DEVELOPING AN EXPLAINABLE DEEP LEARNING MODEL AND REPORTING SYSTEM FOR REAL-TIME CLINICALLY-GEARED SEGMENTATION TASKS

Affiliations

1. EMJMD Program, EDISS, Faculty of Science and Engineering, Information Technology, Abo Akademi University, 20500, Turku, Finland.
2. School of Engineering, University of Melbourne, Parkville, Victoria 3010, Australia.
3. Department of Information and Communication Technology, Technology, Policy and Management, Delft University of Technology, Delft, 2600GA, The Netherlands.
4. Sorbonne Université, Institut du Cerveau—Paris Brain Institute—ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013 Paris, France.

Authors

Saad Abdullah[1], Md Masum Billah[1], Pragati Manandhar[1], Victor Armando Canales-Lima[1], Lameya Islam[1], Alexis Gbeckor-Kove[1], Sarosh Krishan[1], Hergys Rexha[1], Sebastien Lafond[1], Kurt Benke[2], Sepinoud Azimi[1,3], Janan Arslan[4]

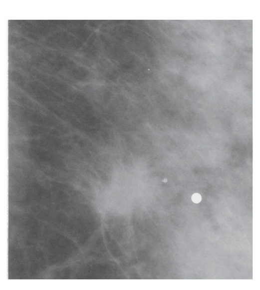
1 Introduction

- **Problem**: Breast cancer is a leading cause of mortality among women, and early detection via mammography improves outcomes.
- **Challenge**: Deep learning (DL) models lack explainability, making clinical adoption difficult.
- **Solution**: We propose explainable models that integrate Texture Analysis to enhance explainability in breast cancer segmentation and classification.

2 Objective

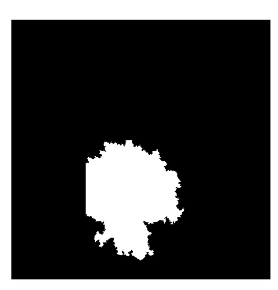
- Develop explainable DL models for breast cancer segmentation and classification.
- Improve multi-scale feature extraction using Laplacian Pyramids.
- Provide interpretable outputs to assist radiologists in decision-making.

3 Dataset



Sample image from CBIS-DDSM dataset

Preprocessed textured image (Levels & Edges)

