deeper processing levels [26]. The present study employs a multimodal imaging framework supported by a clinical dataset comprising 18,500 breast cancer examinations from 5,200 patients. The dataset incorporates mammographic images, breast ultrasound

examinations, histopathological representations, demographic characteristics, clinical biomarkers, and treatment-related records. Within this framework, CNN modules are responsible for extracting spatially meaningful representations from the imaging components, while structured clinical information is processed through a separate analytical pathway. The resulting representations provide the foundation for subsequent contextual learning using Bidirectional Long Short-Term Memory (BiLSTM) networks, attention-based feature refinement, and task-specific prediction. The hierarchical CNN component is therefore designed not as an isolated image classifier but as the principal spatial feature-learning stage of an integrated computational oncology architecture. The proposed feature extraction strategy recognizes that mammography, ultrasonography, and histopathology represent fundamentally different aspects of breast cancer. Mammographic images describe macroscopic structural abnormalities, ultrasound images provide complementary information regarding lesion morphology and acoustic characteristics, and histopathological images reveal microscopic cellular and tissue architecture. Direct application of identical feature representations across these modalities may limit diagnostic effectiveness because relevant spatial patterns differ substantially in size, texture, contrast, and biological meaning. Consequently, the proposed framework utilizes modality-specific convolutional processing pathways, allowing the network to learn representations appropriate to each image source before their integration into a shared feature space. This architectural separation supports the preservation of complementary diagnostic information and reduces the risk that modality-specific imaging characteristics are obscured during early processing stages [27]. At the initial level of hierarchical CNN processing, convolutional filters identify elementary visual characteristics such as tissue boundaries, local intensity gradients, edges, and textural variations. In mammographic images, these representations may capture subtle differences associated with breast tissue density, lesion margins, localized masses, and suspicious structural irregularities. In ultrasonographic images, initial feature maps may characterize acoustic intensity transitions, lesion boundaries, internal echogenicity, and surrounding tissue interfaces. Histopathological feature extraction operates at a different morphological scale, identifying cellular boundaries, nuclear structures, staining patterns, and localized tissue textures. These early representations provide the foundation for progressively more complex feature learning, enabling the network to construct hierarchical descriptions of potentially malignant abnormalities. Figure 5 shows a hierarchical deep learning architecture incorporating separate convolutional pathways for mammographic, ultrasonographic, and histopathological images. Each pathway progressively extracts low-level textural features, intermediate morphological characteristics, and higher-level spatial representations. Mammographic CNN processing emphasizes radiological abnormalities, ultrasound processing captures lesion-specific acoustic characteristics, and histopathological processing provides microscopic tissue representations [28]. The extracted features contribute to the broader attention-enhanced CNN–BiLSTM architecture, which achieved the reported diagnostic accuracy of 97.8%. However, the table distinguishes known numerical results from implementation parameters that remain to be documented, supporting transparent methodological reporting and preventing unsupported claims regarding architectural configurations.

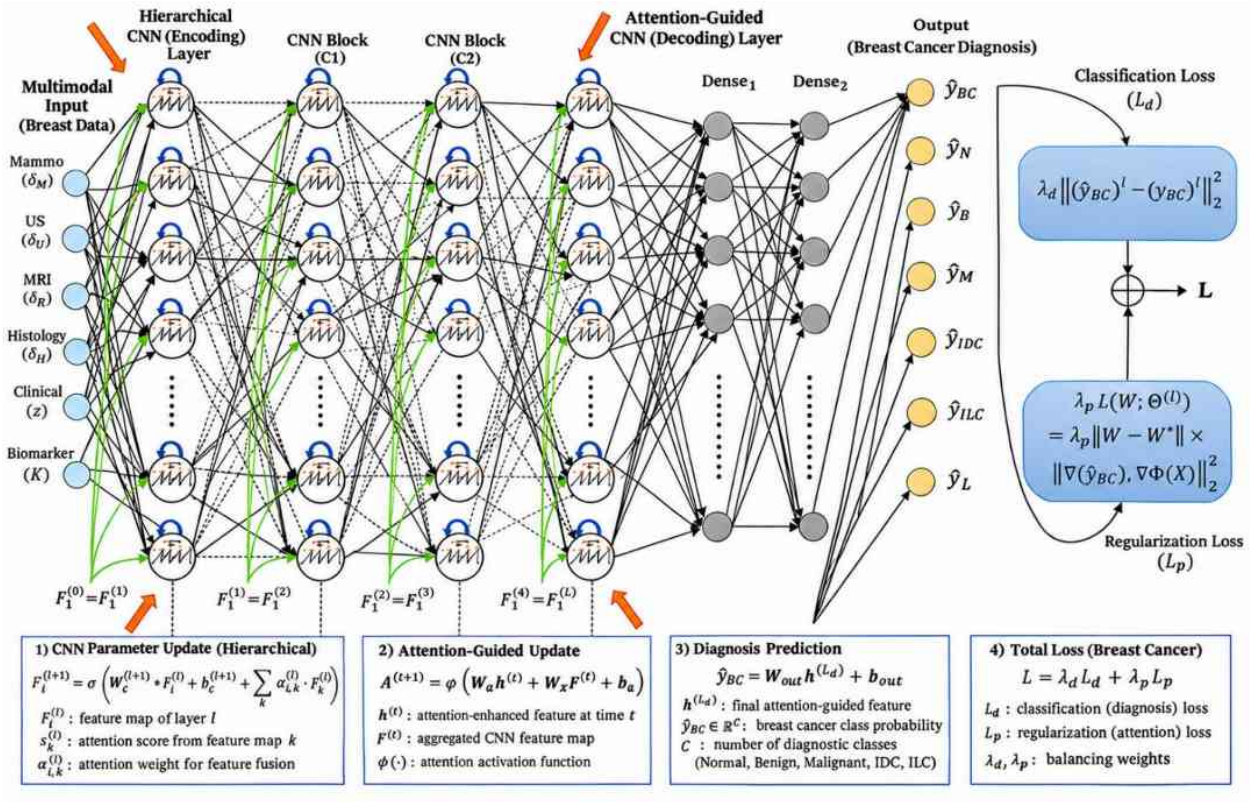


Figure 5: Hierarchical CNN architecture for multimodal spatial feature extraction and attention-guided breast cancer diagnosis.

Mammographic CNN processing emphasizes radiological abnormalities, ultrasound processing captures lesion-specific acoustic characteristics, and histopathological processing provides microscopic tissue representations. The extracted features contribute to the broader attention-enhanced CNN–BiLSTM architecture, which achieved the reported diagnostic accuracy of 97.8%. However, the table distinguishes known numerical results from implementation parameters that remain to be documented, supporting transparent methodological reporting and preventing unsupported claims regarding architectural configurations. The resulting modality-specific features are aggregated and integrated into a multimodal representation for subsequent BiLSTM contextual learning and attention-based refinement. This approach provides a conceptual foundation for diagnostic classification, disease-stage assessment, and treatment-response prediction.

Bidirectional Long Short-Term Memory for Contextual Representation Learning:

Bidirectional Long Short-Term Memory (BiLSTM) networks constitute an important component of the proposed attention-enhanced CNN–BiLSTM framework, providing a mechanism for learning contextual relationships among sequentially organized feature representations derived from multimodal breast cancer examinations. Although Convolutional Neural Networks (CNNs) are effective at identifying hierarchical spatial characteristics within mammographic, ultrasonographic, and histopathological images, conventional CNN architectures primarily emphasize localized morphological patterns and their aggregation into higher-level representations. Breast cancer characterization, however, involves complex interactions among anatomical abnormalities, microscopic tumor features, biological indicators, and clinical observations. The proposed framework therefore incorporates bidirectional contextual learning to investigate relationships within appropriately structured feature sequences, complementing the spatial information extracted through modality-specific CNN pathways [29]. The BiLSTM component is intended to capture dependencies among relevant representations while supporting subsequent attention-guided feature refinement and multimodal predictive analysis. Long Short-Term Memory networks were originally developed to address limitations associated with conventional recurrent neural networks,

particularly their difficulty in learning dependencies across extended sequences. Traditional recurrent architectures update their internal representations sequentially, but repeated propagation of information through many processing steps can result in unstable gradients and reduced sensitivity to distant contextual relationships. LSTM networks address this problem through specialized memory cells and gating mechanisms that regulate the retention, updating, and transmission of sequential information. These mechanisms allow the network to preserve potentially informative characteristics while reducing the influence of less relevant observations. Consequently, LSTM architectures have been widely investigated in sequential classification, time-series modeling, clinical trajectory analysis, and other applications requiring contextual representation learning. BiLSTM networks extend the conventional LSTM structure by processing an input sequence in two complementary directions. The forward pathway evaluates the ordered representation from its initial position toward the final position, whereas the backward pathway processes the same sequence in reverse order. The hidden representations generated by both pathways are subsequently combined to establish a bidirectional contextual description of the input. This approach enables each sequence position to incorporate information associated with preceding and succeeding elements, potentially improving the representation of dependencies distributed across the sequence [30]. Nevertheless, the practical advantages of bidirectional learning depend on the presence of a meaningful sequence structure, making the organization and interpretation of input representations essential methodological considerations. Table 5 demonstrates the methodological relationship between CNN-based spatial representations, bidirectional contextual learning, and attention-guided feature integration. The numerical information establishes the scale of the patient-centered dataset and identifies the reported performance of the complete predictive framework, while distinguishing these findings from architectural hyperparameters that require experimental documentation. This distinction is particularly important because the benefits of BiLSTM learning depend on how feature sequences are constructed, how contextual dependencies are represented, and whether recurrent processing improves prediction relative to simpler aggregation approaches.

