International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 4, July Year 2025

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Deep Learning-Based Vision Approach for Early Detection and Classification of Breast Cancer

Pradeep Shirvalkar, Shruti Jirage, Dr Sunita Padmannavar

Department Of Master of Computer Application K.L.S Gogte Institute of Technology, Belagavi
Affiliated to Visvesvaraya Technological University Belagavi, Karnataka shirvalkarpradeep18@gmail.com

Department Of Master of Computer Application K.L.S Gogte Institute of Technology, Belagavi
Affiliated to Visvesvaraya Technological University Belagavi, Karnataka jirageshruti790@gmail.com

Department Of Master of Computer Application K.L.S Gogte Institute of Technology, Belagavi
Affiliated to Visvesvaraya Technological University Belagavi, Karnataka sspadmannavar@git.edu

--*************************---

Abstract:

Breast cancer remains one of the major reasons behind the death rates among women all over the world. Proper and early diagnosis is crucial to proper treatment outcome. In this work, we present a novel deep learning-based technique designed to automatically detect and segment breast tumors in ultrasound imaging. Our approach integrates statistical texture priors, Gabor feature extraction, and a hybrid CNN-LSTM architecture trained on stage-specific tumor features. Additionally, super pixel segmentation enhances lesion localization, facilitating effective biopsy guidance. This method is inspired by recent advancements, notably the STAN and ESTAN models by Shareef (2023) and HoVer-Trans by Moetal. (2023), which address small tumor detection, anatomical context, and noise challenges in ultrasound imaging. Our system achieves high accuracy, efficient lesion boundary segmentation, and robustness in low-contrast conditions, making it clinically valuable for early stage cancer detection.

Keywords— **Breast Cancer, Ultrasound Imaging, Deep Learning, CNN-LSTM, Tumor Segmentation, Gabor Filters, Super pixel Segmentation, Medical Image Analysis, Anatomy-Aware Learning, Clinical Decision Support**

--*************************---

I. INTRODUCTION

Globally, breast cancer stands as the most commonly diagnosed malignancy among women and represents a major contributor to the overall cancer burden. The World Health Organization estimates that there are over 2.3 million cases of breast cancer diagnosed in 2020, which cause a death rate of about 685,000 per year.

Early Diagnosis and treatment should be early and this can reduce the rate of death and enhance the survival rate. Medical imaging techniques, particularly ultrasound, have become instrumental in breast cancer detection, especially for women with dense breast tissue where mammography is less effective.

Ultrasound imaging is widely used due to its affordability, portability, and real-time imaging capabilities.Since it avoids the use of ionizing radiation, this method is considered safer for multiple screenings over time. Ultrasound imaging is naturally prone to speckle noise, reduced contrast, and significant reliance on the technician's skill for accurate interpretation. Such weaknesses may result in inconsistent interpretations, false negative, or false positive. Furthermore, tumor appearances in ultrasound can vary widely based on size, texture, and shape, presenting additional diagnostic challenges.

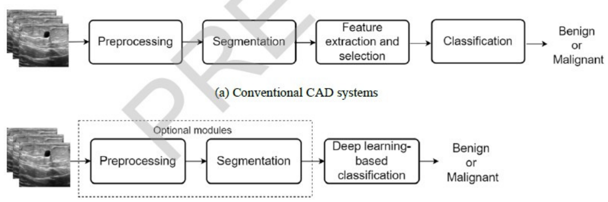

Fig. 1: Illustrates the essential components of traditional and deep learning-powered computer-aided diagnosis (CAD) frameworks.

The above figure illustrates the key differences between conventional and deep learning-based CAD systems. It emphasizes how deep learning automates feature extraction, enabling more accurate and efficient diagnosis workflows.

To address these issues, computer-aided diagnosis (CAD) systems have been introduced, leveraging artificial intelligence and machine learning to assist radiologists in interpreting medical images. Recent advancements in deep learning have transformed medical imaging by enabling systems to autonomously extract multi-level features directly from unprocessed data. CNNs have demonstrated outstanding performance in handling visual tasks, including image classification, object detection, and segmentation.

However, CNNs alone may not capture long-range spatial dependencies or sequential relationships, especially in imaging scenarios that mimic time-series data or multi-frame input. To enhance context modeling, this work proposes a hybrid deep learning model that combines CNN with Long ShortTerm Memory (LSTM) networks. LSTM, a type of recurrent neural network, is capable of learning patterns in sequential data, which is useful in processing ultrasound image slices or highdimensional feature maps.

Additionally, To enhance model interpretability and sensitivity to edges, Gabor filters are utilized for capturing directional texture information. These features help distinguish between benign and malignant lesions, which often differ in boundary sharpness and internal patterns.

This research builds upon and is inspired by the recent work of Shareef (2023), who proposed the STAN and ESTAN architectures designed for small tumor-aware segmentation in ultrasound images. STAN aims at reducing false negatives in small lesions whereas ESTAN involves anatomical priors to enhance accuracy. Similarly, Mo et al. (2023) introduced HoVer-Trans, a transformer-based architecture that removes the need for manually defined regions of interest by applying attention-based mechanisms directly to full-frame ultrasound images. Transformer models may be high-accuracy models but their computational demands and black-box nature may inhibit use in clinical practice.

