Project 5
Semi-Supervised Breast Cancer Ultrasound Classification

Project Overview

Breast cancer ultrasound imaging plays a critical role in early detection and diagnosis, particularly for distinguishing between benign and malignant lesions. Deep learning–based classification models have shown promising results in this domain; however, their success typically depends on large, well-annotated datasets curated by expert radiologists. In practice, such annotations are expensive, time-consuming, and often limited in availability.

This project investigates how breast cancer ultrasound classification performance degrades under limited annotation availability, and how different convolutional and transformer-based architectures cope with varying proportions of labeled data. By systematically reducing the percentage of labeled samples in a breast ultrasound dataset, this study aims to benchmark classical CNN models and modern Vision Transformer (ViT) models, establishing a clear performance–annotation trade-off. The project will first target an international conference and then be extended toward a Q2 journal publication, where classification will be further integrated with segmentation in a multi-task learning framework.

Research Objectives

- Quantify classification performance under different proportions of labeled data (e.g., 1%, 5%, 10%, 25%, 50%, 100%).
- Benchmark popular CNN-based and transformer-based classification models.
- Analyze robustness and generalization under annotation scarcity.
- Provide qualitative and quantitative insights into misclassification patterns.
- Produce a conference paper and an extended journal submission.

Methodology

1. Dataset Preparation:
- Use a breast cancer ultrasound dataset with benign and malignant labels.
- Apply standardized preprocessing (normalization, resizing, augmentation).
- Construct multiple training subsets by sampling different proportions of labeled images.
- Keep validation and test sets fixed to ensure fair comparison.

2. Limited Annotation Simulation:
- Systematically vary the proportion of labeled data used for training.
- Maintain consistent data splits and training protocols across experiments.

3. Model Benchmarking:
- Train and evaluate CNN-based models (e.g., ResNet, DenseNet, EfficientNet).

- Train and evaluate transformer-based models (e.g., ViT, hybrid CNN–ViT models).
- Use consistent optimization and evaluation settings for fair comparison.

4. Evaluation Metrics:
- Accuracy
- Precision, Recall, F1-score
- ROC-AUC
- Sensitivity–specificity trade-offs

5. Visualization and Analysis:
- Performance–annotation curves
- Confusion matrices under different annotation ratios
- Failure case and misclassification analysis

Timeline (6 Months)

Phase 1 (Months 1–3): Conference Target
- 1 breast ultrasound dataset
- Binary classification (benign vs. malignant)
- Benchmark CNN and ViT-based models under limited annotation
- Conference submission (e.g., ICISN, CITA, MAPR)
- Compute: Google Colab, Kaggle

Phase 2 (Months 4–6): Journal Target
- Extend to additional datasets or splits
- Introduce multi-task learning combining classification and segmentation
- Explore semi-supervised and self-supervised learning strategies
- Q2 journal submission
- Compute: AIVN GPUs, Private servers