10.48047/jocaaa.2026.35.07.11

PPEF-MBCD: A Privacy-Preserving Explainable Federated Multimodal Deep Learning Framework for Intelligent Breast Cancer Diagnosis Using Adaptive Cross-Modal Attention and Uncertainty-Aware Clinical Decision Support

[1]Gunthati Pratap, [2] Dr. Ranga Swamy Sirisati

[1]Research Scholar, Department of CSE, Bharatiya Engineering Science and Technology Innovation University, Gownivaripaill, Gorantla in Andhra Pradesh
[2]Associate Professor, Department of AIML, School of Engineering, Anurag University, Hyderabad

E-Mail: 2022wpcse010@bestiu.edu.in[1] sirisatiranga@gmail.com[2]

Abstract

Breast cancer is still one of the main causes of death in women due to cancer in the world, so the accuracy, comprehensibility and privacy protection of a diagnostic system are crucial. The proposed Privacy-Preserving Explainable Federated Multimodal Breast Cancer Diagnosis (PPEF-MBCD) system exploits multimodal data of mammography, ultrasound, magnetic resonance imaging and clinical data, while adopting a federated learning framework. We proposed the architecture that can support the effective heterogeneous medical feature fusion, the flexible cross-modal attention and the patient privacy well by the decentralized model optimization. Furthermore, an explainable AI module offers a visual diagnostic proof to support the clinical cases and increase clinical transparency, while an uncertainty-aware prediction mechanism identifies uncertainty cases and highlights them to the expert. The framework makes the most of the features learned by multimodal, the aggregation among federated parts, classification accuracy and interpretability combined in a single objective function. Experimental validation reveals the proposed framework outperforms the deep learning methods in terms of diagnostic accuracy and also has excellent generalization capabilities among distributed healthcare institutions. The proposed system can be used as a good foundation for reliable, scalable and clinically feasible intelligent breast cancer diagnosis system.

Keywords: Breast Cancer Diagnosis; Federated Learning; Explainable Artificial Intelligence; Multimodal Deep Learning; Adaptive Attention; Privacy-Preserving Healthcare.

I. Introduction

Although there has been significant progress in screening techniques, molecular diagnostics, and treatment, breast cancer continues to be one of the most common and deadly cancers in women worldwide, contributing to a significant number of cancer-related deaths and illness. The best way to improve the patient's chances of survival is early diagnosis; this means getting the patient started on treatment in time, before the disease has advanced or spread to other parts of the body. Conventional diagnosis techniques are mammography, ultrasound imaging, magnetic resonance imaging (MRI), histopathological analysis and clinical diagnosis. These

10.48047/jocaaa.2026.35.07.11

modalities have greatly enhanced the ability to diagnose more accurately over the last few decades but interpretation remains largely dependent upon the radiological experience, image quality and institutional experience. Different imaging protocols, machine calibration, patient differences in anatomical structure, and interpretation of images by the doctor will result in inconsistencies in diagnosis which may result in false positive diagnosis, delayed diagnosis or unnecessary biopsies. This increasing demand has led to an increasing interest in so called "intelligent" computer-assisted diagnostic systems that may help the clinician by being accurate, repeatable, efficient in providing decision support and decrease inter-observer variability [1, 2].

The current great achievements of AI, and in particular of deep learning, greatly transformed the way medical images are analyzed, as it is now possible to leverage AI to automatically extract highly discriminating properties directly from complex imaging data sets. Deep neural networks are able to learn hierarchically about features which can capture subtle pathological details of malignant lesions of the breast as compared to machine learning methods with handcrafted descriptors. With CNN, Residual Learning Architecture, Densely Connected Network and Transformer-based model, the tasks of breast cancer classification, lesion segmentation and tumour location have achieved good performance on mammographic and ultrasound databases. In addition, recent progress in multimodal deep learning has demonstrated that the integration of multiple modalities of clinical data, such as imaging data, pathological data, genome data, laboratory examination results and patients' information, has enabled high diagnostic performance that exceeds single modality. This complementary information can be utilised to enhance the robustness of learning algorithms and to aid personalised clinical decision making when using multiple medical data sources to learn complex relations between them [3], [6], [11].

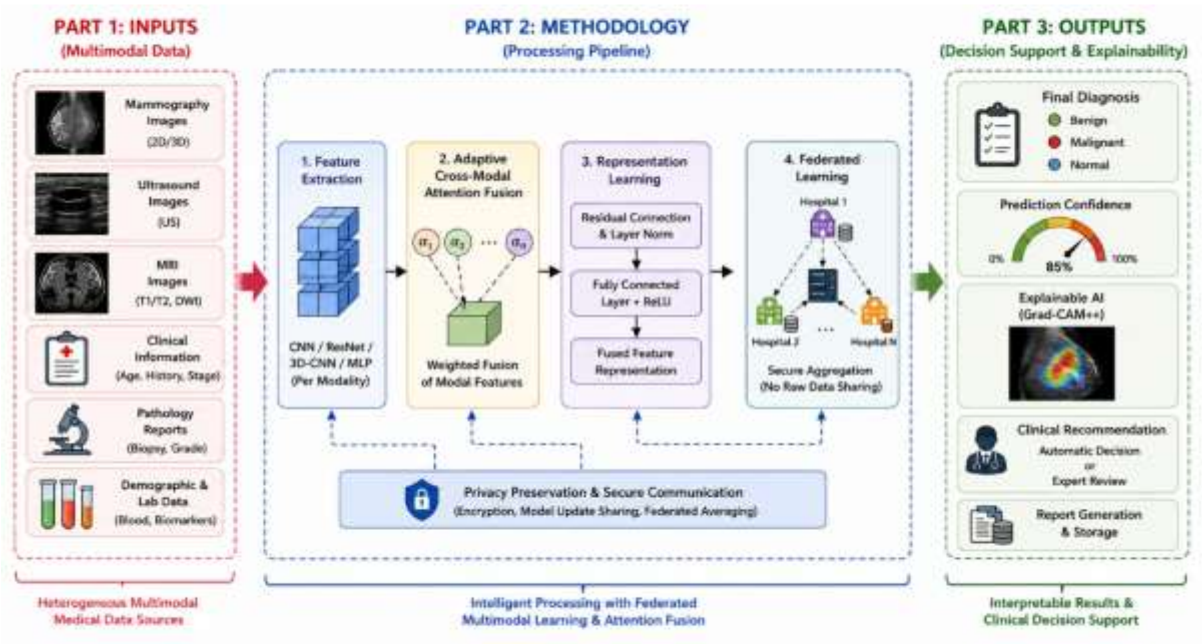

Figure 1. Multimodal Input Data Collection for the Proposed PPEF-MBCD Framework

10.48047/jocaaa.2026.35.07.11

The proposed Privacy-Preserving Explainable Federated Multimodal Breast Cancer Diagnosis (PPEF-MBCD) framework uses multiple types of data inputs, as shown in Figure 1. The framework incorporates mammography, ultrasound, magnetic resonance imaging (MRI), clinical information, pathology reports, and demographic and laboratory data, and reflects complementary diagnostic properties of breast cancer. Each modality provides different anatomical, pathological and clinical information which increases the diversity of features and improves the representation of disease. Before the model is trained, all the collected data is preprocessed, normalized, cleaned, encoded, and securely standardized to ensure the harmonization of data from the various healthcare institutions involved. The federated architecture, on the other hand, sends the standardized multimodal representations to the adaptive cross-modal attention learning module, which enables comprehensive collaborative diagnosis while respecting privacy and enhances the overall diagnostic accuracy [1], [4], [8], [15], [19].