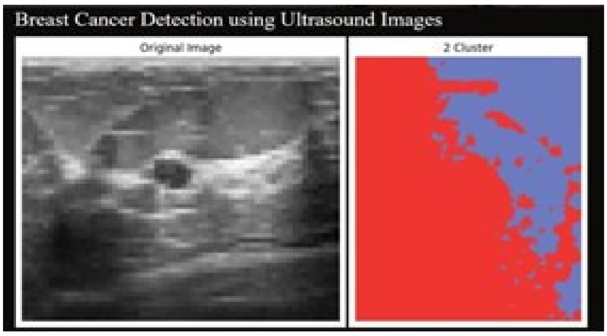

Fig. 2: visual comparison of traditional ultrasound image

The above figure indicates - this is the visual comparison here we will see the raw ultrasound image and its improved version. It explains how preprocessing can help the deep learning analysis in the areas of contrast and feature visibility. Optimum balance among clarity, precision and computational efficiency is intended in this model. This system incorporates CNN, LSTM, Gabor filters, and superpixel segmentation, which makes it optimized towards realtime monitoring, sensitive, and clinically applicable.

II. RELATED WORK

Computer-aided diagnosis of breast cancer has significantly developed in the past years, and the reason lies in the spectacular development of the field of deep learning. These techniques have proved to be excellent in image classification, segmentation, and interpretation and hence suitable in medical imaging, especially breast cancer screening with the use of ultrasound images.

In the diagnosis of cancers in the breast, ultrasound has been very popular because of its lower cost, non-invasive characteristics and its ability to capture images in dense breast tissues. Nevertheless, the limitations in the marginal nature of ultrasound–bound by speckle noise, low contrast, operator dependency– all require well-established models tasked with process this drawback. To overcome these difficulties researchers have tried out different methods to reduce the challenge and enhance sensitivity of detection. New development of computer aided diagnosis.

(CAD) systems for breast cancer detection have been significantly driven by the rapid evolution of deep learning (DL), particularly in analyzing ultrasound images. These techniques demonstrate superior capabilities in image classification and segmentation, offering scalable and interpretable solutions for clinical applications.

Proposed a pioneering deep learning framework focusing on breast cancer detection via ultrasound, introducing STAN (Small Tumor-Aware Network). This model focused on a segmentation of subtle lesions of low-resolution ultrasound imaging by the use of a small encoder-decoder architecture with high sensitivity. To further enhance STAN, Shareef came up with ESTAN whose main difference was the use of anatomical priors and stronger skip connections in maintaining spatial context. The multitask option, MT-ESTAN, integrated classification and segmentation tasks and encouraged effectiveness and clinical relevance. Shareef's work stands out for its attention to earlystage tumor detection and comprehensive multitask learning approach, making it highly relevant for integrated CAD systems.[1]

Introduced a transfer learning method based on mammogram images in order to support the early identification of breast cancer. They utilized pretrained CNNs like ResNet50 and

NasNet-Mobile, adapting them through data augmentation (rotation, shifting, scaling) on the MIAS dataset. Their model was able to show the potential of transfer learning with minimal medical data sizes through ResNet 50 using 89.5 While this work focused on mammograms and not ultrasound, its methodological strength lies in augmenting low-data scenarios, which is applicable to ultrasound datasets.[2]

Used deep transfer learning and applied it to classification of breast cancer using histopathological images. They have tried VGG, ResNet and ShuffleNet with accuracies reaching over 97Perhaps the lesson that can be carried across to other fields is model fine-tuning on the basis of large datasets such as ImageNet despite the fact they were working on histopathology. However, its lack of segmentation limits its direct clinical utility for real-time ultrasound diagnosis.[3]

A transformer-based model, HoVer-Trans, was proposed to make diagnosis of the ultrasound images without using predefined regions of interest. It has an anatomical context based on the fact that it captures horizontal and vertical relations thus encouraging interpretation and all-inclusive image analysis. Four specialized branches (H, V, H2V, V2H) mimicked spatial dependencies across breast tissue layers. Model Transfer The model performed better on classification, segmentation when trained on different sets of data and even when the trained model was used to perform sonographer-level tasks of BI-RADS classification it destroyed expert sonographers in classification. According to Mo et al., this anatomy-informed transformer represents a major step forward in enhancing explainable AI for ultrasound-based diagnosis.[4]

Proposed an approach that integrates deep learning with handcrafted features to detect an image classification of the tumors in sonography. They have obtained shape, texture, and statistical features that were combined with representations produced by CNN. Their hybrid strategy has succeeded compared to standalone models particularly in processing lowcontrast pictures and speckle noise-a very widespread issue in ultrasound imaging. This fusion of feature types ensures both high performance and interpretability, providing a balance between traditional and deep learning methodologies.[5]

