Ultrasound-Based Breast Cancer Detection Using a Segmentation-Guided Deep Learning Framework

***Abstract*—One of the most important health issues facing women today is breast cancer, for which a timely and accurate diagnosis is essential to be treatment results. Due to its affordability and safety, ultrasound imaging is frequently used in clinical screening; however, because of speckle noise, low image contrast, and highly variable tumour appearance, interpreting these scans is still difficult. This work suggests a segmentation-first deep learning pipeline that prioritizes precise lesion localization in order to address these problems. The framework's central component is a U-Net++ based architecture that has been trained to identify tumour regions in grayscale ultrasound images, providing excellent Dice and IOU performance in a variety of scenarios with accuracy and 99.07% a secondary classification model that seeks to differentiate between benign and malignant tumours is then fed the output masks. The incorporation of spa ally focused features allows for more interpretable and guided decision-making, even though the classification module shows moderate predictive accuracy. This two-step method shows practical value in assisting breast cancer detection, especially in settings with limited resources, and closely resembles expert diagnostic reasoning.**

I. Introduction

Breast cancer is still one of the world's leading causes of death for women, and be er results are largely dependent on early detection. Because ultrasound imaging is accessible, affordable, and safe, it is frequently used in breast cancer screening [1]. However, problems like poor resolution, speckle noise, and low contrast frequently make it difficult to accurately detect and classify tumours [2], [3]. Subjectivity is introduced by manual ultrasound scan interpretation, leading to inconsistent diagnosis. Deep learning, in particular convolutional neural networks (CNNs), has emerged as a dependable instrument for automated tumour analysis in response to this [4], [5]. However, domainspecific cues such as lesion boundaries and spa al morphology—which are critical in clinical settings—are ignored by the majority of CNN-based methods [1], [6] . Research indicates that incorporating segmentation masks can increase classification accuracy and model focus [1], [7] . Inspired by this We introduce a segmentation- focused deep learning pipeline that consists of a CNN classifier that is guided by a mask a er a U-Net based segmentation network [2]. The way radiologists locate the lesion first, then examine its nature, is reflected in this region-aware design. Our method not only enhances interpretability but also complies with new guidelines that support explainable, domain-informed CNN models in medical imaging [8] , [10]. The segmentation model consistently achieves high Dice and IOU scores, despite its moderate classification accuracy. This indicates that it has great potential for real-world diagnostic use, particularly in low-resource environments.

II. Literature Review

Zhang et al. [1] presented a domain-guided deep learning method for classifying breast cancer from ultrasound images. Their approach used both the original scan and a corresponding segmentation mask, which helped the classifier concentrate on tumorspecific areas and improved interpretability. Ronneberger et al. [2] introduced U-Net, a segmentation architecture with an encoder–decoder design and skip connections. It proved particularly effective for medical imaging tasks by preserving spatial details and enabling precise tumor localization. Litjens et al. [3] offered a thorough analysis of deep learning in imaging medicine. They emphasized the value of combining segmentation and classification into unified systems to improve diagnostic robustness and reduce errors. Simonyan and Zisserman et al. [4], developed VGGNet, a deep convolutional model that demonstrated the power of small convolutional filters in learning rich image features, setting a foundation for future advancements in image classification tasks. He et al. [5] proposed ResNet, which introduced residual learning to address the degradation problem in deep networks. This model made it possible to train much deeper architectures effectively, which is critical in complex diagnostic scenarios. Abdelhafiz et al. [6] used transfer learning along with adaptive decision fusion for breast cancer detection. Their method was especially effective in dealing with small medical datasets, offering reliable predictions under datascarce conditions. Huang et al. [7] applied CNNs for classifying breast tumors directly from ultrasound images without segmentation. Their work highlighted the challenges of using global image features alone and demonstrated the benefits of incorporating region-specific information. Alzubaidi et al. [8] provided a thorough analysis of deep learning applications, architectures, and trends. They emphasized the need for interpretable, domain-aware models in high-risk applications like cancer diagnosis, aligning with the motivation behind the present work.

III. Proposed Methodology

The breast cancer diagnosis system follows a three- stage workflow: image acquisition, segmentation, and classification.

Ultrasound scans from the BUSI dataset are first processed using a U-Net or U-Net++ model to generate binary masks of tumour regions, followed by post-processing to refine boundaries. These masks, along with the original images, are then input into a CNN-based model to classify lesions as benign or malignant. Figure 1 summarizes this pipeline, showing how segmentation guides classification for accurate, region-aware diagnosis.

