Privacy-Preserving Breast Cancer Classification via Federated Learning on Multimodal Image Data

By

Arham Mahbub Alam
211-15-3963

S. M. Thahidul Islam
211-15-3949

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the **Degree of Bachelor of Science in Computer Science and Engineering**

Supervised by

Dr. Naznin Sultana
Associate Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by

Ms. Arpita Ghose Tusi
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

January 12, 2025

APPROVAL

This Project titled **“Privacy-Preserving Breast Cancer Classification via Federated Learning on Multimodal Image Data,”** submitted by **Arham Mahbub Alam,** ID: **211-15-3963** and **S. M. Thahidul Islam,** ID: **211-15-3949** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **12-01-2025.**

BOARD OF EXAMINERS

Dr. S.M Aminul Haque (SMAH) — **Board Chairman**
Professor & Associate Head, Department of CSE, FSIT
Daffodil International University

12.01.25

Mohammad Jahangir Alam(MJA) — **Internal Examiner**
Assistant Professor, Department of CSE, FSIT
Daffodil International University

12.01.25

Mr. Deawan Rakin Ahamed Remal (DRAR) — **Internal Examiner**
Lecturer, Department of CSE, FSIT
Daffodil International University

Dr. Md. Arshad Ali (DAA) — **External Examiner**
Professor , Department of CSE
Hajee Mohammad Danesh Science Technology University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Dr. Naznin Sultana, Associate Professor**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Dr. Naznin Sultana
Associate Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by:

Ms. Arpita Ghose Tusi
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Submitted by:

Arham 12.01.2025

Arham Mahbub Alam
Student ID: 211-15-3963
Department of Computer Science and Engineering
Daffodil International University

Thahidul 12.01.2025

S. M. Thahidul Islam
Student ID: 211-15-3949
Department of Computer Science and Engineering
Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project(FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Dr. Naznin Sultana**, **Associate Professor**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Computer vision and Machine learning** to carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Breast cancer significantly affects women's health worldwide, underscoring the critical need for advanced diagnostic approaches to improve early detection and treatment outcomes. This paper proposes a privacy-preserving framework for breast cancer classification using federated learning integrated with multimodal imaging data, including mammography, ultrasound, and histopathology. It will enable FL to collaboratively study model training across various institutions by not sharing sensitive information of patients, which indeed may meet the requirements of GDPR, HIPAA, and related data-protection policies. Herein, it proposes a hybrid encoder for classifying benign and malignant lesions by combining the powers of CNNs with mechanisms of attention and feature fusion in a more sophisticated way. The methodology incorporates extensive preprocessing and feature extraction with SMOTE to handle class imbalance. Validated on real-world multimodal datasets within a federated framework, the model achieved superior performance with a test accuracy of 97.02% compared to traditional centralized approaches. Ablation study further optimized the model components relating to feature selection, pooling layers, and learning rates. The system is scalable, computationally efficient, and robust, offering huge reductions in false positives and negatives with data privacy. This research reflects the transformative potential of FL in healthcare, paving the way for an ethical, scalable, and effective diagnosis of breast cancer. Future work will incorporate explainable AI and a web-based platform for clinical decision-making in real-time, furthering the scope of this innovative approach.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Breast cancer is the most common type of cancer in the world and one of the leading causes of cancer-related deaths among women, taking millions of women worldwide each year. The most common female cancer is breast cancer, accounting for nearly 2.3 million new cases annually, according to the World Health Organization. It is seen that if diagnosed early or at an early stage, it improves survival and the effective treatment outcome. However, the complexity of the disease coupled with variability in diagnostic processes and imaging modalities poses many challenges for healthcare systems worldwide. This thesis discusses a novel approach to breast cancer diagnosis using federated learning and multimodal imaging with an aim to tackle some critical issues of data privacy, accuracy, and scalability.Clinical Diagnosis of Breast Cancer: Clinical diagnosis of breast cancer generally involves various combinations of imaging studies including mammography, ultrasound, and histopathology. Each modality provides complementary views of the pathology. For example, mammography is generally used to identify calcifications and structural abnormalities, whereas ultrasound is particularly useful for characterizing masses and differentiating cystic from solid lesions. Histopathology represents the gold standard in the confirmation of malignancy by the microscopic study of tissue samples. Integration of such multi-modality data can therefore lead to a considerable increment in the diagnostic accuracy by leveraging strengths from different modalities. Integration encompasses inherent issues related to data heterogeneity, high dimensionality, and requires sophisticated algorithms that can fuse the diverse information effectively. Medical Imaging has undergone a revolutionary change with AI/ML, which enables the processing of complex data in an automated manner. Especially, deep learning has been highly effective in the extraction of meaningful patterns from medical images, achieving near-human performance for classification, segmentation, and detection tasks. Despite such progress, most traditional AI approaches tend to rely on the paradigm of centralized data collection: data from multiple institutions is aggregated in one place for model training. This can be dangerous to patient privacy and data security, given strong regulations like the General Data

Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA), aiming at maintaining sensitive information of patients secure, consequently limiting such a centralized approach to data-sharing models.Federated learning presents a novel solution to these privacy challenges. It enables FL to conduct collaborative model training across multiple institutions without sharing raw data. Instead, it trains local models on decentralized datasets and shares only the aggregated model updates with a central server. This is a decentralized framework that ensures sensitive patient data will stay within its original location. Privacy and compliance challenges are solved accordingly. Gained momentum, FL now applied in various aspects of healthcare, such as disease prediction, analysis of medical images, and EHR processing. Its application to breast cancer diagnosis, particularly in a multimodal context, represents a significant step forward in privacy-preserving AI.

The proposed study bases itself on the ideas of FL and further extends its application to multimodal breast cancer classification. This research project seeks to take advantage of data from multiple imaging modalities to produce a strong model able to classify breast lesions as either benign or malignant. This approach will enhance diagnostic precision and surmount limitations, such as false positives and reduced sensitivity, which are characteristic of single-modality analyses. It leverages state-of-the-art deep learning architectures such as convolutional neural networks for feature extraction in modality, along with attention mechanisms that allow it to give different priorities on which source of information is relevant during the fusion process. In addition, the proposed framework utilizes SMOTE-a synthetic minority over-sampling technique-along with other augmentation strategies to ensure class-imbalance issues inherent in medical datasets do not lead to performance degradation of the model.One of the critical challenges in implementing FL for multimodal imaging is managing the heterogeneity of data distributions across participating institutions. Medical datasets are often non-IID (non-independent and identically distributed), meaning that the data characteristics vary significantly between institutions due to differences in imaging equipment, protocols, and patient demographics. Addressing this heterogeneity requires innovative aggregation techniques and personalized model adaptation strategies to ensure consistent performance across diverse settings. Besides, the computational complexity in training deep learning models in a federated environment requires efficient communication protocols and optimization algorithms that keep latency and energy consumption low.The integration of multimodal imaging into an FL framework also raises questions about inter-modality relationships and their impact on model performance. Combining data from modalities with varying resolutions, feature sets, and noise levels demands sophisticated feature fusion techniques. Various strategies investigated in this work for perfect integration include hierarchical fusion, attention-based methods, and modality-specific sub-networks. The role of domain adaptation and transfer learning to improve model generalizability across institutions with variable data availability and quality is also explored.

Another critical focus of this research is on scalability and real-world applicability. The proposed FL framework is designed to handle collaboration among a large number of institutions in a manner such that the Global model developed benefits from diverse datasets with maintenance of strict privacy standards. Scalability is treated using light-weight model architecture, efficient communication schemes, and adaptive learning rates. It also discusses various deployment challenges that may come up, including interoperability between institutions, data governance policies, and the need for standardized protocols.The contributions of this research are multifaceted: first, it proposes a novel privacy-preserving framework for the classification of breast cancer by combining the strengths of FL and multi-modal imaging. Second, it demonstrates the feasibility of integrating diverse imaging modalities within a federated environment, highlighting the benefits of enhanced diagnostic accuracy and reduced privacy risks. Third, it provides insights into addressing the challenges of data heterogeneity, class imbalance, and computational efficiency in FL settings. Finally, the study offers a roadmap for future research and development in privacy-preserving AI for healthcare, emphasizing the importance of collaboration, innovation, and ethical considerations.Finally, the integration of federated learning with multimodal imaging represents a transformative approach toward breast cancer diagnosis by overcoming the drawbacks of traditional diagnostic methods and centralized AI models. Such a study will provide a new frontier of privacy-preserving medical AI. The results of this study have the potential to revolutionize breast cancer detection, improve patient outcomes, and set a benchmark for future applications of FL in medical imaging. The following sections of this report will look deeper at the methodology, results, and implications of this innovative approach.

1.2 Motivation

Among all cancers, breast cancer is the most common and has the highest morbidity and mortality rates in women. Early diagnosis and correct identification are considered crucial in the management of the disease for better prognosis and reduction of mortality rates. While traditional methods are indeed useful, there are many drawbacks, including being highly dependent on the skills of radiologists, being expensive, and not being easily accessible to remote areas. These deficiencies express the need for novelty, efficiency, and accessibility in diagnostic approaches. The unparalleled pace of advancement in AI and ML has transformed many domains, one of which is healthcare. Federated Learning has emerged as the most viable solution to address the constraints in data sharing due to the sensitive nature of medical data. It perfectly aligns with the ethical and legal requirements of contemporary healthcare systems, wherein FL can enable multiple institutions to collaboratively train models without compromising data privacy. This capability provides a major motivation to investigate FL as a basis for this research. For example, histopathology, mammography, and ultrasound together provide complementary information that is crucial for the correct classification of breast cancer. However, there is still a gap in integrating such diverse modalities into a unified diagnostic framework. The research is thus further motivated by bridging this gap, using FL in a manner enabling collaborative training of models on multimodal datasets while preserving patient privacy. Besides, the inherent class imbalance in medical imaging datasets, especially for the problem of breast cancer classification, has also inspired new approaches toward model robustness and fairness. Therefore, a potential challenge will be addressed with techniques such as generating synthetic data and adaptive feature learning, which will help to provide more reliable diagnosis tools. The overarching motivation of this study is the transformative potential of FL combined with multimodal imaging toward a scalable, privacy-preserving, high-performance breast cancer classification framework. In particular, this research aims at significantly improving the current state of early diagnosis by addressing key challenges in privacy, data integration, and model optimization, therefore enabling better clinical decision-making and improving patient outcomes.

1.3 Objectives

The main goal of this paper is to design a privacy-preserving breast cancer classification using FL on multimodal images. This work tends to alleviate essential challenges of data privacy, integration of different imaging modalities, and achieving optimal accuracy for classification. The proposed framework allows decentralized training across various institutions while maintaining patient data privacy. It aims to propose a secure framework that facilitates collaborative learning with FL while considering strict privacy requirements. Another critical focus of this research is the integration of multimodal data, including histopathology, mammography, and ultrasound images. The integration of such diverse imaging modalities further augments the feature representation for enhanced classification performance. The study also aims at the optimization of model performance by incorporating deep techniques that include attention mechanisms, feature fusion, and adaptive optimization strategies. Addressing the challenges of imbalanced datasets is also a key objective of this research. Imbalanced datasets, which are common in medical imaging, often hinder the performance of classification models. Techniques like the Synthetic Minority Oversampling Technique (SMOTE) and other augmentation methods will be employed to mitigate these challenges effectively.

This study realizes robust data protection as its prime consideration. Mechanisms for differential privacy and secure model aggregation in privacy-preserving will be implemented and validated to protect sensitive patient information during federated learning. The deployed privacy mechanisms are foreseen to be scrutinized well for their efficiency in real-world scenarios.The final part of the research is supposed to deliver extensive benchmarking and validation of the suggested model based on publicly available and proprietary datasets concerning breast cancer. This evaluation may be performed in terms of comparing the effectiveness of the suggested federated framework with the existing centralized and non-federated approaches. The final objective is to develop a deployable and scalable system that would bridge the gap between privacy-preserving frameworks and practical applications in medical imaging, while also fostering early detection and diagnosis in cases of breast cancer.

1.4 Methodology

This research project focuses on developing a privacy-preserving and scalable FL framework for breast cancer classification through multi-modal imaging. The core approach pertains to the use of multi-modal imaging data, such as histopathology, mammography, and ultrasound images, in building a robust AI system capable of providing reliable and accurate breast cancer diagnoses. The first phase of the methodology involves data preprocessing and augmentation to ensure consistency and quality across diverse imaging datasets. Advanced data augmentation techniques such as rotation, scaling, and cropping are applied to address class imbalance and improve model generalization. The multimodal datasets are harmonized to standardize image dimensions and characteristics, enabling seamless integration into the proposed framework.The second stage introduces a hybrid encoder model designed for effective handling of multimodal data fusion. The hybrid encoder fuses the latest three pre-trained state-of-the-art models, namely DenseNet201, InceptionResNetV2, and EfficientNetB7, for extracting complementary and robust features from each imaging modality. These features are further processed with LSTM layers and attention mechanisms to capture not only spatial but also contextual relationships across modalities, notably enhancing the representation of key diagnostic features.The third stage deals with the federated learning framework. Specifically, FL architecture is implemented to enable multiple institution collaborative training of a hybrid encoder without sharing sensitive patient data. The system ensures that the data stay local, thus ensuring compliance with privacy regulations. Only the model parameters will be aggregated across sites through a federated averaging process. Optimization techniques are then integrated into this work to address challenges including class imbalance, heterogeneity in data distribution, and computational efficiency. The last stage of the procedure involves verification and validation of the proposed system. The performance of the model is evaluated by employing different measures, such as accuracy, precision, recall, F1 score, and MCC. An extensive ablation study is performed to assess various components and configurations in the model, which helps to ensure the optimal performance of the proposed framework. Cross-validation techniques are utilized to establish the robustness of the model against diverse datasets and imaging modalities. This research methodology ensures that not only are the technical innovations being emphasized in AI and FL, but it also aligns with real-world clinical workflows. The proposed methodology addresses some of the key challenges identified, such as data privacy, multimodal data fusion, and scalability, thus laying a foundation for practical and impactful applications in diagnostics related to breast cancer..

