BREAST CANCER DETECTION

With Neural Networks.

PRESENTED BY:

23BCE8566 - Panidapu Meghana
23BCE9684 - Chiluvuri Sai Sri Vaishnavi

PRESENTED TO:

Prof. Beebi Naseeba

Agenda

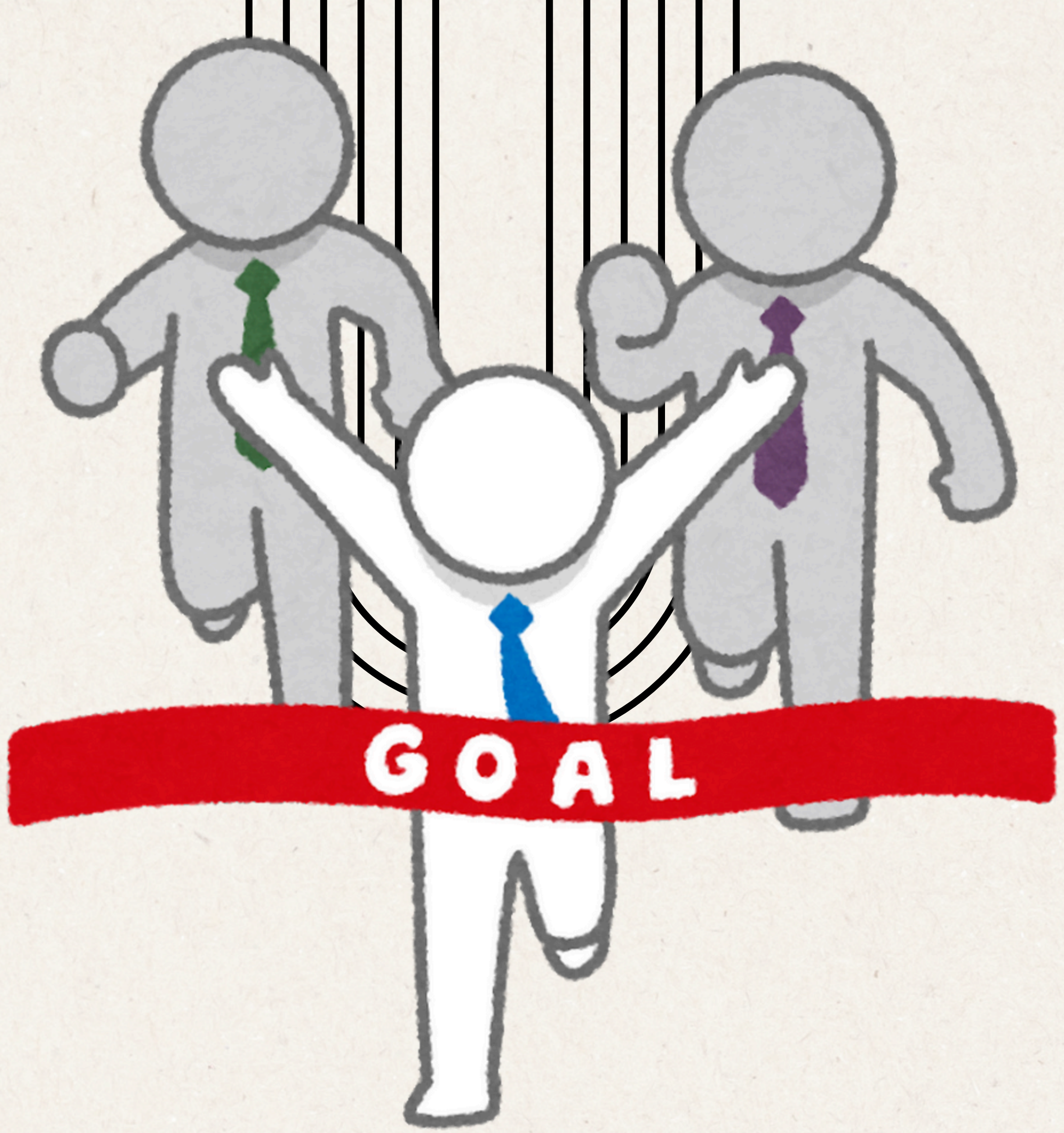
01	Abstract
02	Introduction
03	Literature Survey
04	Existing works
05	Proposed Works - Methodology
06	Dataset

Abstract

Breast cancer is a leading cause of death among women, especially in low- and middle-income countries due to delayed diagnosis. Deep learning, particularly convolutional neural networks (CNNs), has proven highly effective in automated breast cancer detection through image analysis. This study reviews recent advancements in CNN-based methods, comparing CNNs, recurrent neural networks (RNNs), and hybrid models. Using a dataset of 569 instances with 33 tumor features, models like CNNs, LSTMs, and MLPs achieved accuracies of 89–98%. Techniques such as data augmentation, transfer learning, and feature selection significantly improved performance. Hybrid CNN-based models performed best by capturing both spatial and sequential dependencies. The study concludes that AI-driven methods can enhance early detection and reduce diagnostic errors, while future research should focus on transformer models, federated learning, and explainable AI for better robustness and interpretability.

Objectives and Goals

To evaluate and improve deep learning models for accurate and reliable breast cancer detection.



Goal 1

To develop accurate deep learning models for breast cancer detection.

Goal 2

To reduce diagnostic errors through automated image analysis.

Goal 3

To support early detection and improve clinical decision-making.

