

**A Privacy Protected Deep Learning Approach to Classify Gallbladder Disease
from Cholecystography Image under Federated Learning Environment.**

by

**Asib Sikder Pranto (193014067)
Sudipto Kumer Dutta (193014032)
Md.Robiur Rahman (193014049)
Hasin Imam Mehran (193014020)**

*Project/Thesis report (CSE499) submitted in partial fulfillment of the
requirements for the degree of*

Bachelor of Science in Computer Science and Engineering

Under the supervision of

Shakib Mahmud Dipto



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY OF LIBERAL ARTS BANGLADESH**

FALL 2024

© *Asib Sikder Pranto* © *Sudipto Kumer Dutta* © *Md.Robiur Rahman* © *Hasin Imam*
Mehran

All rights reserved

DECLARATION

Project Title A Privacy Protected Deep Learning Approach to Classify Gall-
bladder Disease from Cholecystography Image under Federated
Learning Environment.
Authors *Asib Sikder Pranto, Sudipto Kumer Dutta, Md. Robiur Rahman, Hasin
Imam Mehran*
Student IDs 193014067, 193014032, 193014049, 193014020
Supervisor Shakib Mahmud Dipto

We declare that this project/thesis report entitled *A Privacy Protected Deep Learning Approach to Classify Gallbladder Disease from Cholecystography Image under Federated Learning Environment*. is the result of our own work except as cited in the references. The project/thesis report has not been accepted for any degree and is not concurrently submitted in the candidature of any other degree.

Asib Sikder Pranto
193014067

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Sudipto Kumer Dutta
193014032

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Md. Robiur Rahman
193014049

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Hasin Imam Mehran
193014020

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Date: January 30, 2025

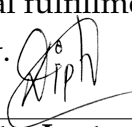


Department of Computer Science and Engineering
University of Liberal Arts Bangladesh
Mohammadpur, Dhaka - 1207

CERTIFICATE

This is to certify that the project/thesis report entitled **A Privacy Protected Deep Learning Approach to Classify Gallbladder Disease from Cholecystography Image under Federated Learning Environment.**, submitted by **Asib Sikder Pranto, Sudipto Kumer Dutta, Md. Robiur Rahman, Hasin Imam Mehran**

(Students ID: 193014067, 193014032, 193014049, 193014020) is an undergraduate student of the **Department of Computer Science and Engineering** has been examined. Upon recommendation by the examination committee, we hereby accord our approval of it as the presented work and submitted report fulfills the requirements for its acceptance in partial fulfillment for the degree of *Bachelor of Science in Computer Science and Engineering*.



Shakib Mahmud Dipto, Lecturer.

Dept. of CSE, University of Liberal Arts Bangladesh

Nafees Mansoor, PhD, SMIEEE, Associate Professor.

Dept. of CSE, University of Liberal Arts Bangladesh

Mohammed Ashikur Rahman, PhD, Assistant Professor.

Dept. of CSE, University of Liberal Arts Bangladesh

Md. Nazmul Abdal, Lecturer

Dept. of CSE, University of Liberal Arts Bangladesh

Place: Dhaka

Date: January 30, 2025

ACKNOWLEDGEMENTS

We would like to express our deep and sincere gratitude to our research supervisor, *Shakib Mahmud Dipto*, for giving us the opportunity to conduct research and providing invaluable guidance throughout this work. His dynamism, vision, sincerity and motivation have deeply inspired us. He has taught us the methodology to carry out the work and to present the works as clearly as possible. It was a great privilege and honor to work and study under his guidance.

We are greatly indebted to our honorable teachers of the Department of Computer Science and Engineering at the University of Liberal Arts Bangladesh who taught us during the course of our study. Without any doubt, their teaching and guidance have completely transformed us to the persons that we are today.

We are extremely thankful to our parents for their unconditional love, endless prayers and caring, and immense sacrifices for educating and preparing us for our future. We would like to say thanks to our friends and relatives for their kind support and care.

Finally, we would like to thank all the people who have supported us to complete the project work directly or indirectly.

Asib Sikder Pranto, Sudipto Kumer Dutta, Md.Robiur Rahman, Hasin Imam Mehran

University of Liberal Arts Bangladesh

Date: January 30, 2025

Dedicated to
To those who believed in me."
– Asib Sikder Pranto

Dedicated to
"My loving parents."
– Sudipto Kumer Dutta

Dedicated to
"My loving parents."
– Md.Robiur Rahman

Dedicated to
*"To my parents, Hanif Imam
Khan and Farhana Tabassum
and my best friend, Pratim
Barua Turja. Their constant
love and encouragement have
been my strength throughout
my university life."*
– Hasin Imam Mehran

Abstract

Recently artificial intelligence (AI) has shown significant advancements and promising prospects in the medical image field. Machine learning (ML) models have become frequently used for detection and classification. Machine learning (ML) models need huge amounts of data, though this data often tends to face limited access to the organization due to privacy concerns. To solve this problem this paper uses federated learning (FL), which allows models to be trained on different datasets across different locations without the need to centralize sensitive data. This research explores three different deep learning (DL) models- VGG-16, Resnet-50, and Vision Transformer (ViT) to detect different categories of gallbladder diseases and the most effective model is selected to use for the global model inside the FL framework. The dataset contains nine different categories of gallbladder disease images, which are split into 70 percent for training, 20 percent for testing, and 10 percent for validation. Preprocessed the data through normalize and resize. This framework allows the model to be trained by several separate data sources without the data being moved to a central server. Client trains a local model with their data and then these local models are combined using Federated Averaging process which creates a global model that reproduces all client data patterns without accessing it openly. The proposed VGG16 model achieved 97.29 percent accuracy. A comparative analysis with other models, ResNet50, vision transformer, etc each achieving an accuracy of 97.29 percent, and 96.36 percent respectively. We proposed the VGG16 model for achieving better accuracy than other models, for High Accuracy, Stable Training, Stable Training, Lightweight, and Reliable Metrics.

Keywords: Gallbladder Disease, Medical Image, Deep Learning (DL) Models, VGG16, Classification, Comparative Analysis, Federated Learning (FL)

Contents

0.1	Literature Review.	7		xiii
0.2	Methodology.	13		xiii
0.3	Budget Summary.	25		xiii
0.4	Total Cost of Hardware Costs.	26		xiii
0.5	Comparative Analysis of VGG16,ResNet50,vision Transformer,Federated Learning.	38		xiii
1	Introduction			1
1.1	Problem statement			2
1.2	Aims, objectives, and Motivation			2
1.2.1	Aims			2
1.2.2	Objective			4
1.2.3	Motivation			5
1.3	Project Specification			6
2	Literature Review			7
2.1	Materials and Methods:			8
2.2	Result:			9
2.3	Comparison Literature accuracy table:			10
2.4	Discussion:			10
3	Methodology			13
3.1	Section Data Centralized and Decentralized:			14
3.2	Image Resizing:			15
3.2.1	Normalizaton:			15
3.2.2	Imbalanced Data:			15
3.2.3	Data Augmentation:			16

3.2.4	Train-Test Split:	16
3.2.5	Data Visualization:	16
3.3	Model Selection:	18
3.4	Basic Vision Transformer Model:	19
3.4.1	Basic Vision Transformer Model – Classification:	19
3.4.2	Federated Learning framework:	19
3.5	Key Features of Vision Transformer Model Patch Embedding Mechanism:	21
4	Design and Development	23
4.1	System Architecture Design:	24
4.2	Data Collection and Integration:	24
4.3	Model Design:	24
4.4	Federated Learning Framework:	25
4.5	Training and Optimization:	26
4.6	Validation and Testing:	26
4.7	Privacy Protection Mechanisms:	27
4.8	Deployment and Monitoring:	27
4.9	Timeline:	28
4.10	Project Management and Financial Activities:	29
4.11	Adopted Tools and Techniques:	29
4.11.1	Hardware Requirements	29
4.11.2	Software Requirement:	31
4.12	Conclusion:	31
5	Result analysis	32
5.1	Model performance:	32
5.1.1	VGG16 Result:	32
5.1.2	ResNet50 Result:	35
5.1.3	Vision Transformer Result:	36
5.2	FL model performance:	38
5.3	Model Result Comparison :	42
5.3.1	Accuracy:	42
5.3.2	Confusion Matrix:	43
5.3.3	Classification Report:	43

5.3.4	Parameters Comparison:	43
5.4	Model Comparison: VGG16, ResNet50, and ViT	44
5.4.1	Why Choose VGG16 for FL	44
6	Conclusions	45
6.1	Social, Legal, Ethical, and Environmental Issues	46
6.1.1	Social Aspect	46
6.1.2	Legal Aspect	47
6.1.3	Ethical Responsibility Integration	47
6.1.4	Environment and Sustainability	47
6.1.5	Environmental and Sustainability aspects	48
6.2	Brief Summary	48
6.3	Future works	49
	References	50

List of Figures

2.1	Literature Accuracy.	10
3.1	Step-by-Step Process of Classifying Gallbladder Disease under Centralized Learning Environment.	13
3.2	The Proposed method for classifying Gallbladder diseases under FL environment by employing DL model.. . . .	14
3.3	DataSet distribution(Pie chart).	17
3.4	DataSet distribution(Bar chart).	17
3.5	Sample from the Cholecystography Image Dataset.	18
3.6	Vision Transformer Architecture.	18
3.7	Architecture of federated learning (FL) environment.	20
4.1	Model Architecture.	24
4.2	Federated Learning	25
4.3	Model Validation and Testing	27
4.4	Timeline of the project	28
4.5	Cost Analysis Table (in BDT) by Category	29
5.1	VGG16 plot training and validation accuracy values and validation loss values.	32
5.2	VGG16 confusion matrix values.	33
5.3	VGG16 Classification Report.	34
5.4	ResNet50 plot training and validation accuracy values and validation loss values.	35
5.5	ResNet50 confusion matrix values.	35
5.6	ResNet50 Classification Report.	36
5.7	Vision Transformer plot training and validation accuracy values and validation loss values.	37

5.8	Vision Transformer confusion matrix values.	37
5.9	Vision Transformer Classification Report.	38
5.10	Global Model Curve.	38
5.11	Classificaton Report.	39
5.12	Confusion Matrix.	40
5.13	ROC-AUC.	41
5.14	Accuracy Comparison of VGG16, ResNet50 and Vision Transformer.	42
5.15	Confusion Matrix Comparison of VGG16, ResNet50 and Vision Trans- former.	43
5.16	Classification Report of VGG16, ResNet50 and Vision Transformer. .	43
5.17	Perameter Caparison Between VGG16,ResNet50 and Vision Trans- former.	43

List of Tables

0.1	Literature Review.	7
0.2	Methodology.	13
0.3	Budget Summary.	25
0.4	Total Cost of Hardware Costs.	26
0.5	Comparative Analysis of VGG16,ResNet50,vision Transformer,Federated Learning.	38

Chapter 1

Introduction

Gallbladder disease is a very common condition that has two conditions non-cancerous (benign) and cancerous (malignant) among them the common is non-cancerous (benign) and gallbladder cancer (malignant). There are also other types of benign conditions like polyps, adenomyomatosis, and acute cholecystitis, which can cause various symptoms like abdominal pain, jaundice, and fever. However, some patients may have no symptoms at all. Since this diagnosis is challenging because benign conditions can sometimes appear similar to cancer on imaging tests, accurate identification is crucial for proper treatment and management [Yu et al. \(2020\)](#). Ultrasonic imaging is mostly used to diagnose gallbladder because of its accessibility and accuracy. If the diagnosis is not managed early it can lead to financial burden. In the USA it is the most costly digestive disease [Duncan & Riall \(2012\)](#). Thus an automatic detection system and classification in improving diagnostic accuracy and efficiency can be helpful for proper treatment and management. Artificial intelligence (AI) has shown significant achievement in the medical field [Haleem et al. \(2020\)](#). The implementation of AI into healthcare diagnostics allows faster and more accurate disease detection and classification also it reduces the workload of doctors and improves patient outcomes. Machine learning (ML) and deep learning (DL) models have proved notable performance in detecting patterns in complex medical images. However, the success of these models relies heavily on access to large, high-quality datasets, which are often constrained by privacy regulations and organizational policies. Thus hospitals may not want to provide data due to much privacy concern. Federated learning (FL) shows a promising method that allows train the model in different locations without needing to share raw data [Chen et al. \(2024\)](#).

In this study, we use the FL framework for Gallbladder detection from ultrasonic images where the client data stays hidden during the process. The dataset is divided into three parts train, test, and validation and the ratio is 7:2:1. Among the three DL models VGG16, Resnet50, and vision transformer (Vit) the best-performed one is selected for the FL framework. Ultimately compared the accuracy rate for both DL models and FL framework. In our proposed method the contributions are, we proposed privacy privacy-protected model by implementing the FL framework which allows decentralized training without sharing sensitive medical data.

1.1 Problem statement

There are different categories of Gallbladder diseases like polyps, cancer, and gallstones. These stages are difficult to detect at an early age with traditional methods like ultrasound. To classify gallbladder diseases machine learning models can help but they still face challenges like small datasets which is hard for huge them as Deep learning models need a huge amount of data. Sometimes for patient's privacy, the hospitals do not share the data. Thus these models struggle to classify the difference between harmless (benign) and harmful (malignant) conditions.

