Project Proposal

EEE 385L

Machine Learning Laboratory

Section: 1

Spring 25

Group No: 03

Group Members:

ID	Name
22121092	MOHAMMAD FASIUL ABEDIN KHAN
21321007	FARRDIN NOWSHAD
22121076	ABRAR MAKSUD NAHEAN
22121024	MD. ABU ANAS MRIDUL

Project Proposal

Deep Learning for BI-RADS Classification of Mammograms Using the King Abdulaziz University Dataset: A Saudi Medical Imaging Study

1. Introduction

Breast cancer remains one of the most prevalent and life-threatening cancers among women worldwide. Its early detection is vital for improving patient outcomes. Mammography is the gold standard for breast cancer screening, capable of detecting tumors that are too small to be felt. However, the diagnostic accuracy of mammography is limited by human factors such as radiologist fatigue, variability in experience, and subjectivity in interpretation.

Recent advances in deep learning (DL) have transformed image analysis across fields, including radiology. Convolutional Neural Networks (CNNs), a subclass of DL models, have shown promise in detecting and classifying abnormalities in medical images with performance that can rival or exceed that of human experts [1], [2].

However, AI development is hindered in Saudi Arabia due to the lack of publicly available, locally relevant datasets. The King Abdulaziz University Mammogram Dataset fills this gap, providing the first digitized, annotated breast cancer dataset from the region. It is categorized according to BI-RADS (Breast Imaging-Reporting and Data System), the global standard for mammogram assessment.

This project proposes to develop and evaluate a deep learning-based diagnostic tool using this dataset, aiming to assist Saudi radiologists in automated BI-RADS classification of mammograms.

2. Problem Statement

Field: Healthcare AI - Medical Imaging

Classifying mammograms into BI-RADS categories is often done manually by radiologists. This process can be subjective, and different experts may give different opinions for the same image. The new Saudi dataset provides a great opportunity for training AI models, but it has a major challenge: the number of images in each BI-RADS category is not balanced. In particular, high-risk categories like BI-RADS 4 and 5 have fewer samples, even though they are very important for early cancer detection.

Because of this imbalance and the complexity of reading mammograms, we need advanced machine learning models. These models should be able to learn from limited data and still perform well across all classes. Our goal is to build an AI system that can accurately detect high-risk cases while also performing reliably for all BI-RADS categories.

3. Problem Description and Project Objectives

3.1 Challenges:

- **Data Scarcity & Imbalance:** The dataset has fewer samples for BI-RADS categories 4 and 5 compared to category 1, which can bias the model.

- **Noise & Real-World Variability:** Mammograms are affected by image quality, artifacts, and differing anatomical features.

- **Computational Constraints:** Optimizing model performance while ensuring inference time is within acceptable limits for clinical use.

3.2 Main Objective

To design and evaluate a deep learning model capable of classifying mammograms into BI-RADS categories 1, 3, 4, and 5 using the King Abdulaziz University dataset.

3.3 Specific Objectives

- Perform extensive preprocessing, including image enhancement and metadata parsing.
- Balance the dataset using synthetic oversampling and loss engineering.
- Develop CNN-based models using transfer learning (e.g., ResNet50, EfficientNet).
- Fine-tune models with advanced optimization and regularization techniques.
- Evaluate models with class-sensitive metrics (AUC, F1-score, confusion matrix).
- Visualize and explain model decisions using Grad-CAM and saliency maps.
- Ablation studies will be conducted to identify impactful components of the model pipeline.
- Create a user-ready code repository and complete a technical report.
- Deliver a live demonstration and presentation by 15 May.

4. Dataset Description

- **Name:** King Abdulaziz University Mammogram Dataset [3]
- **Images:** 6,109 images (CC and MLO views of left and right breasts)
- **Cases:** 1,521 annotated patients
- **Categories:** BI-RADS 1 (normal), 3 (probably benign), 4 (suspicious), 5 (highly suggestive of malignancy)
- **Challenge:** Severe class imbalance BI-RADS 4 and 5 are underrepresented
- **Format:** JPEG images + Excel metadata
- **DOI:** https://dx.doi.org/10.21227/a4cs-ax02

5. Methodology

5.1 Data Preprocessing

- **Importing and Matching:** Parse Excel metadata to associate each image with its BI-RADS label.
- **Cleaning:** Remove corrupted or incomplete records.
- **Normalization:** Resize all images to 224x224 pixels, convert to grayscale (if needed), and normalize pixel values.
- **Enhancement:** Apply histogram equalization and CLAHE to improve contrast and edge visibility.
- **Augmentation:** Use geometric and photometric transformations (rotation, flipping, noise injection) to expand minority classes.