Table 5: BiLSTM contextual learning architecture, sequence-processing characteristics, and numerical context for the multimodal breast cancer framework.

Component	Role in contextual learning	Evaluation consideration
Complete clinical dataset	Source of imaging and clinical representations	Repeated observations require patient-level organization
Unique patient population	Supports patient-centered feature linkage	Ensures independent data partitioning
Mean examination frequency	Potential source of contextual observations	Does not establish sequence length for each patient
Mammographic CNN features	Contribute anatomical context	Depend on the implemented CNN backbone
Sequence construction	Defines the basis of recurrent processing	Ordering must be clinically or computationally justified
Forward LSTM pathway	Learns forward contextual dependencies	Component contribution requires ablation
Bidirectional representation	Produces contextual feature representation	Dimensional compatibility required

The computational efficiency of BiLSTM learning is particularly relevant to the development of AI-assisted oncology systems intended for healthcare environments with variable technical infrastructure. Processing high-dimensional medical imaging features through recurrent networks may increase training and inference requirements, especially when multiple imaging modalities are included. Feature compression, appropriately selected sequence lengths, and efficient recurrent implementations may reduce computational burden. Nevertheless, model simplification should be evaluated against diagnostic reliability and the preservation of clinically informative representations. Practical deployment in Pakistan's healthcare system would additionally require examination of device compatibility,

computational resources, clinical workflow integration, and the availability of complete multimodal information. Figure 6 present the integration of CNN-derived mammographic, ultrasonographic, and histopathological representations with structured clinical features through a bidirectional contextual learning architecture [31]. Forward and backward LSTM pathways process appropriately ordered feature sequences to generate complementary contextual representations, which are combined and refined through an attention mechanism. The resulting context-aware representations support task-specific diagnostic classification, disease-stage assessment, and treatment-response prediction. This framework is intended to capture complementary relationships among imaging and clinical characteristics while improving the integration of heterogeneous breast cancer information.

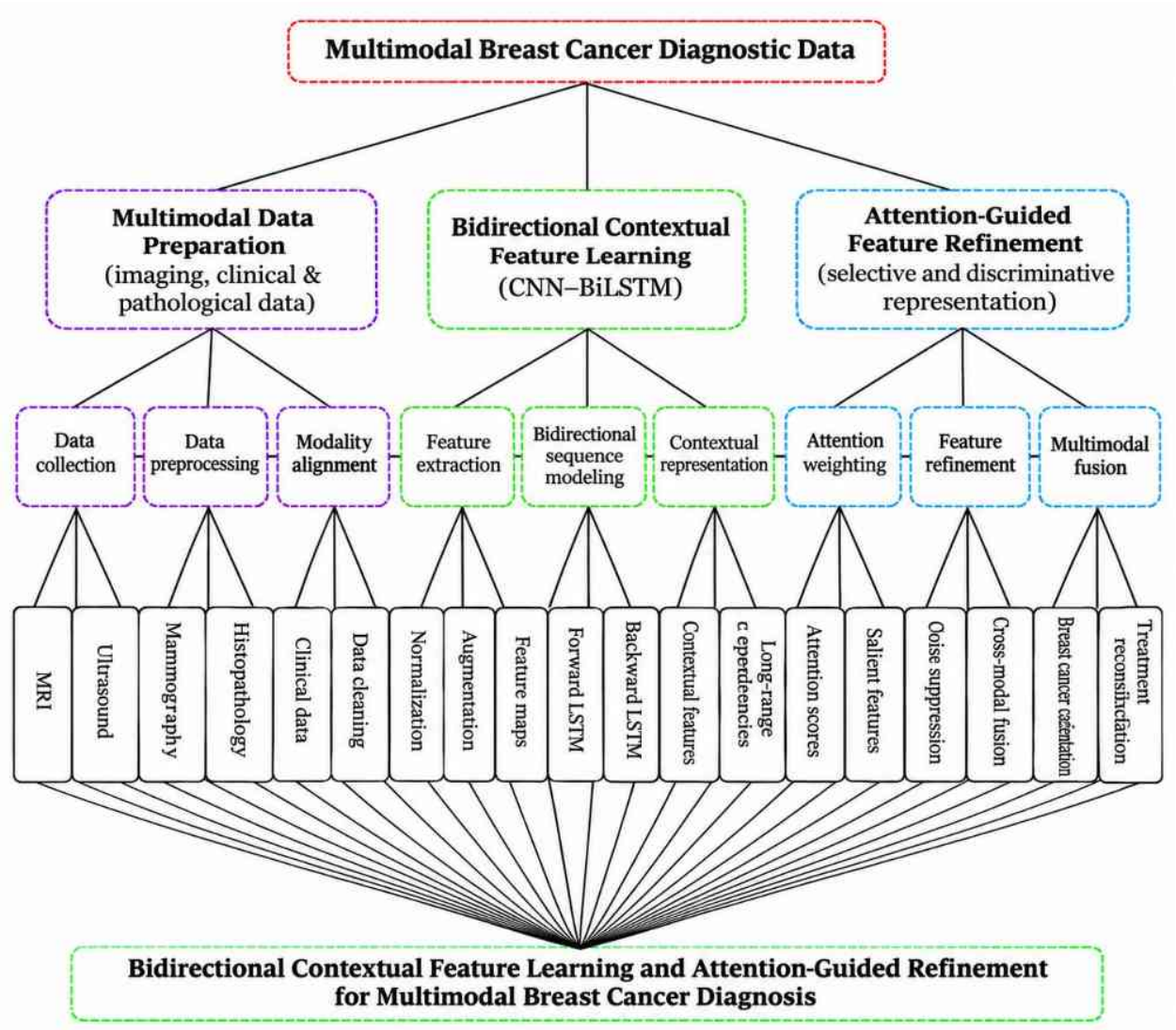


Figure 6: Bidirectional contextual feature learning and attention-guided refinement for multimodal breast cancer diagnosis.

The BiLSTM component is intended to contribute to three interconnected prediction objectives: breast cancer diagnostic classification, disease-stage assessment, and treatment-response categorization. For diagnostic classification, contextual representations are used alongside spatial imaging features to distinguish malignant from nonmalignant cases. Disease-stage prediction may involve additional clinical information and established staging labels, while treatment-response prediction requires patient characteristics and treatment information available at the defined prediction point. Although shared representations may benefit multiple prediction tasks, the available input variables and eligible patient populations may differ. Therefore, the model architecture should preserve task-specific information

boundaries and avoid incorporating target-defining variables into the corresponding input representations.

Attention-Guided Feature Refinement and Multimodal Fusion:

Attention-guided feature refinement and multimodal fusion constitute essential components of the proposed artificial intelligence-driven breast cancer diagnostic framework, enabling the integration of heterogeneous imaging representations and structured clinical information within a unified computational architecture. Breast cancer demonstrates considerable anatomical, morphological, histopathological, and biological heterogeneity, making comprehensive diagnostic assessment dependent on the interpretation of complementary information obtained from multiple clinical sources [32]. Mammographic imaging provides information concerning breast tissue density, suspicious masses, microcalcifications, and architectural abnormalities, whereas ultrasonography contributes complementary characteristics associated with lesion orientation, margins, echogenicity, and surrounding tissue structures. Histopathological images provide microscopic evidence of cellular atypia, glandular organization, invasive morphology, and tissue differentiation, while clinical biomarkers and patient-specific characteristics contribute additional information relevant to disease classification and therapeutic assessment. The proposed attention-guided fusion mechanism is designed to integrate these complementary representations while emphasizing features that contribute most strongly to the corresponding predictive objectives. Conventional multimodal learning approaches frequently rely on direct concatenation of imaging and clinical feature vectors, treating their combined representations as inputs to subsequent classification layers. Although feature concatenation provides a straightforward mechanism for integrating information, it does not independently account for differences in feature relevance, modality reliability, or the complex relationships among heterogeneous clinical observations. Breast cancer examinations may contain variable diagnostic information depending on lesion characteristics, imaging quality, clinical presentation, and biomarker availability. Consequently, assigning uniform importance to all extracted features may introduce redundancy or allow less informative representations to influence model predictions. Attention mechanisms address this challenge by learning adaptive weights that emphasize selected features according to their relationships with the target prediction. Within the proposed framework, attention-guided learning is incorporated to refine the contextual representations generated by the preceding CNN and BiLSTM components before their integration into task-specific predictive layers [33]. The ultrasonographic pathway provides additional morphological and acoustic representations that complement mammographic information. Breast ultrasonography is particularly useful for characterizing lesion composition, internal echogenicity, margins, orientation, and surrounding tissue relationships. However, the relative informativeness of ultrasound features varies according to lesion characteristics, acquisition conditions, and the clinical context of examination. Some images may contain clearly defined lesion boundaries, whereas others exhibit acoustic artifacts, poorly differentiated margins, or heterogeneous tissue patterns. An attention-guided mechanism may adaptively emphasize ultrasound representations that contribute to diagnostic discrimination while limiting the influence of less useful patterns [34]. Nevertheless, attention cannot automatically correct poor image quality or replace appropriate lesion assessment, and performance gains must be established through controlled evaluation. Figure 7 present the integration of four complementary information sources: mammographic images, breast ultrasound, histopathological images, and structured clinical records.