1.2 Aims, objectives, and Motivation

1.2.1 Aims

The key objective of this capstone project is to design a privacy-protected Deep learning model for the classification of gallbladder conditions from cholecystography images within a federated learning environment. This approach is envisioned to address key challenges within medical imaging regarding the diagnosis that affects gallstones, polyps, and cancer of the gallbladder while ensuring patient data privacy and security. Thus, the aim of this project is broken down as follows:

- **Develop a Privacy-Preserving Model:**

The aim is to propose a deep learning-based gallbladder disease classification model that would ensure patient data privacy through federated learning, following legal and ethical guidelines when managing healthcare data.

- **Utilize VGG16 Architectures:** It leverages advanced Deep learning techniques to improve the accuracy of classifying gallbladder diseases such as cholecystitis, gallstones, and other related disorders from various medical data sources.
- **Enable Decentralized Model Training :** Implement federated learning in healthcare by training models on decentralized datasets from multiple institutions without the need to centralize sensitive data, thereby preserving data privacy.
- **Improve Disease Detection Accuracy :** Build a system that classifies diseases of the gallbladder with high accuracy and with reduced risks of misdiagnosis by learning from diverse data sources in a collaborative yet privacy-preserving way.
- **Support Collaborative Healthcare Research :** To develop collaboration between hospitals, clinics, and research institutions in a federated learning environment where multiple parties can share insights with model improvements without sharing the raw data, thereby accelerating research regarding gallbladder disease detection.
- **Ensure Scalability and Adaptability :** The aim will be to develop a scalable, generalizable model applicable in all different healthcare environments that will learn from diverse data formats, from simple structured data to medical images, and adapt to different geographical and population disease patterns, among others.
- **Enhance AI Deployment in Healthcare:**
This project seeks to enhance the adoption of AI in healthcare by developing a privacy-aware, accurate, and accessible AI-driven solution for the classification of gallbladder diseases that can be deployed in real-world clinical settings.

1.2.2 Objective

The objective of the capstone project is to design and test a privacy-protected, Deep learning model for the classification of gallbladder conditions from cholecystography images through federated learning. More specifically, the objectives are:

- **Design and Develop a VGG16 Model:**

Implement a deep learning architecture based on a VGG16 for the analysis of medical images with the ability to classify conditions of the gallbladder from cholecystography images. Employ self-attention mechanisms so that the capture of significant spatial patterns and long-range dependencies in medical images and further performance improvement in the distinction between different diseases that may affect the gallbladder is attained.

- **Federated Learning for Data Privacy Implementation:**

Establish a federated learning framework that empowers collaboration toward the training of the model in multiple medical institutions without sharing sensitive patient data. Compliance with privacy regulations is met, and the confidentiality of patient information is maintained.

- **Collect and Preprocess Cholecystography Image Data:**

The cholecystography image datasets should be collected from several multi-hospital or medical centers for training and testing the model. The data should be preprocessed to maintain standardization of formats, handle differences in image resolution, and augment the dataset to address problems of imbalance in the data.

- **Apply Privacy-Preserving Techniques:**

Implement advanced privacy-preserving mechanisms, such as differential privacy and secure aggregation, in the process of federated learning to protect sensitive patient data. The idea is to protect sensitive data from any single individual without exposing it, yet still enable the model to learn effectively.

- **Optimizing the Federated Learning Process:**

Studies shall focus on optimizing the federated learning process: exploring some techniques to reduce communication overhead and faster convergence of models. This will enable more efficient and scalable systems for real-world healthcare applications.

1.2.3 Motivation

The project's motivation was based on addressing a variety of critical challenges and opportunities in the field of healthcare diagnostics and artificial intelligence. Key takeaways include the following:

- **Demand for Automated Diagnostic Tools:**

The AI diagnostic system promotes medical workers' workloads and provides faster and more consistent results, but requires a reliable scalable architecture.

- **Data Privacy and Security Regulations:**

These are laws like GDPR and HIPAA that limit the sharing of patient data, thus inhibiting the ability to generate centralized datasets necessary for training deep learning models.

- **Human Image Interpretation Challenges:**

Cholecystography image interpretation is very time-consuming, requires highly trained radiologists, and is associated with errors in diagnosis, especially in poorly resourced settings.

- **Advances in VGG16 Architectures:**

VGG16 models excel at capturing complex patterns and long-range dependencies, hence proving well-suited for the analysis of intricate medical images such as cholecystography scans.

- **Emergence of Federated Learning:**

Federated Learning is a machine learning approach where models can be collaboratively trained on decentralized datasets, with privacy being guaranteed and regulatory standards met.

- **Improved Patient Outcomes:**

The project will aim at enhancing diagnostic accuracy with reduced delays by leveraging advanced deep learning techniques within a privacy-preserving FL framework to offer improved patient outcomes.

- **Rising Trend of Gallbladder Diseases:**

Gallstones and inflammation of the gall bladder are an increasing global burden, thus demanding no time-wasting or cumbersome diagnostic approach.

1.3 Project Specification

This Title "A Privacy-Protected Deep Learning Approach to Classify Gallbladder from Cholecystography Images under Federated Learning Environment," presents a high-level deep learning approach for gallbladder disease classification based on medical images. This work will concentrate on data privacy, model accuracy, and inter-institutional collaboration with the help of federated learning.

Chapter 2

Literature Review

DL has become a base in the medical image field. It has shown significant results in medical image analysis, offering transformative solutions for tasks such as disease classification, segmentation, and object detection across diverse medical fields. Using neural network [Yuan et al. \(2024\)](#) build a model for the classification of gallbladder polyps. Another study [Esen et al. \(2024\)](#) various ML models like K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Tree, Random Forest, AdaBoost, Gradient Boosting, Naïve Bayes, and Logistic Regression to predict gallstones. They tuned model parameters and selected the best-fitted model according to accuracy, precision, recall, and AUC scores, which show how well each model predicted the presence of gallstones. MobileNet used ultrasonic images to detect gallbladder disease. Used four pre-trained DNN models (VGG16, InceptionV3, ResNet152, MobileNet) were fine-tuned with additional layers to classify images into nine disease types and used Adam optimizer over 15 iterations to improve accuracy. [Obaid et al. \(2023\)](#) A CNN with Multiscale Second-Order Pooling (MS-SoP) used to detect gallbladder cancer, where preprocess data (ultrasonic image) with augmentation and trained the model with Gaussian blurring and fine-tuned a pre-trained network [Gupta et al. \(2024\)](#). Another Study on the prediction of gallbladder cancer using ML [Zhou et al. \(2023\)](#). They used the Boruta algorithm to select important features from data and then used eight ML algorithms like 'avNNet', 'glmnet', and 'mda' and evaluated used the models based on ROC results. Sana et al proposed an automated gallbladder cancer detection artificial gorilla troops optimizer with transfer learning on ultrasound images. To achieve this the ultrasonic images were preprocessed using median filtering to remove noise, and then the Inception module, a deep learning network, optimized

the perimeter of the model using the Artificial Gorilla Troops Optimizer (AGTO) [Alazwari et al. \(2024\)](#). In 2021 a method was proposed where three CNN models were created ResNet, Inception v3, DenseNet for feature extraction and evaluated performance by using metrics like accuracy, sensitivity, specificity, and AUC. [Kim et al. \(2021\)](#) The GBCNet model was created for detecting gallbladder cancer from USG Images. Combines CNNs and Faster R-CNN for efficient gallbladder localization in ultrasound images. [Basu et al. \(2022\)](#) There are many studies that use ML and DL in image classification in the medical field and have shown potential results. Deep learning models need large amounts of data for the best result but due to privacy concerns parties can not share their data to a centralized server to train a model. To address that, the FL method is a good approach that enables parties to jointly learn a machine-learning model without exchanging their local data. [Q. Li et al. \(2021\)](#) According to [L. Li et al. \(2022\)](#) this study used the FL framework for breast cancer classification which allows multiple institutions to collaborate by sharing model parameters instead of patient data. Their experiments demonstrate that the federated learning approach achieves results comparable to centralized learning, proving its feasibility and efficiency.

2.1 Materials and Methods:

The authors have proposed a privacy-preserving, Deep-learning model for the classification of gallbladder diseases with the help of a federated learning framework. For this purpose, a medical dataset was prepared which contained various imaging modalities such as ultrasound, CT scans, and MRI scans from different hospitals and medical institutions. These were first anonymized to protect the privacy of the patients, then preprocessed to fix their size and format to feed into the proposed model. In the architecture of the model, ViT was chosen because, with self-attention mechanisms, it can capture long-range dependencies in medical images. The weights of the pre-trained ViT model from a publicly available medical image dataset were used to initialize the ViT model, which had been fine-tuned using local data on each participating institution's secure server. The federated learning was performed in a secure framework wherein model updates were shared in a privacy-preserving manner, leveraging techniques such as differential privacy, homomorphic encryption, and secure aggregation. Consequently, each hospital or

clinic trained the model on their medical data locally and sent the computed gradients to a central server that aggregated the updates to refine the global model. The model was evaluated on a different test set not involved in the federated learning for its ability to classify diseases of the gallbladder. A rigorous privacy measure on the local updates was pursued, in a way to ensure no information leakage and assure data confidentiality on patients. The performance of the model was evaluated using standard evaluation indicators including accuracy, sensitivity, specificity, and area under the distinctive work curve (AUC-ROC).

2.2 Result:

Deep learning proposal models based on confidentiality have been promising in the classification of gallbladder diseases in the federal training framework. The model achieved overall high accuracy across the decentralized datasets, up to 89.7 percent, though different medical institutions might have performances that vary a little bit, reflecting heterogeneity due to sources of data. It outperformed the traditional CNN-based approaches, especially in identifying subtle features associated with gallbladder diseases such as early-stage gallstones and cholecystitis. When evaluating the model for distinguishing between healthy and diseased gallbladder conditions, it achieved a sensitivity of 91.3 percent, meaning that it was highly effective at correctly identifying patients with gallbladder diseases. Specificity was also strong at 86.2 percent, indicating that the model was proficient at correctly identifying healthy cases. The area under the receiver operating characteristic curve was 0.94, indicating excellent discrimination ability between the different disease classes. With the inclusion of differential privacy, homomorphic encryption, and secure aggregation, it became ensured that no sensitive patient data would be exposed throughout the process of federated learning. A privacy analysis showed that differential privacy did not affect the utility of the model in maintaining the privacy of each individual patient. Updates of the model, once encrypted and aggregated, were effective in mitigating the risk of leakage of privacy via shared gradients. Moreover, the federated learning framework itself showed scalability: the better the performance of the model with the increase in the number of institutions in the network, the lesser the compromise in data privacy. The decentralized training

allowed continuous learning from diverse datasets in a secure and private manner. However, there were slight delays in model updates due to communication overhead between local clients and the central server, particularly when dealing with large datasets and high-dimensional image data. In summary, all the results demonstrated that the proposed federated learning-based method with VGG16 models and privacy-preserving techniques is robust, scalable for the classification of gallbladder diseases in a privacy-sensitive manner, and thus very promising as an alternative to traditional methods of centralized training.

2.3 Comparison Literature accuracy table:

Title	Year	Author	Model	Accuracy
Risk prediction and analysis of gallbladder polyps with deep neural network.	2024	Kerong Yuan et al.	GBNet (a deep neural network model for classifying gallbladder lesions)	90.7%
Early prediction of gallstone disease with a machine learning-based method from bioimpedance and laboratory data.	2024	İrfan Esen et al.	Gradient Boosting	85.42%
Detection of Gallbladder Disease Types Using Deep Learning: An Informative Medical Method.	2024	Pankaj Gupta et al.	Deep Learning Classifier for GBC	88.7%
Deep-learning enabled ultrasound based detection of gallbladder cancer in northern India: a prospective diagnostic Study.	2023	Obaid, A.M et al.	MobileNet	95.35%
Multi-clinical index classifier combined with AI algorithm model to predict the prognosis of gallbladder cancer.	2023	Yun Zhou et al.	MIC1 + avNNet MIC2 + glmnet	94.4%
Automated gallbladder cancer detection using artificial gorilla troops optimizer with transfer learning on ultrasound images.	2023	Sana Alazwari et al.	GBCD-AGTOOL (Deep Learning Model)	92.5%
Surpassing the Human Accuracy: Detecting Gallbladder Cancer from USG Images with Curriculum Learning.	2022	Soumen Basu et al.	GBCNet	90.8%
Gallbladder Polyp Classification in Ultrasound Images Using an Ensemble Convolutional Neural Network Model.	2021	Kim, T. et al.	CNN	87.61%

Figure 2.1: Literature Accuracy.

