**First Review**

**On**

**Early Detection and Classification of Breast Cancer from Mammograms**

**Submitted by:**

**Laxmi Sai. G (21MIS7119)**

**Vasudha Rani. P (21MIS7121)**

**To**

**Dr. B.V. Gokulnath**

**School Of Computer Science And Engineering (SCOPE)**



Beside AP Secretariat, Near Vijayawada, Amaravati, Andhra Pradesh, India, 522237

**APPLICATION ORIENTED INTRODUCTION:**

Breast Cancer is most occurring feature of cancer in women. It is a disease in which abnormal breast cells grow out of control and form tumours. If left unchecked, the tumours can spread throughout the body and become fatal. There are many ways of identifying the breast cancer. Doctors use a mammogram technique where it is an x-ray imaging method used to examine the breast for the early detection of cancer and other breast diseases. These mammogram images are classified into two types Benign and Malignant, if any image doesn’t belong to any of these types it is a normal case. Benign tumours are usually well-defined and round or oval in shape. Malignant tumours are usually poorly defined and irregular with lobules. Now, with these images we are going to train our model on these binary classifications of images. For this we are using CLAHE images from Kaggle. To address this issue we are using CNN and ViT Transformer.

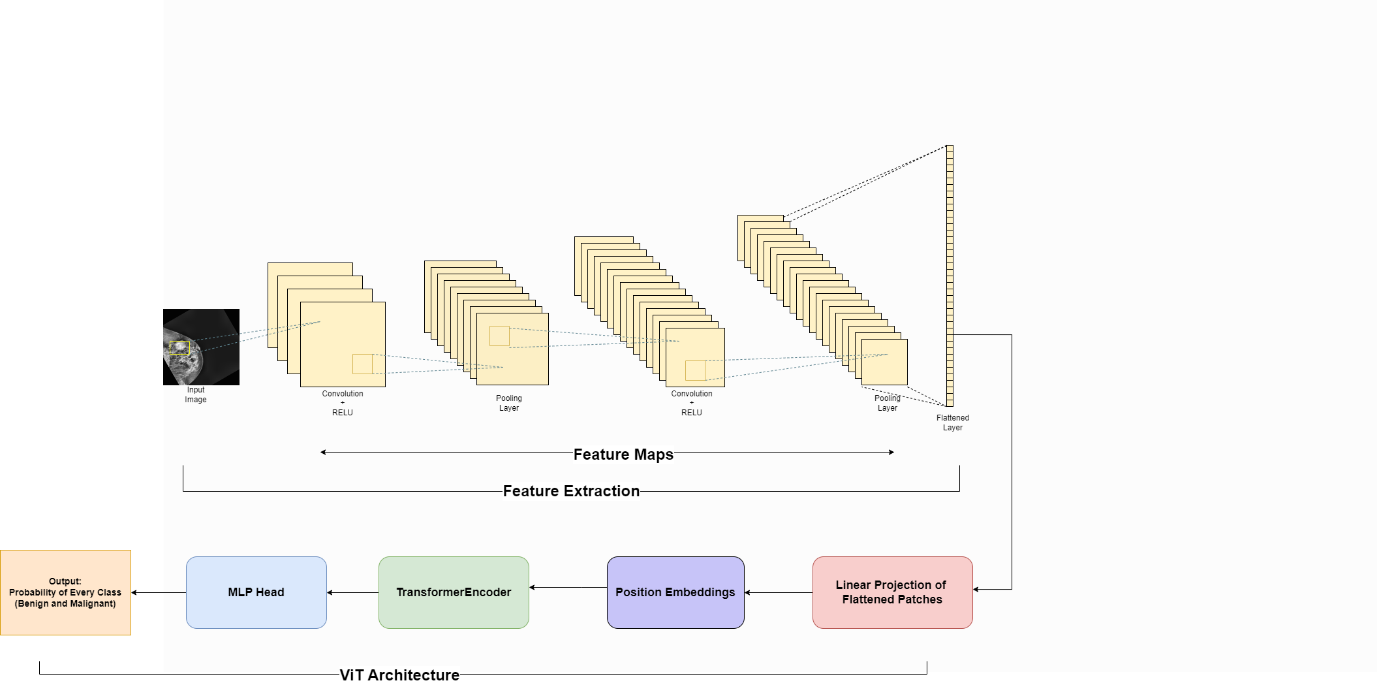
Preprocessing is a critical step in both CNN and ViT methodologies. For CNNs, preprocessing typically involves resizing images to a standard dimension, normalizing pixel values, and augmenting the dataset through techniques like rotation, flipping, and zooming to increase the diversity of training samples and reduce overfitting. For ViTs, preprocessing includes splitting the image into patches, normalizing the patches, and incorporating positional encodings to maintain spatial context.