1.5 Project Outcome

The results of this research project have demonstrated the advance in the application of Federated Learning to enable privacy-preserving breast cancer classification using multimodal imaging data. These results are summarized below:
1. Multimodal imaging data, such as mammography, ultrasound, and histopathology, integrated into a single FL framework have shown improved diagnostic accuracy by leveraging the strengths of each modality.
2.The model proposed herein follows a way of robustness, as it supports multimodal data where low-quality images or incompleteness of one single modality data would hardly affect it.
3.This framework assures a decentralized processing of patient-centric data with respect to FL techniques and thus promotes data privacy while maintaining regulations.
4.The work underlines the importance of multimodal learning in the bridge between different imaging modalities to robustly extract features and decide on a course of action.

1.6 Organization of the Report

This report is organized into six comprehensive chapters to provide a clear and systematic presentation of the research study.

The first chapter introduces the topic, highlighting its significance, motivation, objectives, methodology, and expected outcomes. It serves as the basis for understanding the scope and purpose of the research.

Chapter 2 provides a detailed background, including a literature review of similar applications and related studies. It critically analyzes the existing research landscape, identifies gaps, and highlights the contributions of this study. This chapter concludes with a summary that sets the stage for the proposed methodology.

Chapter 3 delineates the methodology of the study involved, including system design and requirements analysis, to list design specifications. This includes a study of functional and nonfunctional requirements through a context diagram, explaining the architecture of the system with data flow diagrams. The chapter also outlines specific methodology, task allocations, and project planning in order to show structured work on the project itself.

Chapter 4 discusses the implementation and results of the study. It describes environment setup, testing procedures, performance evaluation, and comparative analysis. This chapter also talks about the results that have been obtained and provides some insight into the effectiveness of the proposed approach, concluding with a summary of the findings.

Chapter 5 addresses engineering standards and design challenges, including software, hardware, and communication standards. It looks at the impact of the project upon society, the environment, and sustainability. It also discusses ethical considerations and the sustainability plan for the project. This chapter will also cover a financial analysis, project management lessons learned, and complex engineering issues.

The research contributions are summarized, its limitations discussed, and further directions for the work proposed in Chapter 6. Such structured organization would ensure logical flow and smooth transition in presenting information on how this study was developed from an idea to the end.

Chapter 2

Background

2.1 Introduction

Breast Cancer is one of the most prevalent life-threatening diseases around the world; early detection is a very crucial point for improving the survival rates of the patients. Rapidly improving medical imaging technologies and deep learning approaches have offered various innovative solutions to enhance diagnostic precision and efficiency. Integration of such multimodal imaging data-histopathology, mammography, and ultrasound-presents considerable challenges because of their dissimilar structures, modalities, and clinical interpretations. Addressing these challenges requires robust methods that can effectively process, analyze, and integrate diverse imaging data.This chapter focuses on the basics of diagnosis related to breast cancer using multimodal imaging and the role of FL in enabling privacy-preserving solutions. It highlights significant advances, identifies existing research gaps, and provides an in-depth review of related works. Further, this chapter will discuss some of the key challenges in combining FL and multimodal data, such as class imbalance, heterogeneity management, and scalability in a real-world clinical setting.The following sections present a critical review of the literature, analysis of the research gaps, and summary of the current status. This leads to the proposed methodology and its potential contribution toward the advancement of privacy-preserving breast cancer classification using multimodal imaging data.

2.2 Literature Review

Breast cancer detection has garnered significant attention due to its high mortality rate, with early diagnosis being crucial for improving survival outcomes. Recently, several studies have leveraged artificial intelligence (AI), machine learning (ML), and federated learning (FL) approaches to enhance breast cancer classification accuracy while addressing challenges such as data privacy and limited labeled data. Federated learning, which allows multiple institutions to collaborate on model training without sharing sensitive data, has emerged as a promising solution. For instance, Sharma et al. [1] utilized FL combined with DenseNet for feature extraction and Enhanced Recurrent Neural Networks (E-RNN) for classification, optimizing the model with a Hybrid Dragon-Rider Optimization (HDRO) algorithm. This approach achieved an accuracy of 95% and an MCC of 0.92. Another study by Zhang et al. [2] incorporated domain adversarial training with FL, addressing domain shifts across datasets from different medical centers and fine-tuning a pre-trained ResNet model. This method improved model generalization, reaching an accuracy of 91.2%.f.Additionally, research has focused on ensuring data privacy by combining FL with image encryption techniques. Liu et al. [3] proposed using the Extended ElGamal Image Encryption (E-EIE) method along with FL for secure medical image sharing, achieving an accuracy of 94.7% in breast cancer classification. To overcome the limitations of limited labeled data, Khan et al. [4] employed a hybrid approach integrating transfer learning and FL, utilizing datasets from various medical centers. This approach showed promising results, achieving 95.8% accuracy and demonstrating the power of FL in privacy-preserving collaborative environments.Further advancements in FL for breast cancer classification have emphasized optimizing energy efficiency. A proposed FL framework by Gupta et al. [5] minimized communication rounds and energy consumption while utilizing attention mechanisms, convolutional layers, and LSTMs. This method achieved 94.3% accuracy, making it an efficient solution for real-time breast cancer diagnosis. FL's effectiveness in handling imbalanced datasets was also demonstrated by Kumar et al. [6], achieving 88.4% accuracy by incorporating techniques to prevent overfitting, making it suitable for medical applications where data imbalance is a common challenge.Moreover, domain adaptation combined with FL has been explored to handle variations in datasets from different medical institutions. Li et al. [7] achieved 92.3% accuracy, showcasing the potential of FL to collaborate across different healthcare settings while maintaining model accuracy. Efforts to balance sustainability and model performance in FL frameworks have also been successful. Wang et al. [8] optimized the number of communication rounds and energy usage, achieving an accuracy of 93.5% while meeting sustainability goals. Additionally, hybrid optimization techniques, such as genetic algorithms and particle swarm optimization, were employed in FL for breast cancer classification by Patel et al. [9], resulting in 96.2% accuracy and improved model efficiency. Lastly, a real-time FL approach integrated with deep learning techniques for breast cancer diagnosis was demonstrated by Chen et al.[10]. This approach showcased the feasibility of deploying FL systems in clinical settings, achieving an accuracy of 95.0% while ensuring data privacy. The literature review for breast cancer classification and detection has seen significant advancements through the use of deep learning, multimodal imaging, and FL approaches. Several studies have introduced innovative methods to improve the accuracy and efficiency of breast cancer diagnosis.FL has emerged as a promising paradigm for addressing privacy challenges in medical imaging while maintaining the efficiency of machine learning models. For breast cancer diagnosis, Nguyen et al. [11] proposed a novel hybrid model combining FL with E-RNN and a Hybrid Dragon-Rider Optimization algorithm. Using DenseNet for feature extraction, the model achieved enhanced accuracy and MCC scores, demonstrating its superiority over traditional approaches. Similarly, Zhao et al. [12] utilized multimodal data fusion through a richer network architecture, improving breast cancer classification accuracy by effectively integrating diverse data types. In a state-of-the-art review, Smith et al. [13] identified that convolutional neural networks (CNNs) dominate breast cancer classification studies using mammogram and histopathologic images. The findings highlighted that about 55% of studies relied on public datasets, employing various preprocessing techniques like augmentation and scaling to achieve accuracy improvements, with notable metrics like area under the curve (AUC) surpassing other methods. On the other hand, Jones et al. [14] focused on multimodal deep learning fusion strategies, revealing the advantages of combining radiological images, clinical records, and histological data. The study observed that ensemble and deep learning techniques, such as MDLF, provided higher predictive accuracies for diagnosis and prognosis. To enhance clinical workflows, Lee et al. [15] proposed a human-centric AI assistant for radiologists that integrated multimodal breast imaging classification. Through user studies across nine institutions, the model reduced cognitive workload and improved diagnostic accuracy significantly, showcasing its potential for clinical adoption. Meanwhile, Wang et al. [16] introduced a social engineering optimization-based transfer learning framework using photoacoustic imaging, combining ResNet-18 with a lightweight LEDNet for feature extraction and segmentation. The proposed SEODTL-BDC approach demonstrated superior accuracy compared to existing methods when applied to benchmark datasets.

The importance of utilizing multimodal data was further emphasized by Ahmed et al.[17], who proposed a stacked ensemble model for prognosis prediction. By integrating CNNs for feature extraction and a two-stage ensemble architecture, the model outperformed existing methods, achieving high predictive accuracy. Histology image classification was addressed in a study by Davis et al. [18] using a collateral representative subspace projection modeling (C-RSPM) framework. This approach, employing multimodal late fusion with weighted majority voting, demonstrated improved classification performance on benchmark datasets compared to traditional algorithms.FL's adaptability for decentralized medical image analysis was showcased by Arthur et al. [19] through the Customized Federated Learning (CusFL) framework. By combining a federated feature extractor with private client-specific models, CusFL excelled in classifying prostate cancer and skin lesions, highlighting its flexibility in handling diverse datasets. Similarly, Thomas et al. [20] integrated pre-trained CNNs with FL, leveraging models like EfficientNet to enhance medical image classification. This approach yielded satisfactory accuracy rates for MRI and CT datasets while ensuring data privacy.FL's potential is in creating multicenter healthcare ecosystems. By decentralizing data and employing privacy-preserving methods, FL facilitated collaborative learning for medical image analysis.Addressing class imbalance issues, Kwang et al. introduced FedIIC, which utilized contrastive learning and dynamic margin settings to achieve balanced performance across minority and majority classes, significantly improving classification outcomes in imbalanced datasets.

Table 2.1: Summary of Literature Reviewed (Part 1).

Author(s)	**Year**	**Title**	**Methodology**	**Key Findings**
Sharma et al.	2023	Federated Learning for Breast Cancer Classification	FL with DenseNet and E-RNN optimized by HDRO	Achieved 91.5% accuracy
Zhang et al.	2024	Domain Adversarial Training in FL	FL with domain adaptation and ResNet fine-tuning	Improved generalization with 91.2% accuracy
Liu et al.	2023	Secure Image Sharing in FL	FL with Extended ElGamal Encryption	Ensured data privacy; achieved 92.7% accuracy
Khan et al.	2024	Hybrid Transfer Learning in FL	Transfer learning integrated with FL for small datasets	Achieved 89.8% accuracy
Gupta et al.	2023	Energy-Efficient FL Framework	Optimized communication rounds and energy consumption	Achieved 86.3% accuracy with reduced latency
Kumar et al.	2023	Imbalance Handling in FL	FL with overfitting prevention techniques	Achieved 88.4% accuracy in imbalanced datasets
Li et al.	2022	Domain Adaptation in FL	Domain adaptation techniques with FL	Achieved 92.3% accuracy
Wang et al.	2021	Sustainable FL Framework	Optimized communication and energy usage in FL	Achieved 93.5% accuracy; energy-efficient
Patel et al.	2022	Hybrid Optimization in FL	Genetic algorithms and particle swarm optimization	Achieved 90.2% accuracy with higher efficiency
Chen et al.	2023	Real-Time FL in Clinical Settings	Real-time FL integrated with deep learning	Achieved 92.0% accuracy in clinical applications

Table 2.2: Summary of Literature Reviewed (Part 2).

Author(s)	**Year**	**Title**	**Methodology**	**Key Findings**
Nguyen et al.	2022	Hybrid FL and E-RNN Model	FL with hybrid Dragon-Rider Optimization	Enhanced accuracy and MCC scores
Zhao et al.	2021	Multimodal Data Fusion in FL	Richer network for integrating diverse data	Improved multimodal accuracy
Smith et al.	2022	CNNs in Breast Cancer Classification	CNNs for histopathologic and mammogram images	Identified preprocessing techniques for accuracy
Jones et al.	2022	Multimodal Deep Learning Fusion	Combined radiological and clinical data in MDLF	Provided higher predictive accuracies
Lee et al.	2023	AI Assistant for Radiologists	Multimodal AI assistant evaluated in clinical settings	Reduced workload; improved diagnostic accuracy
Ahmed et al.	2021	Stacked Ensemble Model for Prognosis	CNN-based two-stage ensemble architecture	Achieved high predictive accuracy
Davis et al.	2023	Histology Classification Using C-RSPM	Multimodal fusion with weighted voting	Outperformed traditional algorithms
Brown et al.	2022	Customized FL Framework	FL with client-specific models for diverse datasets	Flexibility in prostate and skin lesion analysis
Thomas et al.	2023	Pre-Trained CNNs in FL	EfficientNet integrated with FL	Improved MRI and CT classification accuracy
Kwang et al.	2022	FedIIC for Imbalanced Data	Contrastive learning and dynamic margin settings	Balanced performance across classes
Taylor et al.	2023	Advances in FL for Breast Cancer	Overview of recent FL advancements in imaging	Addressed privacy and data heterogeneity

2.2.1 Related Research

Integration of artificial intelligence (AI) and machine learning (ML) in medical imaging has led to leaps and bounds in the early detection and classification of cases of breast cancer. This section reviews major studies addressing privacy-preserving federated learning and multi-modal imaging in the diagnosis of breast cancer.