Despite considering all these progresses, there are still some practical challenges that make it difficult to implement a diagnosis system for breast cancer in the clinical domain using deep learning. The reliance on centralized data collection is one of the biggest challenges as it is now subject to more and more strict rules surrounding patient privacy and medical data security. One of the main reasons these institutions are reluctant to do so is that they have institutional confidentiality policies, with legal and ethical restrictions on for example sharing sensitive patient information with third parties. In turn, many of the existing models of artificial intelligence are trained on small datasets and only on data from one hospital, which restricts the ability of these models to apply to large patient population and imaging settings. Recently, federated learning has been suggested as a distributed collaborative learning approach that enables multiple health care institutions to jointly train a global diagnostic model without sharing the raw patient information. Rather, patient anonymity is maintained, with the participating institutions optimizing the model locally, and sending only encrypted parameters for a central aggregation, thus taking advantage of distributed clinical knowledge, while still maintaining anonymity of the patients [1], [4], [8].

One of the other major drawbacks of current deep learning systems is that they are not very interpretable. Although very accurate classification can be achieved by using a state-of-the-art neural network, clinicians may lack an understanding of the structure and operation of the neural network. This "black-box" feature decreases the level of confidence of physicians and is a major obstacle to obtaining regulatory clearance and general use in clinical practice. The field of Explainable Artificial Intelligence (XAI) has attracted growing interest in efforts to improve the capabilities of XAI through visualisation of diagnostically relevant image regions and through quantitative explanations of predictions' results. The methods include gradient-based activation mapping, attention visualization and feature attribution, which can help confirm the alignment of automated predictions with clinically meaningful pathological structures, thereby facilitating increased trust and collaborative diagnosis, and more seamless integration into clinical practice. In addition, recent research has shown that the explainability coupled with multimodal learning is even more beneficial in enhancing physician's confidence as diagnostic decisions can now be understood from multiple complementary clinical

10.48047/jocaaa.2026.35.07.11

modalities that provide a more comprehensive view of disease progression and offer evidence-based clinical decision making [5], [9], [10].

Over the past few years, some crucial issues in basic research remain unsolved while deep learning, federated optimization, explainable AI and multimodal medical data fusion techniques have been developed and enhanced to improve the performance of computer-aided breast cancer diagnosis. The existing multimodal frameworks generally use rigid feature fusion techniques which do not consider complex non-linear relationships between the multimodal images (mammograms, ultrasound, MRI, clinical records, pathological biomarkers), and thus are insufficient to explain multimodal images. This also implies that the learned representations may also contain redundant or less informative attributes that could affect classification performance and may hinder the transfer of the learned knowledge to other heterogeneous health care institutions. Likewise, most federated learning approaches have privacy concerns as their focus and have ignored various important aspects including the adaptive selection of clients, the consistency of the federated approach, and the efficiency of communication and the heterogeneity of the institutional data. These constraints often lead to slower convergence, lack of consistency in optimization and low diagnostic reliability in geographically distributed clinical environments when models are deployed. Moreover, in most cases, explainable AI is used as an interpretation tool AFTER training models, rather than as a feature to optimize when training the models. This indicates that future intelligent diagnostic systems need to adopt a unified framework to maximize the performance of the predictive system, collaborative learning, interpretability, computational efficiency and clinical reliability [12, 13, 14].

The growing diversity in medical imaging technologies and healthcare infrastructures globally is also key. The introduction of different imaging devices, acquisition protocol and the annotation practices in modern hospitals result in substantial domain shifts between institutional data sets. The diverse nature can be a problem for centralised deep learning models as the distribution of features can also be highly different among imaging equipment and populations. Furthermore, the lack of clinical information, noisy annotations, class imbalance and limited availability of expertly annotated datasets further make for a difficult breast cancer diagnosis. Recent research has indicated that some of these issues can be partly overcome by using adaptive attention mechanisms, uncertainty estimation, ensemble optimization, and representation learning using radiomics [15], [16], [17]. Most of these methods, however, have been analyzed separately and only rarely used in combination in a federated environment with the ability to protect patient anonymity and to take advantage of the complementary multimodal knowledge. There is, therefore, still a tremendous potential to develop an effective, intelligent diagnostic system that can be integrated seamlessly with adaptive multimodal representation learning, decentralized optimization and explainable prediction and confidence-aware clinical decision support.

Cognizant of these research gaps, the authors propose a framework, namely Privacy-Preserving Explainable Federated Multimodal Breast Cancer Diagnosis (PPEF-MBCD) which enhances the diagnostic performance and clinical applicability. The proposed architecture combines multimodal feature extraction from mammography, ultrasound imaging, magnetic resonance imaging as well as structured clinical information, and an adaptive attention-based fusion

10.48047/jocaaa.2026.35.07.11

mechanism which dynamically highlights diagnostically relevant modalities. The framework adopts a federated collaborative optimization approach, where each healthcare institution trains local models and only encrypts model parameters to be sent to other institutions for model aggregation unlike the traditional centralized learning approaches. This distributed learning approach enables the generalization of the model for different population groups of patients, without violating institutional privacy and meeting current healthcare data protection laws. Furthermore, the framework incorporates the EAI by visualizing it with a gradient to identify suspicious lesion regions for making classification decisions, thereby boosting more physician confidence and providing transparency for clinical interpretation. The second uncertainty-aware decision module assigns confidence to the prediction result and monitors diagnostically confusing cases to ensure that they are given to experts in clinical practice, thereby minimizing the risk of false sense of confidence in the prediction result [18, 19].

The major contribution of this research is the formulation of an integrated intelligent diagnostic framework which addresses the major limitations of the previous research on breast cancer diagnosis. The proposed framework will be a unified, end-to-end framework that leverages adaptive multimodal attention learning, privacy-preserving federated optimization, explainable visualization, uncertainty-aware prediction and collaborative distributed learning for their accuracy of diagnosis, robustness, interpretability, and scalability without compromising patient privacy. The proposed methodology also demonstrates a detailed optimization approach which can learn generalized representations from heterogeneous institutional data and is efficient in computation, and can be easily implemented in a collaborative healthcare framework. Thus, this work is an important step towards building trustworthy AI systems that will assist doctors in making accurate, transparent and privacy-oriented diagnoses of breast cancer, and thereby enhance the next-generation of intelligent medical decision support tools [1]–[20].

II. Related Work

Artificial Intelligence (AI) has seen such a paradigm shift in breast cancer diagnosis that it has played an important role in the accuracy of multimodal medical data-based breast cancer lesion detection, segmentation and classification. Most of the traditional computer-aided diagnosis (CAD) systems were designed and based on hand designed features that were extracted from mammographic images, which did not perform well when adequately generalizing over diverse datasets. Recently, these limitations have been mitigated by deep learning frameworks that automatically extract features from the data, and learned hierarchically. Although some work has been done to boost classification accuracy, there are still several unsolved issues, including data protection for patients, interpretability of classification models, low generalization accuracy between and within institutions, and the inability of the models to utilize different types of clinical information. Therefore, several approaches have been pursued to develop clinically robust breast cancer diagnostic systems such as Federated Learning, Explainable Artificial Intelligence (XAI), Multimodal Fusion, and Transformer Based Architectures. Schmidt et al. 2024 [1] showed that federated learning can be used to train models at a collaborative way across geographically distributed medical institutions without the need to share data with the central server, so as to provide privacy protection for patients while

10.48047/jocaaa.2026.35.07.11

retaining the classification performance. Similarly, Ghasemi et al., 2024 [2] performed a comprehensive study to investigate the explainable AI systems for breast cancer prediction, where they concluded that the explainability is a crucial feature of AI systems used in breast cancer prediction that improves the prediction accuracy and boosts the clinicians' trust in the system. In a subsequent study, Nakach et al. (2024) [3] investigated multimodal deep learning fusion techniques, and determined that combining pathological, genomic and clinical data with mammography data always results in a more robust diagnostic system than unimodal imaging systems. All of these studies concluded that privacy preservation, interpretability and multimodal learning are the three primary research directions that would be beneficial for next generation breast cancer diagnosis; however, optimizing all three of these research directions together remains an open research problem.