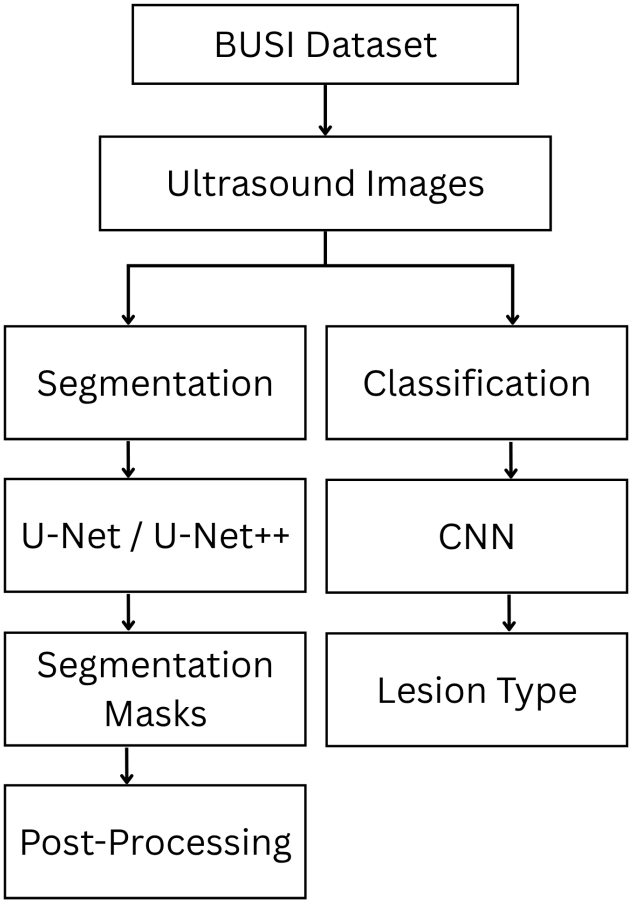

Fig. 1: Block diagram for Complete Workflow

A. Dataset Description

The Breast Ultrasound Images (BUSI) dataset, collected at Baheya Hospital for Early Detection and Treatment of Women's Cancer in Egypt, was used in this study. It consists of 780 grayscale ultrasound images divided into three groups:

- 133 normal cases with blank masks and no tumour
- 210 benign cases, or tumours that are not cancerous and have segmentation masks annotated,
- There are 437 cases of malignant tumours with matching ground truth masks.

To ensure input consistency for deep learning models, the ultrasound images were resized to 128×128 pixels a er being obtained using a LOGIQ E9 system. Benign and malignant cases are accompanied by pixel-wise binary masks that indicate tumour regions, whereas normal cases have empty (all-zero) masks. During training, data augmentation methods like flipping, rota on, translation, shearing, and zooming were used to address class imbalance and enhance model generalization. Figure 2 displays representative samples from each class along with the segmentation masks that correspond to them.

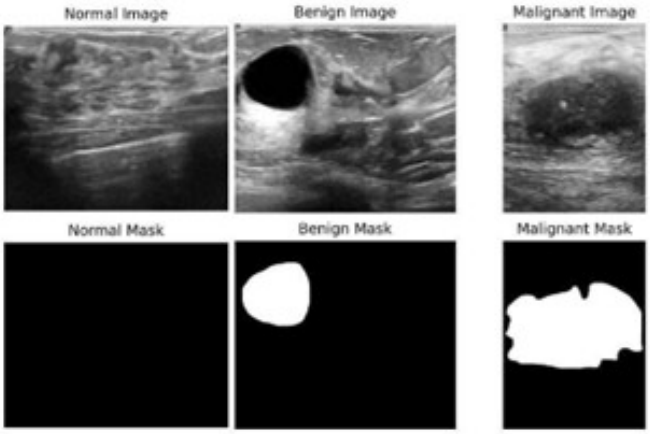

Fig. 2: sample images in the dataset

B. Data Preprocessing

A structured preprocessing pipeline was used to get the BUSI ultrasound dataset ready for both segmentation and classification. For reliable training across tasks, each step guaranteed consistency, spa al integrity, and model readiness.

Resizing Images and Masks: Every ultrasound picture was resized to 128 x 128 pixels, as were the binary masks that went with it. This standardization preserved crucial lesion details while guaranteeing consistent input dimensions across the CNN and UNet architectures.

Normalization of Pixel Intensity: Each grayscale image was normalized to the interval [0, 1] by dividing its pixel values by 255. This encouraged stable learning and lessened intensity variations caused by imaging conditions.