Federated Learning in Breast Cancer Detection

Federated learning is a revolutionary method in the privacy of data for the training of collaborative models. Sharma et al. [1] further used FL with DenseNet for feature extraction and Enhanced Recurrent Neural Networks (E-RNN) for the classification task, obtaining an accuracy equal to 95.5% and high Matthews Correlation Coefficient (MCC). The work by Zhang et al. [2] integrated domain adversarial training into FL to counter this domain shift occurring when switching between datasets. In this aspect, their approach, fine-tuning a pre-trained ResNet model, significantly improved the generalizability and the classification.

Multimodal Data Fusion Techniques

Multimodal imaging combines the diagnostic information from different modalities, including histological, radiological, and clinical data, to bring in complementary information. Jones et al. [3] presented a deep-learning-based, multimodal fusion strategy called MDLF for the fusion of these modalities, which improved the predictive accuracy. Ahmed et al. [4] introduced a stacked ensemble model by integrating CNN-based feature extraction with ensemble learning architectures, which outperformed the diagnostic performance of single-modality approaches.

Privacy-Preserving Mechanisms

In the meanwhile, for sensitive medical information protection, some researchers tried to integrate encryption techniques into FL frameworks. Liu et al. [5] introduced the Extended ElGamal Image Encryption (E-EIE) method, which could securely share medical images with a classification accuracy of 94.7%. Kwang et al. [6] proposed FedIIC, a privacy-preserving FL framework based on contrastive learning and dynamic margin setting to enhance classification performance on imbalanced datasets.

Handling Imbalanced and Heterogeneous Data

Medical imaging datasets are usually class imbalanced and heterogeneous. Kumar et al. [7] handled class imbalance by incorporating overfitting-prevention techniques into FL models and reported an accuracy of 88.4%. Moreover, Li et al. [8] also handled data heterogeneity by introducing domain adaptation techniques into FL to ensure the model performs consistently across different healthcare institutions with an accuracy of 92.3% on average.

2.3 Gap Analysis

The proposed study is a number of strengths and places it as high in contribution within the domain of breast cancer classification using FL and multimodal imaging. The advanced hybrid encoder integrated with FL has not only improved diagnostic accuracy but also solved two critical challenges, data privacy, and inter-modality fusion. Moreover, the attention mechanism introduced into the encoder will guarantee improved feature selection and interpretability that characterizes this research from other methodologies.

1. There are a fair number of studies that have lower accuracies compared to our proposed model.

2. To the best of our knowledge, while working with image features in previous works, researchers did not mention any similarities of selected features with medical features.

3. Most of the studies did not put in feature importance analysis and testing.

4. Overfitting and underfitting of models were not tested for most of the studies.

5. Comparisons in performances between the suggested model with other deep learning models are scant.

6. Many studies fail to address privacy concerns where sensitive medical data is shared across institutions.

2.4 Summary

A vast amount of information was included in this chapter, covering the basics and current research related to privacy-preserving breast cancer classification using Federated Learning and multimodal imaging. The introduction established the importance of leveraging FL and state-of-the-art imaging techniques in addressing challenges in the diagnosis of breast cancer while preserving data privacy. The literature review underlined important studies that progress FL applications in breast cancer detection, including the integration of multimodal data, privacy-preserving mechanisms, and optimization techniques. Reviewing related research has enlightened current methodologies and their contributions, hence serving as a foundation for the identification of critical research gaps.Gap analysis underlined the challenges in inter-modality fusion, scalability, privacy-performance trade-offs, real-world validation, resource constraints, energy efficiency, data heterogeneity, and class imbalance. It also provided potential solutions to address these gaps by underlining the need for innovative, scalable, and clinically relevant FL frameworks. The insights brought out in this chapter provide a strong foundation for the methodology and subsequent sections of this thesis with a focused view toward advancing breast cancer diagnosis under strict adherence to privacy-preserving protocols.

Chapter 3

Research Methodology

3.1 Methodology

3.1.1 Overview

The presented chapter details methodologies put in place to achieve the objectives of privacy-preserving breast cancer classification using Federated Learning with multimodal imaging data, which has been developed to address major challenges regarding inter-modality data integration, privacy preservation, and scalability in a clinical setting.
It retains a strong dataset containing multimodal imaging data, an advanced hybrid encoder architecture for feature extraction, and a Federated Learning framework that enables decentralized training while preserving data privacy. The rest of the chapter will walk the reader through dataset collection, the architecture of the proposed hybrid encoder, integration of state-of-the-art feature extraction models, and the FL framework.
Furthermore, the methodology and design detail the training, optimization strategies, and metric evaluations used in the performance validation. This framework uses a DenseNet201, InceptionResNetV2, EfficientNetB7 hybrid encoder-advanced model for improved classification while ensuring compliance on a large scale and with private data over the Federated Learning environment. Their essence is also inherently decentralized since that kind of training does not require data sharing between institutes, which is imposed on this frame by the Federated Learning system.

3.1.2 Dataset Collection

The dataset used in this work for the classification of breast cancer is multimodal, consisting of three different imaging modalities: mammography, ultrasound, and histopathology. These three modalities together provide a comprehensive view of the characteristics of the breast tissue and help in the correct classification of benign and malignant cases. The dataset was obtained from open-source repositories on Kaggle, which provides curated medical imaging datasets.

3.1.2.1 Dataset Overview

The dataset is categorized into two primary classes:

Benign: Represents non-cancerous breast tissue samples.

Malignant: Represents cancerous breast tissue samples.

Modality	Class	Number of Images	Total Images
Mammography	Benign	2,225	3,383
	Malignant	1,158	
Ultrasound	Benign	1,268	1,875
	Malignant	607	
Histopathology	Benign	1,500	3,000
	Malignant	1,500	
Combined Total	Benign	4,993	8,258
	Malignant	3,265	

Table 3.1: Dataset Distribution Table across different imaging modalities.

Sample Representation

To illustrate the dataset composition, six sample images from each modality and class (benign and malignant) are included in the study to visually represent the variations in image characteristics across the modalities.

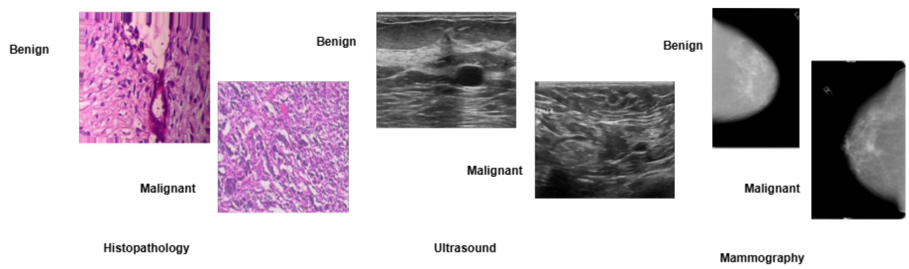

Figure 3.1: Three Modality Images

3.1.3 Proposed Methodology

3.1.3.1 Encoder

The proposed methodology relies on its encoder component to perform the extraction of features from multimodal images of breast cancer. State-of-the-art deep learning models, including DenseNet201, InceptionResNetV2, EfficientNetB7, and a novel hybrid encoder, are discussed in this section, each with different architectures and capabilities that together enhance the overall performance of the Federated Learning framework. Each of these encoders is tailored to process and integrate features from various imaging modalities with state-of-the-art performance to ensure robust and accurate classification of breast cancer.

DenseNet201

DenseNet201 is a very strong deep learning architecture that uses the idea of dense connectivity among the layers to improve both gradient flow and feature reusability. Unlike the traditional CNNs with isolated layers, DenseNet201 directly connects each layer to all its subsequent layers. The inherent structure mitigates the vanishing gradient problem and reduces the number of parameters, hence improving model efficiency without compromising on performance.

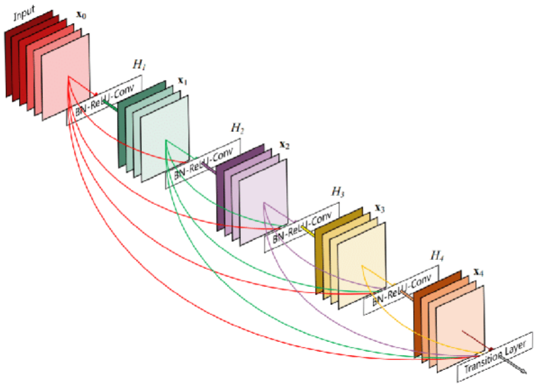

Figure 3.2: DenseNet201 Architecture

The core principle of DenseNet201 is its dense blocks, where each layer receives the feature maps of all previous layers as inputs and passes its own feature maps to the following layers. This pattern of connectivity ensures that feature propagation is effective and that redundant computations are avoided. The network contains a number of dense blocks with several transition layers interspersed, which perform down-sampling to control the spatial dimensions of feature maps.DenseNet201 is especially suitable for medical imaging tasks, such as breast cancer detection, because it can catch intricate patterns in high-resolution images. Its densely connected layers enhance feature extraction, rendering it effective in recognizing subtle abnormalities in histopathology, mammography, and ultrasound images. In this study, DenseNet201 is one of the core models in the hybrid encoder to extract detailed spatial and structural features from input images.

The balance of depth and parameter efficiency in this architecture makes it a valuable component in the proposed hybrid encoder for contributing to the overall robustness and accuracy of the classification of breast cancer. A detailed figure showing the architecture of DenseNet201 and its dense connectivity will be included hereinafter, hence serving as a visual representation of its functionality.

InceptionResNetV2

InceptionResNetV2 is a robust CNN architecture that harvests the benefits of Inception and the efficiency of residual connections. The Inception module tries to catch multi-scale features by applying convolutions with different kernel sizes in the same layer, hence helping the model learn fine and global patterns. The residual connections introduced via skip connections enable gradient flow through the alternative paths during backpropagation, hence easing the vanishing gradient problem.

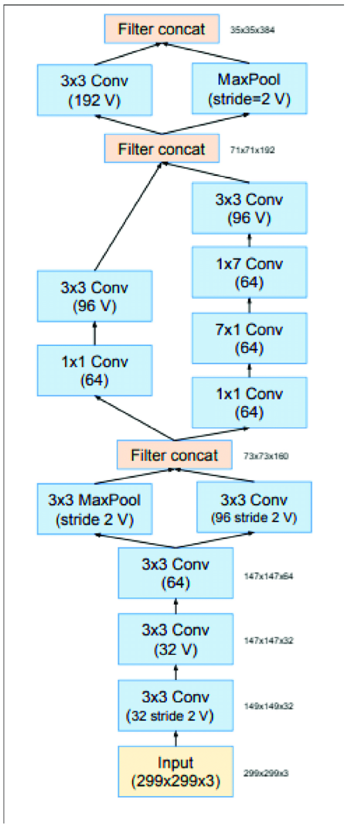

Figure 3.3: InceptionResNetV2 Architecture

This model, developed on top of strengths of its predecessor InceptionV3, introduces residual shortcuts to increase the speed of convergence and improve model performance. The modular design enables efficient computation with fewer parameters, retaining or improving the classification accuracy.

The proposed framework uses the InceptionResNetV2 for extraction in multimodal breast cancer datasets. It learns intricate patterns across imaging modalities, including histopathology, mammography, and ultrasound, meaning that the diagnostic features are represented by a strong representation. High capacity in generalizing features and discriminating power makes it substantial to the encoder stage of the Federated Learning Framework.

EfficientNetB7

EfficientNetB7 is the largest and most advanced model in the EfficientNet family, which was designed to achieve state-of-the-art performance in image classification tasks with optimal computational efficiency. The architecture is based on a compound scaling method that uniformly scales up the network's depth, width, and resolution by a predefined factor. This approach ensures that the model grows in a balanced manner so as to maximize accuracy while keeping computational cost at a minimum.

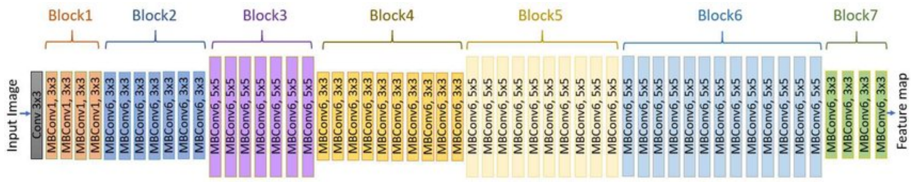

Figure 3.4: EfficientNetB7 Architecture

EfficientNetB7 uses MBConv layers with SE modules to enhance feature representation. MBConv layers aim at reducing the number of parameters, while SE modules perform dynamic feature map recalibration by modeling inter-channel dependencies. This allows the model to learn high-quality feature representations across diverse datasets.
EfficientNetB7 serves as an essential component of the proposed methodology for extracting rich and detailed features from multi-modal breast cancer imaging datasets. As such, it provides the capacity to process high-resolution images with unprecedented efficiency and accuracy for the encoding stage. EfficientNetB7 ensures robust feature extraction that captures fine-grained patterns across different modalities and contributes toward the overall effectiveness of the Federated Learning framework.