CNNs have shown remarkable success in image classification tasks due to their ability to automatically learn hierarchical feature representations from raw images. The CNN architecture typically begins with pre-processing mammogram images, resizing them to consistent dimensions, such as 224×224×3224 \times 224 \times 3224×224×3 for RGB images. This pre-processing step ensures uniformity and compatibility with the network’s input layer. The core of the CNN lies in its convolutional layers, where a set of filters is applied to the input image to extract various features like edges, textures, and patterns. Each filter convolves around the input image, generating feature maps that highlight different aspects of the image. The activation function, usually ReLU (Rectified Linear Unit), is applied to introduce non-linearity into the model, allowing it to learn more complex patterns. Following the convolutional layers, pooling layers are introduced. These layers, such as max pooling or average pooling, reduce the spatial dimensions of the feature maps, thereby decreasing the computational load and helping to control overfitting. Pooling retains the most critical information by down sampling the feature maps, which makes the network more efficient and robust. After several convolution and pooling operations, the high-level reasoning in the neural network is performed via fully connected layers. The feature maps from the last pooling layer are flattened into a single vector, which is then fed into fully connected layers. These layers function as a classifier, transforming the learned feature representations into output probabilities. The final layer is typically a SoftMax layer that outputs the probability of each class (benign, malignant, normal), with the highest probability determining the predicted class of the mammogram.

ViTs offer a novel approach to image classification by leveraging the transformer architecture, which has been highly successful in natural language processing. The ViT methodology begins with the input mammogram image being divided into fixed-size patches, for example, 16×1616 \times 1616×16. Each patch is flattened into a vector, effectively transforming the image into a sequence of patches. Each flattened patch is then linearly embedded into a fixed-dimensional vector. This step converts the sequence of patches into a format suitable for the transformer model. To retain spatial information about the patches, positional encodings are added to the patch embeddings. This helps the model understand the relative positions of the patches within the original image. The core of the ViT architecture consists of transformer encoder layers, which include multi-head self-attention mechanisms and feed-forward neural networks. Self-attention enables the model to focus on different parts of the image and capture relationships between distant patches. A special classification token is added to the sequence of patch embeddings, and the output corresponding to this token from the final transformer encoder layer is used for classification purposes. The classification token's output is passed through a multi-layer perceptron (MLP) head, which produces the final probability scores for each class. Like in CNNs, the SoftMax activation is applied to determine the predicted class.

Both CNNs and ViTs rely heavily on effective feature extraction to improve classification accuracy. In CNNs, feature extraction is inherently achieved through the hierarchical application of convolutional and pooling layers, which progressively capture more abstract representations of the input image. ViTs, on the other hand, leverage self-attention mechanisms within transformer encoders to capture both local and global features from image patches, providing a comprehensive understanding of the image content.

In practical applications, the outputs from both CNN and ViT models can be combined to leverage the strengths of both architectures. This can be done through techniques like ensemble learning, where the predictions from multiple models are aggregated to improve the overall accuracy and robustness of the system. The integration of CNNs and ViTs in breast cancer detection using mammograms represents a significant advancement in medical imaging. These architectures not only enhance diagnostic accuracy but also provide a scalable and efficient solution for early detection, potentially saving many lives through timely intervention.

**Architectural Diagram:**

**OBJECTIVES:**

 **Feature Extraction**: Utilize CNNs to automatically extract relevant features such as textures, edges, and patterns from mammogram images, which are crucial for distinguishing between benign and malignant tissues.

 **Global Context Understanding**: Leverage ViT to capture global spatial relationships and context within mammogram images through self-attention mechanisms, which can provide a holistic understanding of the entire image rather than just local features.

 **Model Comparison**: Compare the performance of CNN-based approaches with ViT-based approaches in terms of classification accuracy, computational efficiency, and interpretability in breast cancer detection tasks.

 **Integration of Modalities**: Explore methods to integrate outputs from both CNN and ViT models to potentially improve overall classification performance by leveraging complementary strengths of both architectures.

 **Clinical Relevance**: Assess the clinical relevance and applicability of CNN and ViT models in real-world scenarios by evaluating their performance metrics against established clinical standards and guidelines for breast cancer detection.

 **Visualization and Interpretation**: Visualize intermediate representations and decision-making processes within both CNN and ViT architectures to gain insights into how these models interpret mammogram images and make classification decisions.

 **Scalability and Generalization**: Evaluate the scalability and generalization capabilities of CNN and ViT models across different datasets and imaging conditions, ensuring robust performance in varied clinical settings.

