Explainable AI in Computer Vision via Learned Textured Kernels and Energy-Based Reports

Manuel Padilla, Ossama Essfadi, Rajshree Rai

1. Introduction

Breast cancer remains a leading cause of cancer mortality worldwide. While Deep Learning (DL) models, like FCBFormer, have achieved state-of-the-art accuracy in lesion segmentation, they suffer from the **Black Box** problem. Radiologists cannot blindly trust a *Malignancy Probability* score without understanding the morphological basis of the decision. Saliency maps (heatmaps) are often noisy and fail to correlate with clinical BI-RADS descriptors.

2. Objectives

Novel Kernel Discovery: Systematically identify and integrate high-performance convolution kernels into the FCBFormer architecture to maximize segmentation precision.

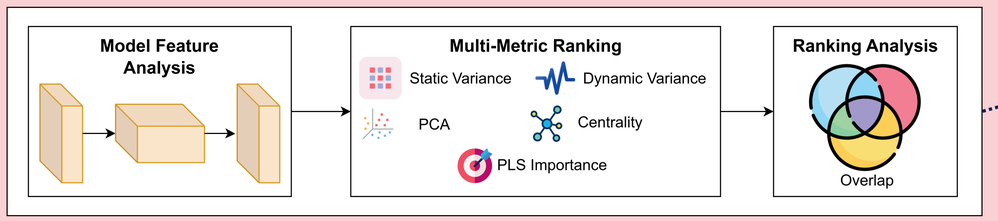
Automated Texture Reporting: Verbalize feature maps by mapping specific kernel activations to standardized BI-RADS descriptors (e.g., Spiculated), bridging the gap between pixel-level AI and clinical diagnosis.

3. Dataset

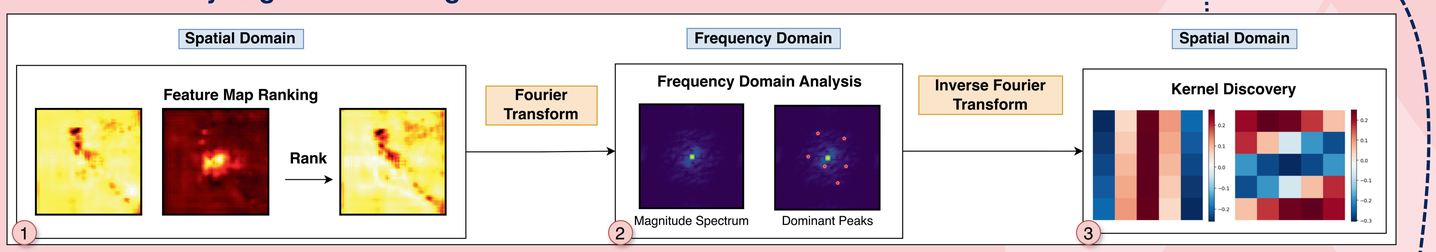
We leverage the gold-standard CBIS-DDSM mammography database. Our study utilizes a specific cohort of 723 meticulously annotated images, curated expressly to train AI models for precise lesion segmentation and diagnosis.

4. Methodology

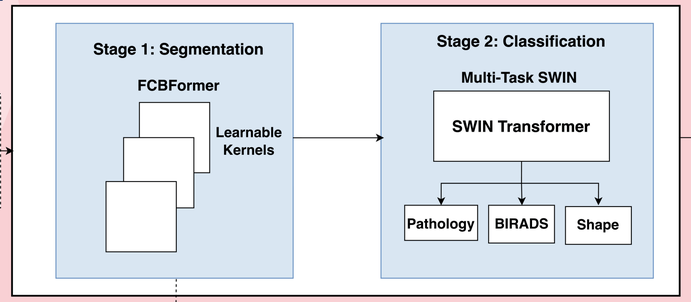
Kernel Discovery: Kernel Ranking

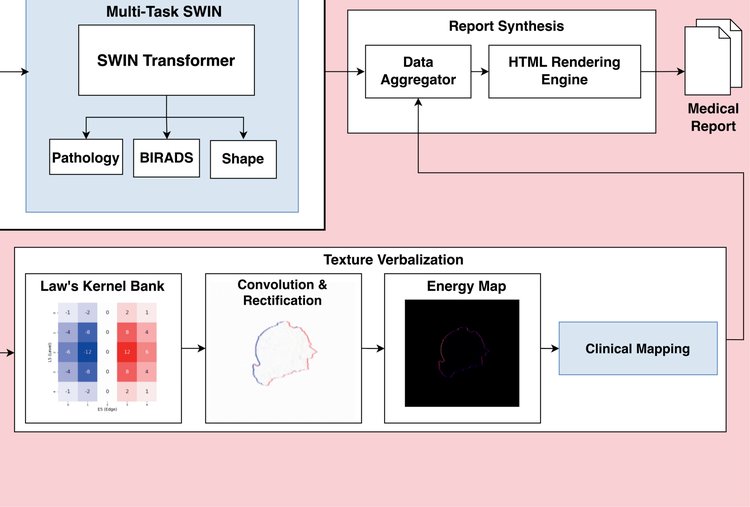


Kernel Discovery: Signal Processing



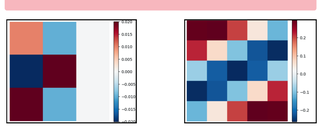
Texture Verbalization Pipeline





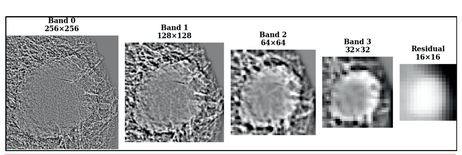
5. Preliminary Results

Discovered Kernels



Segmentation Results on FCBFormer

Approach	IoU
Kernel Ranking	0.855
Signal Processing	0.854
Original Baseline	0.842



Laplacian decomposition of an ROI patch

6. Conclusion

Accuracy Preservation: Our results demonstrate that restricting the model to a discovered set of texture-enhanced kernels maintains segmentation accuracy equivalent to the baseline. This confirms that explicit texture-guidance does not sacrifice predictive performance while significantly reducing model opacity.

Energy-Based Verbalization: By leveraging the mathematical properties of Law's Texture Energy Measures (LTEM), we established a descriptive bridge between handcrafted filters and learned kernels. This allows us to "verbalize" the model's internal logic, mapping high-level activations to clinically understood traits.

7. Future Work

- **Kernel Grafting**
- **Clincal Validation**
- **Adaptive Description**

EDISS MASTER'S PROGRAMME · ICM Institut du Cerveau · TU Delft · The University of Melbourne