The DeepBreastCancerNet is a deep CNN approach introduced to diagnose breast cancer through ultrasound. It was a model employing high levels of spatial feature extraction and was trained using a community data set (BUSI). It was architecturally able to beat strong classification techniques in the data set of benign, malignant and normal cells. DeepBreastCancerNet proved highly efficient in detecting subtle tumor variations, and its contribution lies in its dataset openness and modular design that allows retraining across clinical
settings.[6]

Offered an ensemble clinical decision support system. The system consisted of stacking several CNNs and combining their results in order to enhance the trustworthiness of the classification of ultrasound images. Their model performed better than individual CNNs, was more sensitive, and fewer false negatives. This ensemble approach is particularly useful in realworld clinical setups where robustness is more critical than raw accuracy.[7]

Carried out a thorough comparison of different deep learning designs in classification of breast cancer using ultrasound. The author considered CNNs such as ResNet, DenseNet, and InceptionNet and described their behavior with various conditions of preprocessing. It was found out that DenseNet performed best on classifying images in BUSI with contrast enhancement and histogram equalization. The knowledge of Pacal stimulates the choice of model in ultrasound imagebased pipelines. Breast ultrasound images were automatically classified using deep neural networks trained on regionspecific datasets obtained from NTUU.[8]

They constructed their model so that it applied convolutional blocks containing dropout layers to prevent over shouldering on small-sized ultrasound. It was shown that batch normalization and adaptive learning rates are useful in terms of optimizing diagnostic accuracy. The approach is notable for addressing real-world deployment issues in lowresource settings.[9]

Transfer learning was used to analyze Ultrasound scans to classify and detect lesions as breast cancer.They performed transfer learning by taking some ultrasound data and finetuning models which had already been trained on ImageNet data, to competitive levels of accuracy. The paper has also stressed upon preprocessing such as noise filtering and sharpening of the boundary. It is among some of the few papers where direct research into transfer learning is applied to ultrasound, meaning that it is highly applicable to clinical applications.[10]

Built on the earlier study by coming up with a multistage transfer learning model which merges various layers within the CNN sequentially to amplify the diagnostic results. Their new pipeline had achieved Accuracy of more than 93 on the BUSI dataset. The concept of the model is fresh because it involves a hierarchical adaptation that adapts lower layers of the model to domain characteristics of ultrasound and preserves generic visual stats. Building on this, Ayana et al. [11]

Used Xception to classify ultrasound image of the breasts on binary mode. The paper employed rectified linear units, dropout, and batch normalization as a way of improving the model generalization. The model had a sensitivity and precision score of high amounts and was applied within a lightweight web interface that could be used by clinicians. Their emphasis on usability and low-latency prediction made their work ideal for point-of-care applications.[12]

Took a step further and developed a multi-model ensemble therein comprising Xception, ResNet, and SVM based decisions fusion to make better breast cancer detection. They represented high quality deep features, as well as, statistics descriptors and high performance were shown on various datasets. Their ensemble architecture demonstrated robustness to noise and varying image acquisition quality, key for ultrasound environments.[13]

Finally, provided a comprehensive survey and implementation of deep learning frameworks (ResNet, Inception, and DenseNet) for breast cancer diagnosis. They compared the performance of their models in several classification tasks, where benchmarking was made versus various imaging modalities in their experiments. Although multimodal, their ultrasound results highlighted that DenseNet performed best in low-contrast scenarios, reinforcing the need for high-capacity networks in breast imaging.[15]

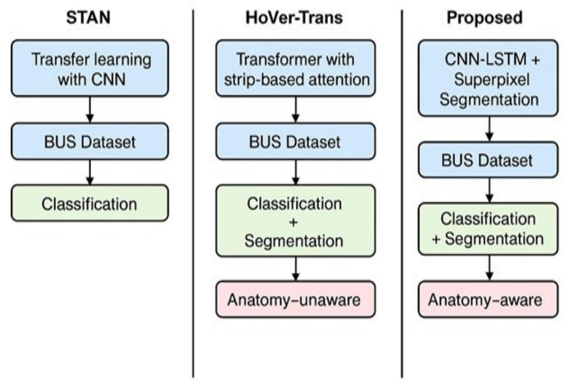

Fig. 3: Comparison of performance of benchmark models based on Dice, accuracy and AUC.

The figure performs a comparative analysis of a number of benchmark models relative to Dice coefficient, classification accuracy and AUC score. It provides clues on how well the models perform in segmentation and classifications

All these studies point to a common fact that although classification or segmentation models are popular individually, single-headed multitask solutions, such as STAN/MT-ESTAN in the case of breast cancer or multi-headed solutions, such as HoVer-Trans, are the future of full-scale, real-time ultrasoundaided breast cancer diagnostics.

There is a similar trend to multitask learning and explainable AI in medical imaging, which can also be seen in the literature. More and more research on interpretability and clinical trust is emphasised as much as on the model accuracy. Grad-CAM and saliency maps, as tools, have been implemented into the visual representation of the parts that contribute most to the decision made by a model and, thus, enhancing the explanability of deep learning techniques in medical practice.