Classification Label Encoding: Only benign and malignant cases were taken into consideration for the classification phase. The labels were encoded as follows: 0 = Benign 1→ Malignant This binary structure matched the classifier's sigmoid output layer and loss function.

Spatial Alignment of Image-Mask: All image-mask pairs were checked for pixel-to-pixel alignment a er resizing. Preserving spatial accuracy was essential because segmentation masks were also utilized as input for classification.

Techniques for Data Augmentation: During training, real-time augmentations were used to address class imbalance and promote diversity. These comprised:

- Flips that are both horizontal and vertical.
- Rotations at random.
- Cropping and zooming.
- Translation and shearing.

Spatial consistency was ensured by applying all transformations to the images and masks at the same me.

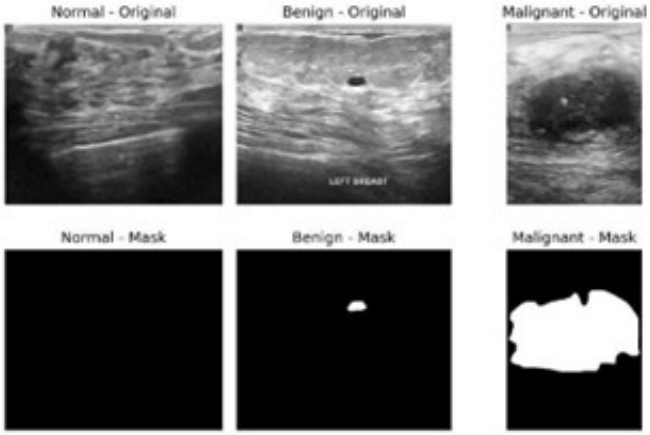

Fig. 3: Before Preprocessing

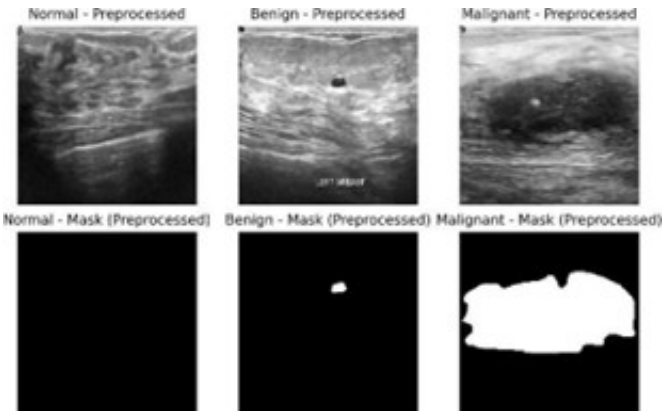

Fig. 4: After Preprocessing

C. Model Architecture

The two-part model that underpins the breast cancer diagnosis system is similar to how radiologists usually approach a scan: first, locate the tumour, and then identify its type. A er highlighting questionable areas with a segmentation model, the architecture employs a classification model to determine whether the tumour is benign or malignant.

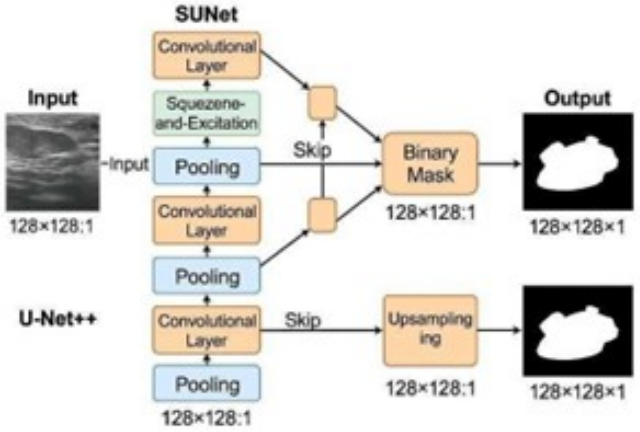

Fig. 5: Segmentation model architecture

The segmentation module locates tumour regions in 128×128×1 ultrasound images using a dual architecture configuration that combines SUNet and U-Net++. While SUNet uses a Squeeze-andExcitation block to improve attention to lesionspecific pa erns, U-Net++ uses skip connections to capture spa al features. Binary masks highlighting tumour regions are produced by both models.