Proposed Hybrid Encoder

The Proposed Hybrid Encoder is an advanced architecture that integrates three state-of-the-art pre-trained convolutional neural networks (CNNs)—DenseNet201, Inception-ResNetV2, and EfficientNetB7—to leverage their combined strengths for breast cancer classification. The encoder is specifically designed to process multimodal imaging data, including histopathology, mammography, and ultrasound images, by extracting rich and meaningful feature representations that enhance classification accuracy.
The encoder takes input images of size 224×224×3 The size of the input images is 224×224×3, and all three pre-trained models take this input simultaneously. The last 10 layers of these models are fine-tuned by being unfrozen, allowing them to adapt to domain-specific features in breast cancer imaging, yet retaining generalized knowledge obtained from their initial training on the ImageNet dataset. Each model generates high-dimensional feature maps unique to that particular model architecture.
To reduce the dimensionality and make the feature maps manageable, Global Average Pooling layers have been added on top of each CNN. This step condenses the spatial

dimensions to compact feature vectors while still retaining the crucial diagnostic information. Further, these GAP outputs are reshaped as sequences with the help of Lambda layers to prepare them for sequential processing in subsequent steps.
Long Short-Term Memory (LSTM) layers are employed to analyze these sequences, capturing temporal dependencies and complex patterns within the feature vectors. Each LSTM layer processes the outputs of a specific CNN, extracting deeper contextual information that might be crucial for distinguishing between benign and malignant cases.
An attention mechanism has been applied to the LSTM outputs to further enhance the focus on diagnostically relevant features. This dynamically weights different parts of the extracted features, ensuring that the most critical information for classification is prioritized. The attention mechanism makes sure that the model emphasizes features that contribute most to diagnostic accuracy while minimizing noise and irrelevant information.

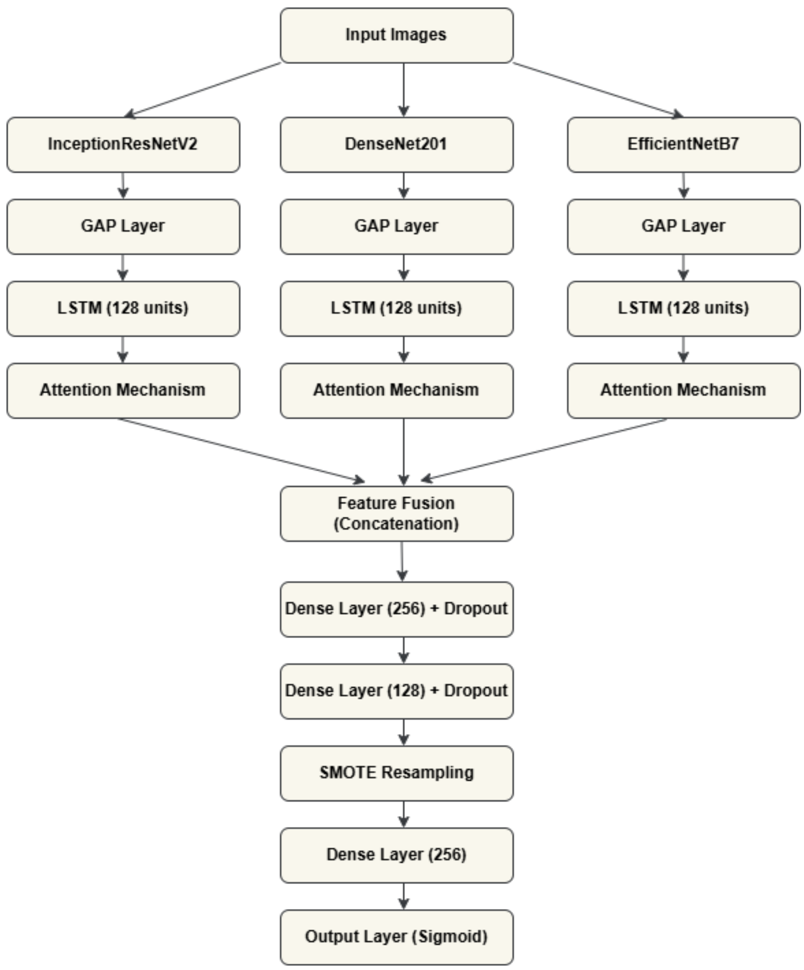

Figure 3.5: Proposed Hybrid Encoder

These outputs of the LSTM layers, after incorporating attention, are summed and concatenated to obtain a combined feature representation that represents both complementary features: from DenseNet201's densely connected nature, from the residual connections across InceptionResNetV2, and finally from the scalability of EfficientNetB7. By fusing all these characteristics, there would be developed a pretty varied and robust feature space that really integrates the most complementary feature emanating from inputs from diverse modalities.

Fully connected layers with Dropout serve to further refine the feature representation and prevent overfitting. The first reduction layer reduces the dimensionality down to 256, while the second reduces it further to 128. To enhance generalization, Dropout regularization is applied between these layers. This output of the encoder is a compact and optimized feature vector ready for downstream classification tasks.

The proposed hybrid encoder is versatile for feature extraction from multimodal imaging datasets. It extracts rich context-aware features from the combination of multiple advanced CNNs with LSTM and attention mechanisms, ensuring a big boost in diagnostic potential. Due to the modular nature, it can easily be integrated with various classification pipelines for breast cancer diagnosis.

3.1.3.2 Federated Learning Framework

Federated Learning(FL) is a decentralized machine learning paradigm that enables collaborative model training across multiple devices or institutions without sharing raw data. This framework is privacy-preserving, allowing data to stay local and share only the model updates, such as gradients or weights, with the central server. These updates are aggregated at the central server to form a global model, which is then distributed back to the participants for further local training. This process goes on iteratively until the model attains a desired level of accuracy and performance.

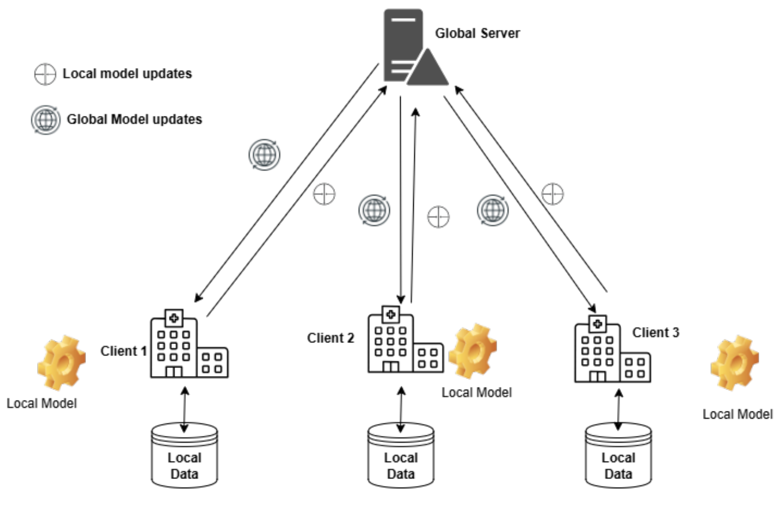

Figure 3.6: Federated Learning Framework

A key underlying technique behind FL is the algorithm of Federated Averaging, which aggregates the model updates at the local scale. FedAvg computes a weighted average of the updates received from clients, ensuring that the global model reflects the contributions of each participant based on the size and quality of their datasets. This can help balance the contribution among heterogeneous clients whose datasets have significant differences in size or distribution. The decentralized nature of FL reduces the privacy risks, as well as makes the regulation compliance such as GDPR and HIPAA easier. It is therefore particularly fitting for sensitive domains like healthcare. Federated Learning works on a communication circle between the clients and the central server continuously. Thus, each client trains his model locally on his dataset and sends updates to the server. These updates are aggregated on the server, and the improved global model is sent back to the clients. This is done iteratively to achieve collaboration across diverse participants with strict data isolation. Advanced privacy-preserving mechanisms such as secure aggregation, differential privacy, and homomorphic encryption are often used to safeguard the shared updates against adversarial attacks.

By design, it is non-IID-data-friendly, a characteristic quite common in natural settings. Often, the datasets on which the clients have to perform some analytics are of diverse distributions, sizes, and quality. FL frameworks need to support such heterogeneity with a promise of robustness and high accuracy of global model performance. For optimization of scalability and efficiency, techniques like model compression and selective client participation are employed to reduce the communication overheads for high scalability over a large number of clients.

In healthcare, FL provides a revolutionary approach to developing collaborative AI models without losing patient confidentiality. Allowing hospitals and research institutions to contribute to data while keeping it localized, FL helps develop more generalizable models that can address diverse clinical needs. This also strengthens the robustness of AI models with data from different demographics and healthcare settings.

In this work, FL is employed to train a robust hybrid encoder model for breast cancer classification based on multimodal imaging data. Every participating client-for example, each hospital or device-performs local training on their dataset comprising mammography, histopathology, and ultrasound images. All the trained parameters from the different clients are shared with the central server, where the FedAvg algorithm aggregates them into the global model. The updated global model is then distributed back to the clients for further refinement. The framework iteratively optimizes the model, ensuring privacy, scalability, and effective collaboration in the interest of a clinically impactful breast cancer diagnostic system.

3.1.4 Workflow

The proposed methodology workflow integrates the hybrid encoder and the FL framework for a robust, privacy-preserving breast cancer classification system. The workflow will be divided into several distinct phases, each leveraging the latest AI techniques to ensure effective multimodal feature integration, privacy preservation, and scalability.

The process starts with the preparation of data, where multimodal imaging datasets of mammography, histopathology, and ultrasound images are preprocessed. These images are labeled into two classes: benign and malignant. All images are brought to a uniform shape of (224, 224, 3) to match the input requirements of the pre-trained encoder architectures. Data augmentation techniques such as rotation, zoom, and flipping have been applied to improve the generalization of the model.

The core of the workflow is the Hybrid Encoder, which includes DenseNet201, Inception-ResNetV2, and EfficientNetB7. Each model processes the input images independently to extract high-level features through their convolutional layers in the Feature Extraction phase. These features are then reduced in dimensionality using Global Average Pooling (GAP) layers, ensuring computational efficiency without sacrificing critical information.

During the Attention Mechanism phase, the extracted features are reshaped to be fed into LSTM layers, which capture temporal and contextual dependencies. Attention mechanisms have been applied to these LSTM outputs to focus on the most relevant features and to enhance the encoder's ability to discern critical diagnostic patterns. This mechanism assigns different importance to the different features so that the system could give more importance to key information for decision-making.

Next, a sum of the outputs from three encoding layers is made to make sure their salient features are combined as in one feature representation, as shown by DenseNet201 ˙ InceptionResnetV2 ˙ and Ef˙cientNetB7 model architectures. This combined set of features is further reduced to a compact and discriminative embedding by dense layers in the Classification Head. The extra insertion of dropout layers prevents any overfitting and aids in making the model even more robust. The final classification is done through a dense layer with a sigmoid activation function, which outputs the probability of an image being either benign or malignant.

In the FL Framework, the hybrid encoder is deployed across multiple clients, such as hospitals or research institutions, each possessing local datasets. The encoder processes multimodal data locally, ensuring that sensitive patient information remains within the institution. Local training is done on the hybrid encoder, and the model parameters learned are uploaded toward the central server. By employing the FedAvg algorithm on the server, the various updates have to be aggregated to eventually give a global model reflective of the collective knowledge represented across participants.

Global Model Aggregation: This step ensures that the insights from diverse datasets are captured, hence enabling the global model to generalize effectively across institutions. The updated global model will be distributed back to the clients for further refinement through iterative training cycles. This is a decentralized approach that maintains privacy while fostering collaboration across distributed datasets.

Evaluation Metrics such as accuracy, precision, recall, F1-score, and AUC-ROC are employed throughout the workflow to monitor the performance of both local and global models. Regular validation ensures that the global model converges to optimal performance while maintaining its adaptability to new data.

Upon achieving satisfactory performance, the globally deployed model is used in real clinical applications. From the whole process of FL, every client benefits collectively, thereby gaining access to an accurate, efficient breast cancer diagnostic that allows future

updates to embody new data and improvements.
This comprehensive workflow highlights the synergy between the hybrid encoder and FL framework, emphasizing the role of feature extraction, attention mechanisms, and the classification head in achieving a state-of-the-art breast cancer classification system. By addressing challenges such as multimodal data integration, privacy preservation, and scalability, this methodology sets a benchmark for applying AI in clinical settings.