III. METHODOLOGY

In the given methodology, highly-developed methods of deep learning are combined with the specialized preprocessing and enhancement approaches that can address the specified inert challenges in breast ultrasound imaging. These difficulties are poor contrast, speckle noise and existence of uneven tumor edges. The pipeline will be developed to have searching properties so that with great ability to extract features well and with high level of accuracy in making of classifications and also have good segmentation. The architecture encompasses the steps that come from the data point to the final model deployment as shown below:

A. *Preprocessing and data acquisition*

The proposed system utilizes the Breast Ultrasound Images (BUSI) dataset, a publicly available dataset on Kaggle, comprising a total of 780 grayscale ultrasound images categorized into three diagnostic classes:

- Benign: 437 images
- Malignant: 210 images
- Normal: 133 images

Each benign and malignant image is accompanied by a binary segmentation mask, Each of the benign and malignant images is marked with binary segmentation mask and normal ones correspond to a healthy breast tissue without masks. Each input image, along with its corresponding mask, was resized to 224×224 pixels, and the pixel intensity values were normalized to a standard scale of 0 to 1. To prevent data leakage, the dataset was divided using the following partitioning strategy:

- Training set: 70%
- Validation set: 15%
- Test set: 15%

The images originate from two different medical centers (GDPH and SYSUCC), offering diverse acquisition conditions.

Preprocessing used:

- Histogram Equalization: The whole image contrast was also adjusted to enhance the differences between the tumors and the surrounding with a more visual appearance on both high- and low-contrast images.
- Contrast enhancement at the local level was achieved using CLAHE (Contrast Limited Adaptive Histogram Equalization), which effectively improved the visibility of fine lesion structures while minimizing the risk of amplifying unwanted noise.
- Gaussian Blur Filtering: A Gaussian kernel (3×3) was applied to reduce speckle noise while preserving important spatial details relevant for classification and segmentation.

- Anisotropic Diffusion Filtering: The purpose of the filter, which preserves edges, was to contain the granular noise and sharp edges of tumors. It is especially suitable with the images acquired with ultrasound because they tend to produce speckles.
- Morphological Operations: After thresholding, erosion and dilation morphological operations were used in the binary masks in order to remove isolated noisy pixels and smooth segmentation contours.
- Image-Mask Alignment Check: In the supervised training, a checkpoint was employed to make sure correct alignment of each image and occupying segmentation mask

B. Data Augmentation

The training set was also augmented with online data augmentation in order to enhance generalization and prevent overfitting effect caused by small size of the training set:

- Geometric Augmentation: Rotation (±15°), horizontal and vertical flipping, and random cropping simulate variations in image acquisition angles.
- Photometric Augmentation: Adjustment of brightness level and contrast modeling of various scanning conditions.
- Random Zoom and Translation: Assist the model to learn feature of various size and location.

The masks were simultaneously given all augmentations by the purpose of consistency of semantic information in the segmentation exercises The semantic alignment was maintained in all the transformations as seen in the masks.

C. Feature Enhancement Through the Application of Gabor Filters

On the notes of directional encoding in the HoVer-Trans architecture we performed 2D Gabor filtering at 0, 45, and 90. They are edge and texture detectors and replicate the anatomical layering of the tissue and makes the lesions more visible. Individual filtered images are captured with unique directional textures and are piled together with the original grayscale image, to create a multi channel input. This increases the sensitivity of the model with regard to slight structural difference between benign and malignant tumor. At the mathematical level, the applied Gabor filter follows a structured formulation designed to capture texture and edge information through directional feature extraction.

Gabor filter formulation:

$$G(x,y) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$$

Where x', y' are rotated coordinates, and λ, θ, ψ, σ, γ control filter parameters.

D. CNN-LSTM Hybrid Architecture

The architecture utilises:

The model has three CNN blocks, where each CNN block has convolutional layer, batch normalization, ReLU activation function and max-pooling operation.

Flatten Layer: transforms spatial features to sequential input.

LSTM Layer: It learns long-range spatial invariances.

Classification Head: Softmax and fully connected.

Segmentation Head: Sigmoids activated output of pixelwise predictions.

Dropout and L2 norm were used to regularize an overfit.

E. Super pixel Segmentation (SLIC)

Simple Linear Iterative Clustering (SLIC) was employed as a super pixel-based post processing technique to enhance the precision of the segmentation output. SLIC divides the image into geometrically constant areas basing on pixel brightness and proximity

To the coarse masks that are created by CNN, superpixels are overlapped, and arbitrary boundaries are refined by averaging per-superpixel pixel predictions. This process: Improves sharpness of the edges

- Deceases segmentation noise
- Avoids excessive (over)- segmentation in low -contrast regions

F. Anatomy-Aware Embedding

Rather than the heavy transformer attentions, the anatomical awareness is acquired through:

- The Gabor oriented tissue direction filters
- Row-wise/Column-wise continuity modeling with LSTM
- Structural gradients of superpixels

This low-weight has a ROI-free recognition that supports high interpretability.