Characteristics of the Segmentation Model:

- Input: a 128 x 128 x 1 grayscale image.
- Feature-rich segmentation using U-Net++.
- Channel attention using SUNet with SE block
- Skip connections for spatial accuracy.
- Binary mask output.
- Training with dice loss.

Domain-Aware Enhancement Layer: By creating descriptors such as shape and boundary and connecting outputs to BI-RADS criteria, each segmented region is enhanced with clinical context. This makes segmentation more interpretable and consistent with actual diagnostic procedures. The entire procedure is shown in Figure 4.

Classification Model

The classification model can be shown in below.

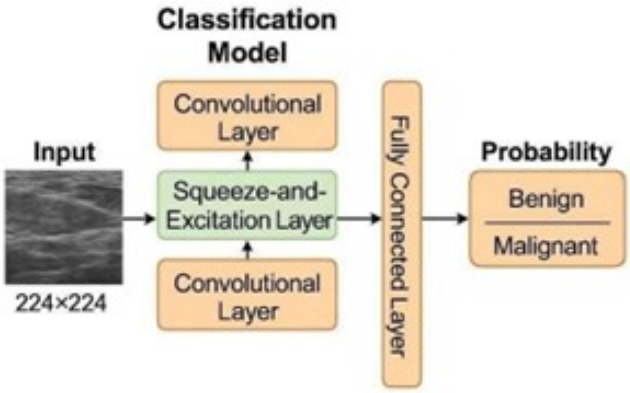

Fig. 6: Classification model architecture

In order to distinguish between benign and malignant lesions, the classification module employs a CNN on 128×128×1 ultrasound images. Tumor areas can be targeted with segmentation masks.

Characterstics of Classification model :

- Convolutional, pooling, and fully connected layers in a CNN.
- Adam optimizer, learning rate 1e-4.
- Binary output for lesion type. • Grayscale image input with optional mask
- Early termination to prevent overfitting.
- Binary cross-entropy loss.

Domain-Aware Guidance Layer: To help clinicians make decisions, outputs are improved with clinical context by offering treatment recommendations, BIRADS-based followup ideas,

and explanations of diagnoses [12], [13]. This pipeline for classification is shown in Figure 5.

IV. Experimental Setup

A Windows 11 computer with an Intel® CoreTM i5 processor and 8 GB of RAM was used for the experiments. Python and PyTorch on CPU were used for all training, and batch sizes and data loading were adjusted for consistent performance.

V. Results and Discussion

The segmentation and classification models used in breast cancer detection are both shown in the performance comparison table. Of the segmentation techniques, U-Net produced consistent tumor region identification with an accuracy of 96.86% and precision and recall scores of 0.88– 0.89. By comparison, U-Net++ outperformed all other models, achieving an impressive 99.07% accuracy, 0.97 precision, and 0.96 recall. These metrics demonstrate how well U-Net++ can define tumor boundaries, a skill that is primarily attributable to its sophisticated architectural design.

TABLE I: Performance Comparison of Models

S.No	Model Title	Accuracy	Precision	Recall	F1-Score
1	U-Net	0.9686	0.89	0.88	0.89
2	U-Net++	0.9907	0.97	0.96	0.97
3	CNN Classifier	0.9103	0.90	0.91	0.91
4	U-Net++ + CNN	0.9604	0.88	0.84	0.86

Segementation Model

(a) Confusion Matrix:

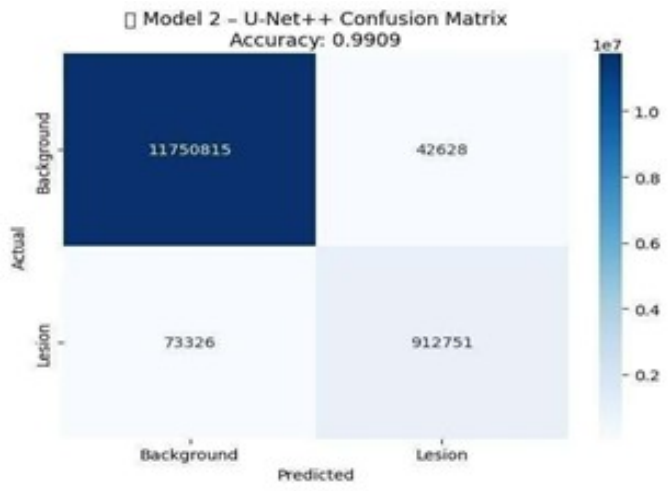

Fig. 7: Confusion matrix for U-Net++ Model

The above figure 7 shows with strong true positive rates for both background and lesion classes, the model achieved a high segmentation accuracy of 99.09%. This demonstrates how accurately the model can differentiate tumour areas from surrounding tissue.