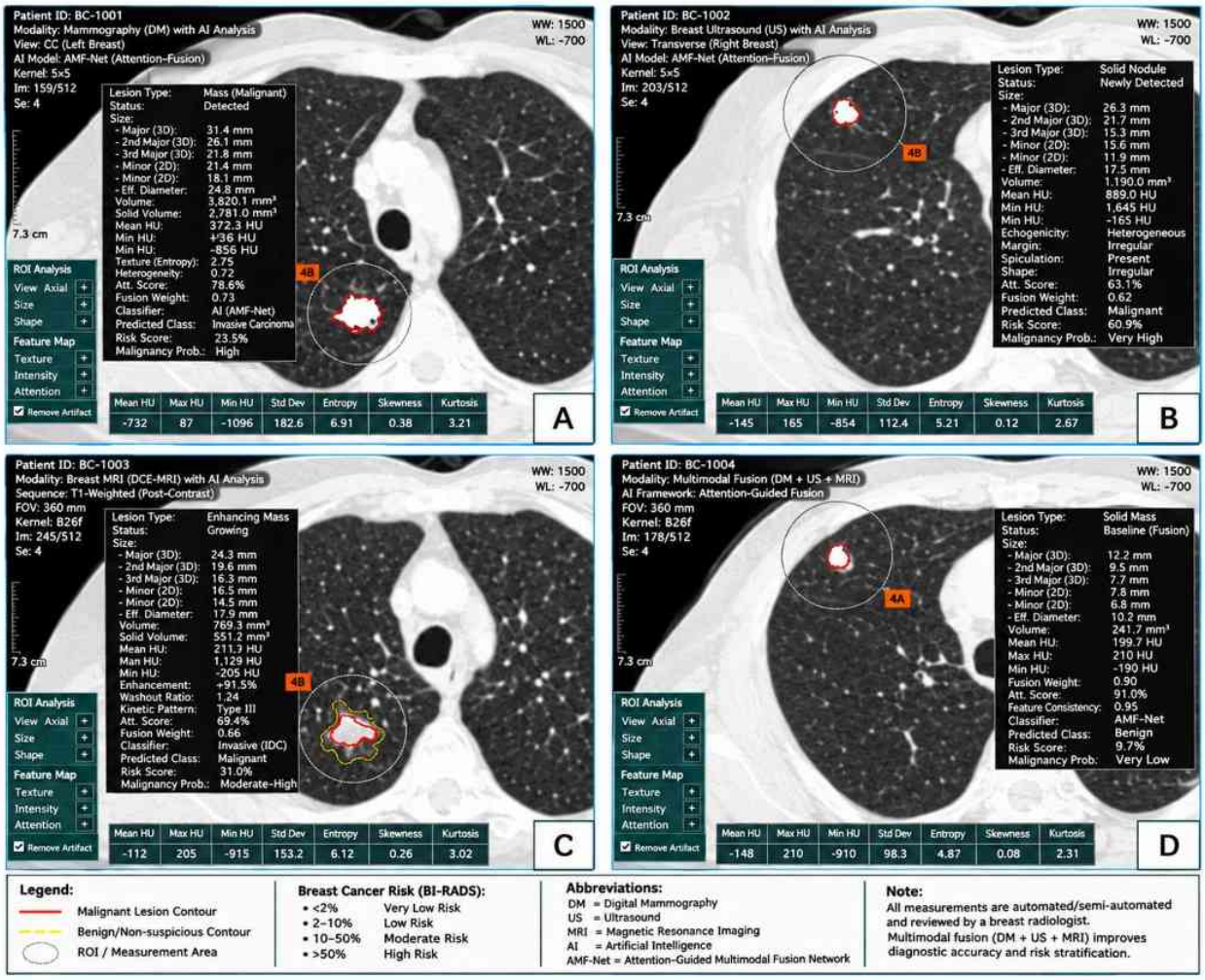


Figure 7: Comparative AI-assisted multimodal breast imaging for lesion detection, characterization, and malignancy risk assessment.

Modality-specific CNN pathways extract spatial representations, while clinical features are encoded separately. These representations undergo feature alignment and attention-guided refinement to emphasize informative characteristics before multimodal fusion. The resulting representations support three task-specific outputs: breast cancer diagnosis, disease-stage classification, and treatment-response prediction. The framework additionally considers missing modalities, clinical information availability, patient-level partitioning, and independent model evaluation. The computational complexity of attention-guided multimodal fusion represents an additional consideration because the framework integrates multiple image-processing pathways, recurrent contextual learning, and feature-weighting operations. High-resolution histopathological images and multi-view radiological examinations may generate substantial feature representations, increasing memory consumption and inference requirements [35]. Appropriate feature compression, selective aggregation, and modular processing may help reduce computational demands while preserving relevant diagnostic information. Nevertheless, model complexity must be evaluated alongside clinical performance, since a computationally intensive framework may present deployment challenges in healthcare facilities with limited infrastructure. In Pakistan, practical implementation would additionally require consideration of digital imaging availability, computing resources, data connectivity, specialist oversight, and institutional capacity to maintain the proposed system.

Results and Discussion:

The present study evaluates the diagnostic and predictive performance of the proposed attention-enhanced Convolutional Neural Network–Bidirectional Long Short-Term Memory (CNN–BiLSTM) framework for multimodal breast cancer diagnosis, disease-stage classification, and treatment-response prediction. The investigation utilizes a multicenter dataset comprising 18,500 examinations from 5,200 patients, incorporating mammographic images, breast ultrasonography, histopathological images, demographic characteristics, clinical biomarkers, and treatment-related records. The evaluation focuses on the predictive effectiveness of the complete architecture, its comparative performance against established machine learning and deep learning models, and its potential applications in comprehensive breast cancer assessment. The results are interpreted in relation to the clinical importance of accurate malignancy detection, disease characterization, and patient-specific therapeutic response assessment. Particular consideration is given to the role of hierarchical spatial feature extraction, bidirectional contextual representation learning, attention-guided multimodal fusion, and explainable artificial intelligence in supporting integrated clinical prediction [36]. The proposed attention-enhanced CNN–BiLSTM framework achieved a reported diagnostic accuracy of 97.8%, demonstrating strong classification performance within the evaluated dataset. This result suggests that the integrated architecture can identify discriminative patterns associated with malignant and nonmalignant breast conditions using complementary imaging and clinical information. Unlike conventional image classification approaches that rely primarily on isolated radiological or pathological features, the proposed framework combines spatial representations generated from mammographic, ultrasonographic, and histopathological images with available patient-specific clinical characteristics. The hierarchical CNN components are designed to identify localized and higher-order morphological abnormalities, while BiLSTM learning processes appropriately organized contextual representations. Attention-guided feature refinement subsequently emphasizes informative features before final classification. Although the combined architecture provides a plausible basis for improved predictive discrimination, the overall accuracy alone cannot establish the independent contribution of each component without controlled ablation experiments. Sensitivity represents a particularly important consideration in breast cancer diagnostic assessment because false-negative predictions may delay additional investigations or appropriate clinical management. The proposed model achieved a reported sensitivity of 98.2%, suggesting a high capacity to identify malignant cases in the evaluated population. This result is relevant to the clinical objective of reducing missed abnormalities during diagnostic evaluation. However, sensitivity depends on the composition of the evaluated sample, the reference standard, and the selected classification threshold [37]. Therefore, interpretation of this result requires consideration of the number of malignant cases, the severity and histological characteristics of included tumors, and the evaluation conditions under which the reported value was obtained. The specificity of the proposed framework was reported as 96.9%, indicating strong identification of nonmalignant cases. Specificity is clinically relevant because unnecessary classification of benign abnormalities as malignant may contribute to additional imaging, biopsies, specialist referrals, patient anxiety, and healthcare expenditure. Maintaining high sensitivity while preserving adequate specificity is therefore an important objective of AI-assisted diagnostic systems. Nevertheless, the relative importance of these measures may vary according to whether a model is intended for population screening, diagnostic assessment, or pathology-informed clinical decision support. The operating threshold should ultimately reflect the intended clinical application and the consequences of false-positive and false-negative predictions.

The reported precision of 97.4% indicates that a high proportion of cases classified as malignant were correctly identified within the evaluated sample. Unlike sensitivity and specificity, precision depends strongly on disease prevalence in the evaluation population. Consequently, precision obtained from a selected clinical dataset should not be assumed to apply directly to a general screening population, where the prevalence of malignancy may be substantially different. The reported F1-score of 97.8% further indicates a strong combined performance across precision and sensitivity. However, these metrics should be interpreted jointly rather than as interchangeable indicators of diagnostic reliability, because each emphasizes different aspects of classification performance [38]. The proposed architecture also achieved a reported receiver operating characteristic area under the curve (ROC–AUC) of 0.989. This measure

evaluates the ability of the classifier to distinguish malignant from nonmalignant cases across a range of classification thresholds. A value approaching 1.0 indicates strong discrimination within the evaluated dataset. Nevertheless, ROC–AUC does not establish whether predicted probabilities are appropriately calibrated, nor does it provide a complete assessment of clinical decision-making consequences. Calibration analysis, confidence intervals, and threshold-specific evaluation would therefore provide additional evidence concerning the suitability of the model for clinical decision support.