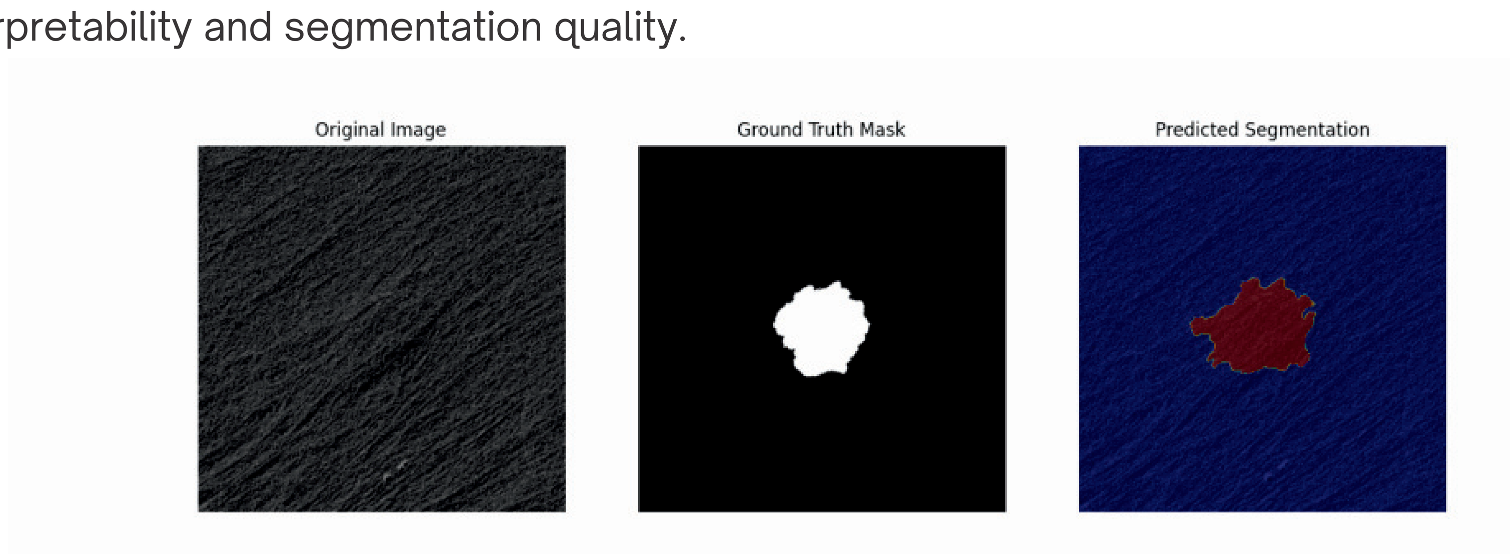


Ground truth mask

The CBIS-DDSM (Curated Breast Imaging Subset of DDSM) dataset is a widely used mammography dataset for breast cancer research. Ours consists of 723 images, carefully curated and annotated to facilitate AI-driven diagnosis and segmentation tasks.

5 Preliminary Results

Our initial results reaffirm the role of texture-based segmentation in enhancing AI explainability. The images demonstrate how texture-driven features guide the segmentation process. However, these results were obtained using Law's Texture Energy Measures (LTEM), primarily L5E5, without incorporating the newly proposed Laplacian Pyramid texture extraction. The next phase of our research aims to evaluate whether these additional multi-scale texture features can further enhance the model's interpretability and segmentation quality.