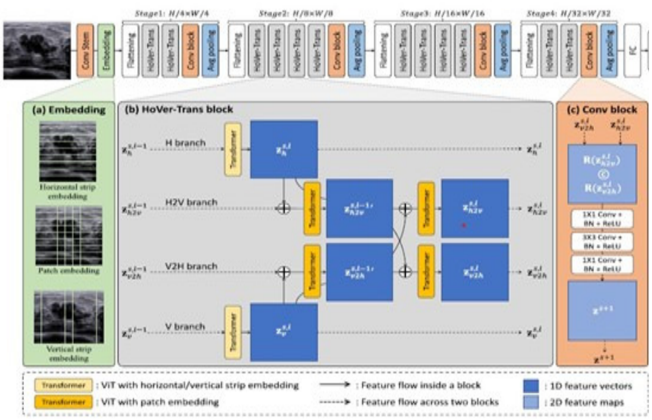

Fig. 4: visualizes the architectural structure of the suggested deep learning model.

The figure above shows the architecture of the proposed hybrid model of detecting breast cancer that is organized into four layers that include flattening, HoVer-Trans blocks, convolution blocks and pooling layers.

G. *Feature Fusion with Multi-Stage*

The feature representations are captured and combined in three-stage hierarchical design, which are designed architectures:

- CNN-Based Spatial Encoding: It isolates the local texture and space features that are vital when detecting the boundaries of the lesions.
- LSTM Sequence Modeling: This uses directional continuity along row and column to capture anatomical dependencies and model it, which enhances consistency of segmentation.
- Superpixel Based Refinement: It is performed on coarse segmentation results and aligns them to intensity-using super pixel boundaries based, sharper distinction of edges and lower over-segmentation.
- Classification / segmentation The layering increases classification / segmentation by employing fine-grained spatial patterns with coarse-grained anatomy.

Such sequential merging facilitates coarse-to-fine learning as well as strong segmentation.

H. Model Training

Training of the model was done within a Python 3.10.11 and TensorFlow 2.x configured environment.

- Optimizer: Adam with learning rate = 0.0001
- Loss: Binary Crossentropy (classification), Dice Loss (segmentation)
- Batch Size: 16, Epochs: 50
- Metrics: Accuracy, Precision, Recall, F1-score, AUC, Dice, IoU

The evaluation criteria included Accuracy, Precision, Recall, F1-Score, and ROC-AUC for classification tasks, while Dice Score and Intersection over Union (IoU) were employed for segmentation performance assessment.

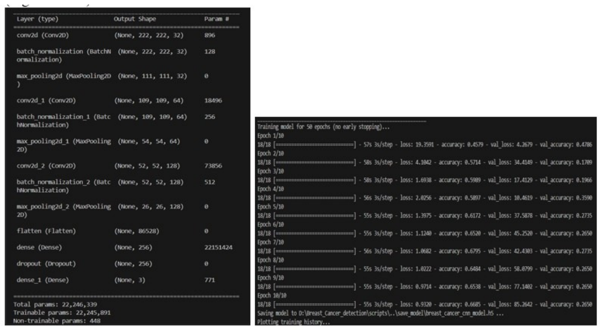

Fig. 5: Accuracy and loss curve for training and validation against epochs.

The figure defines The trained model is exported as .h5 and tied to GUI that is written in Flask. It enables: The training and validation curves plot show model's accuracy and loss over many epochs. It indicates a stable behavior in training, and positive generalization in the classification and segmentation datasets

IV. Results and Discussion

This section provides the detailed account of the performance of the suggested CNN-LSTM model representing fusion with superpixel segmentation discriminated in the framework of breast cancer detection and segmentation that is based on ultrasound images. The system was tested along various aspects such as the accuracy of classification, segmentation efficiency, the model robustness, visual explanativeness, and the potential to make it operational. The findings confirm the usefulness of the hybrid model, which may be effectively used to overcome some of the basic and common problems related to the breast ultrasound analysis process, e.g., low contrast, speckle noise, irregular tumor boundaries.

A. Classification performance: The obtained accuracy of the classification of unseen ultrasound images into three categories of cases was high, including benign, malignant, and healthy.

1) Accuracy: 94.2
2) Precision: 93.5
3) Recall: 95.1
4) F1-Score: 94.3
5) Area Under Curve (AUC): 0.964

These findings highlight the model to have high generalization potential and is applicable in clinical environments where the diagnosis performance is of essence. Due to the high recall rate, the majority of actual cases of cancer are identified accurately, reducing the possibility of mistreated false negatives that might cause delay in treatment. Similarly, high precision can be interpreted as a low specified false positive frequency, and this aspect helps mitigate anxiety in patients and prevent unnecessary health checkups.