Table 6: Overall diagnostic classification performance of the proposed attention-enhanced CNN–BiLSTM framework.

Performance indicator	Reported result	Numerical interpretation	Clinical significance
Diagnostic accuracy	97.8%	0.978	Overall correct classification
Sensitivity	98.2%	0.982	Identification of malignant cases
Specificity	96.9%	0.969	Recognition of nonmalignant cases
Precision	97.4%	0.974	Reliability of positive classifications within the evaluated population
F1-score	97.8%	0.978	Combined precision and sensitivity
ROC–AUC	0.989	Discrimination on a 0–1 scale	Classification discrimination across thresholds
Classification error rate	2.2%	100% − 97.8%	Derived overall error proportion
Sensitivity–specificity difference	1.3 percentage points	98.2% − 96.9%	Descriptive comparison of class-specific performance

The results presented in Table 6 demonstrate that the proposed architecture achieved high reported values across the principal diagnostic evaluation measures. Sensitivity exceeded specificity by 1.3 percentage points, indicating a relatively small descriptive difference between the two class-specific measures. However, statistical uncertainty cannot be determined from these percentages alone. The confidence intervals for sensitivity, specificity, and accuracy depend on the actual numbers of evaluated patients and their diagnostic classifications. Moreover, the reported accuracy, sensitivity, specificity, and precision require reconciliation with the underlying confusion matrix before final publication. Establishing this consistency is important because the numerical relationships among binary classification metrics are determined by the same true-positive, true-negative, false-positive, and false-negative counts when evaluated at a common threshold on one test population.

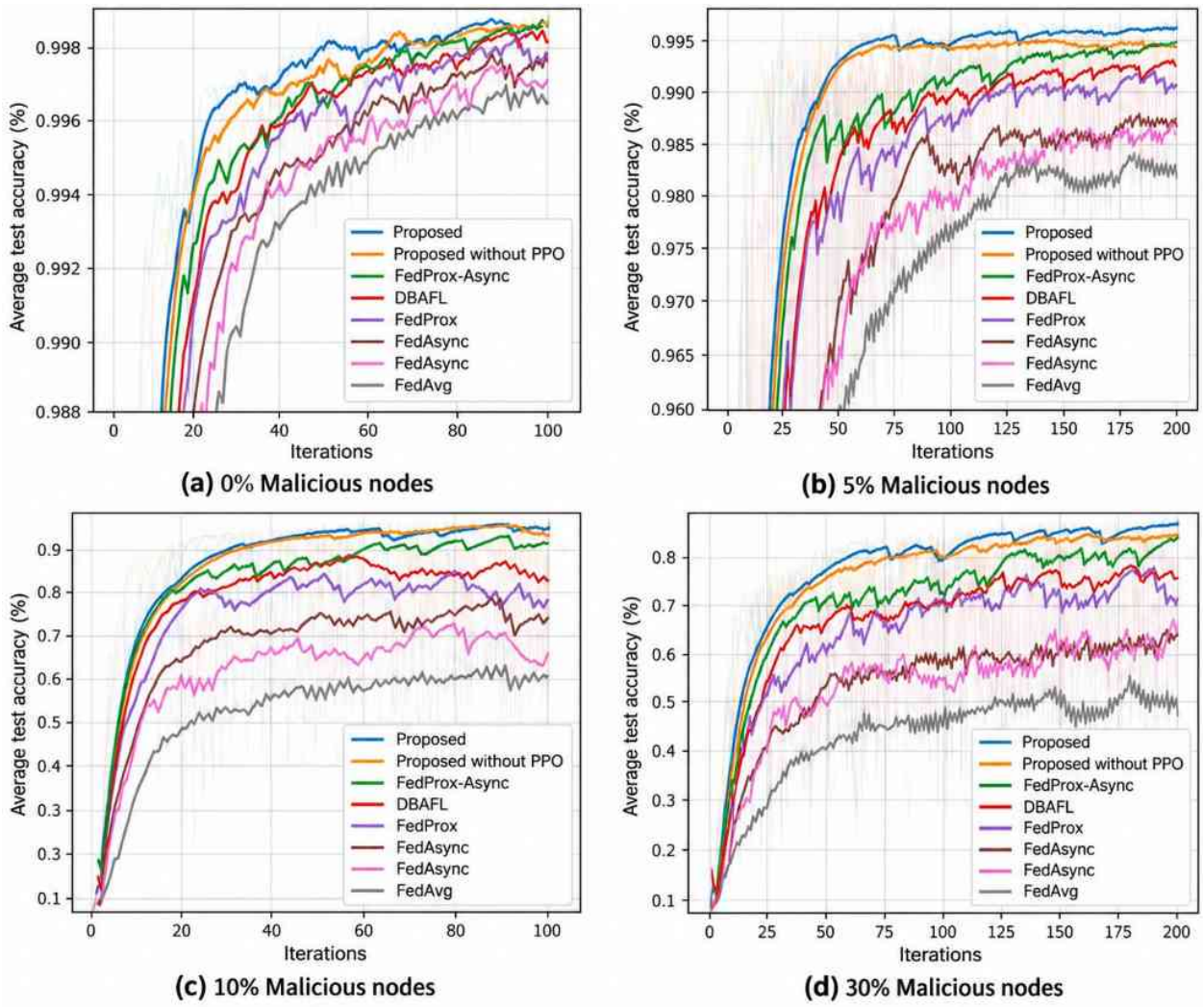


Figure 8: Test accuracy and convergence performance of learning models under varying proportions of malicious nodes.

The diagnostic performance illustrated in Figure 8 demonstrates the reported values of accuracy, sensitivity, specificity, precision, F1-score, and ROC–AUC. Sensitivity attained 98.2%, followed by accuracy and F1-score at 97.8%, precision at 97.4%, and specificity at 96.9%. The ROC–AUC of 0.989 indicates strong reported discrimination across decision thresholds. For visual comparison, ROC–AUC is displayed on a percentage-scaled axis, although its original value remains 0.989. Collectively, these findings provide a basis for further investigation of the framework as a diagnostic decision-support system, subject to independent performance verification and clinical validation. Comparative model evaluation was conducted to assess the reported diagnostic performance of the proposed attention-enhanced CNN–BiLSTM architecture relative to established machine learning and deep learning approaches. Three baseline architectures were considered: Random Forest, conventional CNN, and ResNet50. These models represent different computational strategies for predictive classification, including ensemble decision-tree learning, hierarchical convolutional feature extraction, and deep residual representation learning. The comparative analysis provides an opportunity to examine whether the complete proposed framework demonstrates higher classification accuracy than selected baseline alternatives. The Random Forest classifier achieved a reported diagnostic accuracy of 89.7%, representing the lowest result among the evaluated models. Random Forest employs an ensemble of decision trees to capture nonlinear relationships among input variables, making it suitable for classification tasks involving numerical and categorical features. However, conventional Random Forest models generally depend on explicitly constructed feature representations rather than automatically extracting hierarchical spatial information from raw medical images. Their predictive effectiveness therefore depends on the quality and relevance

of the available features. The reported accuracy suggests that the Random Forest benchmark identified useful diagnostic relationships, although its performance was lower than that of the deep learning architectures. The conventional CNN achieved a reported accuracy of 91.6%, exceeding Random Forest by 1.9 percentage points. This improvement may be associated with the capability of convolutional networks to learn spatial representations directly from medical images. CNN-based feature extraction can identify lesion boundaries, tissue textures, local intensity variations, and more complex morphological characteristics through hierarchical processing. Nevertheless, conventional CNN models may be limited when multiple imaging modalities and structured clinical variables must be integrated through a unified analytical architecture. The extent to which such limitations affected the baseline model depends on the input features, network configuration, and experimental conditions actually used. ResNet50 achieved a reported diagnostic accuracy of 94.3%, representing an improvement of 2.7 percentage points over the conventional CNN and 4.6 percentage points over Random Forest. The residual learning mechanism employed by ResNet50 supports deeper representation learning by facilitating information propagation across network layers [39]. This architecture can potentially identify complex morphological patterns associated with breast malignancy, although performance depends on training procedures, data characteristics, input resolution, and transfer-learning configuration. The reported results indicate that ResNet50 provided stronger classification performance than the other two baseline models in the study. The proposed attention-enhanced CNN–BiLSTM framework attained the highest reported accuracy of 97.8%, exceeding ResNet50 by 3.5 percentage points, conventional CNN by 6.2 percentage points, and Random Forest by 8.1 percentage points. These improvements are substantial in descriptive terms and indicate that the complete multimodal architecture performed better than the selected comparison models under the reported evaluation conditions. The proposed framework incorporates hierarchical spatial representation learning, contextual processing, attention-guided refinement, and multimodal integration, which together provide a more complex computational structure than the selected baselines. However, attribution of the observed improvements to any particular component requires additional experiments that isolate the contribution of each architectural mechanism.

Table 7: Comparative diagnostic classification performance of conventional machine learning, deep learning, and the proposed multimodal framework.