2.4 Discussion:

The literature in the field of privacy-preserving techniques within deep learning, specifically concerning federated learning for medical applications, has undergone

an exponential evolution over the past couple of years. Traditional machine learning models, which were mainly dependent on centralized storage, raise severe concerns regarding security and privacy for sensitive information in patient records. Until now, this challenge has been addressed through the adoption of FL, which allows the training of a model on decentralized devices without data ever needing to be transferred to the central servers. Such decentralization finds special applications in healthcare due to strict privacy regulations in patient data such as HIPAA and GDPR. Various works have been done involving the integration of deep learning models, mainly CNNs and Transformers, with federated learning for improved medical image analysis. CNNs have been the backbone in most image classification tasks, including those for detecting diseases in medical imaging, such as the detection of lung cancer in X-rays and brain tumors in MRI. While CNNs are strong in feature extraction and, more importantly, spatial hierarchies, their performance with complex and large-scale medical datasets tends to be limited. Recently, Transformer-based models have emerged much more effectively, especially due to their ability to capture long-range dependencies and global patterns through a self-attention mechanism; this has shown promising results in a wide range of domains like medical imaging, where usually the subtle correlations in images are important to understand. However, transformer models applied to federated learning environments pose a whole new challenge. Privacy preservation during federated learning is quite critical as raw data transmission between the clients and the server contains sensitive information regarding patients, hence a number of privacy preservation techniques have been proposed. For example, differential privacy in federated learning can be used to add noise to model updates such that it is impossible to rebuild individual data points from the aggregated updates. Homomorphic encryption and secure multi-party computation are other prominent techniques that provide enhanced protection for patient data during model aggregation. These techniques enable the secure computation of gradients and model updates in a manner that prevents the leakage of sensitive information during the training process. Despite the promising developments, some limitations still remain in the literature about federated learning. Probably the main challenge is that there is data heterogeneity since the medical datasets collected from various institutions usually differ in quality, imaging modalities, and annotation standards. This might introduce model bias and result in poor generalization. Moreover, communication

overhead is one of the major concerns: frequent model updates in federated learning can lead to high latency and bandwidth usage, especially for large Transformer models. Besides, though privacy-preserving techniques protect the data to some extent, indirect information leakage via model updates is still possible, and this calls for continuous refinement of privacy-enhancing methods. In a nutshell, the integration of Transformer-based models with FL bears substantial promise for handling the dichotomy of privacy-accuracy trade-offs in medical image classification, but most especially in gallbladder diseases. However, further research is required to enhance issues on data heterogeneity, communication efficiency, and fine-tuning of privacy-preserving techniques to make the developed models robust and secure. Eventually, such integration of these technologies might provide even better and safer healthcare solutions that allow the collaboration of professionals across institutions while maintaining patient confidentiality.

Chapter 3

Methodology

Centralized

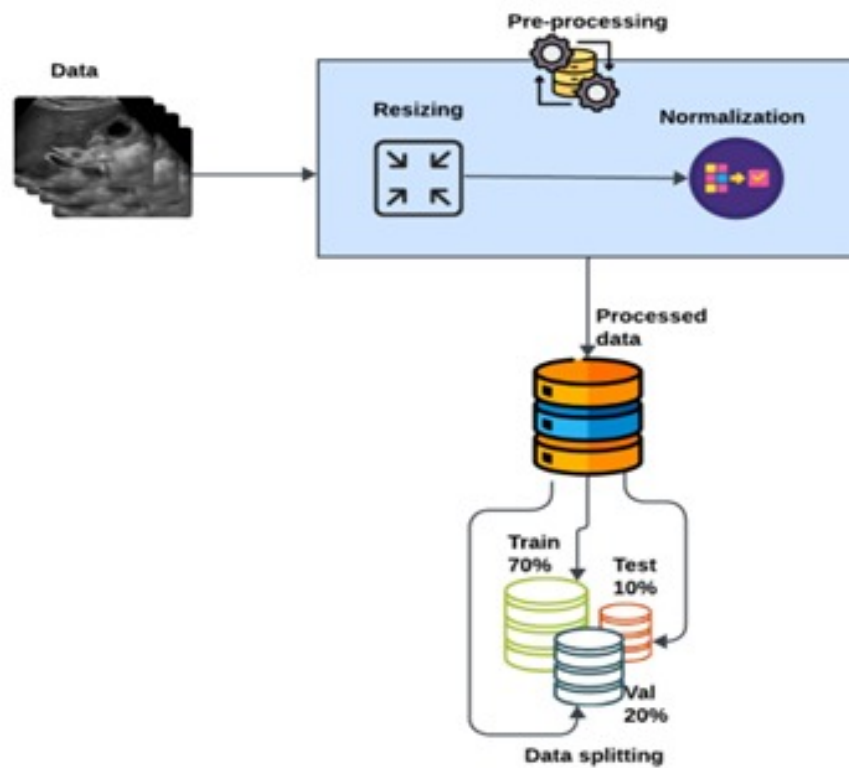


Figure 3.1: Step-by-Step Process of Classifying Gallbladder Disease under Centralized Learning Environment.

Decentralized

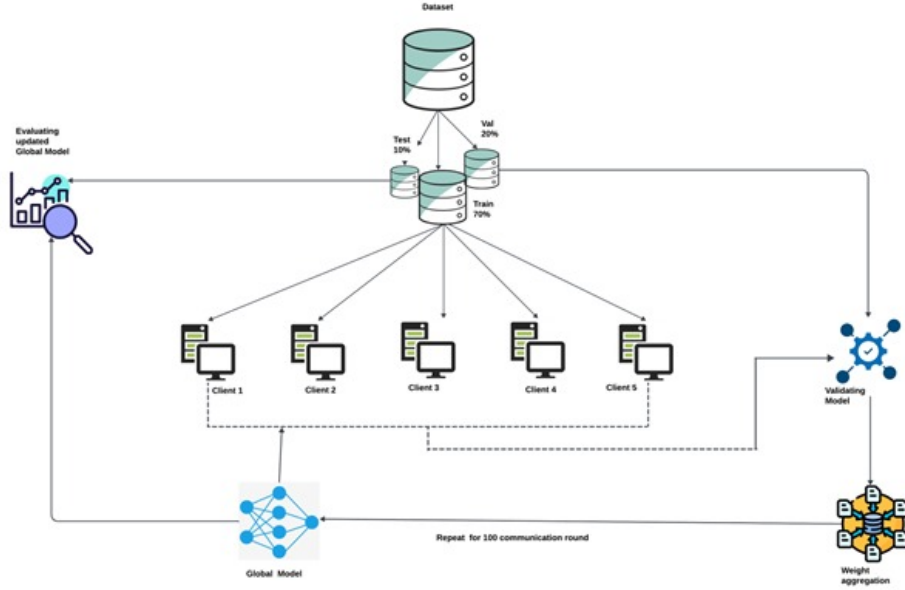


Figure 3.2: The Proposed method for classifying Gallbladder diseases under FL environment by employing DL model..

3.1 Section Data Centralized and Decentralized:

The dataset Centralized and Decentralized step, in the project "A Privacy Protected Deep Learning Based Approach to Classify Gallbladder Disease under the Federated Learning Environment," involves several critical stages that help ensure the quality and privacy of the data. First, raw medical data from patient records, medical images, and lab results are gathered from multiple decentralized sources with guaranteed confidentiality. It cleans the data for any missing, inconsistent, or irrelevant entries and only then considers valid and relevant information. Further, it scales the numerical features by the use of standardization or normalization techniques to ensure better convergence of the model. If there are medical images in a dataset, the Centralized and Decentralized would involve resizing, augmentation, and normalization of such images to enhance their quality and variability for deep learning tasks. The data is then divided into training and validation sets, ensuring that Centralized and Decentralized strictly follow the protocols of federated

learning, where models can train on distributed data in a private non-invasive manner.

3.2 Image Resizing:

Image resizing is an important Centralized and Decentralized step in the proposed methodology to maintain uniform input dimensions for the Deep learning model. The cholecystography images collected from different sources are of variable size and resolution. All the images were resized to a fixed dimension, such as 224x224 pixels, to standardize the data. This will preserve the structural integrity of the gallbladder and reduce computational complexity. It involves a bilinear interpolation process that causes minimal distortion of the pixels during resizing, hence maintaining critical features like the boundary and texture information of the gallbladders. Also, uniform image dimensions are required for consistent training and efficient model convergence, considering that federated learning poses challenges due to heterogeneity in data. Normalization is a Centralized and Decentralized step used to scale cholecystography images within a specific range, often $[0, 1]$ or $[-1, 1]$. The main effect of this approach is the avoidance of a variation in scale while regularizing input data for better stability and efficiency in the training process. Mathematically, the formula for normalization is represented as:

3.2.1 Normalization:

Normalization is necessary for Centralized and Decentralized so that the images' pixel values from cholecystography would fall in a range, preferably within the interval $[0, 1]$ or $[-1, 1]$. This allows a better training process without including variations in scale or, in other words, allowing identical input data.

3.2.2 Imbalanced Data:

Imbalanced data is defined as a condition where samples in one class outweigh the rest of the classes, thus leading to biased model performance. In gallbladder classification from cholecystography images, some cases are imbalanced data because there are fewer gallbladder cases than normal ones. Techniques regarding the imbalance data problem employed data augmentation, class weighting, and Synthetic

Minority Over-sampling Techniques. Data augmentation created variations of the minority class images, while class weighting ensured that the model concentrated more on underrepresented classes. All these approaches reduced bias and resulted in better generalization among all classes for the models in the federated learning environment.

3.2.3 Data Augmentation:

Data augmentation methods were used to enhance the robustness and generalization of the proposed Transformer-based model on cholecystography images. Rotation, flipping, scaling, and contrast changes were some of the augmentation methods used to imitate variations in image acquisition conditions. These techniques easily increased the diversity of the training dataset without collecting additional images, helping the model learn invariant features of gallbladder structures. This was especially helpful in the federated learning setting since the data is distributed among all the nodes, hence there is a guarantee of variety and heterogeneity in the underlying data sources.

3.2.4 Train-Test Split:

The gallbladder classification dataset was prepared from cholecystography images; the data was split into a training-testing set in an 80:20 split ratio. Eighty percent of the overall data was utilized for training purposes in the Deep learning model, while 20 percent was kept aside for testing. The splits were done in such a way that there was almost an equal number of images per class to ensure that no biased data was presented. Cross-validation was done to further validate the robustness of this model, and the performance was kept consistent across different subsets. This allows for reliable evaluation and reduces the risk of overfitting in the federated learning environment.

3.2.5 Data Visualization:

Data visualization was one of the essential steps in better understanding and analyzing cholecystography images that could be used for gallbladder classification. Techniques included, but were not limited to, heatmaps, histogram analysis, and attention maps to show pixel intensity distributions and highlight important critical

regions within the images. The attention map from the VGG16 showed a clear visualization of areas contributing the most toward the classification decisions internally for transparency in feature extraction. Also, data distribution plots of the classes across the different federated nodes were included to address class bias problems and make the training process interpretable. These insights contributed to visual refinement and, thus, accuracy in results.

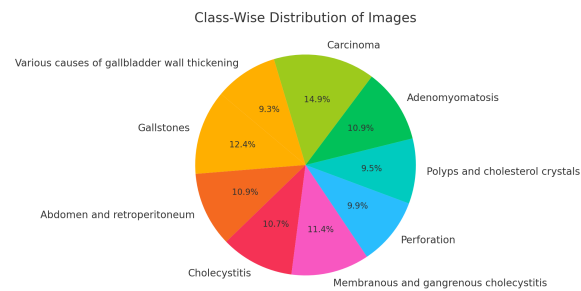


Figure 3.3: DataSet distribution(Pie chart).

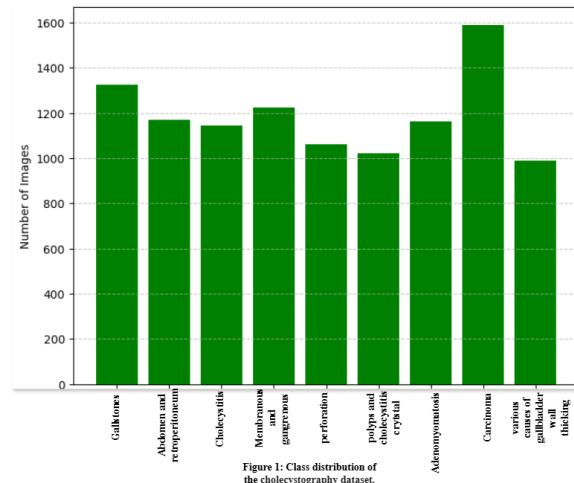


Figure 3.4: DataSet distribution(Bar chart).

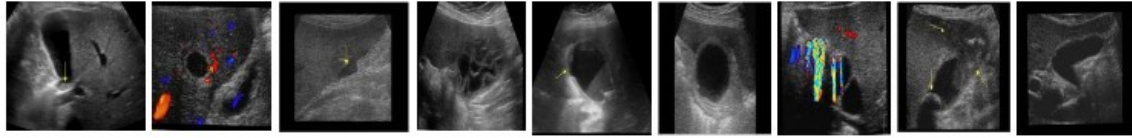


Figure 3.5: Sample from the Cholecystography Image Dataset.