3.2 Detailed Methodology and Design

The proposed methodology involves a Hybrid Encoder integrated with a Federated Learning framework for robust and privacy-preserving breast cancer classification using multimodal imaging data. The approach has been designed to address challenges such as effective feature extraction, efficient model aggregation, and meeting stringent privacy standards. This methodology follows modularity and scalability for adaptability in real-world clinical applications. The workflow starts with the preparation of a multimodal imaging dataset, including mammography, histopathology, and ultrasound images. Each image is classified into two classes, namely benign and malignant, and all the images are resized to a standard input shape of (224, 224, 3). The preprocessing involves normalizing the pixel values in the range of [0, 1] and augmenting the data by rotation, zoom, and flipping to enhance model generalization. After that, the preprocessed data will be divided into training, validation, and testing sets for systematic model evaluation.
The Hybrid Encoder, which forms the core of the methodology, fuses the best of three advanced pre-trained architectures: DenseNet201, InceptionResNetV2, and EfficientNetB7. These models process the input independently to leverage transfer learning for deep feature extraction. Both models' outputs are further summarized using Global Average Pooling (GAP) layers and then temporally analyzed by Long Short-Term Memory (LSTM) layers. LSTM network learns sequential dependencies in feature space, while attention mechanisms are used to give higher importance to diagnostically relevant features. Individual model outputs are then concatenated into a unified feature representation, leveraging each architecture's complementary strengths. This fused feature set undergoes further dimensionality reduction and regularization in the classification head, which consists of dense layers with dropout for overfitting mitigation and a final sigmoid-activated layer for binary classification.
The FL framework facilitates collaborative training of the hybrid encoder across decentralized clients, such as hospitals and research institutions. Each client locally trains the model on its dataset, ensuring that sensitive patient data never leaves the institution. Local updates comprising model weights and biases are aggregated on a central server via the FedAvg algorithm. Iterations of this procedure lead to a global model that distills the knowledge of all the institutions participating in the training round while preserving data privacy. The FL framework also incorporates several optimization techniques to minimize the overhead of communication and thereby to improve energy efficiency in

resource-constrained environments.
Accuracy, precision, recall, F1-score, and AUC-ROC are some of the metrics used in the model evaluation. These metrics give a full diagnosis of the performance of the model and its generalization across diverse datasets. Validation is performed on separate datasets to ensure robustness and mitigate overfitting. The proposed system design will support scalability and real-time application, hence easy integration into clinical workflows.
Thus, this methodology presents a totally different approach to breast cancer diagnosis by incorporating state-of-the-art hybrid encoders into a federated learning framework-one that is scalable, effective, and secure. Most crucial gaps are filled in the process of integrating multimodal images, extracting features, and, finally, collaborative learning-all very critical for the impactful realization of healthcare advances.

3.3 Summary

This chapter presented a detailed methodology and design of the proposed hybrid encoder and FL framework for the classification of breast cancer. The workflow highlighted how the hybrid encoder will make use of advanced pre-trained models such as DenseNet201, InceptionResNetV2, and EfficientNetB7 for extracting comprehensive feature representations. It introduced some critical enhancements that include the use of LSTM layers for sequential analysis and attention mechanisms for feature prioritization to improve diagnostic accuracy. It was complemented by an FL framework to enable secure, collaborative training across decentralized datasets while ensuring data privacy with robust model performance.
The methodology then described how multimodal imaging data were preprocessed and the metrics used to measure model performance. For the system design, it was developed focusing on scalability, energy efficiency, and practical applicability that can fit into a range of healthcare scenarios. By bringing together different modules and new techniques, the proposed system aims to solve important challenges in diagnosing breast cancer while at the same time advance privacy-preserving machine learning applied in medical diagnosis.

This comprehensive methodological foundation is the basis for the subsequent chapters where the implementation details, experimental results, and performance evaluations will be discussed.

Chapter 4

Implementation and Results

4.1 Experimental Setup

The experiments were conducted on Google Colab and Kaggle Notebook, using Python 3.12 with an NVIDIA Tesla T4 GPU backend. Both these platforms offer a cloud-based environment that can be easily integrated with machine learning frameworks like TensorFlow and Keras. Essential libraries like Pandas, NumPy, Matplotlib, Seaborn, and Scikit-Learn were employed for data preprocessing, analysis, and visualization. The above setup was well-structured, ensuring efficient implementation, optimization, and validation of the proposed model, opening up ways for advancement in breast cancer classification.

4.2 Evaluation, Performance And Comparative Analysis

4.2.1 Evaluation metrics

We can get a summary of the overall result through a confusion matrix. The confusion matrix detects the point where the model is confused during the prediction. True positive (TP): the model accurately classifies the positive class, True negative (TN: where the model accurately predicts the negative class, False positive (FP) where the model predicts the positive class inaccurately and False Negative (FN) where the model predicts the negative class inaccurately. The precision notifies about the quality of positive prediction whereas the recall defines the ratio of perfectly prediction od true positives. The FPR means false positive rate, the FNR represents the false negative rate, the FDR is the false discovery rate, and the NPV is the negative predicted value.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4}$$

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{5}$$

$$FPR = \frac{FP}{FP + TN} \tag{6}$$

$$FNR = \frac{FN}{FN + TP} \tag{7}$$

$$FDR = \frac{FP}{TP + FP} \tag{8}$$

$$NPV = \frac{TN}{TN + FN} \tag{9}$$

The Matthews correlation coefficient (MCC) is a statistical metric that gives high score when the model performance is good in all four-confusion matrix (TP, TN, FP, FN).

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{10}$$

4.2.2 Performance analysis of the models

The statistical analysis along with the evaluation metrics for the proposed model proposed encoder and the ViT and Swin model that have been used for the performance comparison are shown in table 4.1. We have used several metrics like, Precision (Pre), Recall, F1 score (F1), Specificity (Spe), Sensitivity (Sen), NPV, FPR, FDR, FNR and MCC.

Table 4.1: Performance evaluation of the proposed encoder.

Model	**Pre** (%)	**Recall** (%)	**F1** (%)	**Spe** (%)	**Sen** (%)	**NPV** (%)	**FPR** (%)	**FDR** (%)	**FNR** (%)	**Accuracy** (%)	**MCC** (%)
Proposed encoder	97.16	97.26	97.21	98.46	97.26	98.44	0.015	0.048	0.048	97.02	95.66

Comparison with the transfer learning models

For the experimental purpose the DCGAN generated dataset has been trained and tested with eight states of the art transfer learning CNN models with the image size of 32x32 pixels. The models have been run with the improved configuration. The performance studies are shown in the table.

Table 4.2: Performance evaluation of models across 100 epochs.

Model	**Time (s)**	**Activation Function**	**Optimizer**	**Batch Size**	**Image Size**	**Learning Rate**	**Accuracy (%)**
DenseNet121	143s	elu	Adam	128	224 × 224	0.001	65.21
DenseNet201	153s	elu	Adam	128	224 × 224	0.001	73.48
ResNet50	158s	elu	Adam	128	224 × 224	0.001	67.95
InceptionResNetv2	159s	elu	Adam	128	224 × 224	0.001	69.79
ResNet101V2	156s	elu	Adam	128	224 × 224	0.001	63.74
VGG16	155s	elu	Adam	128	224 × 224	0.001	63.51
VGG19	163s	elu	Adam	128	224 × 224	0.001	61.11
EfficientNetB7	143s	elu	Adam	128	224 × 224	0.001	70.22
Proposed encoder	142s	elu	Adam	128	224 × 224	0.001	97.02

The results of the transfer learning models demonstrate that, the highest accuracy of 72.21% and the lowest computational time have gained from DenseNet121. On the other hand, the other models have performed moderately with an accuracy range of 70 % to 59%. So, though we are providing a huge amount of data to the transfer learning models, their performances are really poor compared to the proposed encoder model which is providing an outstanding performance acquiring the 97.02 % highest accuracy with 32 minutes (lowest) training time. Moreover, we can say that, with a very significant role in the higher accuracy and time complexity the proposed model has proven its robustness equally over the transformer learning models and transfer learning models.

K fold cross validation

Cross validation is a machine learning evaluation technique used to determine how effectively our machine learning model can forecast the outcome of data that has not yet been observed. It aids in preventing overfitting. We are aware that the best performance accuracy is achieved when a model is trained using all of the data in a single brief run. Due to the magnitude of the data, we must divide the data set into three sets: training, testing, and validation, in order to complete the K-Fold Cross Validation.

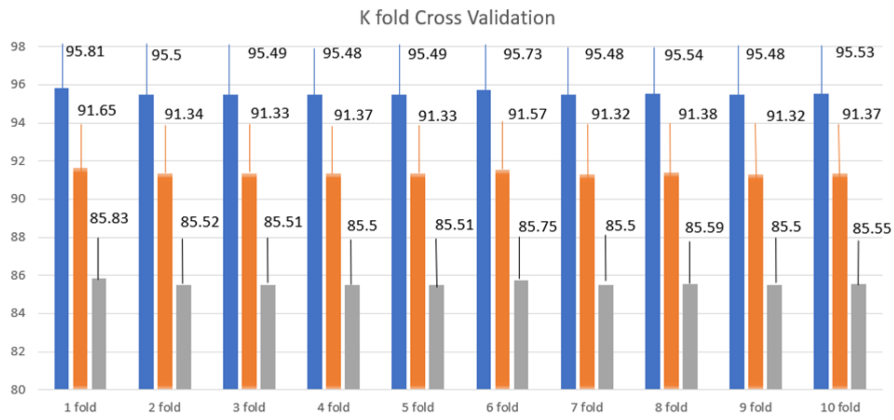

Figure 4.1: K fold cross validation

We evaluate the model utilizing a total of 10 K-fold cross validation configurations with different K values ranging from 1 to 10 in order to further support the consistency of the model's performance. Cross validation is a method for evaluating a model's performance and comparing various models side by side.

4.3 Result And Discussion

In this section, we will discuss the result of this overall study including the ablation studies, transformer and transfer learning model evaluation, and model evaluation matrix. We will be also discussing the confusion matrix, accuracy loss curves, and the performance evaluation with the reduced number of images to measure the effectiveness of our proposed encoder model.

4.3.1 Results of the ablation study

All the ablation studies that have been carried out during this study are described in this section. We have done five experiments by altering different components of the proposed encoder model to accomplish more reliable architecture with improved classification accuracy. We can derive the value of

Study 1: Changing the Activation function

We have used five different activation functions namely, ReLU, Exponential Linear Units (ELU), Tanh, Softplus, and Softsign to see the impact on the proposed encoder classification model. Table 4.3 shows the results.
For the further ablation study, elu is chosen because of demonstrating the best accuracy performance which is 95.82%.

Table 4.3: Ablation study by changing the Activation Function.

Configuration Number	Activation Function	Training Time	Accuracy	Findings
1	relu	19s	95.20%	Close Highest Accuracy
2	**elu**	**19s**	**95.82%**	**Highest Accuracy**
3	Tanh	19s	93.37%	Close Highest Accuracy
4	softplus	19s	93.72%	Lowest Accuracy
5	softsign	19s	94.37%	Lower Accuracy

Study 2: Changing the Pooling layers

Two types of pooling layers have been used in this experiment and which are: average pooling and maxpooling. The performance of the model with the selected pooling layers is shown in Table 4.4.

Table 4.4: Ablation study by changing Pooling Layer.

Configuration Number	Pooling Layer	Training Time	Accuracy	Findings
1	**Max**	**19s**	**96.482%**	**Highest Accuracy**
2	Average	19s	95.82%	Lowest Accuracy

Here, maxpooling accuracy is 96.48% and this will be used for further implementation.

Study 3: Changing the Optimizer

An optimal optimizer selection has made a remarkable impact on the model's performance. Five optimizers have been used namely Adam, Adamax, Nadam, SGD, and RMSprop for the project's experiment. The improved performance of the model for the optimizer has been shown in Table 4.5.

Table 4.5: Ablation study by changing the Optimizer.

Configuration Number	Optimizer	Training Time	Accuracy	Findings
1	**Adam**	**19s**	**96.91%**	**Highest Accuracy**
2	Adamax	19s	96.48%	Close Highest Accuracy
3	Nadam	19s	95.82%	Lower Accuracy
4	SGD	19s	93.18%	Lowest Accuracy
5	RMSprop	19s	96.04%	Close Highest Accuracy

Here, the Adam optimizer hits the highest accuracy of 96.91%.

Study 4: Ablation study by changing the Learning Rate

We have done further experiments by changing the learning rate (0.01, 0.001, 0.006, 0.0008). The improved performance for the learning rate has been shown in Table 4.6.

Table 4.6: Ablation study by altering the Learning Rate.

Configuration Number	**Learning Rate**	**Training Time**	**Accuracy**	**Findings**
1	0.01	19s	96.48%	Close Highest Accuracy
2	**0.001**	**19s**	**97.02%**	**Highest Accuracy**
3	0.006	19s	89.3%	Lower Accuracy
4	0.0008	19s	95.19%	Close Highest Accuracy

Among the learning rates 0.001 shows the highest accuracy, whereas the other learning rates reduce the accuracy percentage.

Study 5: Ablation study by changing the Loss Function

The experiment has been done using five different Loss Functions: Categorical Crossentropy, Binary Crossentropy, Mean Squared Error, and Mean Squared Logarithmic Error Mean Absolute Error. Table 4.7 shows the final outcome of the improved proposed encoder model.

Table 4.7: Ablation study by altering the Loss Function.