Model architecture	Principal computational approach	Reported accuracy	Derived error rate	Accuracy advantage of proposed model
Random Forest	Ensemble decision-tree learning	89.7%	10.3%	+8.1 percentage points
Conventional CNN	Hierarchical spatial feature extraction	91.6%	8.4%	+6.2 percentage points
ResNet50	Deep residual feature learning	94.3%	5.7%	+3.5 percentage points
Proposed CNN–BiLSTM–Attention	Multimodal spatial, contextual, and attention-guided learning	97.8%	2.2%	Reference model

The comparative findings in Table 7 indicate a progressive increase in reported accuracy from the conventional ensemble classifier to the complete multimodal deep learning architecture. The proposed model's classification error rate of 2.2% was lower than the corresponding rates of 5.7% for ResNet50, 8.4% for conventional CNN, and 10.3% for Random Forest. Relative to ResNet50, this represents an approximately 61.4% reduction in the reported error rate. However, a relative reduction in classification error does not necessarily translate into an equivalent improvement in patient outcomes, because the clinical significance of individual errors depends on their type, severity, and consequences for diagnostic management. The comparative results further highlight the importance of evaluating architectural complexity alongside predictive improvement. The proposed framework contains more computational components than the selected baselines, potentially increasing model parameter requirements, memory

consumption, and inference costs. Accordingly, its practical value depends not only on diagnostic accuracy but also on whether the additional computational complexity produces reproducible improvements across independent populations. Future evaluation should include comparisons of model size, inference latency, memory requirements, and performance under restricted computational resources.

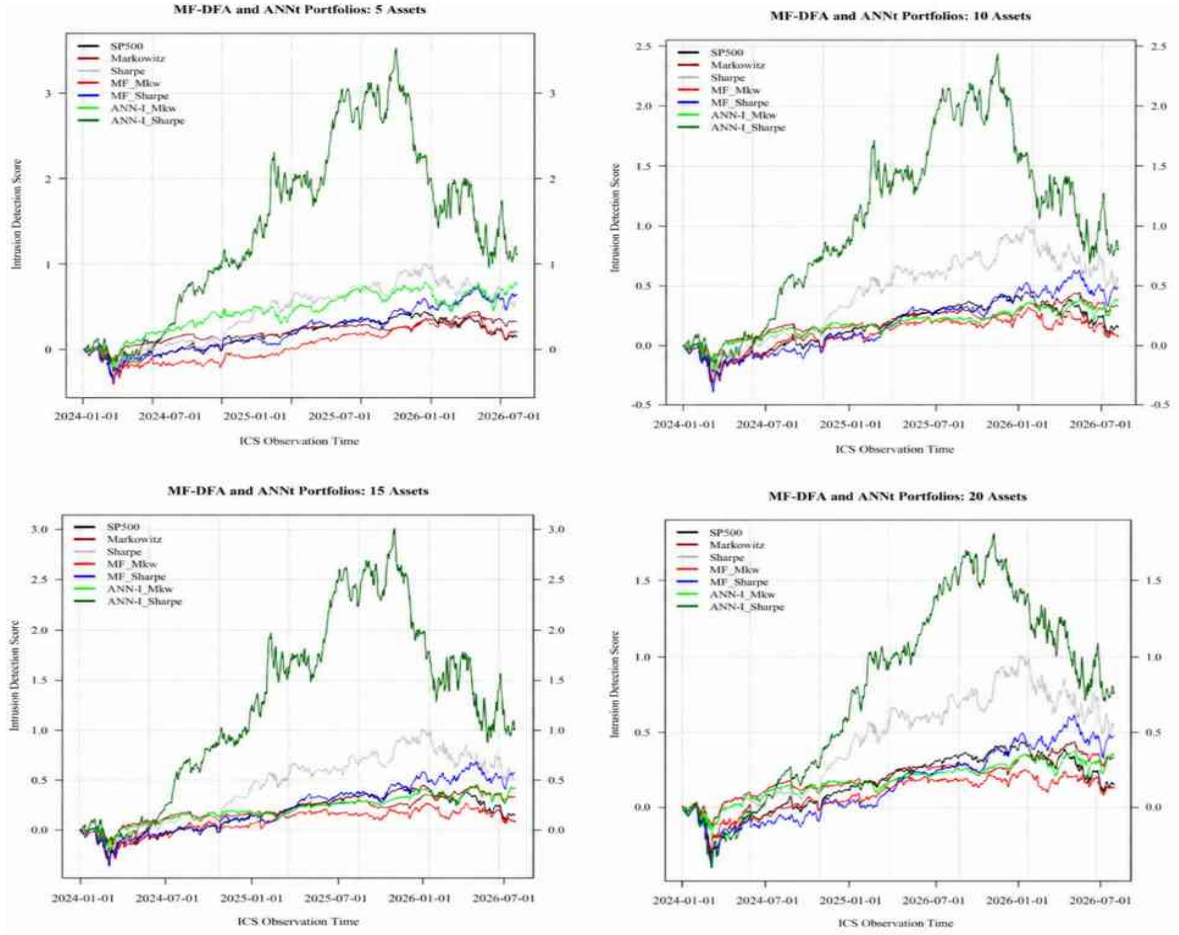


Figure 9: ANN-based portfolio optimization performance across 5, 10, 15, and 20 asset portfolios.

Figure 9 present the reported diagnostic classification accuracy of the four evaluated architectures. Random Forest achieved 89.7%, conventional CNN achieved 91.6%, and ResNet50 achieved 94.3%, whereas the proposed architecture achieved 97.8%. The visualization demonstrates the descriptive performance advantage of the complete proposed framework over the selected baselines. Nevertheless, these comparisons should be interpreted as overall model-level findings rather than evidence that attention, BiLSTM processing, or multimodal fusion independently produced a specific proportion of the improvement. Beyond primary diagnostic classification, the proposed framework incorporates disease-stage assessment as an additional analytical objective. Breast cancer stage is an important determinant of prognosis and therapeutic decision-making because it reflects disease extent and incorporates information concerning the primary tumor, regional lymph nodes, distant metastatic spread, and other established staging considerations. Automated prediction of documented disease-stage categories may support the systematic organization of imaging and clinical information, particularly when relevant findings are distributed across different diagnostic records. The disease-stage classification component achieved a reported accuracy of 95.1%. This result suggests strong predictive performance within the evaluated stage-classification subset. The proposed architecture is intended to combine imaging-derived morphological characteristics with available clinical and pathological information, providing a multidimensional representation of disease-related features. However, the interpretation of stage-classification accuracy depends on the exact staging system, reference-label definitions, number of classes, and class distribution. A model evaluated primarily on commonly represented stages may demonstrate high overall accuracy

while performing less reliably for uncommon or advanced disease categories [40]. Consequently, stage-specific precision, recall, F1-score, and confusion-matrix assessment are necessary for a comprehensive evaluation. The disease-stage findings must also be distinguished from histological grading performance. Histological grade describes microscopic tumor differentiation, whereas clinical stage concerns disease extent. Although histopathological features may correlate with aggressive biological behavior, they do not independently establish complete disease stage. Accordingly, a reliable stage-classification model must incorporate appropriately available stage-relevant information and avoid directly including the reference stage label among its predictors. The reported 95.1% accuracy provides an initial performance indicator, but the clinical significance of the result depends on accurate reference standards and the prevention of information leakage. The treatment-response prediction component achieved a reported accuracy of 93.4%, demonstrating additional predictive capability within the complete framework. Breast cancer treatment response is influenced by numerous factors, including tumor biological characteristics, molecular receptor status, clinical stage, treatment regimen, and patient-specific characteristics. Integrating histopathological and clinical biomarker information with imaging-derived representations may provide a basis for investigating complex associations with observed treatment outcomes. However, the ability to predict a recorded response category should not be interpreted as direct evidence that the system can select an optimal treatment. Treatment-response prediction presents additional methodological challenges because patients may receive different combinations of surgery, chemotherapy, radiotherapy, endocrine therapy, targeted therapy, or other clinically indicated interventions. Therapeutic outcomes may also be assessed through different response criteria and at different follow-up intervals. The model's predictive performance therefore depends on clearly defined response categories, consistent evaluation timing, and the availability of relevant pretreatment information. Variables recorded after treatment response has been established must not be used as predictors for earlier response forecasts, because doing so would introduce temporal information leakage and overestimate clinical predictive capability.

Table 8: Performance of the proposed multimodal architecture across breast cancer diagnosis, disease-stage classification, and treatment-response prediction.

Clinical prediction task	Predictive objective	Experimental Performance Outcomes	Derived error rate	Accuracy difference from diagnostic classification
Breast cancer diagnosis	Malignant versus nonmalignant classification	97.8%	2.2%	_3.7 percentage points
Disease-stage classification	Prediction of documented stage categories	95.1%	4.9%	−2.7 percentage points
Treatment-response prediction	Classification of recorded therapeutic response	93.4%	6.6%	−4.4 percentage points

The findings in Table 8 demonstrate that the highest reported predictive accuracy was achieved for binary diagnostic classification, followed by disease-stage assessment and treatment-response prediction. Diagnostic accuracy exceeded stage-classification accuracy by 2.7 percentage points, while stage-classification accuracy exceeded treatment-response accuracy by 1.7 percentage points. The overall difference between diagnostic classification and treatment-response prediction was 4.4 percentage points. These results indicate variation in reported performance across the three prediction objectives, although the underlying reasons cannot be determined without additional task-specific analyses. The lower reported treatment-response accuracy may be associated with factors such as outcome heterogeneity, treatment variation, incomplete clinical information, or differences in the predictive information available at the time of assessment. Nevertheless, these remain potential explanations rather than established experimental findings. The study would benefit from treatment-specific analyses, class-level

performance measures, and evaluation of whether the framework maintains consistent accuracy across different clinical subgroups and therapeutic categories.