INTRODUCTION

Breast cancer is one of the most common and life-threatening diseases among women worldwide, with early detection playing a crucial role in improving survival rates. Traditional diagnostic methods often face challenges such as delayed diagnosis and reliance on manual interpretation. Recent advancements in deep learning, particularly convolutional neural networks (CNNs), have shown great potential in automating breast cancer detection by analyzing medical images and tumor features with high accuracy. This project focuses on implementing and comparing various deep learning models, including CNNs, LSTMs, and hybrid approaches, to enhance diagnostic accuracy, reduce errors, and support timely clinical decision-making.

Literature survey

Importance of Early Detection

- Breast cancer is among the leading causes of mortality in women worldwide.
- Deep learning methods, especially CNNs, have been widely adopted to automate image analysis, improving detection accuracy and reducing reliance on manual interpretation.

Datasets Used

- Wisconsin Diagnostic Breast Cancer (WDBC): 569 samples with 30 numerical features derived from FNA images.
- Histopathology and mammography image datasets are extensively used for CNN-based detection tasks.

Performance Improvement Techniques

- Data Augmentation to expand limited datasets.
- Feature Selection & Dimensionality Reduction for noise removal and efficiency.
- Class Imbalance Handling (SMOTE, weighted loss).
- Models frequently report 90%+ accuracy when combining these strategies.

Limitations in Current Research

- Small and imbalanced datasets limit model robustness.
- Lack of standard evaluation metrics across studies.
- Models often fail to generalize across different medical centers and imaging devices.

Existing works

Pre-1995 – Early Screening & CAD

Initial computer-aided detection (CAD) systems developed before deep learning

1995 – Shallow CNNs in Breast Cancer

First use of shallow convolutional neural networks for medical image analysis.

2012 – AlexNet Breakthrough

Deep CNNs gained popularity after ImageNet success, paving way for medical imaging applications.

2016 – Metastasis Detection

Deep learning applied to whole-slide images, AUC 0.925; human+AI hybrid reached near-perfect accuracy.

2017 – Transfer Learning in Mammography

CNNs with fine-tuning improved lesion classification performance.

2018 – Contrast-Enhanced Mammography

Shallow-Deep CNN (SD-CNN) improved diagnostic accuracy from 91% to 95%.

2019 – Large-Scale CNNs

Training on >200,000 mammography exams; AI reached radiologist-level AUC (0.895) and outperformed humans when combined.

2022 – Ultra-High Accuracy Models

- Hybrid optimization (DEGDO) and U-Net approaches achieved >98% accuracy on WDBC and mammography datasets.

Proposed works - methodologies

- **Data Preparation –** Collect dataset (569 instances, 33 features), clean, normalize, and split into train/validation/test sets.
- **Feature Engineering –** Apply feature selection (correlation, PCA) and handle class imbalance (SMOTE or weighting).
- **Model Design –** Implement three pipelines:
    - MLP for tabular features.
    - CNN (with transfer learning) for image data.
    - Hybrid CNN + LSTM/MLP to capture both spatial and sequential patterns.
- **Training & Optimization –** Use Adam optimizer, learning rate scheduling, dropout, and early stopping; perform hyperparameter tuning.
- **Evaluation –** Measure accuracy, AUC, precision, recall, F1-score; visualize with ROC/PR curves and confusion matrix.
- **Expected Outcome –** A robust, high-accuracy (90–98%) model supporting early breast cancer detection with interpretability.

Dataset

link:
https://drive.google.com/file/d/1HQqUCF23fihnE3aCbx2MNPmER9o_If_m/view

REFERENCES:

- R. S. Lee, F. Gimenez, A. Hoogi, and D. Rubin, “A curated mammography dataset for use in computer-aided detection and diagnosis research,” Scientific Data, vol. 4, 2017. [Online]. Available: CBIS-DDSM dataset
- H. T. Nguyen et al., “VinDr-Mammo: A large-scale benchmark dataset for full-field digital mammography,” Medical Image Analysis, vol. 87, 2023.
- M. Moreira, I. Amaral, A. Domingues, and P. Cardoso, “INbreast: Toward a full-field digital mammographic database,” Academic Radiology, vol. 19, no. 2, pp. 236–248, 2012.
- F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, “A dataset for breast cancer histopathological image classification,” IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1455–1462, 2016. (BreakHis dataset)
- M. Gadermayr and M. Tschuchnig, “Multiple instance learning for digital pathology: A review of the state-of-the-art and future perspectives,” Comput. Med. Imaging Graph., vol. 108, p. 102212, 2024.
- S.-C. Huang et al., “Self-supervised learning for medical image classification: A survey,” MedRxiv preprint, 2023.
- A. Yala, T. Schuster, J. Miles, C. Barzilay, and R. Lehman, “A deep learning model to triage screening mammograms: A simulation study,” The Lancet Digital Health, vol. 1, no. 7, pp. e307–e316, 2019.
- M. Shen, M. Wu, Z. Gao, and J. Zhu, “Breast cancer diagnosis from ultrasound images using deep learning techniques,” Frontiers in Oncology, vol. 11, p. 546064, 2021. (BUSI dataset reference)
- R. Chen, L. Ma, and X. Zhang, “Deep learning for breast cancer classification: A review,” Frontiers in Oncology, vol. 12, 2022.
- P. Sudharshan, C. Petitjean, F. Spanhol, L. Oliveira, L. Heutte, and P. Honeine, “Multiple instance learning for histopathological breast cancer image classification,” Expert Systems with Applications, vol. 117, pp. 103–111, 2019.

Thank you