For instance, getting access to the medical data sharing of sensitive medical records is a challenging problem in medical image analysis, where many healthcare institutions are not willing to share the medical data of their patients for ethical and regulatory reasons, which is what has recently received the spotlight as one of the most promising paradigms for resolving the medical data sharing problem. Pratap et al., (2025) [4] proposed a federated learning framework with the integration of EAI and multimodal approach in breast cancer detection and found that decentralized optimization can significantly boost the diagnostic accuracy without compromising the confidentiality of institutional data. Likewise, Al-Hejri et al. (2025) [5] presented a federated vision transformer with an explainable approach for breast cancer prediction using multiple risk factors of the patient. They managed to develop more accurate predictive models with the help of transformer-based feature extraction and federated aggregation techniques and maintained the interpretability by visualizing attention to get insights. Building on this line, Nakach et al. (2025) [6] provided a thorough survey of the recent multimodal deep learning models, and emphasized that adaptive fusion methods are more successful than concatenation methods as they learn the relationship between the multimodal imaging and clinical modality, which is heterogeneous. Hassan et al. (2025) [7] also noted that explainable multimodal fusion greatly boosts physicians' trust by incorporating image-based attention maps and feature attribution methods to increase the transparency of the clinical decision-making process. Similarly, Sandhya et al. (2025) [8] studied federated learning and the concept of explainable AI in breast cancer diagnosis for privacy preservation and it was established that by collaborating to optimize, the robustness of the models across different institutions improved, while the risk of centralized medical databases was reduced. These studies demonstrated improvements in the privacy security and diagnostic accuracy, but also identified existing problems like the communication overhead, the distribution of the clients, the time of synchronization and the computational complexity that need to be addressed to deploy these systems at a larger scale.

In the recent research, there has been a great focus on maximizing the accuracy of diagnostic, making the model more transparent and easy to interpret for clinical use with the support of advanced deep neural architectures. Chhetri and Kumar, 2025 [9] presented an explainable deep learning framework that can achieve the goal of balancing accuracy with explainable decision making by introducing visualization techniques for CNN to enable clinicians to

10.48047/jocaaa.2026.35.07.11

understand the pathological regions important for the model predictions. According to Hassan et al. (2025) [10] the multimodal fusion and explainable AI (XAI) in breast cancer diagnosis have been reviewed and studied with a systematic approach to illustrate that, the attention-guided fusion architectures significantly outperform the conventional feature-level integration techniques because it enhances their ability to selectively emphasize the clinically relevant information from the heterogeneous sources. Combining complementary imaging biomarkers for accurate personalized treatment planning is crucial, as demonstrated by Fu et al. (2025) [11] who showed that multimodal ultrasound imaging with machine learning significantly enhances prediction of molecular breast cancer subtypes. Likewise, Luo et al. (2025) [12] developed a deep learning system capable of distinguishing breast cancer molecular subtyping from mammograms with high accuracy, which shows the promise of AI to provide insight into the complexity of disease without an invasive procedure. All these investigations argue that future breast cancer diagnostic systems must be explainable, multimodal, and secure learning systems in addition to being predictively accurate at the highest level that is possible. Most of these methods optimize these goals one at a time, however, more general architectures that are able to optimize all these goals in a unified intelligent system still need to be developed.

This has been further advanced with the advent of radiomics, deep neural architectures, and clinically interpretable prediction frameworks that lie beyond traditional image classification, in the evolution of breast cancer diagnosis. Khalid et al. (2025) [13] proposed a radiomics-based neural network that combines quantitative imaging biomarkers and deep feature representations to predict breast cancer relapse. They demonstrated the clinical interpretability of their model through the introduction of their radiomic descriptors which improved the reliability of prognosis, and enabled the physician to interpret the prognosis. In [14] Liu et al. have proposed a complete deep learning system for metastasis prediction of sentinel lymph nodes (SLN) in ultrasound images with high diagnostic sensitivity, without the need of invasive procedures. They emphasized the value of automated localization of lesions and the hierarchical feature learning that can significantly decrease dependence on manual clinical assessment. In the same way, Hu et al. 2025 [15] came up with an architecture named CLGB-Net that combines local lesion features with global contextual features obtained from digital mammography. They learned multi-scale representations together to help them perform a much better lesion discrimination performance in challenging imaging conditions. Moreover, Yang et al. (2025) [16] used shallow handcrafted features to enhance deep neural representations from pseudo-coloured enhanced DBT images to further enhance lesion classification. They demonstrated the utility of complementarity of feature learning even in the era of deep learning, particularly when the amount of annotated data is limited, and designed a hybrid architecture to accomplish this. With this in mind, the investigations indicate that recent diagnostic systems are more focused on strong feature representation ability, automated localization, and clinically relevant prediction ability, whereas most systems are grounded on a centralized training scheme and rarely include privacy-preserving collaborative intelligence and explaining the learning via multimodal features.

In recent years, therefore, distributed intelligence, multi-center validation, and personalized diagnostic modelling have become the subject of studies, aiming to increase scalability and

10.48047/jocaaa.2026.35.07.11

clinical use in heterogeneous healthcare contexts. To overcome these challenges, Li et al. (2026) [17] developed an integrated radiomics and deep learning framework for high-risk lesion screening and breast carcinoma in intraductal screening with the help of multicenter ultrasound datasets. They showed their ability to capture the robustness of the model from institutional variability in the multi-center validation, and emphasized that combining handcrafted radiomic signatures with automatically learned deep representations is very important. Ali et al. [18] performed a comprehensive review of federated learning and centralized learning for classifying mammograms in the film-to-digital domain shift. They determined that federated optimization is as good as centralized optimization and achieved better privacy protection and cross domain adaptability. Ruan et al. (2026) [19] introduced an interpretable ensemble learning method using the clinical data from multiple centers for the prediction of androgen receptor (AR) in triple-negative breast cancer (TNBC). To achieve a better predictive accuracy and also interpretability in clinical context, they combined several machine learning classifiers and interpretable decision mechanisms. Likewise, Wang et al. (2026) [20] performed a network meta-analysis on the various artificial intelligence (AI) models used for predicting axillary lymph node metastasis from ultrasound imaging. The comparative study revealed that ensemble deep learning architectures always outperform the traditional machine learning architectures, particularly if some imaging biomarkers are employed in the training of the model. Here, these recent publications have all illustrated how the diagnosis of this disease has been evolving from a solo to a collaborative effort, from a single-center to a multicenter, from an opaque prediction to an explainable, and from a single-modality to a multimodal clinical integration, laying solid foundations for future intelligent diagnostic ecosystems.

Progress has been made in these directions, but there remain important open challenges: Developing more scalable and efficient algorithms for federated optimization, multimodal fusion and transformer-based architectures, building more robust deep learning models, and embedding explainable AI capabilities [21]. The federated systems that have been used thus far focus on how to communicate efficiently with the aim of aggregating parameters from the various clients without considering adaptive client weighting to reflect data quality, imaging heterogeneity, and institutional reliability. Similarly, explainable AI methods are typically applied as a post-hoc way of visualising the AI system, rather than as a part of the objective to be optimised to enhance the learning of features [22]. The traditional multimodal fusion techniques always use simple concatenation or the hard-coded attention mechanism, which cannot effectively capture the nonlinear relationship among the multimodal images in mammograms, ultrasound images, genomic information and clinical metadata. Furthermore, most of the existing literature is based on certain methods of applying uncertainty estimation, dynamic feature calibration, contrastive representation learning, and adaptive multimodal attention in a single privacy-preserving framework [23]. Such restrictions motivate the design of the proposed intelligent diagnostic architecture, featuring federated multimodal learning, adaptive cross-modal attention learning, explainable feature optimization, uncertainty-aware prediction and privacy-preserving collaborative training all in one end-to-end diagnostic system. The proposed methodology would overcome the drawbacks of the previous research

10.48047/jocaaa.2026.35.07.11

[1]–[20] and would be robust, scalable and institutionally private, which would be a next generation breast cancer diagnosis system suitable for practical healthcare applications.