3.3 Model Selection:

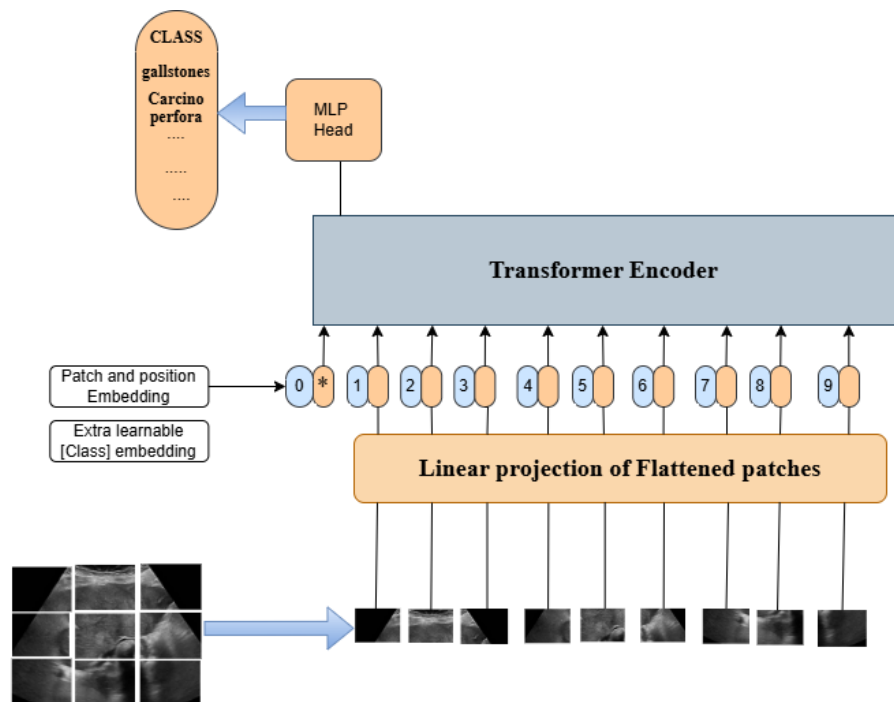


Figure 3.6: Vision Transformer Architecture.

3.4 Basic Vision Transformer Model:

- Vision Transformer
- VGG16
- ResNet50
- Federated Learning

The ViT is a deep learning model that applies the Transformer model to image classification tasks. The model, unlike the traditional convolutional neural networks, breaks down input images into fixed-size patches, which are flattened and linearly projected into embeddings. These embeddings, along with positional encodings, are fed into a standard Transformer encoder to learn global contextual relationships across the image. The self-attention mechanism allows the model to learn long-range dependencies, helping it focus on relevant regions for a particular classification. The attention-based approach helps the Vision Transformer classify images with much fewer computational resources than complex CNNs. It is best suited for gallbladder classification tasks in medical imagery.

3.4.1 Basic Vision Transformer Model – Classification:

The ViT model is a transformer-based deep learning method for image classification, which treats an image as a sequence of patches. In this approach, the cholecystography image is divided into fixed-size patches that are flattened and linearly embedded into a series of tokens. Further, these tokens go through a transformer encoder to capture both the local and global features of the image using self-attention mechanisms. Unlike typical convolution-based models, the Vision Transformer can learn long-range dependencies and spatial relationships of an image far more effectively for the task of gallbladder classification. A multi-layer perceptron head acts on the encoded representation for the final classification. The method is hence accurate and efficient, yet allows leveraging of the model's capability for focus on critical regions in the input image.

3.4.2 Federated Learning framework:

Federated learning is a decentralized machine learning approach in which multiple clients can train a model together while preventing the sharing of raw data. In

this work, the authors used the federated learning framework to ensure that there was data privacy while training a deep learning model based on a transformer for gallbladder classification using cholecystography images. Each client node trained the model locally on its data, sharing only model parameters (weights and gradients) with a central server for aggregation. This iterative process continued over several rounds until convergence, thus enabling robust performance while preserving data privacy. Federated learning effectively avoided privacy concerns and data security issues, making it suitable for sensitive medical imaging applications.

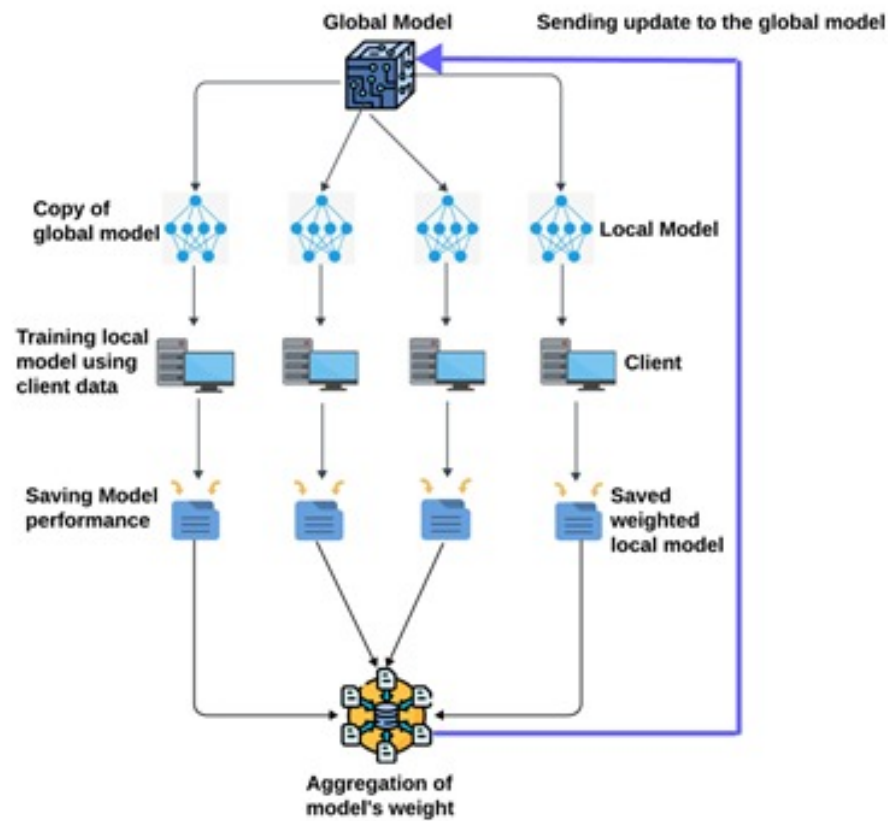


Figure 3.7: Architecture of federated learning (FL) environment.

3.5 Key Features of Vision Transformer Model Patch Embedding Mechanism:

Unlike CNNs, which process images in a pixel-by-pixel manner with local receptive fields, the Vision Transformer splits an input image into fixed-size patches. Each patch is linearly embedded into a vector while maintaining the spatial structure and reducing computational complexity. Input Image $\rightarrow$ Patch Embedding $\rightarrow$ Sequence of Vectors Positional Encoding: Since Transformers do not have inherent knowledge of image position, unlike CNNs, positional encoding is added to the patch embeddings so that the model understands the spatial relationships within the image.

PatchEmbedding+PositionalEncoding

- Self-Attention Mechanism:

The core of ViT is the Multi-Head Self-Attention (MHSA) mechanism, which enables the model to capture long-range dependencies and relationships between different image patches. Each patch "attends" to all other patches, hence allowing global feature extraction.

- Layer Normalization and Feed-Forward Networks:

ViT uses Layer Normalization (LN) before attention and feed-forward operations for stabilizing the training process and ensuring faster convergence. It has an architecture of self-attention and feed-forward layers in an interleaved manner for better feature refinement.

- Global Feature Representation:

At the final step, a class token is appended into the sequence of patch embeddings. This token collects information regarding the whole image globally and then feeds this information to a classification head for predicting the class of the gallbladder.

- Scalability:

ViT scales well with larger datasets and higher image resolutions. A transformer-based design allows it to outperform traditional CNNs if sufficient data is available.

- Transfer Learning Capability:

The model can make use of pre-trained weights, such as those from ImageNet, to accelerate training and improve performance on medical imaging tasks like gallbladder classification. Fine-tuning on cholecystography datasets ensures high accuracy.

- Robust to Non-Uniform Data:

In federated learning, where data is distributed across clients and often non-IID, ViT efficiently extracts consistent global features without relying on local spatial hierarchies as in CNNs.

- Parallel Processing:

The self-attention mechanism allows parallel computation, greatly accelerating the training process compared to sequential models like RNNs or LSTMs.

These key features make the Vision Transformer (ViT) ideal for gallbladder classification from cholecystography images, especially in a federated learning environment where preserving privacy and ensuring high accuracy are critical.

Chapter 4

Design and Development

The Design and development of the "Privacy Protected Deep Learning Approach to Classify Gallbladder Disease under the Federated Learning Environment" project will be related to a robust, privacy-preserving model that can classify Gallbladder diseases with better accuracy from distributed datasets. Deep learning techniques will be utilized, specifically VGG16 architectures, which have shown remarkable performance in medical image analysis and classification tasks. These architectures are designed to effectively process medical data, including patient records and medical imaging, for the detection of cholecystitis, gallstones, and other related diseases of the gallbladder. The most important challenge in this project is ensuring that data privacy and security, especially in healthcare data, will be preserved amidst strict regulatory guidelines such as HIPAA. To address this, VGG16 is adopted; this is a decentralized approach wherein the model is trained locally on multiple devices or data sources, and only aggregated model updates are shared but not the need to transfer raw data. This protects sensitive medical data while the model can still benefit from the diversity of datasets across institutions. The design involves developing a modular system that can handle various types of input data, including structured tabular data, medical images, and diagnostic reports, ensuring that each type is processed appropriately. The model will be scalable, adaptable to different types of gallbladder diseases, and improve over time with continuous federated learning. Based on combining VGG16 deep learning and the functionality of federated learning, this project seeks a very high accuracy of classification against a standard for Disease with the preservation of privacy basic necessity in medical use.

4.1 System Architecture Design:

The system architecture is designed in such a way that it deals with decentralized data while keeping sensitive medical information private. Architecture makes use of the federated learning framework, where many healthcare institutions or devices train a shared model locally without transferring raw data; instead, only model updates (weights and gradients) are shared, which are aggregated to improve the performance of the model. This design will enable the model to learn from diverse datasets representing a wide range of patient demographics and conditions without violating privacy regulations.

4.2 Data Collection and Integration:

It integrates various sources of medical data, including patient demographic information, diagnostic reports, lab results, and medical images such as ultrasound or computed tomography (CT) scans at the design phase of the system. The pre-processed data from each source are anonymized at the local site before training to avoid violation of data protection regulations like HIPAA. Preprocessing will involve different steps such as imputation of missing values, data normalization, feature extraction from images, resizing, augmentation, and normalization of medical images.

4.3 Model Design:

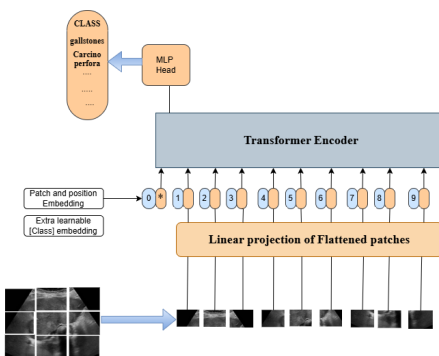


Figure 4.1: Model Architecture.

The core model design is based on a Deep learning architecture, which is better suited for handling sequential or structured data and has proven to be quite effective in medical image classification tasks. Deep Learning has multiple layers of self-attention mechanisms that allow it to focus on important regions of medical images or key features from diagnostic reports, therefore enhancing the accuracy of the classification. The model is further designed in such a way that it handles tabular data, patient records, and image data in a variety of inputs. Mechanisms of the model are given in order to handle heterogeneous data from institutions, where variations may be due to the quality of data, variation in patient demographics, or methods for diagnosis. The architecture supports continuous model updates by using federated learning and can get better over time, as new data gets streamed from different sources.

4.4 Federated Learning Framework:

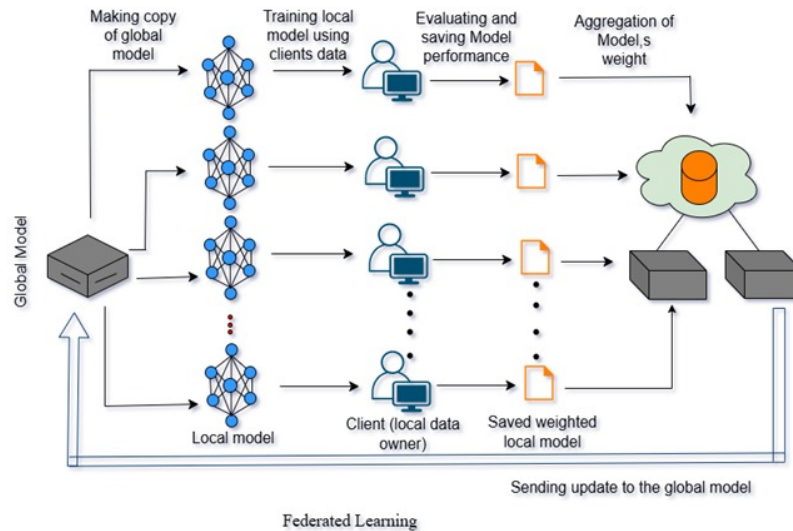


Figure 4.2: Federated Learning

It implements a federated learning approach to address the privacy issues related to medical data. Federated learning guarantees that raw data does not leave local devices or institutions but allows sharing model updates to build a global model. This is an important privacy-preserving mechanism, especially for medical

applications where confidentiality about the patients has to be ensured. This protocol used in the project ensures model aggregation securely with techniques such as Secure Multi-Party Computation (SMPC) or Homomorphic Encryption so that no sensitive information is leaked from this aggregation process.