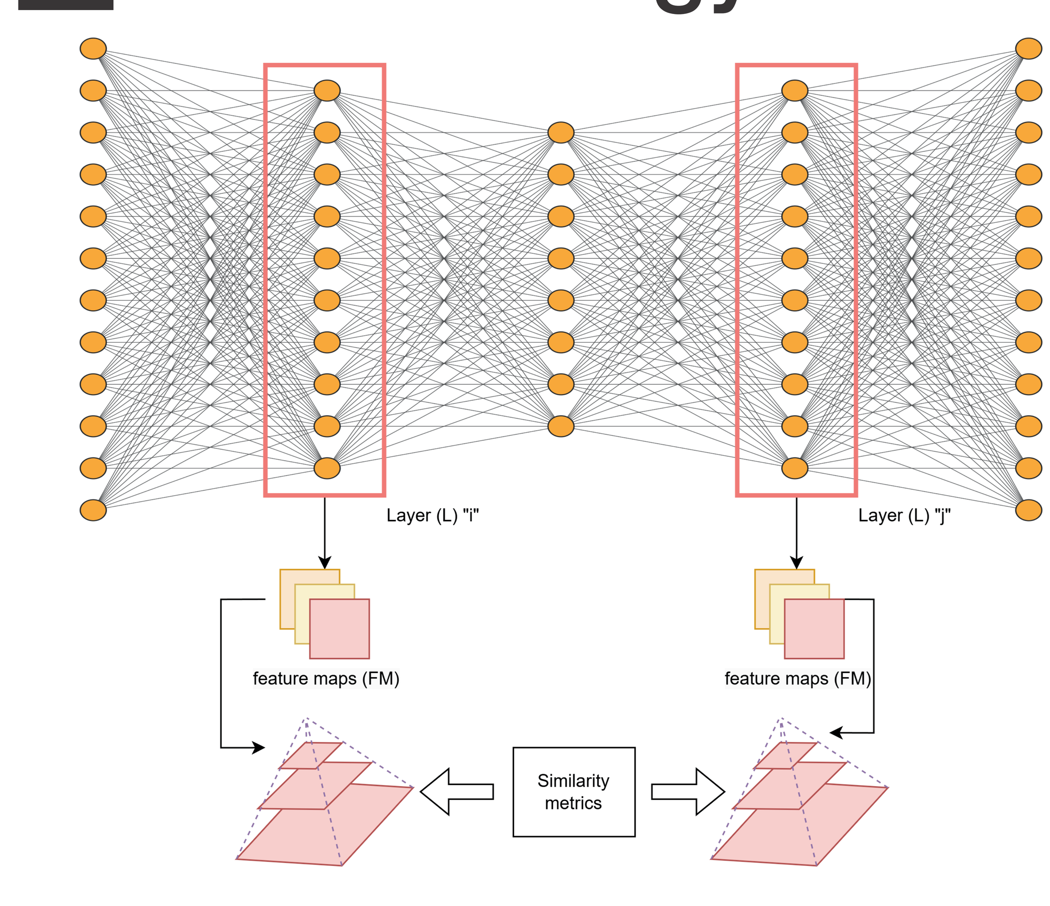


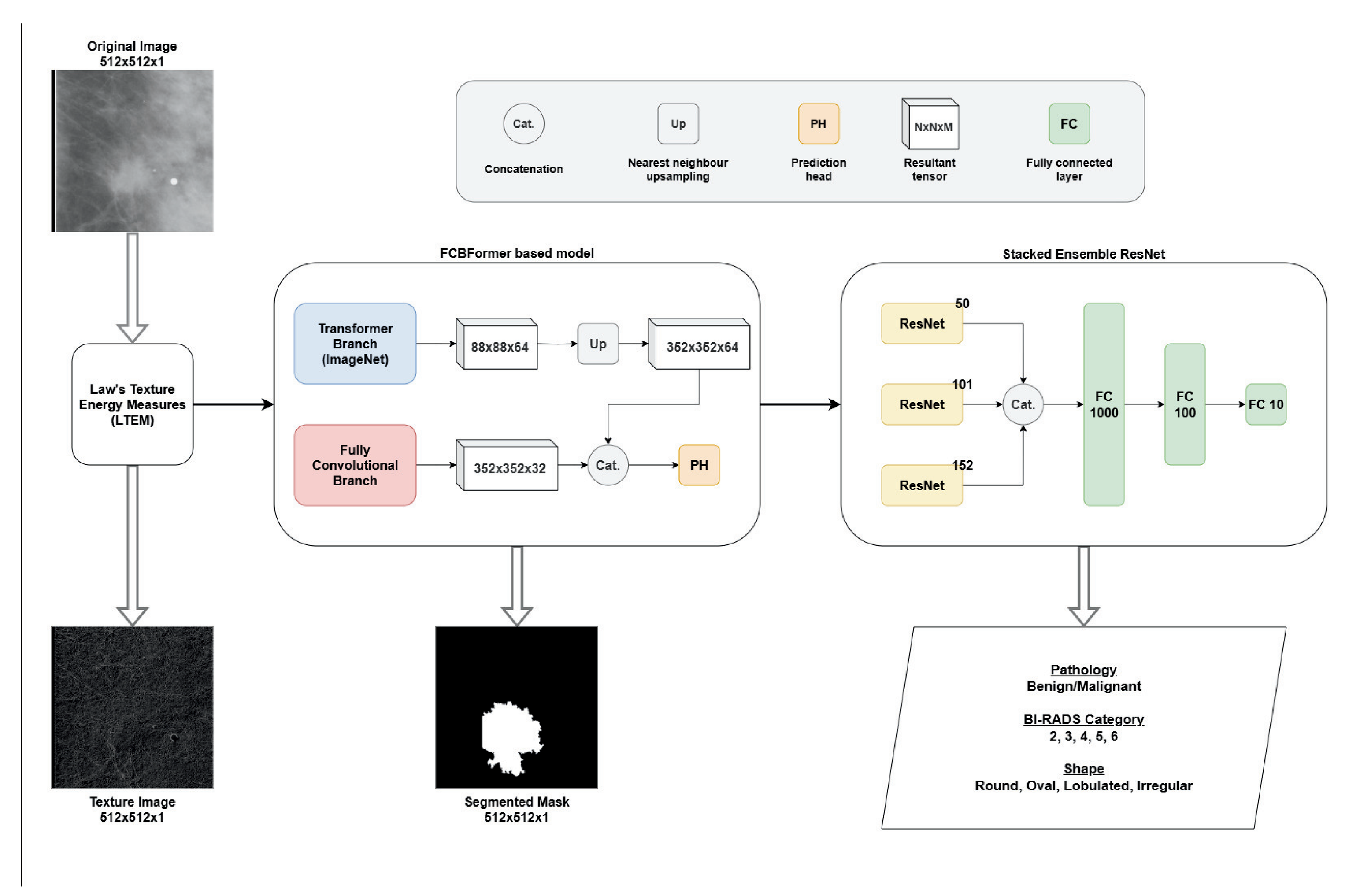
6 Analysis

Our segmentation model was designed to balance efficiency, accuracy, and explainability. The FCBFormer base model achieved 85% accuracy but required the longest training time (34 minutes). In contrast, ViT and Swin Transformer variants significantly reduced training time (22.24 and 25 minutes, respectively) but experienced a minor drop in accuracy (80%). These findings highlight a trade-off between computational efficiency and segmentation performance. Importantly, our approach ensures that segmentation not only achieves high accuracy but is also interpretable, helping radiologists understand why certain regions are highlighted by the model.

Model	Transformer	Training Time	Accuracy
FCBFormer	Base	34 minutes	85%
FCBFormer	VIT	22.24 minutes	80%
FCBFormer	Swin	25 minutes	80%

4 Methodology





The Laplacian pyramid extracts texture features by preserving high-frequency details crucial for distinguishing breast tissue patterns. These multi-scale texture representations enhance the interpretability of deep learning models by ensuring the most relevant diagnostic features are captured. The extracted features are processed through the FCBFormer-based model, where the Transformer Branch captures global contextual relationships, while the Fully Convolutional Branch (FCB) retains spatial details necessary for precise segmentation.

Once processed, the outputs from both branches are concatenated and segmented, producing feature-rich maps that emphasize key diagnostic regions. These segmented outputs are then passed to a stacked ensemble ResNet, which performs multi-task classification, predicting pathology (benign/malignant), BI-RADS category, and lesion shape (round, oval, lobulated, irregular). This combined approach ensures that the model is not only accurate but also explainable, reinforcing trust in AI-driven breast cancer diagnostics.

7 Future Work

Moving forward, we aim to refine our segmentation model by integrating both LTEM-based textures (L5E5) and newly extracted textures from the Laplacian Pyramid methodology. This combined feature set will allow us to assess whether multi-scale texture representations enhance AI-driven medical imaging. Additionally, we will extend our framework beyond segmentation by incorporating a stacked ensemble classification model that predicts pathology (benign/malignant), BI-RADS category, and lesion shape. This will provide a multi-modal, explainable diagnostic system, bridging segmentation and classification for improved clinical relevance.

8 Conclusion

This study demonstrates the potential of texture-based analysis in enhancing the explainability of deep learning models for breast cancer segmentation and classification. By integrating LTEM textures with a hybrid FCBFormer model, we established a baseline for explainability, ensuring that texture-aware features align with clinical insights. As we move forward, our focus shifts toward incorporating additional texture representations from the Laplacian Pyramid method and developing a comprehensive classification framework. These advancements will further enhance model interpretability and reliability, making AI-assisted cancer diagnosis more transparent, trustworthy, and clinically adaptable.

Åbo Akademi
TU Delft — Delft University of Technology
EDISS MASTER'S PROGRAMME
ICM Institut du Cerveau
THE UNIVERSITY OF MELBOURNE