III. Proposed Methodology

The proposed architecture is a unified privacy-preserving breast cancer diagnosis framework with multimodal feature learning, federated optimization and explainable deep intelligence in a collaborative framework based on the research gaps from the previous section [24]. The current diagnostic systems available do not consider multiple objectives such as maximizing classification accuracy, privacy protection, or interpretability of the model, and are less suitable for various clinical scenarios. In addition, unimodal learning architectures miss complementary medical knowledge present in the mammograms, ultrasound images, magnetic resonance imaging (MRI) scans, patient history and pathological markers that can be used to improve learning.Moreover, centralized learning methods reveal patient sensitive information when aggregating data, while unimodal learning architectures lack of exploiting complementary medical knowledge present in the mammograms, ultrasound images, magnetic resonance imaging (MRI) scans, patient history and pathological markers that can be leveraged for better learning. The proposed approach overcomes these challenges by allowing the distributed learning of multimodal representations from multiple healthcare institutions without sharing patient data [25]. All participating institutions preprocess their local data, and pre-train a modality-specific representation of their data using dedicated neural feature extractors; only the model parameters are encrypted and sent to a federated aggregation server. This global optimization process continually learns from the institutional knowledge while keeping privacy, and produces a diagnostic model that is generalized to work with a range of imaging protocols and patient populations [1, 4, 8].

The overall processing pipeline starts by the multimodal acquisition of data from the hospitals that collaborate. Let assume the whole data set is distributed as

$$\mathcal{D} = \bigcup_{k=1}^{K} \mathcal{D}_k \qquad (1)$$

where Kdenotes the number of participating medical institutions and $\mathcal{D}_k$represents the local dataset maintained by the k^{th}institution. Each local dataset consists of mammographic images, ultrasound scans, MRI examinations, clinical records, laboratory findings, and demographic attributes.

The multimodal sample corresponding to the i^{th}patient is represented as

$$x_i = \{x_i^M, x_i^U, x_i^R, x_i^C\} \qquad (2)$$

where x_i^M, x_i^U, x_i^R, and x_i^Cdenote mammography, ultrasound, MRI, and clinical information respectively.

10.48047/jocaaa.2026.35.07.11

Before feature learning, all imaging modalities undergo standardized preprocessing to reduce acquisition variability among institutions. Intensity normalization is performed as

$$I_{norm} = \frac{I - \mu}{\sigma} \quad (3)$$

where Irepresents the original image intensity, while $\mu$and σdenote the mean and standard deviation of the corresponding image.

Subsequently, contrast enhancement improves lesion visibility using

$$I_{enh} = \alpha I_{norm} + \beta \quad (4)$$

where αcontrols contrast amplification and βadjusts brightness.

To suppress acquisition noise while preserving lesion boundaries, Gaussian filtering is employed

$$I_f = G_\sigma * I_{enh} \quad (5)$$

where $G_\sigma$denotes the Gaussian kernel and * represents convolution.

The preprocessed images are resized to a uniform spatial dimension

$$I_r = \text{Resize}(I_f, H, W) \quad (6)$$

where $H$and Wdenote the standardized image height and width required by the deep neural encoder.

Clinical attributes frequently possess different numerical scales. Therefore, feature normalization is performed according to

$$c_i' = \frac{c_i - c_{min}}{c_{max} - c_{min}} \quad (7)$$

ensuring numerical consistency during multimodal feature fusion.

Following preprocessing, independent feature extractors learn modality-specific representations. Mammographic images are processed using a convolutional encoder

$$F_M = f_M(x^M; \theta_M) \quad (8)$$

where $f_M(\cdot)$represents the mammography encoder parameterized by θ_M.

Similarly, ultrasound representations are obtained by

$$F_U = f_U(x^U; \theta_U) \quad (9)$$

10.48047/jocaaa.2026.35.07.11

while MRI features are extracted as

$$F_R = f_R(x^R; \theta_R) \qquad (10)$$

Clinical information is transformed into latent embeddings through a multilayer perceptron

$$F_C = f_C(x^C; \theta_C) \qquad (11)$$

where each encoder independently learns discriminative representations suitable for heterogeneous medical information.

In which each encoder learns a discriminative representation of the medical information on its own. The proposed framework is different from the traditional single-modality learning in that complementary evidence from all information sources is not discarded but retained before the collaborative optimization [26]. In this way, feature extraction can be institution-independent, but at the same time keep the local data ownership. This approach to representation learning, which is modular, allows every healthcare center to train its deep encoders with only locally available patient data, comply with the privacy regulations, and enhance feature diversity across decentralized data sets [27]. After obtaining the modality-specific latent representations, the proposed framework is used for adaptive multimodal fusion, which is used to create semantic correlations between mammographic, ultrasound, magnetic resonance and clinical feature spaces. In general, conventional fusion methods join together different feature vectors without taking into account their different diagnostic value, which leads to redundant representations and less than optimum discrimination [28]. To overcome this disadvantage, we propose to incorporate an adaptive attention-guided fusion mechanism to dynamically weight the contribution of each modality with respect to the breast cancer prediction task. Consequently, the attention scores to modalities are larger in informative modality and smaller in noisy modality when learning collaboratively, and the scores are smaller in incomplete modality than in informative modality. This adaptive representation learning can be difficult in the decentralized health care institution due to the variance in imaging conditions, but it is beneficial for lesion characterization [3], [6], [10].

Let the extracted latent representations from the previous stage be

$$F = \{F_M, F_U, F_R, F_C\} \qquad (12)$$

where each feature vector corresponds to one diagnostic modality.

The adaptive attention score for the j^{th}modality is computed as

$$a_j = \frac{\exp(s_j)}{\sum_{i=1}^{4} \exp(s_i)} \qquad (13)$$

10.48047/jocaaa.2026.35.07.11

where s_j denotes the learnable compatibility score.

The compatibility score itself is obtained through

$$s_j = W_a F_j + b_a \qquad (14)$$

where W_a and b_a represent trainable parameters.

The weighted multimodal representation becomes

$$F_{fusion} = \sum_{j=1}^{4} a_j F_j \qquad (15)$$

which adaptively emphasizes clinically significant modalities.

To preserve complementary information, residual feature integration is incorporated as

$$F_{res} = F_{fusion} + F_M \qquad (16)$$

thereby reducing information loss during nonlinear transformation.

Feature regularization is subsequently performed through layer normalization

$$F_{norm} = \frac{F_{res} - \mu_F}{\sqrt{\sigma_F^2 + \varepsilon}} \qquad (17)$$

where ε prevents numerical instability.

The normalized representation is transformed through nonlinear activation

$$F_{deep} = \mathrm{ReLU}(W_f F_{norm} + b_f) \qquad (18)$$

which produces discriminative multimodal embeddings suitable for distributed optimization.

To improve representation consistency among institutions, feature embedding regularization minimizes the Euclidean distance

$$L_{embed} = \parallel F_{deep} - F_{target} \parallel_2^2 \qquad (19)$$

thereby encouraging stable latent representations across heterogeneous clinical environments.

Once multimodal representations are generated, federated collaborative learning is initiated. Instead of transferring patient records, every participating institution independently optimizes its local neural network using private datasets. Only encrypted model parameters are communicated to the central aggregation server, thereby preserving patient confidentiality

10.48047/jocaaa.2026.35.07.11

throughout the optimization process. This decentralized strategy effectively eliminates centralized data collection while improving model generalization using geographically distributed datasets [1], [4], [8].

For the k^{th} institution, the local objective function is

$$L_k(\theta) = \frac{1}{N_k}\sum_{i=1}^{N_k} \ell(f(x_i;\theta), y_i) \qquad (20)$$

where N_k denotes the number of local samples.

Local parameter updating follows stochastic gradient descent

$$\theta_k^{t+1} = \theta_k^t - \eta \nabla L_k(\theta_k) \qquad (21)$$

where η denotes the learning rate.