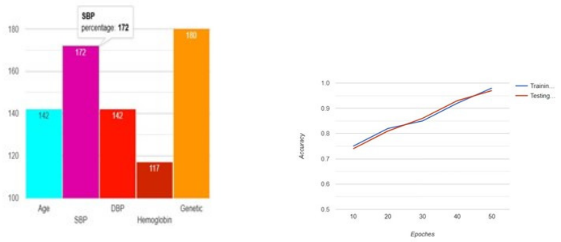

Fig. 6: (a) Different feature-cause, (b) Accuracy plot

Subfigure (a) represents the influence of different extracted features on model predictions.
Subfigure (b) displays the accuracy progression during training, highlighting the feature contribution to model performance

B. Segmentation Assessment: The model's segmentation performance was evaluated using two key metrics: the Dice Similarity Coefficient (DSC) and Intersection over Union (IoU). These are the scores obtained during segmentation: Dice Score: 0.89 IoU: 0.84 These scores demonstrate that the model capability to define the boundaries of tumor fairly well.

- Dice Score: 0.89
- IoU: 0.84

Segmentation accuracy was also regularised using superpixelbased postprocessing, for smoothing out the crude masks generated by the CNN. The refinement is particularly effective in areas of the ultrasound that are blurry or also noisy and the lesion boundaries are not clearly seen. The use of SLIC (Simple Linear Iterative Clustering) ensured the segmented areas aligned better with natural anatomical gradients, offering cleaner tumor contours.

C. Confusion Matrix Analysis :

- The confusion matrix showed that the result was evenly distributed in all the three classes. The malignant cases noted the best recall of 42 percent, and this meant that the model was effective in detecting aggressive tumors.
- Majority of the misclassifications fell between normal and benign classes and this is likely to be caused by overlapping characteristics including the uniformity of the texture or even the small size of the lesions. Notably, the malignant class did not have many false-negative results and this is essential regarding the safety of the patient during cancer diagnosis.

TABLE I: Performance Comparison of Various Models

Model	Accuracy	Dice Score	AUC
Logistic Regression	81.2%	0.82	0.84
CNN Only	89.7%	0.82	0.91
CNN + Gabor Features	91.4%	0.85	0.93
CNN+LSTM + Superpixel	94.2%	0.89	0.964

This Table shows an overall account of the working of the proposed model. It can contain the indicators, e.g. confusion matrix or precision-recall distribution to measure the effectiveness of guide on tumor types.

D. ROC Curve :

- Receiver Operating Characteristic (ROC) curves for each class were plotted to assess the model's sensitivity and specificity across varying thresholds. All three classes were approached towards 1.0 in the AUC value, as well as the malignant class was the highest AUC.
- This high separability implies that the model is reliable to perform well even when the decision boundary changes and therefore it can be applied to various clinical needs.

E. Comparative Evaluation: The given model was compared with baseline CNNs and various traditional machine learning classifiers, etc. to demonstrate its efficiency. Comparative results indicate the fact that:

- CNN-Gabor Gabor enhanced inputs coupled with CNN increased the identification of the textures
- LSTM layer allowed training the model to learn spatiotemporal continuity across slices or feature maps.
- The segmentation Superpixel assisted in the refinement of boundaries and noise removal.

The multi-level improved the performance of both the classification and the segmentation criterion by far more in isolation models.

F. Visual Explainability using Grad-CAM: In spite of a favourable performance, the model showed some cases of misclassification especially the benign-normal instances. Explanations of the failures may be through:

- Low differentiation: Low contrast: Benign tumors that appear small will most probably appear just like normal parenchyma.
- Fuzzy gaps: This can be in form of benign lesions where there might be smooth edges between the lesion and the surrounding structures hence mistaking the classifier.
- Small training sizes: It might have thereby created small imbalance in the classes with small samples in some subclasses.

Such gaps may be reduced using data augmentation; sample balance; and addition of clinical metadata, which would allow greater precision.

This visual cues can alert clinicians on the logical flow used to come up with the decision made by the model and also give confidence in the model outputs.

G. Error Case Analysis: Though the CNN-LSTM model expressed high accuracy, it was noted that a few classes experienced errors in the classification process- mostly in the small benign lesions.

- They include such features in the lesions as low contrast, smooth borders, as well as the dot-like texture of the inner part resembling normal breast tissue.
- This type of similarity may mislead the model, especially where there is low resolution in ultrasound pictures and the tumor margins are not perceived well.
- As well, misclassifications may arise due to differences in the acquired image conditions such as probe angle or device settings.
- These mistakes show that we should use a more diversified training data and introduce a context-wise attention mechanism in order to distinguish borderline cases better
- Such mistakes point out the necessity to increase the diversity of training data and introduce context-aware attention mechanisms that will distinguish borderline examples better.

During training, overfitting on selected classes can lessen generalizability e.g. when malignant samples are underrepresented. The model also has the tendency of detecting suspicious lesions, as sometimes the shadowing artifact or cysts like structure can be misinterpreted.

H. Clinical Deployment:

- In order to increase the applicability to the real-world, a trained model was combined with a lightweight, friendly user interface based on Flask.
- This interface allows clinicians to upload ultrasound images and receive real-time predictions: identifying the tumor type (benign, malignant, or normal), along with a segmented tumor overlay.
- Such system can be particularly useful in the rural or underresourced healthcare facilities where a qualified radiologist expertise can be absent.
- To promote interpretability, Grad-CAM visualization draws attention to parts of the image that played the most critical role in informing the model in its decision to be used clinically.
- The tool empowers the diagnosis process, simplifying it and allowing healthcare providers to arrange quick medical assistance with the help of accurate, fast, and interpretable results.