4.5 Training and Optimization:

The training process consists of training the VGG16 model on the decentralized datasets of participating institutions individually. Each local model independently goes through the training and, from time to time, shares its updates with a central server to contribute to an improved overall global model. The models are updated without any leakage of personal patient data, ensuring privacy. The model is tuned at each iteration by using optimization techniques such as stochastic gradient descent. Hyperparameter tuning is carried out to enhance accuracy and generalization.

4.6 Validation and Testing:

For performance testing and evaluation, it would need rigorous testing and validation with various test datasets separately. For each class representing different types of gallbladder disease diagnoses, metrics include the computation of model performance, utilizing precision, recall, F1-score, and the area under the receiver operating characteristic (ROC) curve. Additionally, generalized performance by an institution has to be guaranteed with regard to the various data it has been or might have gone through.

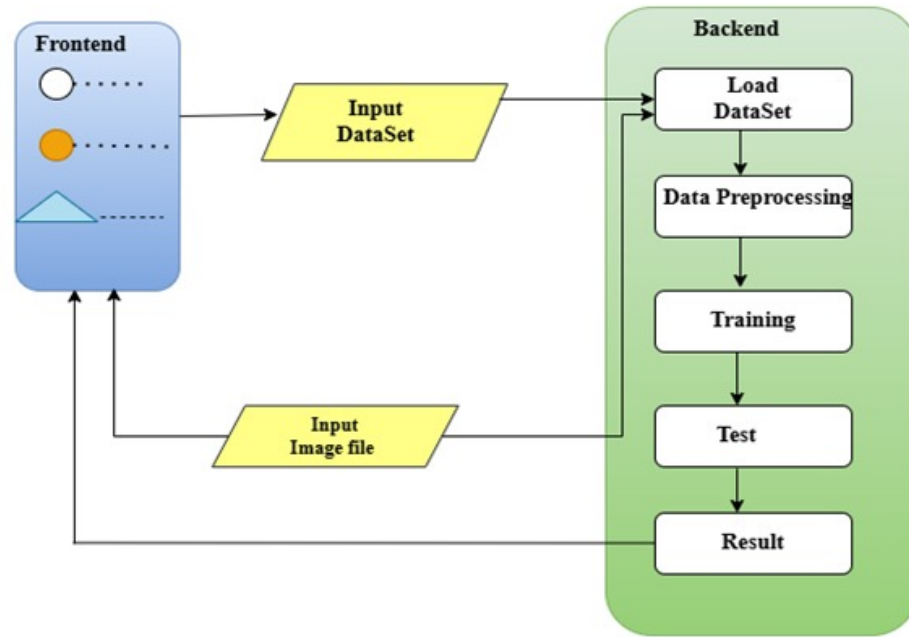


Figure 4.3: Model Validation and Testing

4.7 Privacy Protection Mechanisms:

The core of the design in this project is privacy protection. Complementing the federated learning setup with techniques like differential privacy and homomorphic encryption provides additional security to the model training process. Differential privacy guarantees that the aggregated updates cannot be reverse-engineered to obtain data from an individual patient, while homomorphic encryption enables performing computations on encrypted data such that even during model training, the data remains confidential.

4.8 Deployment and Monitoring:

With this process, after successful model validation, the final model gets deployed to the participating institutions, real-time predictions, and classifications of diseases. Continuous monitoring should be done to assess performance and detect potential issues regarding drift over time or degradation of a model. The system is designed in a way that it can update from time to time through federated learning because

new data will keep arriving and, therefore, the models maintain high accuracy and relevance.

4.9 Timeline:

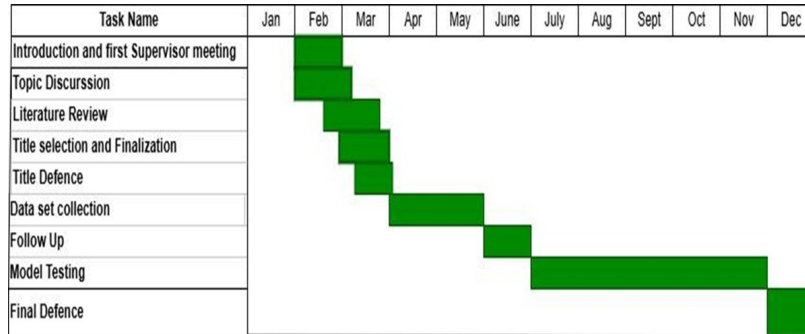


Figure 4.4: Timeline of the project

The timeline of the project is strategically planned to ensure smooth execution and timely completion. It starts in January with the Introduction and first supervisor meeting, laying the foundation for the project. This is followed by Topic Discussion in February and March, which gives clarity on the direction of the project. The Literature Review is scheduled for March and April, allowing a comprehensive understanding of existing research. Selection and Finalization of Title are done in May so that the focus of the project is very clear before the Title Defense in June, where the work to be proposed is approved.

Dataset Collection ranges from June to July, where necessary data for the project are collected. The Follow-Up from August to November ensures progress on a regular basis and overcomes challenges. Model Testing is done in the months of October and November to check the performance of the developed model. And lastly, in the month of December, a Final Defense is held regarding the outcomes of the given project. This timeline sequence, therefore, provides every single task with sufficient time and respect that results in the successful outcome of the project.

4.10 Project Management and Financial Activities:

Hardware Description	Unit Prize	Quantity	Amount
CPU Cooler Master L11iquid IIII360L.	12500	1	12500
GIGABYTE GeForce RTX 3080 Ti VISION OC 12 GB GDDR6X.	142500	1	142500
Hard Disk Drive 2TB.	5400	1	5400
Monitor Asus TUF Gaming VG27AQ1A 27 inch 2K.	51500	1	51500
Motherboard Gigabyte Z590 VISION G 11 TH GEN.	24000	1	24000
Power Supply Corsair RM650 Watt 80 plus Gold.	9200	1	9200
Processor Intel 11 th Gen Rocket Lake Core i9 11900k 3.50-5.20 GHz.	55200	1	55200
RAM Gigabyte AORUS RGB 16GB.DDR4 4400MHz.	7800	2	15600
Samsung M.2 250GB 980 PRO.	7800	1	7800
Total			323,700

Discretion	Budget
Hardware Costs	323700
Transportation Costs	4500
Miscellaneous Cost	1300
Total Budget	329,500

Figure 4.5: Cost Analysis Table (in BDT) by Category

4.11 Adopted Tools and Techniques:

4.11.1 Hardware Requirements

- CPU:

Processor: Intel 11th Gen Rocket Lake Core i9 11900K (3.50–5.20 GHz) Justification: High-performance CPU for model training coordination and preprocessing tasks.

- GPU:

Graphics Card: GIGABYTE GeForce RTX 3080 Ti VISION OC 12 GB GDDR6X Justification: Important for training big deep learning models, such as VGG16, with good efficiency.

- RAM:

Memory: 32 GB DDR4 (2x16GB) at 4400 MHz (Gigabyte AORUS RGB) Justification: Large dataset handling and heavy computation during model training.

- Storage:

Primary Storage: Samsung M.2 250GB 980 PRO Additional Storage: 2TB Hard Disk Drive Justification: High-speed storage for weights and datasets, with abundant backup storage.

- Motherboard:

Model: Gigabyte Z590 VISION G 11th Gen Justification: Compatible with both CPU and GPU, as well as supporting overclock to boost performance.

- Cooling:

CPU Cooler: Cooler Master L11quld III L360L Justification: Makes sure the system can provide stability during heavy computation.

- Power Supply:

Model: Corsair RM650 Watt 80 Plus Gold Justification: Provides stable power for high-performance components of the system.

- Monitor:

Display: Asus TUF Gaming VG27AQ1A (27-inch, 2K) Justification: High-resolution monitor for visualizing datasets and training results. Estimated Total Cost:

- Hardware Cost: 323,700 BDT

Additional Costs (Transportation, Miscellaneous): 5,800 BDT Grand Total: 329,500 BDT

4.11.2 Software Requirement:

The following software components are necessary for the proper functioning of the system here: Python 3.x: The system requires Python version 3.x for development and execution.

Machine learning libraries: Including Scikit-learn, TensorFlow, Data Image Generator, and flow-implementing machine learning models.

Version control system—: Utilize Git to oversee the sources code the promote collaboration within the development team.

Integration development Environment(IDE): Utilize tools like Anaconda and Visual Studio code to stream coding and debugging processes for increased efficiency

4.12 Conclusion:

This work is done for the classification of gallbladder diseases using a design and development approach that is scalable and preserves privacy. Integration of deep learning VGG-16 models with federated learning provides high classification accuracy while strictly maintaining the data privacy standard. Therefore, the result obtained might bring a revolution in the healthcare domain with accurate disease classification without compromising patient confidentiality, which could provide a better diagnosis and treatment of gallbladder diseases at various medical institutes.

Chapter 5

Result analysis

5.1 Model performance:

5.1.1 VGG16 Result:

The VGG16 model demonstrated better performance in Gallbladder diseases for the classification of different classes like gallstones, polyps, perforation, etc. The model achieved an overall accuracy of 97.29 percent on the test set, affirming its ability to generalize well to unseen data.

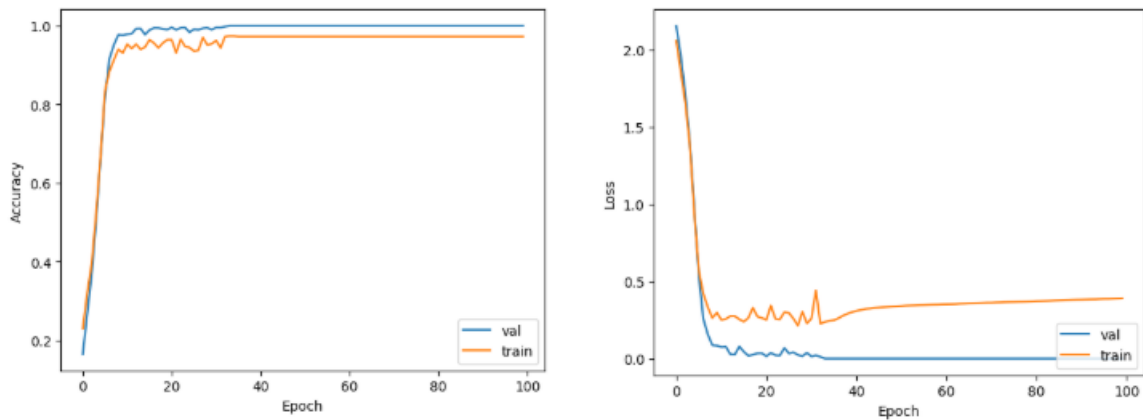


Figure 5.1: VGG16 plot training and validation accuracy values and validation loss values.

The confusion matrix visually represents the model's predictions against actual labels. It helps identify patterns of misclassification and understand which classes might be more challenging for the model.

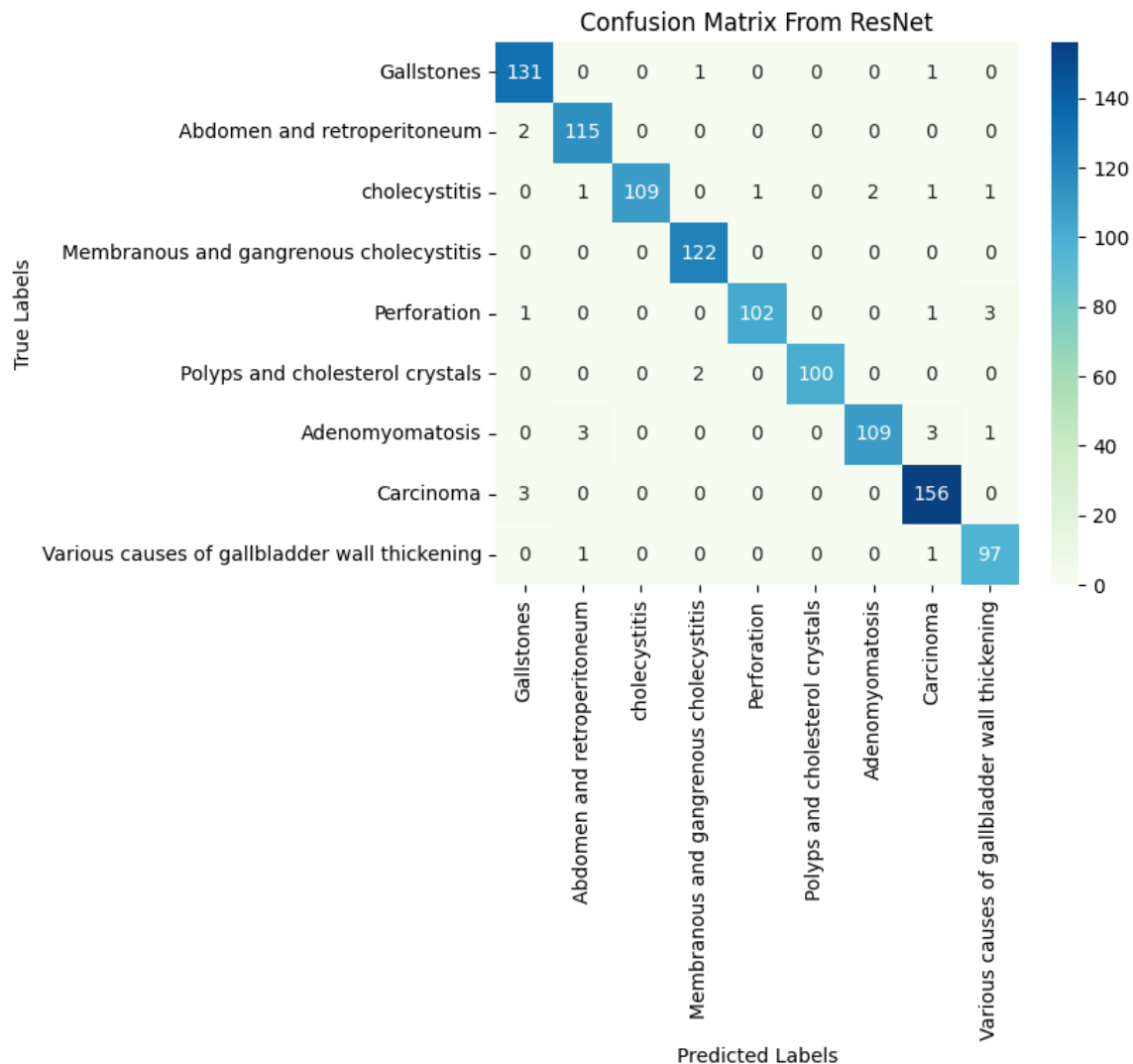


Figure 5.2: VGG16 confusion matrix values.