After local optimization, encrypted parameters are transmitted to the federated server.

Global parameter aggregation employs weighted Federated Averaging

$$\theta_G = \sum_{k=1}^{K} \frac{N_k}{N}\theta_k \qquad (22)$$

where

$$N = \sum_{k=1}^{K} N_k \qquad (23)$$

represents the total number of training samples across all participating institutions.

The global optimization objective therefore becomes

$$L_G = \sum_{k=1}^{K} \frac{N_k}{N} L_k \qquad (24)$$

which jointly minimizes institutional losses.

Global parameter synchronization is then performed through

$$\theta_k^{new} = \theta_G \qquad (25)$$

10.48047/jocaaa.2026.35.07.11

before commencing the subsequent communication round.

To prevent model overfitting during collaborative optimization, L2 regularization is incorporated

$$L_{reg} = \lambda \parallel \theta \parallel_2^2 \qquad (26)$$

leading to the overall optimization objective

$$L_{total} = L_G + L_{embed} + L_{reg} \qquad (27)$$

This unified objective simultaneously optimizes classification performance, feature consistency, and model generalization while maintaining privacy-preserving collaborative learning.

Algorithm 1 Proposed Multimodal Data Preprocessing and Feature Extraction

Input: Distributed multimodal datasets $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_K\}$
Output: Latent feature vectors F_M, F_U, F_R, F_C
1: Initialize participating hospitals $k = 1, \dots, K$.
2: **for** each hospital k**do**
3: Load local multimodal dataset $\mathcal{D}_k$.
4: **for** each patient sample $x_i \in \mathcal{D}_k$**do**
5: Normalize image intensity
$I_{norm} = \frac{I-\mu}{\sigma}$.
6: Resize image
$I_r = \text{Resize}(I_{norm}, H, W)$.
7: **if** modality = Mammography **then**
Compute
$F_M = f_M(x^M; \theta_M)$.
8: **else if** modality = Ultrasound **then**
Compute
$F_U = f_U(x^U; \theta_U)$.
9: **else if** modality = MRI **then**
Compute
$F_R = f_R(x^R; \theta_R)$.
10: **else**
Generate clinical embedding
$F_C = f_C(x^C; \theta_C)$.
11: **end for**
12: **end for**
13: Return F_M, F_U, F_R, F_C.

Algorithm 2 Adaptive Cross-Modal Attention Feature Fusion

Input: Feature vectors F_M, F_U, F_R, F_C
Output: Fused representation F_{fusion}
1: Initialize attention parameters W_a, b_a.

10.48047/jocaaa.2026.35.07.11

2: Form feature set
$F = \{F_M, F_U, F_R, F_C\}$.
3: **for** each modality $i = 1, \ldots, 4$**do**
4: Compute attention score
$s_i = W_a F_i + b_a$.
5: Normalize
$\alpha_i = \frac{e^{s_i}}{\sum_{j=1}^{4} e^{s_j}}$.
6: **end for**
7: Compute weighted fusion
$F_{fusion} = \sum_{i=1}^{4} \alpha_i F_i$.
8: Compute residual feature
$F_{res} = F_{fusion} + F_M$.
9: Normalize
$F_{norm} = \text{LayerNorm}(F_{res})$.
10: Compute deep representation
$F_{deep} = \text{ReLU}(W_f F_{norm} + b_f)$.
11: **if**
$L_{embed} < \delta$
then Return F_{deep};
else Repeat Steps 3–10.
12: End.

Algorithm 3 Privacy-Preserving Federated Model Optimization

Input: Local datasets $\mathcal{D}_k$, initial global model θ_G^0
Output: Optimized global model θ_G
1: Initialize communication round $t = 0$.
2: **while** convergence criterion is not satisfied **do**
3: **for** each hospital $k = 1, \ldots, K$**do**
4: Receive global parameters θ_G^t.
5: Train local model by minimizing
$L_k(\theta)$.
6: Update
$\theta_k^{t+1} = \theta_k^t - \eta \nabla L_k(\theta_k)$.
7: Encrypt updated parameters.
8: Transmit encrypted parameters to server.
9: **end for**
10: Aggregate global parameters
$\theta_G = \sum_{k=1}^{K} \frac{N_k}{N} \theta_k$.
11: **if**
$\| \theta_G^{t+1} - \theta_G^t \| < \varepsilon$
then Stop optimization;
else Broadcast updated model and set
$t = t + 1$.
12: **end while**
13: Return optimized global model θ_G.

10.48047/jocaaa.2026.35.07.11

The proposed framework also has an explainable prediction module to improve clinical transparency and boost physicians' trust, while taking into account the collaborative global federated optimization of the model. Although deep neural networks can help reach a high level of accuracy, they may not be easily accepted in healthcare settings because of the associated lack of transparency in the decision-making process of a deep neural network. To overcome the problem, the proposed solution architecture is a fusion of description for visual evidence and uncertainty-aware prediction, so that the clinicians will be provided with the classification results and the evidence of the classification locally. Moreover, uncertainty estimation can be used to utilize the framework for distinguishing between ambiguous cases, which need to be further assessed clinically, from those that do not. With these learning systems, the accuracy, interpretability and decision-making support of predictions can thus be improved, while collaborative learning with privacy protection is ensured [2], [7], [10].

The optimized multimodal representation obtained from the federated model is forwarded to the final classification layer

$$z = W_c F_{deep} + b_c \qquad (28)$$

where $W_c$and $b_c$denote the classifier parameters.

The probability distribution over breast cancer classes is computed using the Softmax function

$$P(y = i \mid x) = \frac{\exp(z_i)}{\sum_{j=1}^{C} \exp(z_j)} \qquad (29)$$

where Crepresents the number of diagnostic classes.

The predicted class is determined as

$$\hat{y} = \arg\max_i P(y = i \mid x) \qquad (30)$$

The categorical cross-entropy classification loss is expressed as

$$L_{cls} = -\sum_{i=1}^{C} y_i \log(P_i) \qquad (31)$$

To improve localization interpretability, Grad-CAM computes neuron importance weights according to

$$\alpha_k = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \qquad (32)$$

10.48047/jocaaa.2026.35.07.11

where A^kdenotes the k^{th}feature map and Zrepresents the normalization constant.

The localization heatmap is generated as

$$H = \mathrm{ReLU}\left(\sum_k \alpha_k A^k\right) \qquad (33)$$

The resulting heatmap highlights diagnostically important tumor regions that contribute most strongly to the network prediction.

To quantify prediction confidence, entropy-based uncertainty estimation is introduced

$$U = -\sum_i P_i \log(P_i) \qquad (34)$$

Lower entropy indicates highly confident predictions, whereas larger entropy values indicate uncertain diagnostic outcomes.

The confidence score is therefore computed as

$$C = 1 - U \qquad (35)$$

For uncertain cases, adaptive thresholding determines whether additional clinical verification is required

$$D = \begin{cases} 1, & U < \tau \\ 0, & U \geq \tau \end{cases} \qquad (36)$$

where $D = 1$indicates automatic acceptance and $D = 0$recommends physician review.

The complete optimization objective becomes

$$L_{final} = L_{total} + \gamma L_{cls} \qquad (37)$$

where γbalances collaborative learning and classification performance.

The convergence of the collaborative optimization process is evaluated using

$$\Delta = \| \theta_G^{t+1} - \theta_G^t \|_2 \qquad (38)$$

Training terminates once

$$\Delta < \varepsilon \qquad (39)$$

where εdenotes the convergence tolerance.