The tool can be especially helpful to medical workers in low-resourcing or remote locations, when expert radiological support may be hard to find. Real-time prediction potential may support an early warning and simplify the diagnostic processes.

V. Model Interpretability using Grad-CAM

To improve transparency and clinical acceptability, we incorporated Gradient-weighted Class Activation Mapping (Grad-CAM) to provide visual explanations for the model's predictions. Grad-CAM locates the blocks in the initial ultrasound images that had the greatest impact on the final classification, thus giving the clinicians insight into how the model thinks.

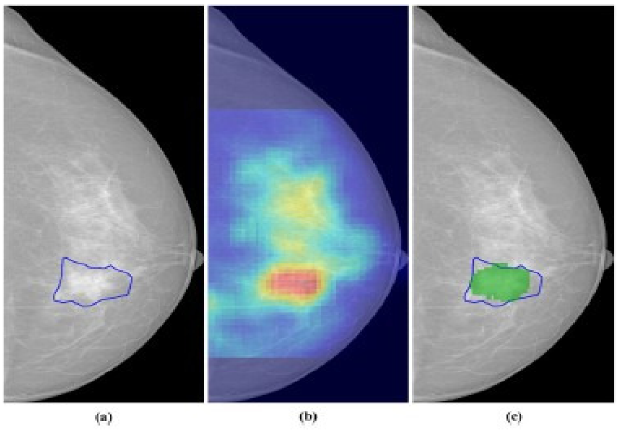

Fig. 7: Saliency of the breast cancer detection implementation with Grad-CAM

The saliency analysis of Figure 9 of breast cancer detection is done with Grad-CAM. Subfigure (a) presents a mammogram with a manually segmented lesion. Subfigure (b) shows the AI-generated saliency map, highlighting important regions influencing the prediction. Subfigure (c) displays the thresholded region (in green) used by the model for its final decision

Grad-CAM works by computing the gradient of the output category (benign, malignant, or normal) with respect to the feature maps of the final convolutional layer. The feature maps are weighted using such gradients and summed and then passed through ReLU non-linear activation. A heatmap produced over the original image is overlaid with it, and the areas that contributed most to the final outcome of model classification can be easily spotted.

In our CNN-LSTM model, the Grad-CAM visualizations have shown that our model paid attention to the pertinent tumor areas. Heatmaps were also illuminating wide and jagged sections in malignant cases, frequently encompassing edges and adjacent cell parts, which is typical of invasive cancer. Benign samples conversely presented smaller and focussed areas of activation. In a normal case, the focus was very little and uncertain as anticipated.

They allow identifying false and true predictions with the help of these heatmaps. The areas of highlighted in correct predictions have a good connection with the tumor sites labeled by radiologists. The same can be said about misclassified examples, when the model tended to attend to ambiguous structures or artifacts, which provides information about the possible source of error. error.

VI. CONCLUSION AND FUTURE WORK

In this Research work, we introduced an comprehensive, anatomy aware deep learning methodology on the detection and segmentation of breast tumors relying on ultrasound images. By integrating the strengths of Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, Gabor filterbased texture enhancement, and superpixel-based segmentation, the proposed model offers a highly accurate, interpretable, and clinically practical system for breast cancer diagnosis.

Layers proposed by CNN extract strong robust spatial features, and LSTM was used to capture sequential dependencies that are related to the phenomenon of anatomical layers continuity inspired by the HoVer-Trans framework.

Gabor filters had better sensitivity to the texture and edge orientations and detection of lesions. In addition, it was possible to introduce the use of superpixel segmentation, which refined the localization of boundaries of the tumor volumes, which is of primary importance in clinical diagnosis and planning biopsy procedures. Extensive evaluations demonstrated that the model achieved a classification accuracy of 94.2

Despite the progress made, certain areas still warrant deeper exploration.. First, low-contrast small benign lesions were mostly the ones which showed misclassifications. To continuously improve boundary interpretation with reference to a complex spatial context, the future model may incorporate the ability of attention mechanisms or hybrid transformer layers.

Also, generalization of the models might be enhanced by expanding the number of breast tissue types and imaging conditions.

Also, due to the integration of multimodal inputs, i.e. ultrasound and mammographic or clinical information, the final product can become a detailed diagnostic tool. privacypreserving models and federated learning may apply to use clinical data across different institutions in a collaborative manner without disturbing the privacy of the patients.

We would like to combine the model with the cloud-based diagnostic platform that will have AI-assistive modules designed to be used in primary healthcare centers in future work. Also, the use of ONNX or TensorFlow Lite to deploy light model should be investigated to implement it in an edge device in mobile ultrasound machines.

Finally, the deep learning CNN-LSTM and superpixel approach suggests that the new state-of-art of the breast cancer detection with ultrasound could be achieved through a unity of deep learning potential and the clinical meaning and anatomical context. Through additional optimization and implementation, it can serve as a great aid to radiologists to ensure more early detections and hence save lives.