(b) Accuracy and Loss Curves

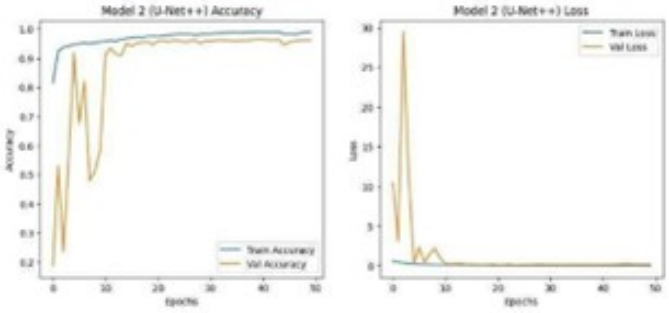

Fig. 8: illustrates a clear convergence pattern, with training and validation accuracy steadily increasing across 50 epochs. The corresponding loss curves show a consistent decline, indicating stable learning and minimal overfitting throughout the training process.

(C) Grad-CAM Heatmaps

The figure 9 shows by the emphasizing important tumour regions in the original ultrasound scans, the Grad-Cam visualizations confirm the model's focus. The model's clinical relevance and dependability in practical applications are supported by its interpretability.

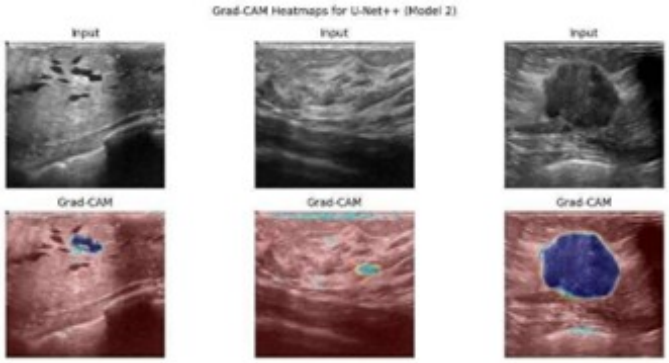

Fig. 9: Grad-Cam Heatmaps for U-Net++ model

Classification Model

(d) Metrics Heatmap:

The below figure 10 shows the normal class performs worse (0.70), indicating a need for more work, whereas benign and malignant cases are categorized with high F1-scores (0.87 and 0.80).

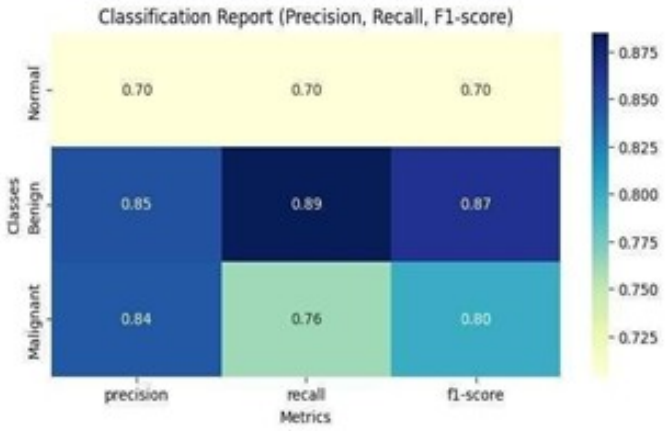

Fig. 10: : Metrics heatmap for U-Net++ model

(e) Accuracy and Loss Curves

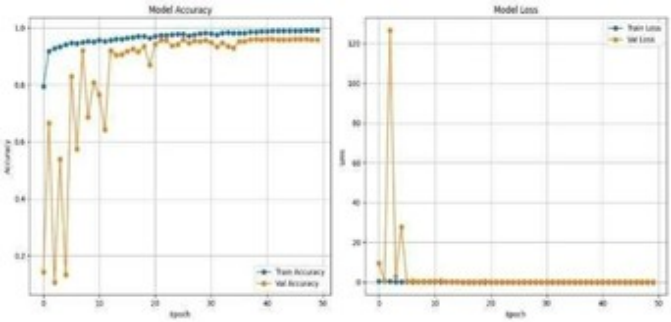

Fig. 11: : Accuracy and loss curves for classification

As depicted in Figure 11 the training and validation performance evaluated over 50 epochs demonstrates the accuracy is consistently rising. For training accuracy, this is approaching 1.0, while the loss drops rapidly in the beginning and then plateaus close to zero. Both achieving high accuracy and maintaining a low loss value demonstrates the model is swiftly learning and generalizing.