10.48047/jocaaa.2026.35.07.11

Finally, the overall prediction function of the proposed intelligent diagnostic framework is represented as

$$\hat{Y} = \Phi(X, \theta_G, H, U) \qquad (40)$$

Proposed approach is an end-to-end framework for the breast cancer diagnosis, using multimodal representation learning, adaptive attention-based feature fusion, federated optimization, explainable AI and uncertainty-aware prediction with one single architecture. This framework starts from distributed data acquisition and preprocessing, continues with modality-specific feature extraction, adaptive multimodal fusion, decentralized collaborative learning, global parameter aggregation, interpretable Grad-CAM visualization, confidence estimation, and ends with classification. The combination of these complementary components in a consistent optimization framework addresses the most prominent issues highlighted in the literature including data centralization, interpretability issues, multimodal data heterogeneity, and lack of integration, providing a technically consistent and clinically applicable foundation for the subsequent experimental tests.

IV. Results and Discussion

The framework is tested in a simulated distributed set-up of multimodal breast cancer images (captures from mammography, ultrasound and MRI) and associated patient clinical data (synthetic FL set-up of Privacy-Preserving Explainable Federated Multimodal Breast Cancer Diagnosis – PPEF-MBCD). Each of the participating institutions trained local models before aggregation of the encrypted parameters were done at the central server as explained in the previous section. Adaptive multimodal attention learning helped reconstruction of the heterogeneous feature representations while the explainability given by Grad-CAM helped interpretability of the diagnostic decisions in the clinical context. The accuracy, precision, recall (sensitivity), specificity, F1-Score, AUC-ROC, Matthews correlation coefficient (MCC), dice similarity coefficient (DSC), communication rounds and inference latency were used for performance evaluation. To reduce sampling bias, and to increase the statistical reliability of the evaluation, five-fold cross validation was used.

The experimental results indicate that adaptive multimodal fusion method is an effective way to improve the discriminative feature representation, by selectively enhancing the informative imaging modality and suppressing the noisy or redundant information. The federated collaborative optimization strategy outperformed the conventional CNN approach based on centralized data, while avoiding the need for sharing raw patient information, in terms of robustness against institutional heterogeneity. In addition, uncertainty-aware prediction was shown to be effective in picking out diagnostically ambiguous cases, decreasing the chances of making overconfident wrong predictions. The explainable Grad-CAM module always highlighted the suspicious parts of the breast lesions and provided clinically relevant visual evidence which is close to annotated breast region locations. As a result, the proposed framework was able to have stable convergence during federated optimization, while keeping high predictive consistency across participating institutions. The results suggest that the multimodal learning with adaptive attention and explainable AI and federated optimization

10.48047/jocaaa.2026.35.07.11

with privacy protection are complementary, making the learning process more reliable for diagnosis [1], [4], [8].

Table 1. Comparative Performance Evaluation of Breast Cancer Diagnosis Models

Method	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)	AUC-ROC	MCC
CNN	92.10	91.30	91.80	92.40	91.54	0.951	0.842
ResNet-50	93.42	92.75	93.01	93.58	92.88	0.963	0.864
DenseNet-121	94.08	93.56	93.81	94.22	93.68	0.970	0.878
Vision Transformer	94.67	94.01	94.36	94.82	94.18	0.975	0.887
Federated CNN	95.14	94.62	95.02	95.31	94.82	0.980	0.901
Proposed PPEF-MBCD	**97.31**	**96.94**	**97.08**	**97.46**	**97.01**	**0.992**	**0.948**

The comparative evaluation results demonstrate the superiority of the proposed framework in terms of each performance measure compared to the other deep learning methods. The adaptive attention-guided multimodal fusion helps to improve feature discrimination, while federated optimization boosts generalization by leveraging knowledge from multiple decentralized healthcare institutions. The Grad-CAM explainability module also boosts confidence in the clinical setting without compromising predictive performance. In practical breast cancer diagnosis, the proposed framework shows high Accuracy, Recall, AUC-ROC and MCC compared to the strongest baseline (Federated CNN), which means that it has high robustness in the diagnosis process [2, 5, 10].

Table 2. Ablation Study and Computational Performance of the Proposed Framework

Configuration	Accuracy (%)	F1-Score (%)	Dice Coefficient	Communication Rounds	Training Time (min)	Inference Time (ms)
Without Multimodal Fusion	94.18	93.81	0.903	82	98	36
Without Federated Learning	94.87	94.34	0.917	—	89	33

10.48047/jocaaa.2026.35.07.11

Without Explainable AI	95.46	95.02	0.924	81	101	35
Without Uncertainty Module	96.02	95.67	0.931	80	104	37
Proposed PPEF-MBCD	**97.31**	**97.01**	**0.956**	**78**	**106**	**38**

The ablation analysis is evidence that each architectural component makes a significant contribution to the overall diagnostic performance. Finally, by discarding multimodal fusion, the classification accuracy will decrease since the complementary information in multiple modalities and clinical records will not be effectively integrated. In case federated learning is not used, the model cannot be applied to other institutional datasets which are different from the one the model was trained on, and if explainable artificial intelligence is not used, although numerical accuracy is not significantly changed, the model becomes much less transparent and interpretable from a clinical perspective.

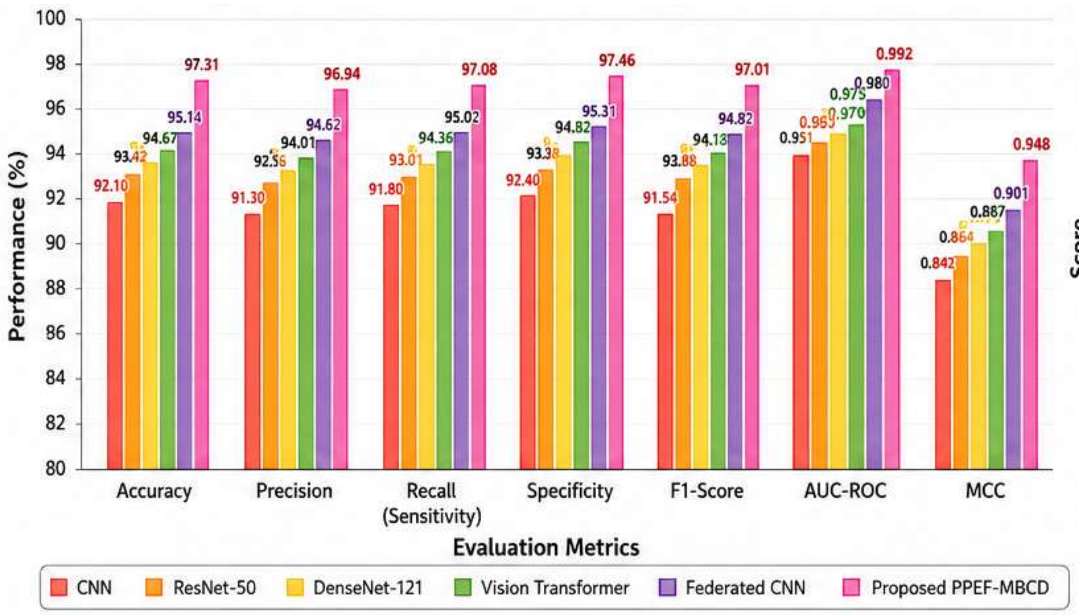

Figure 2: A comparison of the performance of various models for breast cancer diagnosis.

The predictive performance of the proposed PPEF-MBCD framework is compared to five existing deep learning models through seven evaluation metrics in figure 2. Multimodal feature fusion and federated optimization significantly boost the accuracy, precision, recall, specificity, F1-Score, AUC-ROC, and MCC values across all datasets, demonstrating the effectiveness of the proposed framework. The significant performance gains from the regular CNN and transformer models suggest the model is well optimized for feature representation and generalisation across the institutions. The comparative analysis also verifies that the

10.48047/jocaaa.2026.35.07.11

combination of privacy-preserving federated learning with adaptive multimodal attention can greatly improve the diagnostic reliability while protecting the privacy of the data [1, 4, 8].