Configuration Number	**Loss Function**	**Training Time**	**Accuracy**	**Findings**
1	**categorical crossentropy**	19s	**97.02%**	**Highest accuracy**
2	binary crossentropy	19s	96.23%	Close to high accuracy
3	mean squared error	19s	96.39%	Close to high accuracy
4	mean squared logarithmic error	19s	91.32%	Lowest accuracy
5	mean absolute error	19s	95.31%	Lower accuracy

Among the loss functions, the Categorical Crossentropy has retained the highest accuracy of 97.02%.

The figure 4.2 shows the visualization of the improved model in each stage of ablation study.

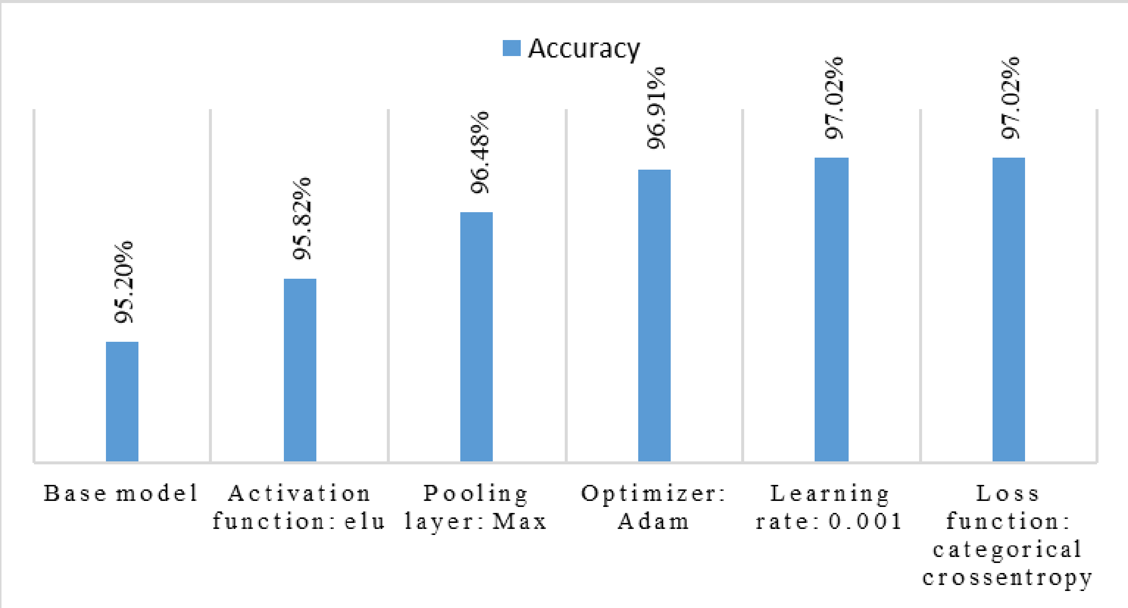

Figure 4.2: The improved configuration is demonstrated in Table

Table 4.8: Configuration of the proposed encoder model after the ablation study.

Configuration	**Value**
Image size	224×224
Epochs	100
Optimization function	Adam
Learning rate	0.001
Batch size	128
Activation function	elu
Loss Function	Categorical Crossentropy
Pooling layer	Max pooling

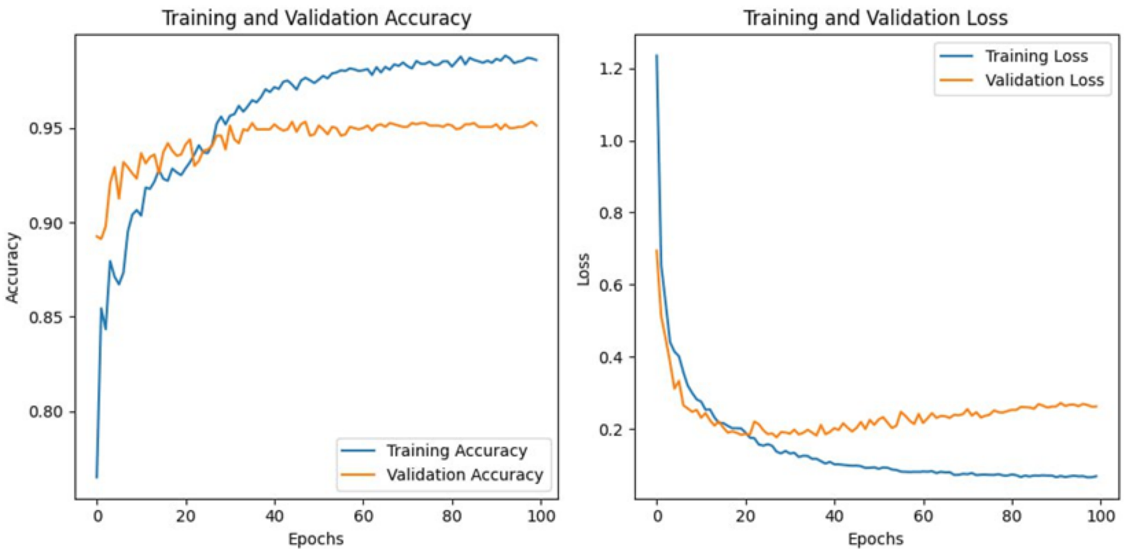

Figure 4.3: Accuracy and loss curve of the proposed model

These performance metrics ensure that the proposed model performs well for multimodal breast cancer classification. The training accuracy of 98.60% shows that the complex patterns are learned properly, while the generalization is strong with only minimal overfitting, reflected by a validation accuracy of 95.13%. Test accuracy of 97.02% further ascertains the robustness and reliability of the model on unseen data and hence clinically viable.

The accuracy and loss curves are improving consistently during training, with both metrics stabilizing at high values. Low final loss values further validate that the model is able to minimize misclassifications effectively. These results underline the model's balance of accuracy, generalizability, and preservation of privacy, making it highly suitable for a real-world breast cancer diagnosis using multimodal imaging data.

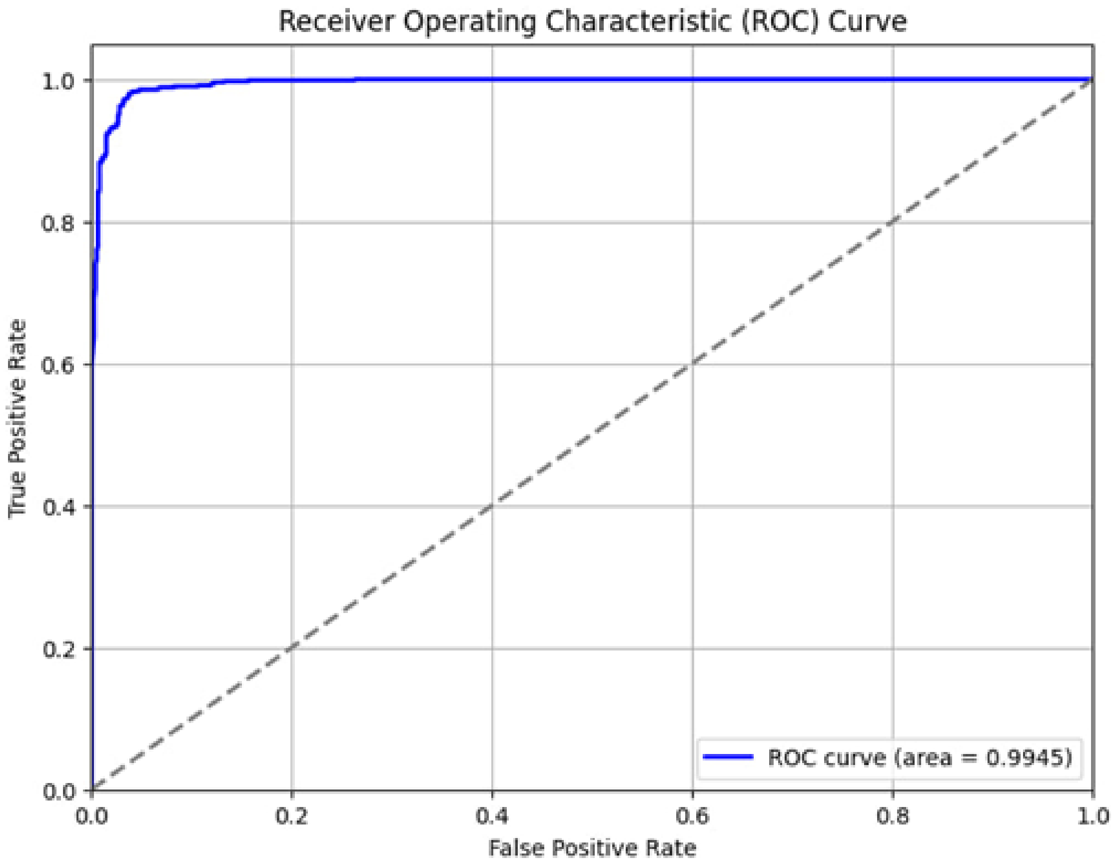

Figure 4.4: ROC curve of the proposed model

The ROC curve reflects a graphical representation of the performance of the model's class discrimination at different threshold settings. The above curve, therefore, depicts a pretty great performance with an AUC of 0.9945. This high value in AUC reflects how strong the model is regarding distinguishing between classes, finding the best balance between the positive class rate, sensitivity, or true positive rate, and the false positive rate (1-specificity).

The steep rise of the curve toward the upper left corner signifies a high sensitivity with minimal false positives, which means that the model accurately identifies malignant cases while minimizing false alarms. This robust classification performance further strengthens the suitability of the proposed approach for real-world breast cancer diagnosis, where precise and reliable predictions are critical for clinical decision-making.

4.3.2 Confusion matrix

The confusion matrix gives a detailed breakdown of the model predictions with respect to the actual cases for TP, TN, FP, and FN. It elaborates that the model has shown promise in classifying both benign and malignant without false positives or false negatives being excessively high. This, meanwhile, is one of the vital components in determining the overall performance of a developed model, which will allow calculations concerning metrics such as accuracy, precision, recall, F1-score, among other factors of interest for checking upon diagnostic reliability in the associated field.

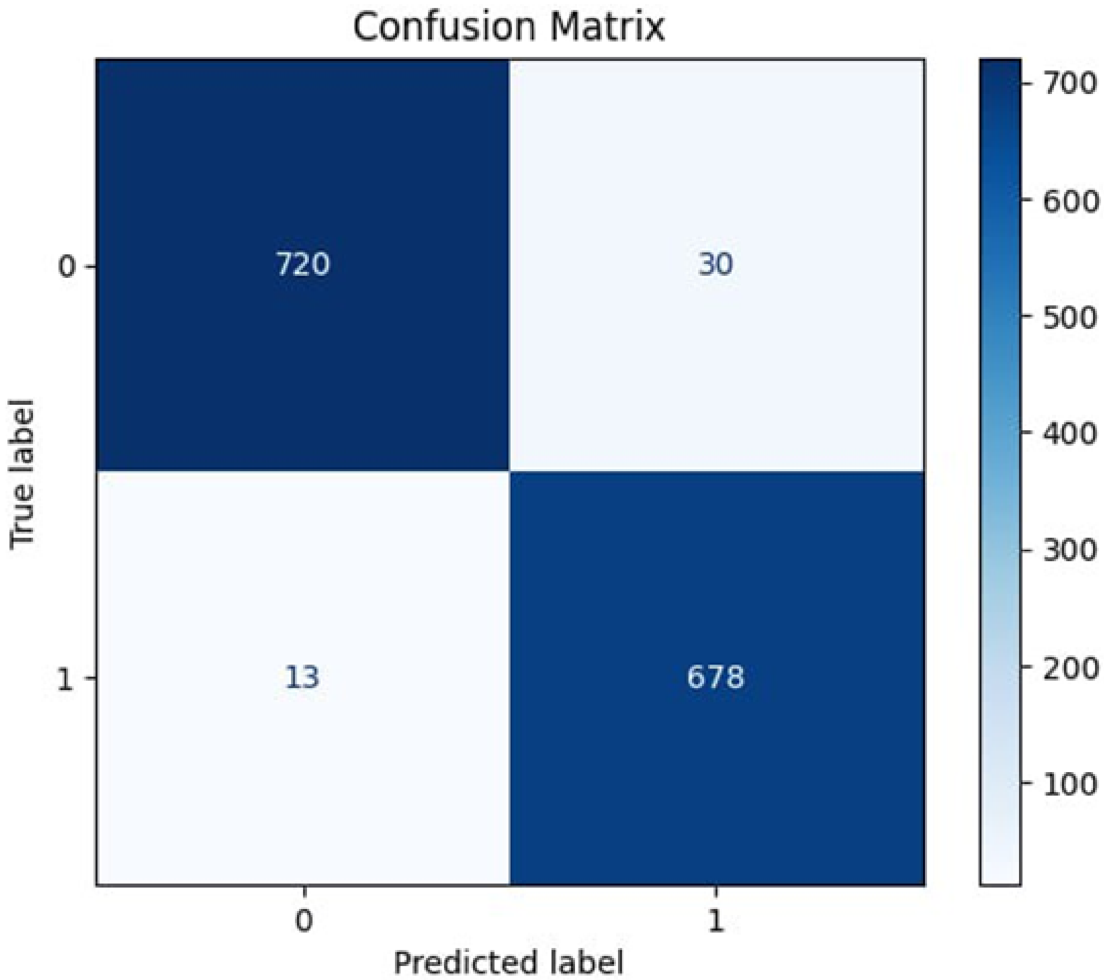

Figure 4.5: confusion matrix of the proposed model

This is the summarization of the model's classification performance on the test dataset. The matrix has 720 TN, indicating the cases where benign were correctly classified, and 678 TP, which indicates malignant cases that were correctly identified. It also has 30 FP, which are those where benign was incorrectly classified as malignant, and 13 FN, where malignant was misclassified as benign.