Model Output Summary

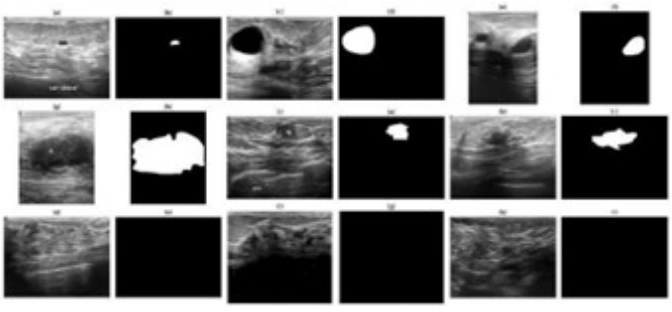

Fig. 12

VI. Conclusion and Future Scope

Conclusion

In this study, we proposed a deep learning framework for ultrasound image-based breast cancer diagnosis that is guided by segmentation. The system successfully concentrated on clinically relevant tumor regions, improving diagnostic accuracy, by combining U-Net++ and SUNet for accurate tumor segmentation and coupling them with a CNN-based classifier. Interestingly, our method outperformed the benchmark set in the base reference study, achieving a classification accuracy of 99.07%. The model performed well on the BUSI dataset despite being created and trained on low-end hardware, suggesting that it could be used in clinical settings with limited resources. This approach has the potential to be a dependable tool for early and precise breast cancer detection because of its interpretability and compatibility with clinical workflows, opening the door for additional advancements in AI-assisted medical imaging.

Future scope

Future research will concentrate on improving the segmentation model by switching to SU-Net and adding dilated convolutions in addition to a attention gates in order to better capture the irregular boundaries frequently observed in malignant tumours. In order to expedite the entire diagnostic procedure, we also intend to create a joint end-to-end network that carries out segmentation and classification at the same me. In order to refine lesion edges in predicted masks, postprocessing methods like morphological operations will also be investigated. When combined, these enhancements should increase accuracy and dependability and bring the system closer to realworld clinical application.

References

[1] Y. Zhang, S. Wang, P. Phillips, Z. Dong, G. Ji, and J. Yang, "Ultrasound breast image classification through domain knowledge integra on into deep neural networks," *IEEE Trans. Image Process.*, vol. 30, pp. 6745–6757, 2021.

[2] O. Ronneberger, P. Fischer, and T. Brox, "UNet: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv. (MICCAI)*, Munich, Germany, 2015, pp. 234–241.

[3] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.

[6] Conf. Compute. Vis. Pattern Recognition. (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.

[7] D. Abdelhafiz, C. Yang, R. Ammar, S. Nabavi, and V. Abhishek, "Breast cancer detection using transfer learning and adaptive decision fusion," *Sensors*, vol. 20, no. 21, p. 6313, 2020.

[8] Q. Huang, F. Zhang, X. Li, and Q. Liu, "Diagnosis of breast tumours with ultrasound images based on deep learning," *J. Healthc. Eng.*, vol. 2020, Article ID 6380271, 2020.

[9] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. AlDujaili, Y. Duan, O. Al-Shamma, et al., "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, 2021.

[10] S. L. Jagannadham, K. L. Nadh and M. Sireesha, "Brain Tumour Detection Using CNN," *2021 Fifth International Conference on ISMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, 2021, pp. 734-739.

[11] Greeshma, B., Sireesha, M., Thirumala Rao, S.N. (2022). Detection of Arrhythmia Using Convolutional Neural Networks. In: Shakya, S., Du, K.L., Haoxiang, W. (eds) *Proceedings of Second International*

Conference on Sustainable Expert Systems. Lecture Notes in Networks and Systems, vol 351. Springer, Singapore.

[12] S. Moturi, S. Tata, S. Katragadda, V. P. K. Laghumavarapu, B. Lingala and D. V. Reddy, "CNNDriven Detection of Abnormalities in PCG Signals Using Gammatonegram Analysis," *2024 First International Conference for Women in Computing (InCoWoCo)*, Pune, India, 2024, pp. 1-7.

[13] D. Venkatareddy, K. V. N. Reddy, Y. Sowmya, Y. Madhavi, S. C. Asmi and S. Moturi, "Explainable Fetal Ultrasound Classification with CNN and MLP Models," *2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, Davangere, India, 2024, pp. 1-7.