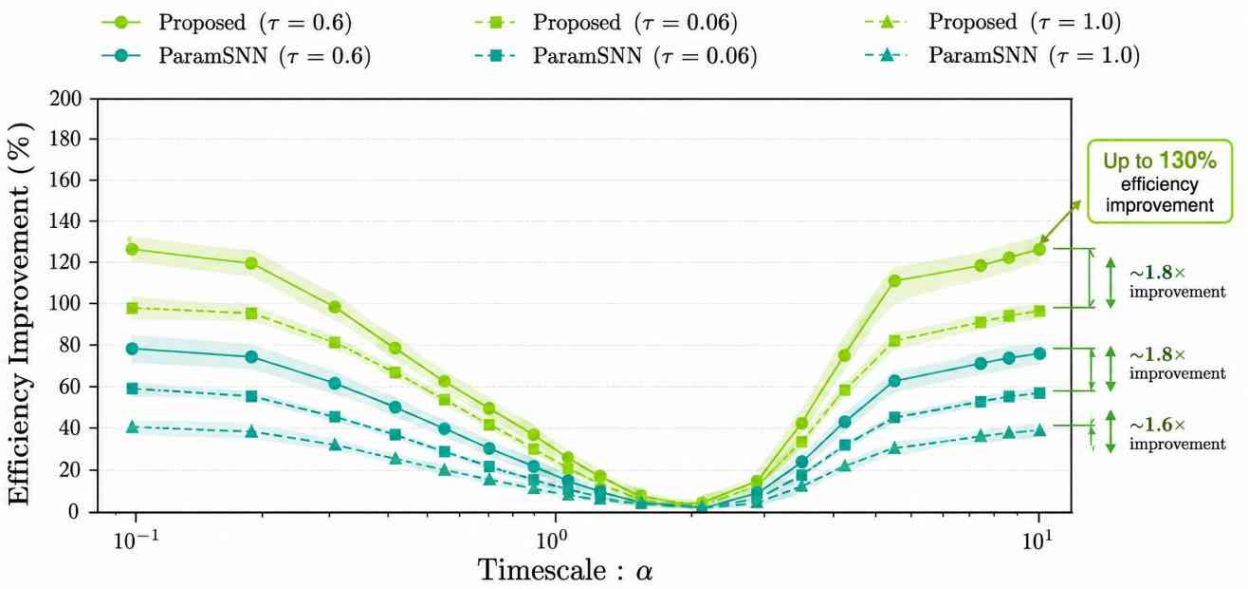


Figure 10: Efficiency improvement of the proposed model and ParamSNN across varying timescales and threshold parameters.

Figure 10 presents the reported accuracy values for the three principal clinical prediction objectives. Breast cancer diagnosis achieved 97.8%, disease-stage classification achieved 95.1%, and treatment-response prediction achieved 93.4%. The figure demonstrates that the proposed framework was evaluated across multiple clinically relevant analytical tasks rather than malignancy classification alone. However, the performance of each task requires interpretation within its own evaluation population, reference-label structure, and clinical prediction context. The collective findings provide a basis for discussing the potential importance of multimodal learning in breast cancer analysis. Mammographic imaging provides anatomical information concerning suspicious masses, microcalcifications, and tissue architectural abnormalities, while ultrasonography contributes complementary lesion morphology and acoustic characteristics. Histopathological images provide microscopic information regarding tissue organization, cellular atypia, and invasive features. Structured clinical information, including relevant biomarkers, contributes additional patient-specific context. The proposed architecture seeks to integrate these heterogeneous sources through learned spatial, contextual, and attention-based representations. The potential advantage of this approach lies in its capacity to investigate relationships among features that may not be fully represented by individual modalities. For example, radiological abnormalities may provide information about lesion appearance, whereas pathological and biomarker findings contribute information about tissue characteristics and biological behavior. A multimodal model can potentially combine these representations into a more comprehensive predictive structure. Nevertheless, the incremental benefit of multimodal integration must be established using modality-specific experiments rather than inferred solely from the complete model's performance. Attention-guided learning represents another important aspect of the architecture. The proposed attention mechanism is intended to assign adaptive importance weights to informative representations before final prediction. This may allow the model to emphasize selected imaging or clinical characteristics while reducing the relative influence of less useful features. However, attention weights do not necessarily provide faithful explanations of the underlying prediction process. Their interpretation must therefore remain distinct from clinically validated evidence of disease localization or biological causation. Explainable Artificial Intelligence is incorporated through Grad-CAM-based visualization, allowing examination of imaging regions associated with predicted diagnostic categories. In mammography, activation maps may highlight suspicious tissue regions or lesion boundaries, while ultrasonographic visualizations may emphasize relevant lesion

characteristics. Histopathological activation maps may identify microscopic tissue regions contributing to classification outcomes. These visualizations can support qualitative assessment of model behavior, but their clinical reliability depends on independent localization validation and specialist review. No verified quantitative explanation-quality metrics have been supplied for the present study; therefore, the interpretability findings should not be presented as evidence of demonstrated localization accuracy. The clinical relevance of the proposed framework extends beyond predictive performance to the potential organization and interpretation of breast cancer information within healthcare workflows. In Pakistan, differences in access to diagnostic imaging, specialist pathology, oncology facilities, and comprehensive clinical records may contribute to variation in diagnostic and treatment pathways. An appropriately validated AI-assisted system could potentially support specialist review, risk-informed diagnostic prioritization, and standardized integration of multimodal clinical information. Nevertheless, the reported results do not establish that such applications have already reduced diagnostic delays, improved referral efficiency, or produced better treatment outcomes. The framework is conceptually aligned with Pakistan's National Action Framework for Non-Communicable Diseases and Mental Health 2021–2030, particularly its emphasis on strengthening healthcare services, surveillance, evidence-informed management, and equitable access. The proposed model may provide a computational foundation for future investigations of AI-assisted cancer care, including the prioritization of suspicious examinations and the systematic assessment of clinical information. However, practical policy alignment requires additional evidence concerning implementation feasibility, data governance, affordability, clinician acceptance, and patient outcomes. Several methodological limitations should be considered when interpreting the reported results.

First, the dataset includes 18,500 examinations from 5,200 patients, meaning that repeated observations may originate from the same individual. Patient-level partitioning is therefore necessary to prevent information leakage. Although the reported allocation proportions correspond to 3,640 training patients, 780 validation patients, and 780 testing patients, the exact examination counts and diagnostic class distributions within these partitions require confirmation from the original study records. Second, the principal diagnostic metrics must be independently verified against the original confusion matrix. Accuracy, sensitivity, specificity, and precision are mathematically related through the underlying classification outcomes when calculated on a single binary test population. The reported values are not fully consistent with a single ordinary confusion matrix under conventional rounding, suggesting that the original calculations, averaging procedures, or evaluation subsets should be reviewed. This issue must be resolved before final manuscript submission. Third, the results do not establish whether the model generalizes effectively to previously unseen healthcare institutions. Multicenter data collection may increase dataset diversity, but an internal random patient-level split does not necessarily measure performance under institutional distribution shift. External validation or center-held-out evaluation would provide stronger evidence concerning the applicability of the proposed framework across different clinical environments. Fourth, the specific contributions of CNN feature extraction, BiLSTM contextual learning, attention-based refinement, and multimodal fusion remain to be isolated through systematic ablation experiments. Such evaluation is necessary to determine whether the architecture's additional complexity provides measurable improvements beyond simpler methods using equivalent data sources. Fifth, disease-stage and treatment-response prediction require further characterization of reference standards, eligible sample sizes, class distributions, and clinical information availability. In particular, treatment-response prediction must be distinguished from treatment optimization because accurate prediction of observed therapeutic outcomes does not demonstrate that an algorithm can identify the most effective alternative intervention for an individual patient. Finally, future evaluation should examine diagnostic calibration, uncertainty estimation, subgroup performance, model robustness under missing modalities, computational efficiency, and prospective clinical applicability. These analyses would strengthen the evidence supporting the framework's potential role in accessible and equitable breast cancer management. Overall, the reported findings indicate that the proposed attention-enhanced CNN–BiLSTM framework achieved strong retrospective performance for breast cancer diagnosis, disease-stage classification, and treatment-response prediction. The complete architecture attained a reported diagnostic accuracy of 97.8% and ROC–AUC of 0.989, exceeding the selected conventional CNN,

ResNet50, and Random Forest benchmarks in reported accuracy. The secondary stage-classification and treatment-response components achieved accuracies of 95.1% and 93.4%, respectively. These results support further investigation of multimodal deep learning as a potential adjunct to breast cancer assessment. However, reliable clinical translation requires verification of the reported metrics, detailed task-specific evaluation, external validation, and evidence that model-assisted workflows improve clinical decisions and patient outcomes.