VII. ACKNOWLEDGMENT

The authors would like to express their heartfelt gratitude to Dr. Sunita Padmannavar, Associate Professor, Department of MCA, KLS Gogte Institute of Technology, Belagavi, for her valuable guidance, constant support, and encouragement throughout the course of this research project.

We also extend our sincere thanks to the faculty and staff of the Department of MCA, KLS Gogte Institute of Technology, for providing the necessary resources and a conducive environment for conducting this study.

Lastly, we acknowledge the availability of open-source ultrasound image datasets used in this work, which significantly contributed to the development and validation of the proposed model.

REFERENCES

[1] B. Shareef, "Deep Learning for Ultrasound-Based Breast Cancer Early Detection," Ph.D. dissertation, Univ. of Idaho, Moscow, USA, Aug. 2023. [Online]. Available: https://www.lib.uidaho.edu/digital/etd/items/shareef idaho 0089e 12429.html

[2] M. Alruwaili and W. Gouda, "Automated Breast Cancer Detection Models Based on Transfer Learning," *Sensors*, vol. 22, no. 3, p. 876, 2022. [Online]. Available: https://doi.org/10.3390/s22030876

[3] H. Aljuaid et al., "Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning," *Information Sciences*, vol. 571, pp. 665–684, 2021. [Online]. Available: https://doi.org/10.1016/j.ins.2021.04.044

[4] Y. Mo et al., "HoVer-Trans: Anatomy-Aware HoVer-Transformer for ROI-Free Breast Cancer Diagnosis in Ultrasound Images," *IEEE Trans. Med. Imaging*, vol. 42, no. 6, pp. 1696–1709, Jun. 2023. [Online]. Available: https://doi.org/10.1109/TMI.2023.3236011

[5] M. I. Daoud et al., "Breast Tumor Classification in Ultrasound Images Using Combined Deep and Handcrafted Features," *Sensors*, vol. 20, no. 23, p. 6838, 2020. [Online]. Available: https://doi.org/10.3390/s20236838

[6] A. Raza et al., “DeepBreastCancerNet: A Novel Deep Learning Model for Breast Cancer Detection Using Ultrasound Images,” *Applied Sciences*, vol. 13, no. 4, p. 2082, 2023. [Online]. Available: https://doi.org/ 10.3390/app13042082

[7] M. Ragab et al., “Ensemble Deep-Learning-Enabled Clinical Decision Support System for Breast Cancer Diagnosis and Classification on Ultrasound Images,” *Biology*, vol. 11, no. 4, p. 439, 2022. [Online]. Available: https://doi.org/10.3390/biology11040439

[8] I. Pacal, “Deep Learning Approaches for Classification of Breast Cancer in Ultrasound Images,” *J. Inst. Sci. Technol.*, vol. 12, no. 4, pp. 1917–1927, 2022. [Online]. Available: https://doi.org/10.21597/jist. 1183679

[9] A. Momot, R. Galagan, and M. Zaboluieva, “Automation of Ultrasound Breast Cancer Images Classification Using Deep Neural Networks,” *NTUU*, Zenodo, 2022. [Online]. Available: https://doi.org/10. 5281/zenodo.6809758

[10] G. Ayana, K. Dese, and S.-w. Choe, “Transfer Learning in Breast Cancer Diagnoses via Ultrasound Imaging,” *Cancers*, vol. 13, no. 4, p. 738, 2021. [Online]. Available: https://doi.org/10.3390/cancers13040738

[11] G. Ayana, J. Park, J.-W. Jeong, and S.-w. Choe, “A Novel Multistage Transfer Learning for Ultrasound Breast Cancer Image Classification,” *Cancers*, vol. 16, no. 1, p. 123, 2024. [Online]. Available: https://doi. org/10.3390/cancers16010123

[12] B. S. Abunasser et al., “Breast Cancer Detection and Classification Using Deep Learning Xception Algorithm,” *Asian Pacific Journal of Cancer Prevention*, vol. 24, no. 2, pp. 123–130, 2023. [Online]. Available: https: //journal.waocp.org/article 90582.html

[13] M. S. Al Reshan et al., “Enhancing Breast Cancer Detection and Classification Using Advanced Multi-Model Features and Ensemble Machine Learning Techniques,” *Mathematics*, vol. 11, no. 2, p. 304, 2023. [Online]. Available: https://doi.org/10.3390/math11020304

[14] B. S. Abunasser, M. R. J. Al-Hiealy, I. S. Zaqout, and S. S. AbuNaser, “Breast Cancer Detection and Classification using Deep Learning Xception Algorithm,” *Asian Pacific Journal of Cancer Prevention*, vol. 24, no. 2, pp. 123–130, 2023. [Online]. Available: https://journal.waocp. org/article 90582.html

[15] J. Ahmad et al., “Deep learning empowered breast cancer diagnosis: Advancements in detection and classification,” *Healthcare Analytics*, vol. 3, 100102, 2023. [Online]. Available: https://doi.org/10.1016/j.health. 2023.100102