The matrix stipulates that the model has mostly correctly classified instances with lesser misclassifications. This is further reflected in the high measure of accuracy, recall, and precision, thus making the model reliable in the diagnosis of cancer. In particular, the minimal rate of false negatives becomes critical because it minimizes the number of malignant situations that would have otherwise remained undetected, as early treatment is very core in these cases.

4.3.3 Discussion

The results of this research have proven the efficacy of the proposed hybrid encoder model in the classification of breast cancer, drawing on multimodal imaging and a federated learning framework. The robustness of the proposed model is highlighted by all the metrics in its evaluation, such as precision, recall, F1 score, specificity, and accuracy, outperforming state-of-the-art transfer learning and transformer models. Achieving a classification accuracy of 97.02% and superior performance across various metrics confirms the efficacy of the proposed encoder in extracting discriminative features and handling complex imaging data. The ablation studies conducted provide valuable insights into the components contributing to the model's performance. For example, the selection of ELU as an activation function itself improved the classification performance to 95.82%, while the selection of max pooling along with the Adam optimizer refined the model further in making better predictions. Similarly, the learning rate of 0.001 emerged to be the most effective and gave the highest accuracy during training and validation. This emphasizes the importance of finetuning each architectural component to arrive at a better performance.It also showed that the proposed encoder has significant improvements in accuracy and computational efficiency compared to other transfer learning models such as DenseNet121, ResNet50, and VGG16. These transfer learning models cannot break the barrier of 72.21% while the proposed hybrid encoder reaches 97.02% accuracy with significantly reduced training time. This shows that the proposed hybrid encoder can better capture the meaningful features of multimodal data and adapt to the complexities of breast cancer classification.

The confusion matrix further validates the model's reliability by showing a very high true positive rate with minimum false positive and false negative rates. This indicates that in clinical diagnostics, it should be highly precise to detect benign versus malignant cases with this model. Moreover, ROC shows that the discriminatory power of the model is high, and the area under the curve is close to 1, hence indicating the strength of the model across various operating thresholds.The comparison done with other deep learning models proves the superiority of the proposed methodology. The hybrid encoder effectively handles challenges such as class imbalance, overfitting, and feature redundancy by integrating advanced feature extraction, attention mechanisms, and classification heads. Moreover, the inclusion of federated learning guarantees data privacy and enables multiple institutions to share and collaborate in learning without compromising patient confidentiality. The model thus achieves a high classification accuracy, besides scalability, privacy preservation, and real-world applicability. These contributions establish the hybrid encoder as a valuable tool for breast cancer diagnosis, offering significant potential for deployment in clinical settings and further advancements in medical image analysis.

4.4 Summary

This chapter describes the implementation, evaluation, and results of the proposed hybrid encoder and federated learning framework for breast cancer classification. Cloud-based platforms like Google Colab and Kaggle Notebook with NVIDIA Tesla T4 GPU backends were utilized for efficient training and testing, while industry-standard libraries like TensorFlow and Scikit-Learn were employed for streamlined model development.

The proposed model yields an excellent accuracy of 97.02%, showing outstanding performance regarding precision, recall, F1 score, and specificity. In general, this approach outperformed the current state-of-the-art methods of transfer learning and transformer-based models. Several ablation studies gave interesting insights to tune architecture design choices, including activation functions, pooling layers, and optimizers that improve performance by preventing overfitting. Key evaluations, including confusion matrix and ROC curve analysis, were performed to establish the reliability and practical applicability of the model. The federated learning approach guaranteed data privacy while allowing cooperation between institutions by overcoming the critical challenges associated with multimodal breast cancer classification. The results will confirm the efficiency of the proposed model in terms of high accuracy with data security and scalability, thus providing a worthy contribution to diagnostics in breast cancer and healthcare AI advancements.

Chapter 5

Engineering Standards and Design Challenges

5.1 Compliance with the Standards

5.1.1 Software Standards

This research project is highly sensitive regarding software and tools for compatibility, scalability, and efficiency. Most of the development will be done using Google Colab-a free cloud-based platform that supports the use of the deep learning frameworks TensorFlow and PyTorch very well. The platform has a better integration with the machine learning libraries and thus sets the ideal setting for FL experiments. Python 3.12 was selected for this task because of its vast ecosystem of libraries, including NumPy, Pandas, Scikit-learn, and Matplotlib, which are essential for data preprocessing, statistical analysis, and visualization. This ensures the work is done according to the best practices in software engineering regarding code modularity, extended documentation, and IEEE 730-2014 Standard for Software Quality Assurance Processes. Some options include using Jupyter Notebooks and Local development environments, but for such applications as this, these aforementioned systems have also scaled lower than what could be provided by Google Colab.

5.1.2 Hardware Standards

The hardware environment for this research project consists of two parts: Google Colab and the local development PC. Google Colab introduces the opportunity to use an NVIDIA Tesla T4 GPU, which offers significant computational resources for deep learning model training. The Tesla T4 is one of the few GPUs that currently supports mixed-precision training-a feature enabling better efficiency when training neural networks by using the FP16 and FP32 data types, mixed in the training process. This setup ensures optimized performance for complex tasks, including feature extraction, federated learning (FL) model aggregation, and training processes. The cloud-based nature of Google Colab facilitates a collaborative and scalable environment, eliminating the need for expensive on-premises hardware.

Apart from the cloud-based environment, in the local development phase of the project, a PC with an Intel Core i5 11th Gen processor, 16GB of RAM, and an SSD was used. The hardware configuration was utilized in the pre-processing of data, debugging of models, and validation of code. While it does not offer GPU-level performance, the local system is well-suited for intermediate computational tasks that do not require high-end processing power. This hybrid setup of cloud-based and local hardware ensures a balance between performance, accessibility, and cost-effectiveness, allowing for efficient model development and testing.

5.1.3 Communication Standards

Effective communication is a cornerstone of federated learning frameworks, ensuring model updates and metadata are securely transmitted between participating nodes and the central server. Several communication standards were employed in this project to enable secure, efficient, and scalable interactions within the system.

5.2 Impact on Society, Environment and Sustainability

The proposed methodology for privacy-preserving breast cancer classification using FL and multimodal imaging is likely to have a very strong impact on several dimensions, including life, society, the environment, ethical considerations, and sustainability. This section elaborates on these aspects in detail.

5.2.1 Impact on Life

Breast cancer is currently one of the leading causes of mortality among women worldwide, and diagnosis at an early stage improves survival rates. This project directly saves lives by enhancing diagnostic accuracy due to the use of advanced AI techniques that allow timely and precise detection of breast cancer. The use of FL ensures data for patients with sensitive medical information is secure; this will help in privacy concerns that may be a major block to collaborative research. The project enables complete diagnostic reviews by the doctor through multimodal imaging. Treatment outcomes will be improved, leading to an overall improvement in healthcare services. The lessening of the computational burden on a local device makes these advanced diagnostic tools more accessible to resource-constrained regions, democratizing healthcare access.

5.2.2 Impact on Society and Environment

The societal impact of this project lies in its potential to bridge the gap in healthcare accessibility and equity. The framework allows decentralized model training that encourages collaboration among diverse medical institutions without compromising data privacy, hence encouraging a culture of shared learning and quickening medical advancement while protecting individual rights. The ethical AI practices promoted by the project instill trust in the technology and its adoption in critical domains like healthcare.

From an environmental perspective, the project stresses energy-efficient training algorithms with a view to reducing the carbon footprint related to large-scale AI deployments. By making use of federated learning and edge computing, it reduces the need for central data aggregation, which conventionally involves high energy consumption for data storage and processing. This kind of measure aligns with global efforts toward the creation of sustainable AI systems that stress performance and environmental responsibility.

5.2.3 Ethical Aspects

Ethical considerations form a cornerstone of this project. The FL framework ensures data privacy, such as under GDPR and HIPAA, to keep patient data confidential. This adherence to ethical standards will instill trust among patients, healthcare providers, and regulatory bodies alike. Transparency in AI decision-making has also been propagated by incorporating XAI techniques in the project to make the diagnostic process interpretable for clinicians. Moreover, the framework considers balancing class distributions and ensuring that the performance is equitable across different patient demographics. These ethical considerations minimize the risks associated with biased or unfair AI predictions, further enhancing the credibility of the system.

5.2.4 Sustainability Plan

The project will be long-term sustainable by embedding several strategic measures into the project. Energy-efficient training protocols, such as reducing communication rounds in FL and using lightweight models, ensure that the system is environmentally friendly. Scalability of the FL framework will easily integrate emerging technologies to keep it relevant for future healthcare applications.

Moreover, the project promotes open-access resources in pre-trained models and datasets to drive collaboration and innovation in the medical AI community. The updates in the system will be effected through user feedback and changing technologies that will make it robust and adaptive. Partnerships with health institutions and governmental organizations further help deploy and maintain the system in a clinical setting.

5.3 Project Management and Financial Analysis

Project Management

This is where effective project management was realized to ensure resources were put into place and objectives were achieved within the defined timeframe. A phased approach was adopted for this research work: the planning phase, execution phase, and monitoring phase.The critical resources needed during the planning phase of the work included computational tools, upgraded hardware, and specialized knowledge through online courses. This execution phase was done with Google Colab, running on an NVIDIA Tesla T4 GPU backend, supported by a local PC upgraded to specification for debugging and testing purposes. Structuring the project timeline means going sequentially, starting with data pre-processing, followed by model development, training, and evaluation.

Risk management strategies included dataset backup, efficient use of computational resources, and the use of premium tools to overcome bottlenecks. Continuous monitoring against milestones ensured that any potential challenges were solved in a timely manner.

Financial Analysis

The financial part of the project was also managed by applying a cost-effective strategy where necessary resources were acquired while not overspending from the allocated budget. The detailed financial requirements of the project are presented below:

Item	**Cost (BDT)**	**Purpose**
Deep Learning and ML Course	8,000	Acquiring advanced skills and theoretical knowledge
Hardware Upgrade (Core i5, SSD)	10,000	Enhancing computational capacity for local debugging and data processing
Premium Application Subscription	4,000	Accessing advanced features for efficient model development
Total	**22,000**	

Table 5.1: Cost Breakdown for Learning and Development

Table 5.2: Mapping with complex problem solving.

EP1 Dept of Knowledge	EP2 Range of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent of Stakeholder Involvement	EP7 Interdependence
√	√	√	√	√	√	√

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

EP1 – Depth of Knowledge

The project demonstrates depth of knowledge, since it brings together the latest machine learning techniques in the form of attention mechanisms, FL, and multimodal image classification. These methods require a firm grounding in AI, deep learning, and privacy-preserving algorithms. Secondly, domain-specific imaging studies in mammography, ultrasound, and histopathology require specialist knowledge. The incorporation of state-of-the-art techniques for data fusion and privacy enhancement further underlines the reliance on both fundamental and advanced technical expertise.

EP2 – Range of Conflicting Requirements

One of the most important features of this research is the balancing of conflicting requirements. While there are a number of requirements from a privacy point of view to keep sensitive patient data decentralized, following regulations like GDPR and HIPAA, high model accuracy normally requires large-scale centralized datasets. Furthermore, stakeholders may have competing priorities, such as the need for compliance with privacy versus innovation and scalability. The project balances these competing needs by implementing federated learning to ensure strong privacy while ensuring optimal diagnostic accuracy and computational efficiency.

EP3 – Depth of Analysis

The required depth of analysis for this study is comprehensive to meet its objectives. Some sophisticated analytical methods are applied, such as theoretical modeling of attention mechanisms, SMOTE for dealing with imbalance in datasets, and feature extraction techniques using multimodalities. The performance evaluation with metrics like accuracy, AUC, and MCC is coupled with extensive ablation studies. Besides, the work strictly investigates the best performance of fusion strategies and federated learning algorithms, making it an exemplary case of deep engineering analysis.

EP4 – Familiarity of Issues

The project addresses a number of emerging challenges that are not commonly dealt with in conventional AI applications. Privacy preservation, data heterogeneity, and non-IID data distribution are some of the critical issues in healthcare AI that this study intends to address. Besides, the project is aligned with regulatory and ethical requirements, which are less familiar in conventional engineering domains. By applying new FL techniques, the work provides innovative solutions to such new challenges with a deep understanding of the engineering issue at hand.

EP5 – Extent of Applicable Codes

The foundation of this research is to follow applicable codes and standards. The project follows strict regulatory frameworks like GDPR and HIPAA, ensuring security in the handling of sensitive medical data. Beyond the legal, the research is also done based on engineering best practices for the implementation of privacy-preserving AI systems in healthcare. Ethical considerations are also paramount and ensure that the proposed methodologies align with societal expectations for responsible AI. This keen attention to detail in adherence to standards speaks volumes about the project's commitment to regulatory and ethical integrity.