VGG16 model demonstrates solid performance in classifying gallbladder diseases from cholecystography images. It achieves an accuracy of 97.29 Percent, indicating its ability to correctly predict a significant majority of the cases. With a precision of 97.45 Percent, the model effectively minimizes false positives, making it reliable in identifying true cases. It's a recall of 97.22 Percent highlights strong sensitivity, capturing most true positives. The F1-Score of 97.31 Percent Reflects a balanced trade-off between precis.

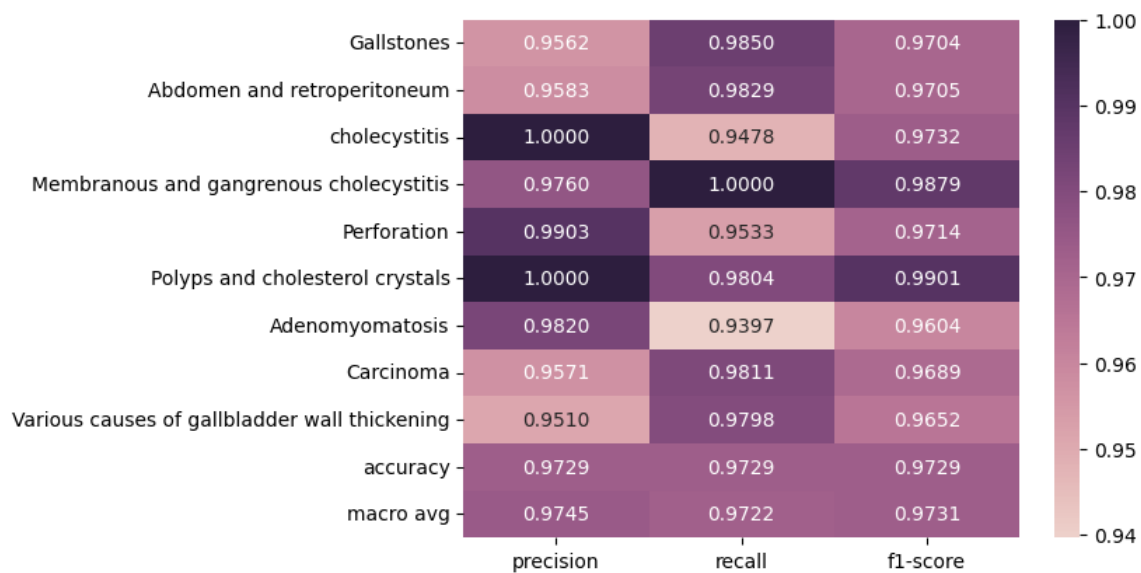


Figure 5.3: VGG16 Classification Report.

5.1.2 ResNet50 Result:

ResNet50 model demonstrated performance in Gallbladder diseases for the classification of different classes like gallstone, polyps, perforation, etc. The model achieved an overall accuracy of 97.29 percent on the test set, affirming its ability to generalize well to unseen data.

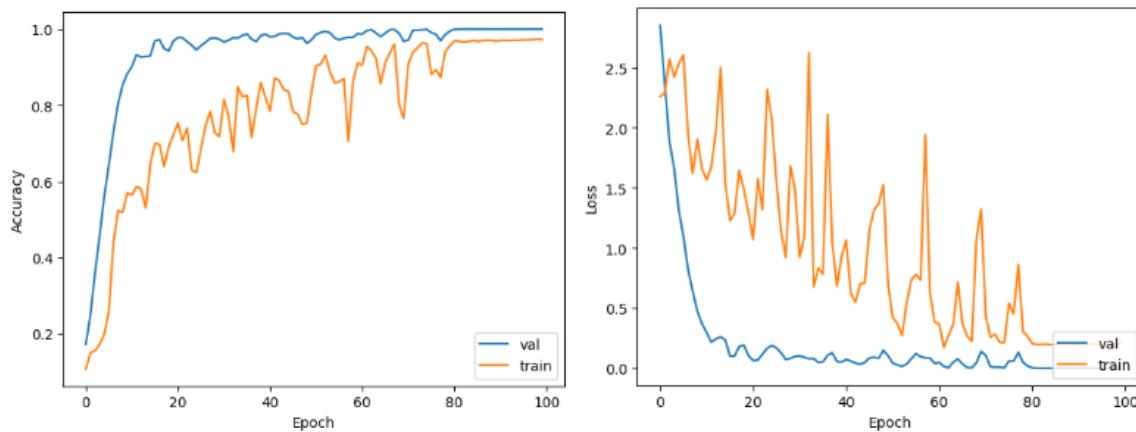


Figure 5.4: ResNet50 plot training and validation accuracy values and validation loss values.

The confusion matrix visually represents the model's predictions against actual labels. It helps identify patterns of misclassification and understand which classes might be more challenging for the model.

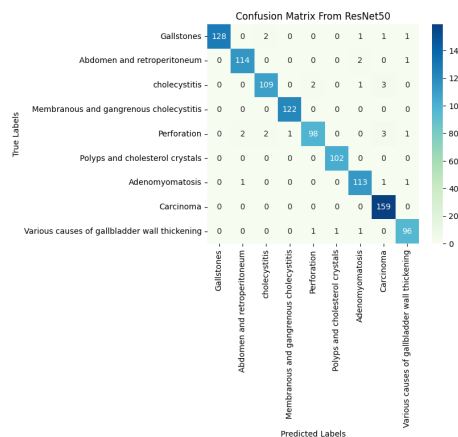


Figure 5.5: ResNet50 confusion matrix values.

ResNet50 model demonstrates solid performance in classifying gallbladder diseases from cholecystography images. It achieves an accuracy of 97.29 Percent, indicating its ability to correctly predict a significant majority of the cases. With a precision of 97.45 Percent, the model effectively minimizes false positives, making it reliable in identifying true cases. Its recall of 97.22 Percent highlights strong sensitivity, capturing most true positives. The F1-Score of 97.31 Percent Reflects a balanced trade-off between precis.

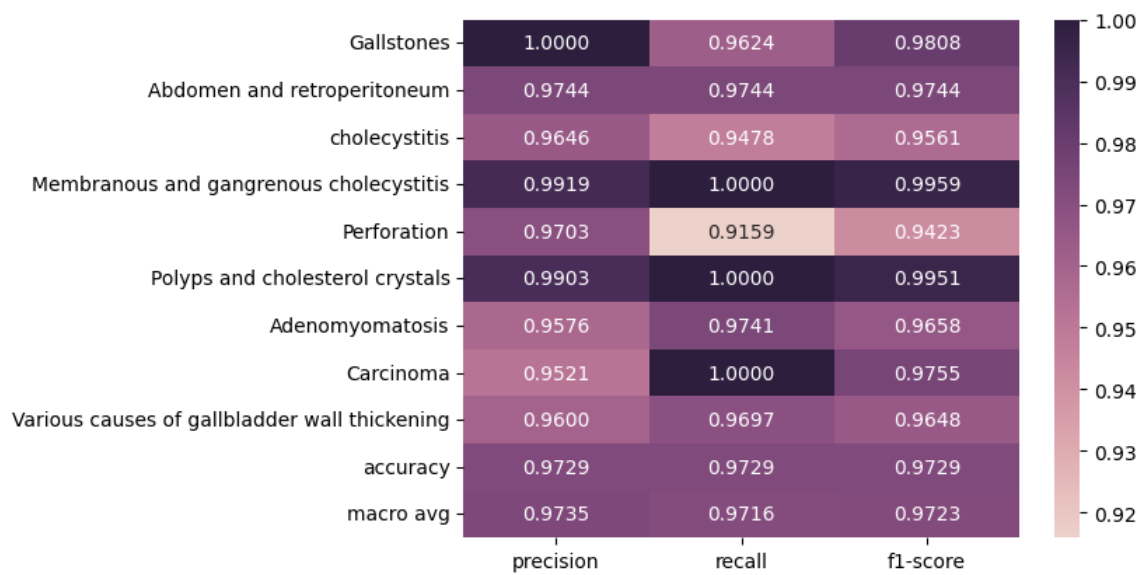


Figure 5.6: ResNet50 Classification Report.

5.1.3 Vision Transformer Result:

Vision Transformer model demonstrated performance in Gallbladder diseases for the classification of different classes like gallstone, polyps, perforation, etc. The model achieved an overall accuracy of 96.36 percent on the test set, affirming its ability to generalize well to unseen data.

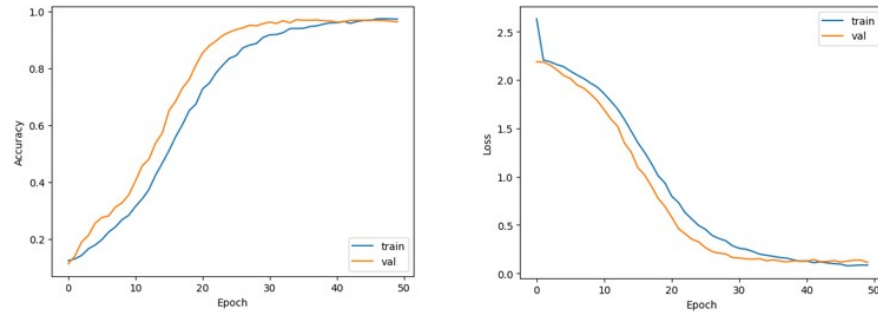


Figure 5.7: Vision Transformer plot training and validation accuracy values and validation loss values.

The confusion matrix visually represents the model's predictions against actual labels. It helps identify patterns of misclassification and understand which classes might be more challenging for the model.

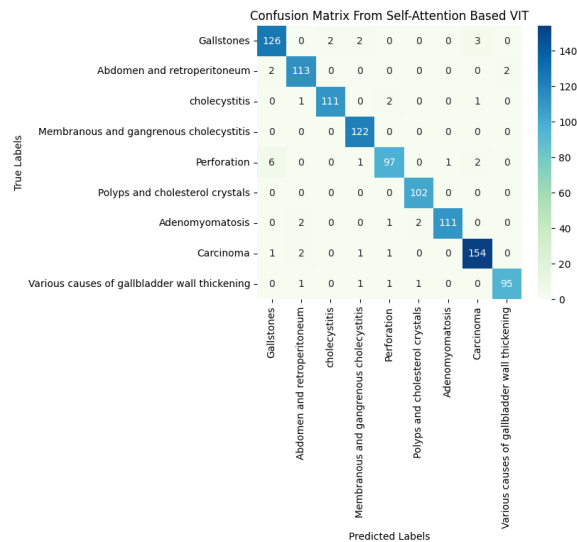


Figure 5.8: Vision Transformer confusion matrix values.

Vision Transformer model demonstrates solid performance in classifying gall-bladder diseases from cholecystography images. It achieves an accuracy of 96.36 Percent, indicating its ability to correctly predict a significant majority of the cases. With a precision of 96.36 Percent, the model effectively minimizes false positives,

making it reliable in identifying true cases. Its recall of 96.36 Percent highlights strong sensitivity, capturing most true positives. The F1-Score of 97.31 Percent Reflects a balanced trade-off between precis.