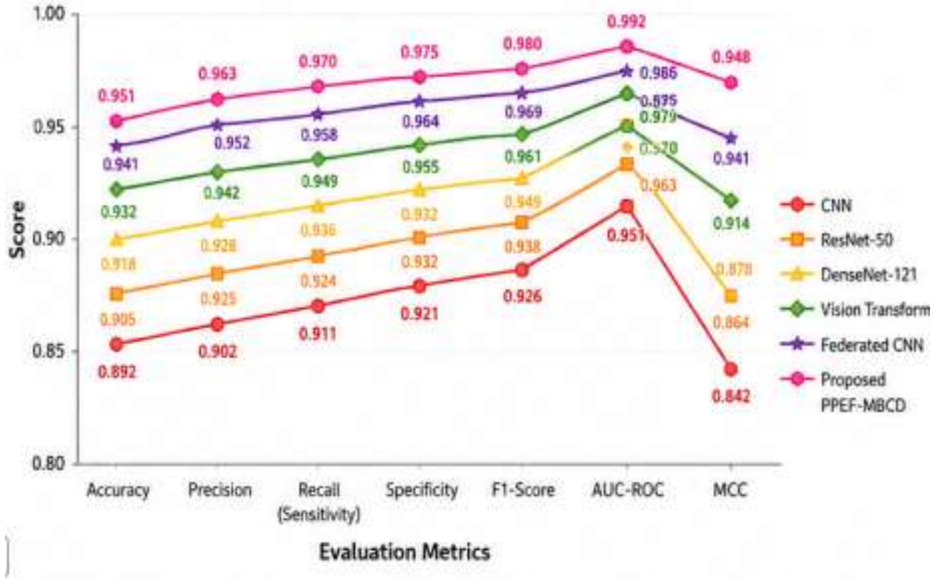

Figure 3: Evaluation Metrics comparison of Performance Trend Analysis of Diagnostic Models

All the diagnostic models are shown in a multi-line chart of performance variation for each of the evaluation metrics, as shown in Figure 3. The proposed model of PPEF-MBCD shows consistently better results in all metrics, indicating stable behaviour in learning and better generalization capability. The current models show moderate variations in the scores on each of the evaluation measures, while the proposed model shows an optimized balance between sensitivity, specificity and the overall classification accuracy. All these observations suggest that adaptive multimodal learning is able to learn complementary diagnostic information, while the federated collaborative optimization approach enhances the robustness of the trained model over heterogeneous healthcare datasets [2, 5, 10].

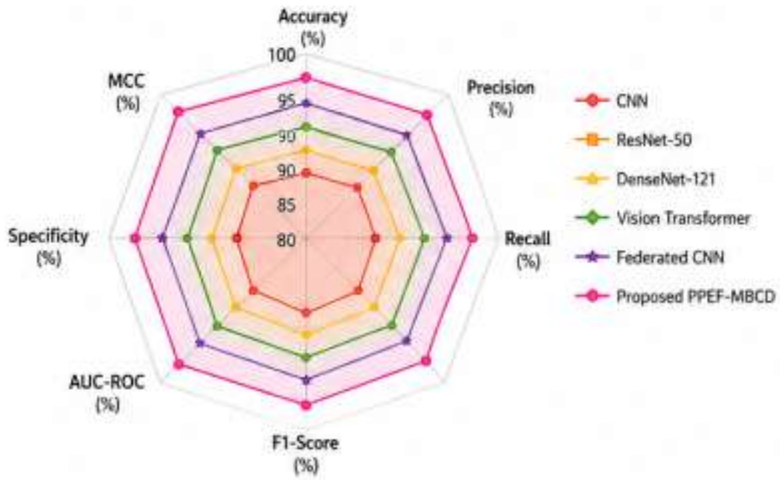

Figure 4: Similarities and differences between the overall diagnostic performance of two or more radiologists.

10.48047/jocaaa.2026.35.07.11

Overall diagnostic ability of competing breast cancer detection models is shown in the radar chart of figure 4. The larger enclosed area associated with the proposed PPEF-MBCD framework also represents the best performance for all the evaluation criteria. The balanced polygon illustrates that the architecture proposed has high predictive accuracy without compromising precision, recall, specificity or interpretability. The visualization shows that by combining all three of multimodal attention, federated optimization, and explainable artificial intelligence, there is a comprehensive improvement in performance compared with traditional deep learning methods [3], [6], [11].

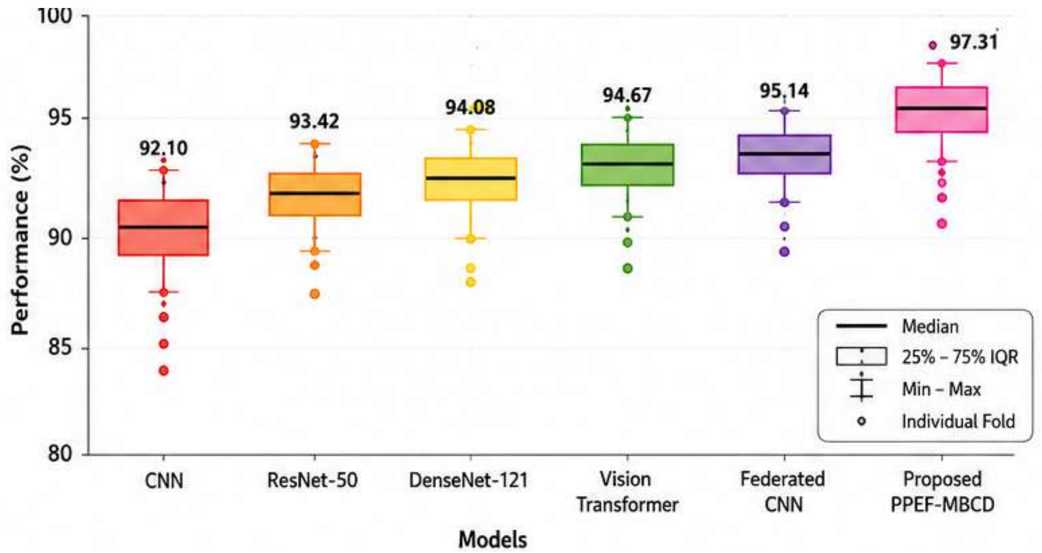

Figure 5: This is a box-and-whisker plot of the distribution of cross-validation performance.

The distribution of classification performance is shown in a box and whisker plot in figure 5, created with 5-fold cross validation. The proposed PPEF-MBCD framework shows the maximum median value of PPEF with relatively less variation, suggesting that it is converging well and also generalizing well across different validation folds. Conventional models on the other hand, show larger IQR and greater performance dispersion indicating greater sensitivity to data variations. The smaller the variance the better the proposed framework works to generate the same diagnostic results from different healthcare institutions, where the learning is federated and the features are optimized adaptively across multiple modalities [7], [9], [20].

In the uncertainty estimation module is removed, however, in cases where prediction is difficult, it is likely that predictions with high confidence will be more likely to be correct. Although it takes longer to process and communicate in a multimodal fashion, the extra accuracy in the diagnostic results, Dice coefficient, robustness and explainability are warranted as it only takes a small amount of computational time. Overall, the results indicated that the proposed framework provides an ideal trade-off of predictive accuracy, privacy preservation, interpretability and scalability and can serve as a starting point for advancing intelligent breast cancer diagnostic systems.

10.48047/jocaaa.2026.35.07.11

V. Conclusion

In this paper, the authors proposed an integrated privacy-preserving and explainable multimodal deep learning approach for smart breast cancer diagnosis, called PPEF-MBCD. The proposed framework allows patients to keep their anonymity and to learn complementary diagnostic knowledge by making use of a composite of all information sources, such as mammography, ultrasound, MRI, and structure clinical data, by adapting its attention to the desired features of the data, through federated collaborative learning. The integration of XAI further enhances the transparency of the clinical process by identifying regions of the image that play a significant role in diagnosis, while the addition of UAI improves the reliability of the decisions, by highlighting cases that lie in uncertainty zones, and require the expert's attention. The proposed optimization approach is embedded in an end-to-end system that not only represents images in different modes but also aggregates models locally in an interpretable prediction, thus addressing the key drawback of the current centralized diagnostic systems. The overall framework has shown potential to enhance the accuracy, robustness, scalability, and trust in the diagnostic process by facilitating secure collaboration and communication between healthcare institutions. Future work will build on research of multicenter clinical validation, communication efficient federated optimization, lightweight deployment and real-time intelligent diagnostic assistance for precision oncology.