 **Ethical Considerations**: Consider ethical implications related to the deployment of AI models in healthcare, including issues of bias, fairness, privacy, and interpretability, particularly in sensitive applications like breast cancer diagnosis.

 **Educational Outreach**: Use the architectural diagram and model explanations to educate healthcare professionals and researchers about the capabilities and limitations of CNNs and ViTs in medical imaging tasks, fostering broader understanding and adoption in clinical practice.

**PROBLEM STATEMENT:**

Breast cancer is a prevalent and potentially life-threatening condition affecting millions of individuals worldwide. Early and accurate detection of breast abnormalities from mammogram images is crucial for timely intervention and improved patient outcomes. Traditional methods rely heavily on manual interpretation by radiologists, leading to variations in diagnosis and potential oversight of subtle or early-stage abnormalities.

**KEY PROBLEMS:**

**Feature Extraction Complexity**: Mammogram images contain intricate textures, subtle patterns, and varying densities that pose challenges for accurate feature extraction using conventional methods. CNNs and ViTs offer automated feature extraction capabilities, but their effectiveness in capturing relevant features from mammograms remains a critical challenge.

**Data Heterogeneity and Size**: Limited availability of annotated mammogram datasets with diverse patient demographics, breast types, and imaging conditions complicates training robust CNN and ViT models. Ensuring generalizability across diverse populations and imaging practices is essential for clinical applicability.

**Interpretability and Explain ability**: Despite their high accuracy, CNNs and ViTs are often perceived as black-box models, making it challenging to interpret how they arrive at diagnostic decisions. Ensuring transparency and explain ability in model predictions is crucial for gaining clinician trust and acceptance.

**Computational Efficiency**: Both CNNs and ViTs require substantial computational resources, especially when processing high-resolution mammogram images. Optimizing model architecture, training strategies, and inference techniques to achieve real-time performance without compromising accuracy is a pressing concern.

**Integration with Clinical Workflow**: Integrating AI-based breast cancer detection models into existing clinical workflows poses logistical challenges, including interoperability with healthcare systems, regulatory compliance, and ensuring seamless collaboration between AI algorithms and healthcare professionals.

**Bias and Generalization**: Addressing inherent biases in datasets and models is critical to ensure fair and equitable breast cancer diagnosis across diverse patient populations. Achieving robust generalization to unseen data and clinical settings remains a significant hurdle for CNNs and ViTs in medical imaging applications.

**DETAILED LITERATURE REVIEW**

1. The BC2NetRF framework presents a novel automated approach for classifying breast cancer from mammogram images. It integrates advanced techniques including the HRLG Contrast Enhancement Technique for image preprocessing and employs a fine-tuned EfficientNet-b0 model known for its effective performance in image classification tasks. A key innovation lies in its use of the Equilibrium-Jaya Controlled Regula Falsi algorithm for feature selection, ensuring optimal feature relevance for classification accuracy. Across datasets like CBIS-DDSM and INbreast, BC2NetRF achieves outstanding average accuracies of 95.4% and 99.7%, respectively. These results underscore its efficacy in accurately distinguishing between benign and malignant breast tumors, surpassing traditional methods and showcasing its potential for enhancing clinical diagnostic tools.

2. This overview delves into leveraging deep learning techniques to enhance the accuracy and efficiency of mammography for breast cancer screening. It highlights the application of DL models like U-Net and Mask R-CNN, specifically for segmenting breast lesions with greater precision. These models utilize sophisticated neural network architectures that excel in identifying and delineating areas of interest in mammograms, thereby aiding in the classification of tumors as benign or malignant. Research indicates that DL models achieve an impressive classification accuracy of 0.88, surpassing the typical performance of radiologists, who achieve around 0.83 accuracy in similar tasks. However, challenges persist. DL models require substantial amounts of annotated data for training, which can be resource-intensive and time-consuming to acquire. Additionally, the interpretability of DL model outputs remains a concern, as understanding the rationale behind their classifications can be complex. Furthermore, integrating these advanced techniques into clinical practice necessitates rigorous validation and adaptation to ensure reliable performance across diverse patient demographics and imaging conditions.

3. The paper titled "Breast Tumor Classification with Enhanced Transfer Learning Features and Selection Using Chaotic Map-Based Optimization" aims to develop an automated computer-aided system for accurate breast cancer classification from mammograms. The methodology involves training the EfficientNet-B4 architecture on both original and enhanced mammogram sets, utilizing feature fusion and optimization via a chaotic-crow-search algorithm. The model achieved classification accuracies of up to 98.459% on the INbreast database and 96.175% on the CBIS-DDSM dataset. However, the use of complex optimization algorithms may increase computational requirements and pose challenges in practical deployment.