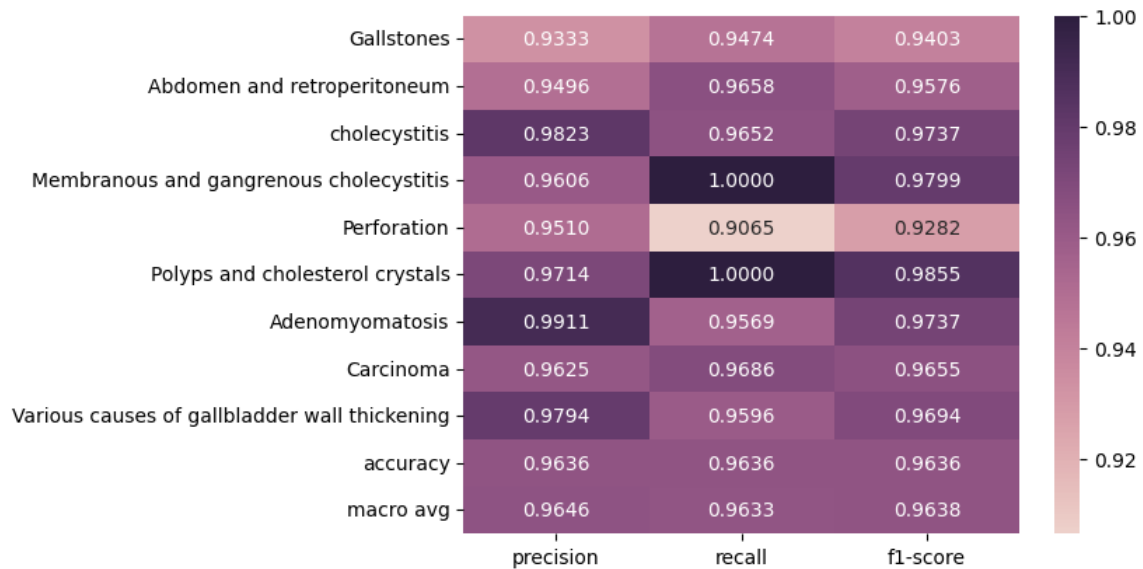


Figure 5.9: Vision Transformer Classification Report.

5.2 FL model performance:

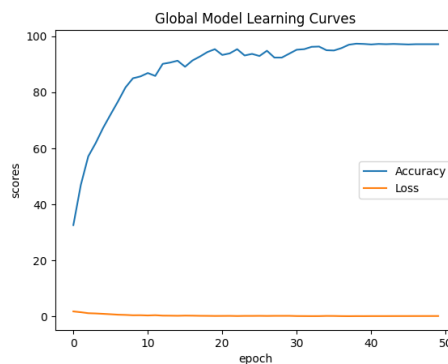


Figure 5.10: Global Model Curve.

This graph illustrates the learning curve of a global machine learning model on different epochs, using two metrics: Accuracy and Loss.

Accuracy Curve (Blue Line): The accuracy starts relatively low and increases steadily, indicating the model's improvement in prediction performance over time. It seems to plateau after around 20–30 epochs, suggesting the model is approaching its optimal performance.

Loss Curve (Orange Line): The loss is higher at the beginning and decreases rapidly in the init.

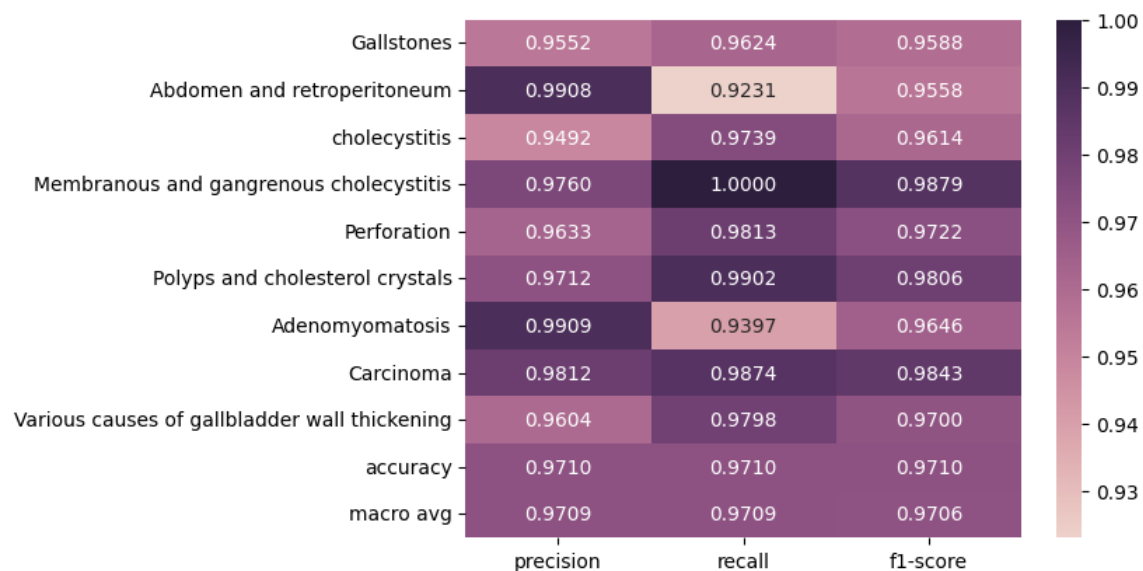


Figure 5.11: Classificaton Report.

This classification report is presented as a heat map and is a detailed performance evaluation of a multi-class classification model. Precision, recall, f1-score, and accuracy are some of the key metrics that have been mentioned for each class, while the macro-average summary can be seen at the bottom. Here's a breakdown of the Classification Report:

Classes and Metrics: Each row is a different metric (precision, recall, f1-score, accuracy), and each column is a class or the macro average.

Color Coding: Dar.

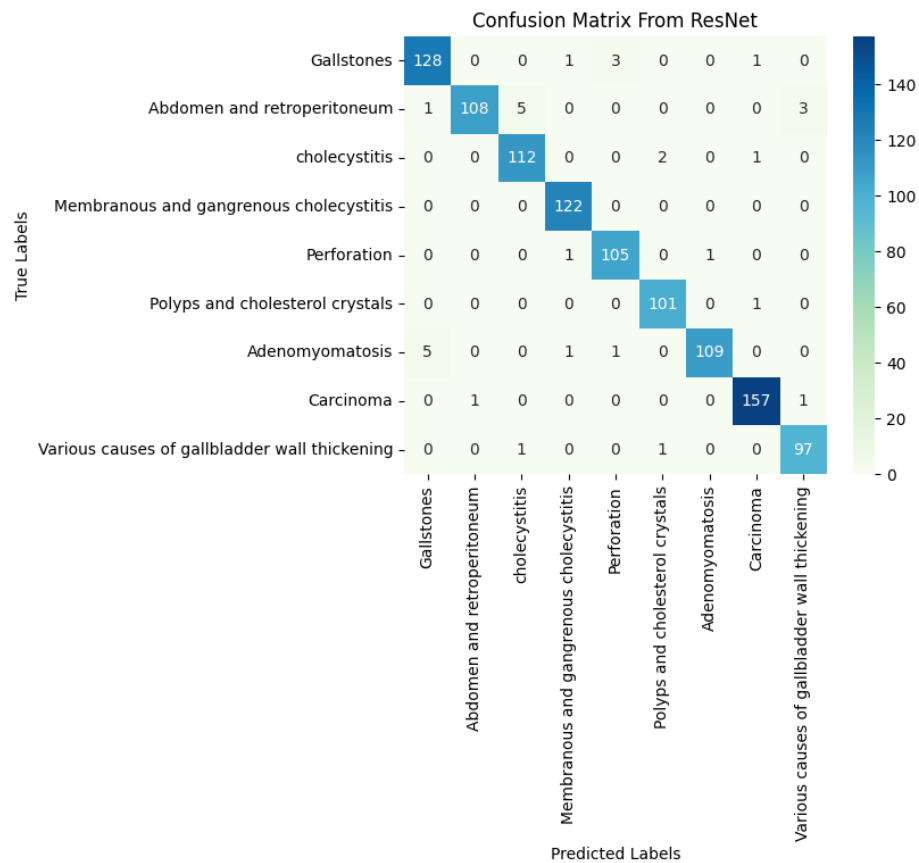


Figure 5.12: Confusion Matrix.

This is a confusion matrix created for a classification model, probably on the ResNet architecture, where performance is evaluated across a number of classes. Now, let me describe this in detail.

Structure: The true labels, Y-axis, represent the actual class labels. The predicted labels, X-axis, are the predicted class labels by the model. **Diagonal Elements:** The diagonal cells contain the number of correct predictions that were made for each class. **Off-diagonal elements** indicate misclassifications-the model predicted the wrong class. **Observations:** High Accuracy:

Most of the values are concentrated on the diagonal, which represents the number of correct predictions for a class. For example, "Gallstones" were correctly predicted 128 times, "Carcinoma" was predicted correctly 157 times, and "Cholecystitis" was predicted correctly 112 times. **Minor Misclassifications:**

Some of the off-diagonal entries show that "Adenomyomatosis" got 5 misclassi-

fications for "Gallstones" and "Abdomen and Retroperitoneum" got 5 misclassifications into some other class, indicating minor confusion among similar classes. These errors might be because some features are similar between classes or the data is unbalanced. Class Performance:

Best Class Performance: "Carcinoma" (157 correct, only 2 misclassifications) and "Gallstones" (128 correct, 5 misclassifications). Relatively Challenging Class: "Adenomyomatosis" and "Abdomen and Retroperitoneum" show a relatively higher percentage of misclassifications in comparison with other classes. Overall Assessment: The matrix shows that the model fits very well with space for improvement on a few aspects. Further analysis of misclassified instances might help in improving the feature extraction or class separation mechanisms, particularly for similar or overlapping categories.

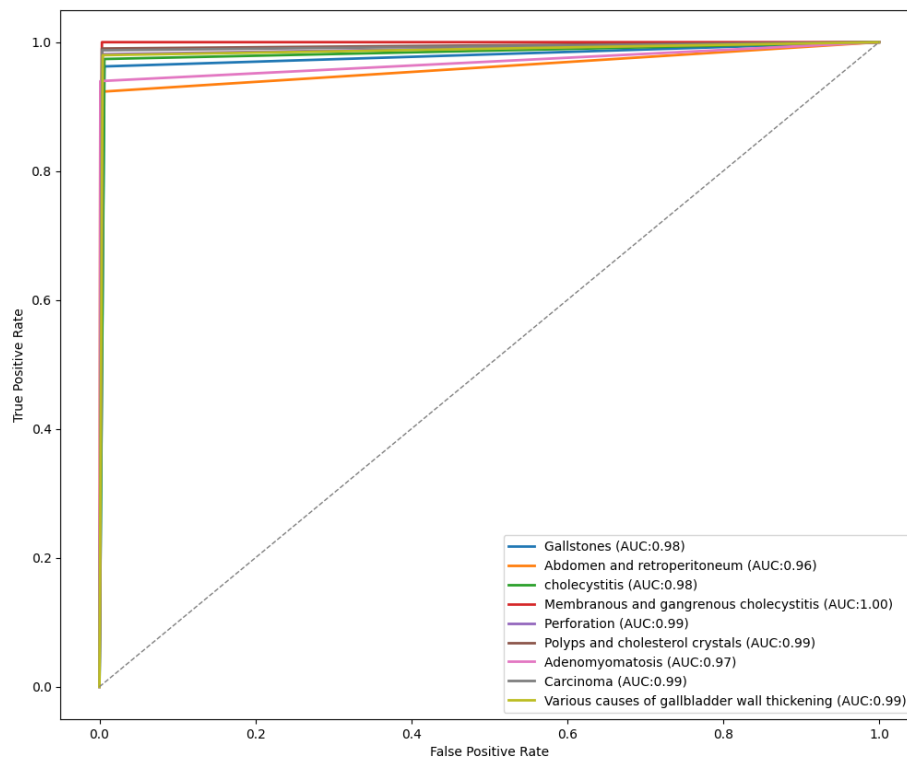


Figure 5.13: ROC-AUC.

This ROC curve plot shows the performance of a federated learning model for

multi-class classification. Each curve represents the performance of the model for a particular class, and the AUC is given for each.

Key Points Axes:

X-axis: The X-axis is the False Positive Rate, which measures the proportion of negatives that are incorrectly classified as positive. The Y-axis represents the True Positive Rate (TPR), which is a measure of the proportion of correctly classified positives. The dashed diagonal line represents a random classifier with no discriminative power, at $AUC = 0.5$. Performance:

The ROC curves of all classes lie closely toward the top-left side of the graph, showing great performance in classification. Higher values of AUC (closer to 1.00) reflect better discrimination between classes. AUC Values:

Most classes have an AUC of 0.98 or higher, indicating very good model performance. "Membranous and Gangrenous Cholecystitis" achieves a perfect AUC of 1.00, which indicates perfect discrimination. The lowest AUC is for "Abdomen and Retroperitoneum" ($AUC = 0.96$), which also is strong but might be improved. Federated Learning Context:

The high AUC scores for all classes indicate the effectiveness of the federated learning model in achieving robust performance across diverse data distributions without centralizing the data. This result underlines the potential of federated learning in handling sensitive healthcare data while maintaining high accuracy.

5.3 Model Result Comparison :

5.3.1 Accuracy:

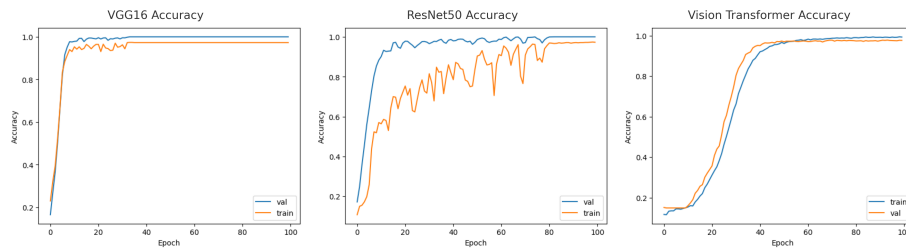


Figure 5.14: Accuracy Comparison of VGG16, ResNet50 and Vision Transformer.