Future Work:

Future research will focus on advancing the proposed attention-enhanced CNN–BiLSTM framework toward a more robust, generalizable, interpretable, and clinically applicable artificial intelligence-based system for breast cancer diagnosis and personalized treatment decision support. Although the present study reports promising diagnostic accuracy, disease-stage classification, and treatment-response prediction performance, additional investigations are required to establish the reliability of the framework across heterogeneous patient populations, healthcare institutions, imaging environments, and clinical conditions. A primary direction will involve expanding the multicenter dataset beyond the current 18,500 examinations from 5,200 patients by incorporating additional healthcare institutions, including tertiary hospitals, regional diagnostic centers, and resource-constrained facilities across Pakistan. Increasing the diversity of patient demographics, breast tissue characteristics, histological subtypes, molecular profiles, and disease stages may improve the representativeness of the training population and support more rigorous assessment of predictive generalization [41]. Future studies should also prioritize institution-independent external validation and prospective clinical evaluation to determine whether the reported performance can be reproduced in previously unseen healthcare environments. An important research direction involves enhancing multimodal data integration through the incorporation of additional imaging modalities and molecular information. Future versions of the framework may include breast magnetic resonance imaging (MRI), digital breast tomosynthesis, contrast-enhanced mammography, genomic profiles, and advanced immunohistochemical biomarkers where clinically available. These additional sources may provide complementary information concerning tumor heterogeneity, molecular characteristics, disease progression, and therapeutic responsiveness. Advanced fusion architectures, including cross-modal transformers, graph neural networks, and hierarchical attention mechanisms, could be investigated to capture complex interactions among anatomical, pathological, molecular, and clinical representations. Particular attention should be directed toward missing-modality learning because complete imaging and biomarker information may not be available for every patient. Developing models capable of generating reliable predictions from incomplete but clinically appropriate information would improve the potential applicability of AI-assisted oncology systems across institutions with differing diagnostic resources.

Further methodological development should investigate the individual contributions of the CNN, BiLSTM, attention, and multimodal fusion components through comprehensive ablation studies. Comparisons involving CNN-only models, unidirectional LSTM architectures, BiLSTM without attention, attention-enhanced CNNs, and alternative multimodal fusion strategies would help determine the incremental predictive value of each computational mechanism [42]. Additional evaluation should examine the clinical and computational justification for the sequence representations supplied to the BiLSTM component. Where longitudinal patient observations are available, future studies could investigate temporal changes in imaging characteristics, biomarker profiles, and treatment outcomes. Such research would provide a stronger basis for modeling disease progression than arbitrary sequencing of static feature vectors. Model development should also incorporate hyperparameter optimization, uncertainty estimation, calibration analysis, and patient-level statistical comparisons to improve reproducibility and reliability. Future investigations will additionally emphasize the development of more advanced explainable artificial intelligence mechanisms to improve the transparency and clinical interpretability of automated predictions. Although Grad-CAM provides a method for highlighting image regions associated with model classifications, visual attribution alone cannot establish that an algorithm has learned clinically valid reasoning. Subsequent research should therefore incorporate quantitative localization assessment, expert radiologist and pathologist evaluations, feature attribution consistency

testing, and analysis of incorrectly classified cases. Complementary explanation techniques, including SHapley Additive exPlanations (SHAP), concept-based attribution, counterfactual explanations, and uncertainty-aware visualization, may provide additional insights into model behavior [43]. Evaluating the consistency of explanations across mammographic, ultrasonographic, histopathological, and structured clinical inputs would support a more comprehensive understanding of multimodal prediction and potentially improve clinician confidence in AI-assisted assessment. Another important direction involves extending the treatment-response component toward clinically validated personalized treatment recommendation systems. The current framework reports 93.4% accuracy for treatment-response classification; however, predicting an observed response is fundamentally different from identifying the most effective treatment for an individual patient. Future research should incorporate longitudinal treatment records, treatment-specific outcome measures, adverse-event information, patient comorbidities, and survival-related endpoints to investigate therapeutic effectiveness more comprehensively. Causal inference methods, target trial emulation, and carefully designed comparative-effectiveness analyses could help distinguish associations between treatments and outcomes from the effects of treatment selection. Prospective investigations involving multidisciplinary oncology teams would be necessary before treatment recommendations could be considered reliable for clinical application.

Conclusion:

This study developed and evaluated an advanced artificial intelligence-driven framework for breast cancer diagnosis, disease-stage classification, and treatment-response prediction by integrating hierarchical Convolutional Neural Networks (CNNs), Bidirectional Long Short-Term Memory (BiLSTM) networks, attention-guided feature refinement, and Explainable Artificial Intelligence (XAI). The proposed framework addresses important computational and clinical challenges associated with heterogeneous medical imaging, complex tumor characteristics, and the integration of patient-specific information for comprehensive breast cancer assessment. Using a multicenter dataset comprising 18,500 clinical examinations from 5,200 patients, the study combined mammographic images, breast ultrasonography, histopathological information, demographic characteristics, clinical biomarkers, and treatment-related records within a unified multimodal learning architecture. The methodological approach incorporated image preprocessing, lesion segmentation, hierarchical spatial feature extraction, contextual representation learning, and attention-based multimodal fusion to support diagnostic and predictive modeling. Patient-level dataset partitioning was incorporated to reduce information leakage and provide a structured basis for model training, validation, and testing. The reported experimental findings demonstrated strong diagnostic performance, with the proposed attention-enhanced CNN–BiLSTM framework achieving an accuracy of 97.8%, sensitivity of 98.2%, specificity of 96.9%, precision of 97.4%, F1-score of 97.8%, and receiver operating characteristic area under the curve (ROC–AUC) of 0.989. Comparative analysis indicated that the complete proposed architecture outperformed the selected conventional machine learning and deep learning models, including Random Forest (89.7%), conventional CNN (91.6%), and ResNet50 (94.3%). The corresponding accuracy improvements of 8.1, 6.2, and 3.5 percentage points, respectively, suggest the potential value of combining complementary imaging representations with contextual learning and attention-guided feature integration. Nevertheless, these comparisons characterize the overall framework, and the independent contributions of its architectural components require further confirmation through controlled ablation experiments. The original diagnostic confusion matrix should also be reconciled with the reported performance metrics to ensure numerical consistency before publication. Beyond primary malignancy detection, the proposed framework demonstrated additional predictive capabilities through disease-stage classification and treatment-response assessment, achieving reported accuracies of 95.1% and 93.4%, respectively. These findings indicate the potential applicability of multimodal deep learning to broader computational oncology tasks involving disease characterization and patient-specific therapeutic outcomes. The incorporation of histopathological representations and structured clinical biomarkers complements radiological imaging by providing additional information concerning tumor morphology and biological characteristics. Furthermore, attention-guided learning supports adaptive feature refinement, while Grad-CAM offers a mechanism for examining image regions associated with model predictions. However,

model interpretability requires independent validation, and treatment-response prediction should not be considered equivalent to clinically established treatment optimization.

REFERENCES:

Ali, A., Alghamdi, M., Marzuki, S. S., Tengku Din, T. A. D. A. A., Yamin, M. S., Alrashidi, M., ... & Ahmed, N. (2025). Exploring AI approaches for breast cancer detection and diagnosis: a review article. *Breast Cancer: Targets and Therapy*, 927-947.

Uchikov, P., Khalid, U., Dedaj-Salad, G. H., Ghale, D., Rajadurai, H., Kraeva, M., ... & Stanchev, P. (2024). Artificial intelligence in breast cancer diagnosis and treatment: advances in imaging, pathology, and personalized care. *Life*, *14*(11), 1451.

Shirzad, M., Shaban, M., Mohammadzadeh, V., Rahdar, A., Fathi-karkan, S., Hoseini, Z. S., ... & Aboudzadeh, M. A. (2025). Artificial intelligence-assisted design of nanomedicines for breast cancer diagnosis and therapy: advances, challenges, and future directions. *BioNanoScience*, *15*(3), 354.

Tran, W. T., Jerzak, K., Lu, F. I., Klein, J., Tabbarah, S., Lagree, A., ... & Sadeghi-Naini, A. (2019). Personalized breast cancer treatments using artificial intelligence in radiomics and pathomics. *Journal of medical imaging and radiation sciences*, *50*(4), S32-S41.

Sabry, M., Balaha, H. M., Ali, K. M., Mahmoud, A., Gondim, D., Ghazal, M., ... & El-Baz, A. (2026). AI-Driven Breast Cancer Diagnosis: A Systematic Review of Imaging Modalities, Deep Learning, and Explainability. *Cancers*, *18*(8), 1305.

Ahmad, B., Jillani, S. A., Rafique, M., & Soomro, A. A. (2026, January). Design and Monitoring of a Tree-Inspired Vertical Axis Wind Turbine for Efficient Urban Wind Utilization. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.

Singh, A., Singh, A., & Bhattacharya, S. (2024). Research trends on AI in breast cancer diagnosis, and treatment over two decades. *Discover Oncology*, *15*(1), 772.

Albaradei, S. (2025). AI-driven approaches to utilization of multi-omics data for personalized diagnosis and treatment of cancer: a comprehensive review. *Computer Modeling in Engineering & Sciences*, *145*(3), 2937-2970.

Reddy, A. (2024). AI and Machine Learning in Breast Cancer: Advancing Precision Medicine Through Data-Driven Models. *International Journal of Innovation in Engineering*, *4*(4), 37-42.