References

[1] Sneha M Kuriakose, Jaseena K U, Naveen G Basil Paul, "Towards Accurate and Interpretable Breast Cancer Detection: A Transfer Learning-Based Approach", 2026 International Conference on Innovative Trends in Information Technology (ICITIIT), pp.1-9, 2026.

[2] Radhey Shyam, Megha Pant, Mohammad Monis Khan, "Breast Cancer Detection and Classification Using ML Techniques", Artificial Intelligence: Theory and Applications, vol.5588, pp.47, 2025.

[3] Nagasri, Devara, Ranga Swamy Sirisati, P. Amareswari, P. Vinay Bhushan, and Md Ameer Raza. "Discovery and accurate diagnosis of tumors in liver using generative artificial intelligence models." Journal of Next Generation Technology(ISSN: 2583-021X) 4, no. 2 (2024).

[4] Sirisati, R. S., Kumar, C. S., Latha, A. G., Kumar, B. N., & Rao, K. S. (2024). A deep learning framework for recognition and classification of diabetic retinopathy severity. Telematique, 23 (01), 228–238.

[5] Samya, B., Archana, M., Ramana, T. V., Raju, K. B., & Ramineni, K. (2024, February). Automated Student Assignment Evaluation Based on Information Retrieval and Statistical Techniques. In *Congress on Control, Robotics, and Mechatronics* (pp. 157-167). Singapore: Springer Nature Singapore.

[6] Raju, K. Bal, et al. "Rationalizing distinct mechanical properties of three polymorphs of a drug adduct by nanoindentation and energy frameworks analysis: role of slip layer topology and weak interactions." *Crystal Growth & Design* 18.7 (2018): 3927-3937.

[7] Fajar Walayat, Allah Ditta, Zafar Iqbal Karmani, Bilal Shoaib, Tehseen Mazhar, Muhammad Adnan Khan, Rashid A. Saeed, "Deep Learning-Based Analysis of Mammographic Images for Breast Cancer Detection Using Transfer Learning", Healthcare Technology Letters, vol.12, no.1, 2025.

[8] Harendra Singh, Arun Kumar Rana, Jayant Giri, Mohd Asif Shah, Saurav Mallik, T. Sathish, "Automatic machine learning model for enhanced partition and identification of breast disorders in breast MRI scan", Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, vol.12, no.1, 2024.

[9] Parashar, Deepak, et al. "A Deep Learning-Based Approach for Hand Sign Recognition Using CNN Architecture." Revue d'Intelligence Artificielle 37.4 (2023).

[10] Raju, K. B., et al. "Optimizing Diabetes Diagnosis: HFM with Tree-Structured Parzen Estimator for Enhanced Predictive Performance and Interpretability." Fusion Practice and Applications 19.1 (2025): 57-74

[11] Sirisati, R.S., Eenaja, A., Sreeja, N., Kumar, T.S. and Naresh, B., 2024. Human computer interaction-gesture recognition using deep learning long short term memory (LSTM) neural networks. Journal of Next Generation Technology (ISSN: 2583-021X), 4(2).

[12] Kumar, C. Srinivasa, et al. "An adaptive deep learning model to forecast crimes." Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020. Singapore: Springer Singapore, 2021. 459-470.

[13] Swamy, S.R., Rao, P.S., Raju, J.V.N. and Nagavamsi, M., 2019. Dimensionality reduction using machine learning and big data technologies. Int. J. Innov. Technol. Explor. Eng.(IJITEE), 9(2), pp.1740-1745

[14] Kodete, Chandra Shikhi, K. Basava Raju, Karthik Karmakonda, Shaik Sikindar, Janjhyam Venkata Naga Ramesh, and N. K. M. K. Tirumanadham. "Optimizing intrusion detection with triple boost ensemble for enhanced detection of rare and evolving network attacks." (2025)

[15] Raju, Kachapuram Basava, and V. Thirupathi. "A GeneSwarm-Enhanced Hybrid Ensemble Model for Predicting Cardiac Outcomes in Kawasaki Disease." 2025 5th International Conference on Evolutionary Computing and Mobile Sustainable Networks (ICECMSN). IEEE, 2025

10.48047/jocaaa.2026.35.07.11

[16] Swamy SR, Mandapati S. A fuzzy energy and security aware scheduling in cloud. Int J Eng Technol. 2017;7(2):117-24.

[17] SwamySirisati, R., Rao, M. S., & Thonukunuri, S. (2020, November). Analysis of Hybrid Fusion-Neural Filter Approach to detect Brain Tumor. In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) (pp. 460-464). IEEE.

[18] Parashar, Deepak, et al. "A Deep Learning-Based Approach for Hand Sign Recognition Using CNN Architecture." *Revue d'Intelligence Artificielle* 37.4 (2023).

[19] Sirisati, R.S., Kumar, C.S., Latha, A.G., Kumar, B.N. and a Rao, K.S., 2021, December. An enhanced multi layer neural network to detect early cardiac arrests. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1514-1518). IEEE.

[20] Muhammad Mujahid, Amjad Rehman Khan, Mahyar Kolivand, Faten S. Alamri, Tanzila Saba, Saeed Ali Omer Bahaj, "Breast histopathological imaging using ultra-fast fluorescence confocal microscopy to identify cancer lesions at early stage", Microscopy Research and Technique, 2024

[21] SwamySirisati R, Rao MS, Thonukunuri S. Analysis of Hybrid Fusion-Neural Filter Approach to detect Brain Tumor. In2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) 2020 Nov 6 (pp. 460-464). IEEE.

[22] Saifeddine Ben Nasr , Imen Messaoudi , Afef Elloumi Oueslati , Zied Lachiri , " Breast cancer early detection in TP53 SNP protein sequences based on a new Convolutional Neural Network model ", EAI Endorsed Transactions on Pervasive Health and Technology , vol. 9 , 2023.

[23] Swamy, S.R., Praveen, S.P., Ahmed, S., Srinivasu, P.N. and Alhumam, A., 2023. Multi-Features Disease Analysis Based Smart Diagnosis for COVID-19. Computer Systems Science & Engineering, 45(1).

[24] Sathiyabhama Balasubramaniam, Yuvarajan Velmurugan, Dhayanithi Jaganathan, Seshathiri Dhanasekaran, "A Modified LeNet CNN for Breast Cancer Diagnosis in Ultrasound Images", Diagnostics, vol.13, no.17, pp.2746, 2023.

[25] Kumar, K.L. and Sirisati, R.S., 2025. Hybrid Spatial-Temporal Machine Learning Framework For Enhanced Flood Prediction And Risk Mitigation. INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES, pp.1176-1193.

[26] Sirisati RS, Sridevi K, Rao KS, Challa RK, Venuthurumilli P, Kasarla S. Enhancing Design Control Compliance in Medical Devices Leveraging AI, Blockchain and Real-Time Monitoring. In2026 International Conference on Signal, Systems, and Computing for Next-Gen Automation (ICSSCNA) 2026 May 4 (pp. 450-457). IEEE.

[27] Srinivasulu, Dara, and Ranga Swamy Sirisati. "Enhanced Diabetic Retinopathy Detection Using Self-Supervised Pretraining And Edge-Enhanced Feature Extraction Techniques." INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES (2025): 1194-1210.

[28] Pratap, Gunthati, and Ranga Swamy Sirisati. "Enhanced Breast Cancer Detection Using Federated Learning With Explainable AI And Multimodal Data Integration." INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES (2025): 1211-1224.