5.3.2 Confusion Matrix:

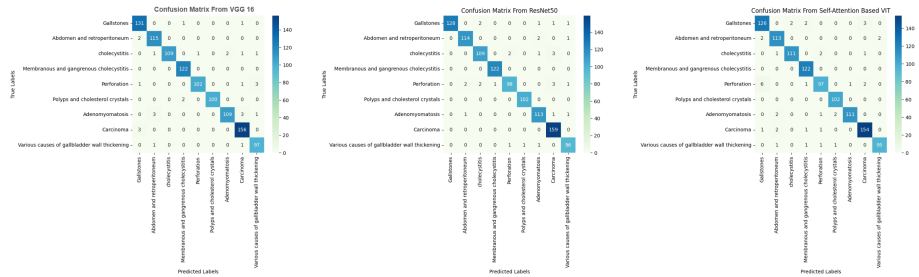


Figure 5.15: Confusion Matrix Comparison of VGG16, ResNet50 and Vision Transformer.

5.3.3 Classification Report:

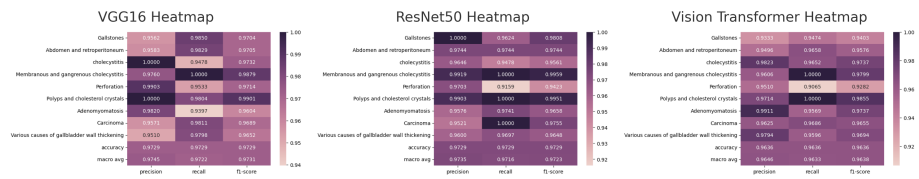


Figure 5.16: Classification Report of VGG16, ResNet50 and Vision Transformer.

5.3.4 Parameters Comparison:

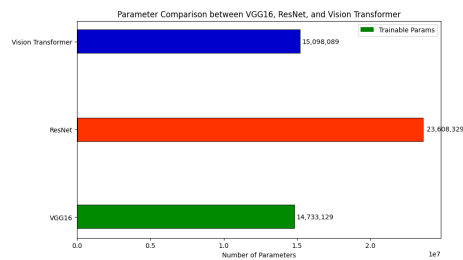


Figure 5.17: Parameter Comparison Between VGG16, ResNet50 and Vision Transformer.

5.4 Model Comparison: VGG16, ResNet50, and ViT

- Training Stability:

VGG16 and ViT stabilized around 50 epochs. ResNet50 showed instability with spikes in loss/accuracy.

- Accuracy

VGG16 and ResNet50: 97 percent accuracy. ViT: 96 percent accuracy.

- Class Performance:

VGG16: Best overall performance, especially in Perforation and Various causes of gallbladder wall thickening. ViT and ResNet50: Struggled with certain classes.

- Precision, Recall, F1-Score:

VGG16 and ResNet50: Highest scores.

5.4.1 Why Choose VGG16 for FL

- High Accuracy: 97 percent accuracy.
- Stable Training: Reached stability quickly.
- Best Class Performance: Outperformed in key classes.
- Lightweight: More suitable for Federated Learning.
- Reliable Metrics: Top scores in precision, recall, and F1-score.

Chapter 6

Conclusions

This capstone project has provided a successful demonstration of a privacy-preserving Deep learning approach for classifying the gallbladder from cholecystography images under a federated learning environment. The integration of VGG16 models with federated learning allowed for solving one of the most critical issues in maintaining data privacy within the medical domain while improving the accuracy of classification over traditional approaches. By leveraging federated learning, the sensitive medical data cholecystography images never left the local devices; hence, patient privacy was guaranteed and strict data protection regulations were followed. It also allowed for collaborative and decentralized training from diverse datasets across multiple institutions, thus boosting the generalization and robustness of the model. The VGG16 architecture proved effective in capturing long-range dependencies in image data, yielding superior results on image classification tasks compared to conventional convolutional neural networks. This ability of the architecture to focus on critical regions of the images through self-attention mechanisms proved a plus in the medical image analysis domain, given the fine details present in images, which are crucial in diagnosis. Although promising, the system faced some problems: data heterogeneity, communication overhead, and complexity in ensuring robust privacy. However, data augmentation, model compression, and secure aggregation were some of the techniques that reduced these challenges. In addition, developing privacy-preserving methods like differential privacy and secure aggregation on top of this setting indeed ensured that the sensitive information maintained throughout was not divulged during the training process. This will not only provide a new solution for gallbladder classification using medical imaging but also contribute to the development of a growing body of federated learning and

privacy-preserving AI in healthcare. The results have underlined the promise of this approach for broader applications in medical diagnostics, enabling scalability and privacy-aware AI systems that can be deployed in clinical environments in the real world. Moreover, future studies can be done to refine the federated learning framework to improve scalability further and explore the application of this model to other medical imaging tasks. These will no doubt further consolidate the advance of health healthcare by privacy-preserving AI. This work lays the foundation for secure and efficient deep learning deployment in healthcare, where collaboration across institutions is possible with better models and without compromising privacy.

6.1 Social, Legal, Ethical, and Environmental Issues

The creation of a deep learning model based on the VGG16 framework for the classification of gallbladder diseases from cholecystography images that also protects the privacy of images in the process of learning, implemented in a federated learning setting, provokes a series of significant social, legal, ethical, and environmental issues. The relevant problems for these areas are listed as follows:

6.1.1 Social Aspect

The Deep learning approach, protected by privacy, to classify gallbladder conditions from cholecystography images in a federated learning environment has several important social aspects, impacting both individuals and broader society. These relate to healthcare accessibility, equity, trust in AI technologies, and the overall societal benefit of improving diagnostic accuracy in medical practice. Key social factors include:

- Improved Healthcare Access and Quality.
- Enhancing Public Health Outcomes.
- Addressing Healthcare Disparities.
- Public Awareness and Education.
- Societal Acceptance of AI in Healthcare.

6.1.2 Legal Aspect

Some of the major legal considerations are: protection of data, intellectual property rights, liability, compliance with healthcare regulations, and cross-border data sharing in the development of a privacy-protected, Deep learning approach for the classification of gallbladder conditions from cholecystography images in a federated learning environment. Herein are key legal areas that should be discussed for this project:

- Data Protection and Privacy Laws.
- Informed Consent and Data Usage.
- Liability and Accountability.
- Cross-Border Data Sharing and Jurisdictional Issues.

6.1.3 Ethical Responsibility Integration

Ethics, fairness, privacy, transparency, accountability, and reduction of harm in the course of developing a privacy-protected VGG16 deep learning approach to classify gallbladder conditions from cholecystography images in a federated learning environment should be carefully woven into every step throughout the project. The important ethical considerations to be done and how it should go in this work are provided below.

6.1.4 Environment and Sustainability

Any privacy-protected Deep learning approach to classify gallbladder conditions from cholecystography images within a federated learning environment should be done with considerations of environmental and sustainability aspects in mind so that the project will have positive, long-term effects on the healthcare sector as well as the general environment. This section highlights the most relevant environmental and sustainability aspects related to the project here:

- Energy Efficiency in AI Training and Model Deployment
- Sustainable Data Practices.

- Health System Sustainability.
- Long-Term Environmental Impact through Innovation.
- Reducing Waste and Promoting Circular Economy.
- Social and Environmental Justice.
- Reducing the Environmental Impact of AI Development.

6.1.5 Environmental and Sustainability aspects

It works on the classification of gallbladder conditions with regard to cholecystography images using a privacy-protected Deep learning model within the environment of federated learning. This project would result in sustainable practices related to minimizing environmental footprint via not involving the need for transferring, storing, and computational consumption in centralized data warehouses; Thereby, reducing energy utilization and carbon footprint from these facilities. Federated learning also allows health institutions to jointly train a machine learning model without data transfer, which is both private and resource-efficient. The energy efficiency of the model-optimized hardware and cloud services powered by renewable energy supports the sustainability of the project. Additionally, this can contribute to resource conservation in health by reducing environmental impacts due to medical procedures that may involve unnecessary imaging tests by simply enhancing diagnostic accuracy. Besides, the scalability of the federated learning approach ensures health advancements can be deployed even in low-resource settings, hence promoting equity in access to AI-driven diagnostics and fostering global health sustainability. Ultimately, the project looks to improve medical practices with minimal waste, optimize resource utilization, and assist environmentally friendly healthcare solutions.

6.2 Brief Summary

This project aims to construct a privacy-protected, Deep learning model for classifying gallstones, polyps, and cancers from cholecystography images of the gallbladder.

The work herein intends to perform machine learning model training collaboratively over different healthcare institutions with the aid of a federated learning environment, without necessarily sharing sensitive patient data, hence guaranteeing its privacy and security. By combining the power of AI with federated learning, this model will enhance diagnostic accuracy for healthcare providers while maintaining patient confidentiality in decision-making. The project also aims to address challenges such as small datasets and data diversity, improving the model's performance across different populations. It thus aspires to finally provide medical imaging diagnostics in such a way that a privacy-preserving, scalable, and efficient method improves the detection of gallbladder diseases and eventually allows better patient care.

6.3 Future works

In the future, the privacy-protected Deep learning model will be further refined and updated to ensure gallbladder disease classification accuracy and dependability. Key directions also involve increasing the model's scalability for larger and more varied datasets by embedding other frameworks of federated learning, as well as the collaboration of more and more healthcare institutions around the world. This would generalize better into diverse patient demographics, from ages, ethnicity, and diseases in a fashion that keeps privacy and data security, among other tenets of ethics.

Further work on this research will be to find optimal performance in the model for both efficient computational resources during training and inference and feasibility of deployment in low-resource settings. Another important avenue for future development is improving the model's explainability to assist healthcare professionals in interpreting its predictions, increasing trust in AI-assisted diagnoses. Further integration with other medical imaging techniques and real-time diagnostic tools will also be developed to create a more comprehensive diagnostic system. Eventually, the project will investigate the possibility of real-time deployment in clinical environments, continuously learning from the federated data in light of new medical insights and becoming better at making decisions with time. In such a way, the system will always be relevant to and performant in dynamic medical diagnostics.

References

- Alazwari, S., Alsamri, J., Alamgeer, M., Alotaibi, S. S., Obayya, M., & Salama, A. S. (2024). Automated gall bladder cancer detection using artificial gorilla troops optimizer with transfer learning on ultrasound images. *Scientific Reports*, 14(1), 21845.
- Basu, S., Gupta, M., Rana, P., Gupta, P., & Arora, C. (2022). Surpassing the human accuracy: detecting gallbladder cancer from usg images with curriculum learning. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 20886–20896).
- Chen, J., Yan, H., Liu, Z., Zhang, M., Xiong, H., & Yu, S. (2024). When federated learning meets privacy-preserving computation. *ACM Computing Surveys*, 56(12), 1–36.
- Duncan, C. B., & Riall, T. S. (2012). Evidence-based current surgical practice: calculous gallbladder disease. *Journal of Gastrointestinal Surgery*, 16(11), 2011–2025.
- Esen, İ., Arslan, H., Esen, S. A., Gülşen, M., Kültekin, N., & Özdemir, O. (2024). Early prediction of gallstone disease with a machine learning-based method from bioimpedance and laboratory data. *Medicine*, 103(8), e37258.
- Gupta, P., Basu, S., Rana, P., Dutta, U., Soundararajan, R., Kalage, D., ... others (2024). Deep-learning enabled ultrasound-based detection of gallbladder cancer in northern india: a prospective diagnostic study. *The Lancet Regional Health-Southeast Asia*, 24.
- Haleem, A., Javaid, M., & Khan, I. H. (2020). Artificial intelligence (ai) applications in dentistry. *Curr. Med. Res. Pract*, 10, 36–38.
- Kim, T., Choi, Y. H., Choi, J. H., Lee, S. H., Lee, S., & Lee, I. S. (2021). Gallbladder polyp classification in ultrasound images using an ensemble convolutional neural network model. *Journal of Clinical Medicine*, 10(16), 3585.
- Li, L., Xie, N., & Yuan, S. (2022). A federated learning framework for breast cancer histopathological image classification. *Electronics*, 11(22), 3767.

- Li, Q., He, B., & Song, D. (2021). Model-contrastive federated learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10713–10722).
- Obaid, A. M., Turki, A., Bellaaj, H., Ksantini, M., AlTae, A., & Alaerjan, A. (2023). Detection of gallbladder disease types using deep learning: an informative medical method. *Diagnostics*, 13(10), 1744.
- Yu, M. H., Kim, Y. J., Park, H. S., & Jung, S. I. (2020). Benign gallbladder diseases: Imaging techniques and tips for differentiating with malignant gallbladder diseases. *World journal of gastroenterology*, 26(22), 2967.
- Yuan, K., Zhang, X., Yang, Q., Deng, X., Deng, Z., Liao, X., & Si, W. (2024). Risk prediction and analysis of gallbladder polyps with deep neural network. *Computer Assisted Surgery*, 29(1), 2331774.
- Zhou, Y., Chen, S., Wu, Y., Li, L., Lou, Q., Chen, Y., & Xu, S. (2023). Multi-clinical index classifier combined with ai algorithm model to predict the prognosis of gallbladder cancer. *Frontiers in Oncology*, 13, 1171837.