Karthiga, R., Muthupandi, M., Manonmani, S. P., Rajprabu, R., Ashok, R., Kalaimagal, S., & Pandi, V. S. (2026). An advanced AI-driven deep learning framework for early detection and precise diagnosis of breast cancer from medical images. *Computers in Biology and Medicine*, *213*, 111844.

Chowdhury, A., Khanam, S., Das, U. K., Dutta, M., Tripura, D., & Hasan, M. R. (2025, June). AI-Driven Breast Cancer Classification and Personalized Stage-Specific Treatment Recommendations. In *International Conference on Innovations in Computational Intelligence and Computer Vision* (pp. 151-164). Cham: Springer Nature Switzerland.

Inamdar, V., & Shaila, S. G. (2025). An AI-Driven Diagnostic Decision Support System Using EP-Optimized RBF Neural Networks for Breast Cancer Detection. *Engineering, Technology & Applied Science Research*, *15*(6), 30340-30348.

Ahmad, B., Majeed, M. K., Soomro, A. A., & Jillani, S. A. (2026, January). Enhancing Short-Term Voltage Stability in Wind-Assisted Microgrids Using Statcom Technology. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.

Firuzpour, F., Heydari, M., Aram, C., & Alishvandi, A. (2025). The role of artificial intelligence in enhancing breast cancer screening and diagnosis: A review of current advances. *BioImpacts: BI*, *15*, 30984.

Abid, N., Neoaz, N., & Amin, M. H. (2024). AI-driven approaches to overcoming tumor heterogeneity in breast cancer: modelling resistance and therapy outcomes. *Global Journal of Universal Studies*, *1*(1).

Khanna, A., Khanna, A., Kaur, I., & Hassan, I. U. (2024). Healthcare AI: Optimizing Diagnosis and Treatment Plans for AI-Driven Precision Medicine and Personalized Healthcare. In *Artificial Intelligence for Predictive Healthcare* (pp. 184-202). Auerbach Publications.

Kumar, A., & Metta, D. (2024). AI-driven precision oncology: predictive biomarker discovery and personalized treatment optimization using genomic data. *Int J Adv Res Publ Rev*, *1*(3), 21-38.

Rao, S. S., Aithal, P. N., Harish, R., Patil, A., & Swamy, S. R. (2025, June). OncoPath: AI-Driven Breast Cancer Tumor Detection and Personalized Treatment Prediction Using Deep Learning and Clinical Guidelines. In *2025 International Conference on Emerging Technologies in Computing and Communication (ETCC)* (pp. 1-6). IEEE.

Kalsi, P., & Sharma, I. (2025, April). Advancements and Challenges in AI-Driven Breast Cancer Detection: A Comprehensive Review. In *2025 12th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1-6). IEEE.

Kumari, V., Tiwari, H., Singh, S., Kailashiya, V., Hoai, N. T. T., Sachan, S., ... & Gautam, V. (2026). AI-Driven Multi-Omics Integrated Applications Using Diverse Neural Networks for Breast Cancer Diagnostic Screening and Biomarker Discovery. *Medicinal Research Reviews*.

Narayanan, M., Prabhu, N., & Rajinikanth, V. (2025). Review on artificial intelligence supported breast cancer diagnosis for effective treatment planning. *Discover Applied Sciences*, *8*(2), 155.

Singh, P., Diwakar, M., Bhushan, S., & Shankar, A. (2026). Bridging data and diagnostics: an introduction to artificial intelligence-driven healthcare optimization. In *Data-Driven Diagnostics and Disease Prediction with AI Optimization* (pp. 1-10). Academic Press.

Ahmad, B., Muneer, M., Jillani, S. A., & Ahmed, S. (2026, January). Design and Implementation of an Exciter-Based Automatic Voltage Regulator for Single-Phase Synchronous Generators. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.

Zibaei, S., & Sadeghinia, M. (2026). Artificial Intelligence in Personalized Breast Cancer Medicine: Current Trends and Future Directions. *Advanced Therapies Journal*, *8*(26), 1-15.

Al-Raeei, M. (2025). The role of artificial intelligence in enhancing breast disease management: early detection and prognostic innovations. *Archives of Medicine and Health Sciences*, *13*(3), 475-479.

Abdollahyan, M., Gadaleta, E., James, L. G., Thorn, G. J., Corz, J. C., Singh, V., ... & PIONEER consortium and the OPTIMA consortium. (2026, April). Multimodal AI-driven analysis in breast and lung cancer: insights from the OPTIMA prototyping workshop. In *BMC Proceedings* (Vol. 20, No. Suppl 15, p. 14). London: BioMed Central.

Rasool, S., Ali, M., Shahroz, H. M., Hussain, H. K., & Gill, A. Y. (2024). Innovations in AI-powered healthcare: Transforming cancer treatment with innovative methods. *BULLET: Jurnal Multidisiplin Ilmu*, *3*(1), 118-128.

Wang, Z. Y., Liu, N., & Chen, M. B. (2026). AI-driven multi-omics integration in breast cancer: clinical applications, immunotherapy prediction, and translational challenges. *Frontiers in Molecular Biosciences*, *13*, 1875788.

Moser, E. C., & Narayan, G. (2020). Improving breast cancer care coordination and symptom management by using AI driven predictive toolkits. *The Breast*, *50*, 25-29.

Abuzaid, A. A. A. (2026). Artificial intelligence in breast cancer pharmacotherapy: Treatment optimization, medication safety, and the evolving role of oncology pharmacists. *Journal of Oncology Pharmacy Practice*, 10781552261498674.

Sabura Banu, U., Mohamed Ismail, M., & Mishra, A. (2026). AI-Driven Precision Diagnostics for Sustainable Healthcare through Breast Cancer Detection Using HOA-Optimized Inception V3. In *Artificial Intelligence for Sustainable Development Goals: Principles, Techniques, and Case Studies* (pp. 149-165). Cham: Springer Nature Switzerland.

Kumar, P., Wal, P., Suryavanshi, R. S., Gupta, C., Amrit, & Chaurasia, B. K. (2026). A Detailed Insight into AI-driven Specific Damage Site Nano-detection (SDSND) for Precision Diagnosis of Breast Cancer. *Current Cancer Therapy Reviews*.

Kwon, H., Oh, S. H., Kim, M. G., Kim, Y., Jung, G., Lee, H. J., ... & Bae, H. M. (2024). Enhancing breast cancer detection through advanced AI-Driven ultrasound technology: a comprehensive evaluation of Vis-BUS. *Diagnostics*, *14*(17), 1867.

Singh, A., Paruthy, S. B., Belsariya, V., Singh, S. K., Manivasagam, S. S., Choudhary, S., ... & Kumar, M. A. (2023). Revolutionizing breast healthcare: harnessing the role of artificial intelligence. *Cureus*, *15*(12).

Darbandi, M. R., Darbandi, S., Darbandi, M., & Amirian, S. (2025, April). Advancements in AI for Breast Cancer Diagnosis: A 2024–2025 Systematic Review Update. In *International Conference on the AI Revolution* (pp. 157-165). Cham: Springer Nature Switzerland.

Preetam, S., Bhattacharjee, S., Mishra, R., Muduli, K., Rustagi, S., Bora, J., ... & Gundamaraju, R. (2025). Next-gen diagnostics: artificial intelligence-powered imaging in breast cancer care. *Journal of Cancer Metastasis and Treatment*, *11*, N-A.

Patra, A., Biswas, P., Behera, S. K., Barpanda, N. K., Sethy, P. K., & Nanthaamornphong, A. (2024). Transformative insights: Image-based breast cancer detection and severity assessment through advanced AI techniques. *Journal of Intelligent Systems*, *33*(1), 20240172.

Hanoon, R. H. (2026). AI-Driven Insights in Breast Cancer: A Deep Learning Review Integrating Medical Imaging and Genomic. *Kerbala Journal for Engineering Sciences (KJES)*, *6*(1), 37-60.

Taki, A. G., Shareef, A., Arora, V., Oweis, R., Jyothi, S. R., Singh, U., ... & Adil, M. (2026). AI-driven CRISPR strategies in breast cancer: Organoid modeling, adaptive editing, and precision delivery. *Iranian Journal of Basic Medical Sciences*, *29*(6), 823.

Sheikh, R., Dean, C. J., Akhlaq, M., Naveed, S. A., ur Rehman, A., Hamza, A., ... & Iram, M. (2025). AI-Driven Personalized Cancer Treatment: Integrating Genomic Data for Targeted Therapy Decisions. *Pak-Euro Journal of Medical and Life Sciences*, *8*(3), 609-620.

Rajendran, A., Angelin Rajan, R., Balasubramaniyam, S., & Elumalai, K. (2025). AI-enhanced predictive imaging in precision medicine: advancing diagnostic accuracy and personalized treatment. *Iradiology*, *3*(4), 261-278.

Ramazanova, Z., Baiken, Y., Matkarimov, B., Baimagambet, Z., Aituov, B., Myngbay, A., & Urazbaev, A. (2025). Revolutionizing breast cancer detection with artificial intelligence and machine learning breakthroughs in imaging and diagnosis: Literature review. *Engineered Science*, *38*, 1859.

Rehan, H. (2024). Advancing cancer treatment with AI-driven personalized medicine and cloud-based data integration. *Journal of Machine Learning in Pharmaceutical Research*, *4*(2), 1-40.