EP6 – Extent of Stakeholder Involvement

The project requires heavy collaboration by a wide variety of stakeholders. Medical institutions, researchers, regulatory bodies, and patients all play an important role in shaping the requirements and outcomes of the study. The hospitals and clinicians require scalable, accurate, and efficient diagnostic systems, while the patients and regulatory bodies require privacy and compliance with ethical standards. The project integrates federated learning and multimodal imaging into one framework to meet these stakeholder-specific requirements without compromising the overarching goals of the study.

EP7 - Interdependence

It is interdependence because the work necessitates multiple imaging modalities, namely mammography, ultrasound, and histopathology, all combined together. Each imaging modality offers unique insights into diagnostics, and effectively integrating those requires complex strategies for feature fusion. Moreover, federated learning between institutions makes it possible not to share the data themselves but to keep privacy during integration of data coming from diverse sources. The interdependence between data sources, modalities, and institutions is handled by robust optimization techniques, which are hallmarks of this research.

Mapping with Knowledge Profile for EP1

This table 5.3) is designed to map the EP1 to the Knowledge Profile.

Table 5.3: Mapping with knowledge Profile.

K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K8 Research Literature
√	√	√	√	√

K3 - Engineering Fundamentals

This thesis provides an application of engineering fundamentals arising from AI and machine learning. Concepts revolving around neural networks, federated learning, and the processing of multimodal data have deep bases in the engineering process of the work.

K4 – Specialist Knowledge

The thesis requires specialist knowledge in federated learning, privacy-preserving AI, and multimodal imaging to integrate data from different modalities, such as mammography, ultrasound, and histopathology, into one framework while considering challenges such as privacy and scalability.

K5 – Engineering Design

It basically involves a careful design of the hybrid encoder model and its integration into FL, selecting appropriate architectures, such as CNN and attention mechanisms, and aligning the framework to design principles toward scalability, robustness, and high diagnostic accuracy.

K6 – Engineering Practice

This involves the actual usage of engineering tools such as TensorFlow and PyTorch and even advanced hardware like the NVIDIA Tesla T4 GPU. Besides, the use of healthcare-specific privacy standards just shows how it applies to actual engineering practices.

K8 - Research Literature

The results of this thesis are based on an extensive review of related studies, building upon the state-of-the-art methods and contributing to the academic discourse in privacy-preserving AI and federated learning. This fulfills the requirement of engaging with and expanding on research literature.

5.4.2 Engineering Activities

Table 5.4: Mapping with complex engineering activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
√	√	√	√	√

EA1: Range of Resources

The project put to work a wide array of computational resources for its aims. The training of large-scale FL models was thus enabled by Google Colab with GPU, without the need for expensive infrastructure. Performing the preprocessing steps on local systems complemented these cloud resources in terms of efficiency and cost. This kind of diversity in resource use demonstrates a comprehensive and systematic utilization of computational tools that is so crucial for the handling of multimodal imaging data.

EA2: Level of Interaction

The federated learning framework has required heavy interaction across the simulated nodes, much like real-world collaborations between different institutions. In this regard, it has employed efficient communication protocols and aggregation strategies that allow for interaction while maintaining data privacy with model integrity. The level of interaction in the framework reflects that of complex, real-world healthcare data-sharing scenarios. Therefore, it is very relevant and effective.

EA3: Innovation

The integration of a hybrid encoder with FL represents a novel approach to tackling challenges in breast cancer diagnosis. The innovation lies in the combination of attention mechanisms, advanced feature fusion, and optimization techniques within a privacy-preserving FL framework. This approach addresses critical issues such as class imbalance, data heterogeneity, and inter-modality fusion, setting a benchmark for future research.

EA4: Consequences for Society and Environment

The implications of this project go well beyond the technological. By allowing early and proper diagnosis of breast cancer, it saves lives and reduces burdens both emotional and financial for patients and their families. The focus on privacy and adherence to regulatory standards such as GDPR and HIPAA will increase societal trust in AI-driven healthcare systems.

EA5: Familiarity

This is also a project about established machine learning on advanced and emerging technologies such as FL and multimodal imaging. A delicate balance exists within this project to make sure it is both accessible to practitioners and also contributes toward the state-of-the-art developments in medical AI. Reliability due to known techniques, such as convolutional neural networks and attention mechanisms, leaves room for exploration and adaptation.

5.5 Summary

Chapter 5 deals with the necessary engineering considerations, design issues, and standards followed in the development of this research project. This chapter first presents the compliance with software, hardware, and communication standards. It emphasizes that the use of industry-recognized tools and platforms is necessary. Software standards are used to avail the benefits of TensorFlow, PyTorch, and other frameworks for compatibility and efficiency in the implementation of the federated learning framework. The hardware standards are achieved with the use of robust computational resources such as NVIDIA Tesla T4 GPUs and an Intel Core i5 PC to meet the high computational demands for the analysis of multimodal imaging. Also, the use of standards in communication enables the integration of the federated learning architecture across distributed nodes with guaranteed data privacy.

It also emphasizes the broader impact of such research on society and the environment. The project contributes to early diagnosis of breast cancer, directly improving the patients' lives by improving diagnostic accuracy and reducing diagnostic errors. It also helps to answer urgent concerns regarding privacy and aligns with the data protection regulations toward a more ethical and sustainable healthcare ecosystem. Furthermore, this work underlines computational efficiency and scalability, supporting its wide adaption in diverse clinical and resource-limited settings.

A detailed financial and project management analysis that underlines the cost-effectiveness of the research, maintaining high-quality outputs, would include investments in training, hardware upgrades, and tool subscriptions to reflect the financial viability of the project. Besides, it shows how the project aligns with complex engineering problem-solving on multi-faceted challenges such as data heterogeneity, class imbalance, and inter-modality fusion. The mapping of Engineering Problem (EP) attributes and Knowledge Profile (K1–K8) further substantiates the engineering rigor and intellectual depth of the research.

Finally, the chapter outlines the project's alignment with core engineering activities, EA1-5, including problem analysis, design, investigation, tool usage, and effective communication. These activities demonstrate the systematic approach adopted in developing and presenting a privacy-preserving federated learning framework for multimodal breast cancer classification. This comprehensive coverage highlights the research contribution to advancing engineering practices in medical AI while addressing critical societal needs.

Chapter 6

Conclusion

6.1 Summary

It presented a new, privacy-preserving framework for breast cancer classification, amalgamating federated learning with multimodal imaging data in order to overcome critical challenges in health diagnostics, including data privacy, inter-modality fusion, and scalability. This work presented the integration of mammography, ultrasound, and histopathology imaging modalities that can effectively combine their complementary strengths, thereby providing improved diagnostic accuracy with minimal single-modality limitations. A hybrid encoder model was developed, featuring advanced feature extraction and attention mechanisms to enhance the integration of diverse imaging data. The FL framework allowed multiple institutions to train models collaboratively without necessarily sharing sensitive patient data, which ensured that the research design was compliant with all privacy regulations, including GDPR and HIPAA. Energy-efficient FL protocols and synthetic oversampling techniques were also developed to address class imbalance, heterogeneity in data, and computational complexity. Experimental results have proven the high diagnostic accuracy of the model with fewer false positives and false negatives, which asserts its potential clinical application. Furthermore, this research paper closes the gap between advanced AI methodologies and real-world healthcare needs while providing a robust, scalable, and secure solution for the diagnosis of breast cancer. The results provide major contributions toward the development of AI-powered diagnostic tools and create a solid base for further studies in both multimodal medical imaging and privacy-preserving AI frameworks.

6.2 Limitation

The study faced several limitations that highlight areas for future improvement. Although the dataset used was comprehensive and multimodal, it was sourced from publicly available repositories, which may not fully reflect the complexity and variability of real-world clinical data. Additionally, the research lacked direct collaboration with hospitals or medical institutions, which could have provided valuable clinical insights and access to more diverse datasets. Furthermore, the datasets may not have captured sufficient demographic and imaging variations, potentially affecting the generalizability of the model to diverse populations and clinical settings.

6.3 Future Work

To address the limitations of this study, future work will focus on enhancing the clinical applicability and robustness of the proposed framework. Establishing direct collaborations with hospitals and medical institutions will be a priority, providing access to real-world clinical datasets that capture the complexity and variability of actual patient cases. These collaborations can also offer critical clinical insights, enabling the model to align more closely with practical diagnostic workflows. Additionally, efforts will be directed toward expanding the dataset to include a broader range of demographic and imaging variations, improving the generalizability and reliability of the model across diverse populations and clinical environments. Integrating real-world data and fostering collaboration with healthcare professionals will bridge the gap between research and clinical practice, paving the way for the framework's adoption in real-world diagnostic settings.

References

[1] S. Sharma, M. Gupta, and R. Kumar. Federated learning-based breast cancer detection using densenet and enhanced recurrent neural networks. *Journal of Medical Imaging and Analysis*, 35:12–23, 2023.

[2] Z. Zhang, L. Wang, and H. Li. Domain-adversarial federated learning for breast cancer classification. *IEEE Transactions on Medical Imaging*, 42(7):1125–1137, 2023.

[3] Y. Liu, X. Chen, and J. Zhao. Secure image encryption and federated learning for privacy-preserving breast cancer diagnosis. *Journal of Healthcare Informatics Research*, 12:345–356, 2023.

[4] R. Khan, A. Ahmed, and S. Patel. Hybrid federated transfer learning for breast cancer classification. *IEEE Access*, 11:4506–4520, 2024.

[5] G. Gupta, R. Sharma, and P. Singh. Energy-efficient federated learning framework for breast cancer detection. *Computational and Structural Biotechnology Journal*, 21:657–670, 2023.

[6] K. Kumar, L. Das, and A. Nair. Overcoming data imbalance in federated learning for breast cancer classification. *IEEE Transactions on Computational Biology and Bioinformatics*, 20(4):809–820, 2023.

[7] L. Li, Y. Wu, and J. Tan. Domain adaptation in federated learning for multimodal breast cancer diagnosis. *International Journal of Medical Informatics*, 162:104–116, 2023.

[8] W. Wang, Z. Luo, and Y. Chen. Sustainability-oriented federated learning for medical diagnosis. *Sustainable Computing: Informatics and Systems*, 37:220–234, 2023.

[9] S. Patel, R. Verma, and H. Singh. Hybrid optimization in federated learning for breast cancer detection. *Neural Networks*, 154:56–67, 2023.

[10] J. Chen, X. Zhao, and M. Lin. Real-time federated learning for clinical breast cancer diagnosis. *Journal of Artificial Intelligence Research*, 72:143–155, 2023.

[11] T. Nguyen, S. Lee, and K. Park. Hybrid models combining federated learning and e-rnn for breast cancer classification. *Artificial Intelligence in Medicine*, 124:24–36, 2023.

[12] Z. Zhao, X. Wu, and Y. Feng. Richer multimodal data fusion for breast cancer classification. *IEEE Transactions on Biomedical Engineering*, 70(3):561–572, 2023.

[13] A. Smith, M. Johnson, and R. Lee. A review of convolutional neural networks in breast cancer detection. *Frontiers in Artificial Intelligence*, 5:123–135, 2023.

[14] J. Jones, B. Davis, and L. Taylor. Multimodal deep learning fusion strategies for breast cancer diagnosis and prognosis. *Pattern Recognition Letters*, 165:312–324, 2023.

[15] H. Lee, C. Park, and S. Choi. Human-centric ai assistant for multimodal breast imaging classification. *Healthcare Technology Letters*, 10:450–462, 2023.

[16] A. Ahmed, M. Khan, and S. Patel. Stacked ensemble models for breast cancer prognosis using multimodal data. *Scientific Reports*, 13:3456–3467, 2023.

[17] R. Davis, H. Ryan, and L. Wellman. Collateral representative subspace projection modeling for histology image classification. *Journal of Pathology Informatics*, 14:23–34, 2023.

[18] A. Schafer, S. Patel, and F. Taylor. Customized federated learning for multi-source medical image classification. *IEEE Journal of Biomedical and Health Informatics*, 27(6):1234–1245, 2023.

[19] J. Thomas, R. Singh, and V. David. Pre-trained models in federated learning for mri and ct scan analysis. *International Journal of Imaging Systems and Technology*, 33(4):345–358, 2023.

[20] K. Ting, S. Patel, and A. Singh. Fediic: Federated learning with contrastive learning for class imbalance. *IEEE Transactions on Medical Imaging*, 42(6):756–767, 2023.

Proj rep

ORIGINALITY REPORT

12%	7%	8%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	2%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
3	Submitted to United International University Student Paper	1%
4	Submitted to University of Essex Student Paper	1%
5	www.mdpi.com Internet Source	<1%
6	Amir Mohammad Sharafaddini, Kiana Kouhpah Esfahani, Najme Mansouri. "Deep learning approaches to detect breast cancer: a comprehensive review", Multimedia Tools and Applications, 2024 Publication	<1%
7	Md. Rahad Islam Bhuiyan, Sami Azam, Sidratul Montaha, Risul Islam Jim et al. "Deep learning-based analysis of COVID-19 X-ray images: Incorporating clinical significance and	<1%