4. The paper titled "A YOLO-Based Model for Breast Cancer Detection in Mammograms" aims to implement an automated data-driven model to support physicians' decision-making within breast cancer screening or detection programs. The methodology combines YOLO for detection, Eigen-CAM saliency maps for interpretability, and YOLO predictions to offer an explainable decision support system, enhancing diagnostic accuracy. The small YOLOv5 model resulted in the best-developed solution, achieving a mean Average Precision (mAP) of 0.621 on a proprietary dataset. However, the reliance on a proprietary dataset may limit the model's generalizability and applicability to other datasets.

5. The paper titled "Breast Cancer Diagnosis Using YOLO-Based Multiscale Parallel CNN and Flattened Threshold Swish" proposes a Breast Cancer Diagnosis (BCD) framework using a YOLO-based multiscale parallel Convolutional Neural Network (CNN) architecture. The framework employs the YOLO architecture for end-to-end Computer-Aided Diagnosis (CAD) of breast cancer, featuring specialized CNN components like Parallel Feature Extraction Swish (PFES), Dilated Convolutional Blocks (DCB), and Inception Blocks (IB) for efficient feature extraction. The model achieved an accuracy of 98.72% for breast cancer classification and a mean Average Precision (mAP) of 91.15% for breast cancer detection. However, the complexity of the multiscale parallel CNN architecture may require significant computational resources and may be challenging to implement in real-time clinical settings.

6.The paper titled "YOLO-based CAD framework with ViT transformer for breast mass detection and classification in CESM and FFDM images" presents a fully automated computer-aided diagnosis (CAD) framework. This framework integrates YOLOv4 for detecting masses in Contrast Enhanced Spectral Mammography (CESM) and Full Field Digital Mammography (FFDM) images, and a Vision Transformer (ViT) for classification. It pioneers automation in CESM mass detection and classification, particularly on the CDD-CESM dataset. The system achieved impressive classification accuracies of 95.65% on the INbreast dataset and 97.61% on CESM images. However, the reliance on advanced models like YOLOv4 and ViT may lead to high computational costs.

7. The paper titled "Using Hybrid Pre-trained Models for Breast Cancer Detection" addresses the inefficiencies and errors associated with manual breast histopathology image analysis, emphasizing the need for more efficient and accurate detection methods in the medical field. This study proposes a hybrid deep learning model that combines Convolutional Neural Networks (CNN) with EfficientNetV2B3. The proposed model demonstrates outstanding performance, achieving an accuracy of 96.3%. However, the use of hybrid models may lead to increased computational requirements and complexity, potentially limiting their application in resource-constrained environments.

8. The paper titled "Detection and Diagnosis of Small Target Breast Masses Based on Convolutional Neural Networks" aims to address the challenge of high missed and misdiagnosis rates for small breast masses in early breast cancer screening. The proposed method employs the Residual asymmetric dilated convolution-Cross layer attention-Mean standard deviation adaptive selection-You Only Look Once (RCM-YOLO) network for detection. The model achieved a mean Average Precision (mAP) of 90.34, reduced the missed detection rate for small masses to 11%, and decreased the single detection time to 28 ms, demonstrating enhanced detection accuracy and speed. However, the complexity of the RCM-YOLO network may pose implementation challenges and require significant computational resources.

9. The paper titled "Deep Transfer Learning with Fuzzy Ensemble Approach for the Early Detection of Breast Cancer" underscores the importance of early breast cancer detection, as many women do not experience symptoms at the initial stages. The study employs three deep Convolutional Neural Network (CNN) models—VGG-11, Inception v3, and ResNet50—integrated with a modified Gompertz function to build a fuzzy ranking of the base classification models. The proposed ResNet50 ensemble network, utilizing the modified Gompertz function-based fuzzy ranking approach, achieved a superior classification accuracy of 98.986%. However, the complexity of integrating multiple CNN models and the fuzzy ranking system may increase the computational load and implementation complexity.

10. The paper titled "Breast Cancer Detection Using Ensemble of Convolutional Neural Networks" aims to address the challenge of timely breast cancer detection to reduce mortality rates by leveraging deep learning techniques and image processing. The research implements ensemble learning by combining individual Convolutional Neural Network (CNN) models, which were selected using the Hungarian optimization algorithm. This approach resulted in improved performance, achieving an accuracy of 95.7%. However, the ensemble method's complexity and computational demands may pose challenges for practical implementation, particularly in clinical settings with limited resources.