5.2 Model Development

- **Baseline:** Simple CNN architecture with 3 convolutional layers.
- **Advanced:** Fine-tune pre-trained models (ResNet50, EfficientNetB0) with a custom classification head [4].
- **Loss Engineering:** Use focal loss and weighted cross-entropy to penalize majority class over-representation.

5.3 Training Strategy

- **Splits:** 70% train, 15% validation, 15% test

- **Optimizer:** Adam with cosine annealing scheduler
- **Validation:** 5-fold cross-validation
- **Early Stopping:** Triggered by no improvement in validation loss for 10 epochs

5.4 Model Evaluation

- **Primary Metrics:** AUC, F1-score, Recall, and Specificity
- **Secondary Metrics:** Accuracy, Precision, Confusion Matrix
- **Error Analysis:** Use Grad-CAM to visualize areas of interest in misclassified images [5].

5.5 Ablation Studies

- Assess the effect of:
  - Data augmentation
  - Loss functions (categorical vs focal loss)
  - Preprocessing pipelines (contrast enhancement, resizing)
  - Model backbone (ResNet vs EfficientNet)

6. Project Management

6.1 Tools and Technologies

- **Programming Language:** Python 3.10
- **Frameworks:** TensorFlow, PyTorch
- **Environment:** Google Colab Pro / Kaggle Kernels / Local GPU
- **Libraries:** Pandas, OpenCV, Albumentations, Matplotlib
- **Version Control:** GitHub

6.2 Timeline

Task	**Dates**	**Deliverable**
Data analysis and cleaning	Apr 24 – Apr 27	Cleaned dataset
Baseline CNN and augmentation setup	Apr 28 – Apr 30	Baseline model and initial results
Proposal Submission	**May 1**	Final PDF/Word proposal

Advanced model development	May 2 – May 6	Fine-tuned DL models (ResNet, etc.)
Evaluation and ablation studies	May 7 – May 10	Metrics, confusion matrices
Error analysis and visualization	May 11 – May 12	Saliency maps and Grad-CAM outputs
Documentation and report writing	May 12 – May 13	Technical report draft
Slide deck and demo setup	May 13 – May 14	Presentation ready
Final Submission and Presentation	**May 15**	Code, report, presentation delivery

7. Expected Outcomes

- A robust, interpretable model that classifies mammograms into BI-RADS categories with high F1 scores for BI-RADS 4 and 5.
- A documented codebase with reproducible experiments and trained weights.
- A detailed report describing methodology, evaluation, and limitations.
- A presentation for medical professionals and AI researchers.

References

[1] T. Mahmood, J. Li, Y. Pei, F. Akhtar, M. U. Rehman, and S. H. Wasti, “Breast lesions classifications of mammographic images using a deep convolutional neural network-based approach,” *PLOS ONE*, vol. 17, no. 1, p. e0263126, Jan. 2022, doi: 10.1371/journal.pone.0263126.

[2] A. F. A. Alshamrani and F. Saleh Zuhair Alshomrani, “Optimizing Breast Cancer Mammogram Classification Through a Dual Approach: A Deep Learning Framework Combining ResNet50, SMOTE, and Fully Connected Layers for Balanced and Imbalanced Data,” *IEEE Access*, vol. 13, pp. 4815–4826, 2025, doi: 10.1109/ACCESS.2024.3524633.

[3] S. Sapkale, “King Abdulaziz University Breast Cancer Mammogram Dataset.” IEEE, Apr. 10, 2024. Accessed: Apr. 25, 2025. [Online]. Available: https://ieee-dataport.org/documents/king-abdulaziz-university-breast-cancer-mammogram-dataset

[4] J. Jiang, J. Peng, C. Hu, W. Jian, X. Wang, and W. Liu, “Breast cancer detection and classification in mammogram using a three-stage deep learning framework based on PAA algorithm,” *Artif. Intell. Med.*, vol. 134, p. 102419, Dec. 2022, doi: 10.1016/j.artmed.2022.102419.

[5] B. Hinton *et al.*, “Deep learning networks find unique mammographic differences in previous negative

mammograms between interval and screen-detected cancers: a case-case study," *Cancer Imaging*, vol. 19, no. 1, p. 41, Jun. 2019, doi: 10.1186/s40644-019-0227-3.