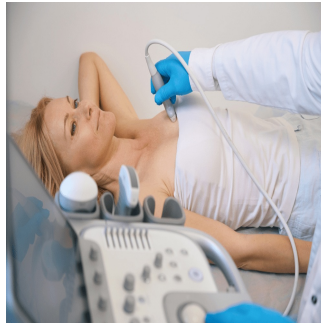


Improving Breast Ultrasound Analysis with Transfer Learning



Breast cancer remains the most prevalent malignancy among women and the second leading cause of cancer-related mortality. Timely and accurate diagnosis is paramount for improving outcomes. Ultrasound is crucial as a supplementary imaging tool due to its wide availability, non-invasiveness and cost-effectiveness. However, operator dependency and variability in interpretations often compromise its effectiveness, leading to false positive results and unnecessary biopsies. The emergence of artificial intelligence (AI) in medical imaging has provided a new pathway to enhance diagnostic accuracy. Specifically, transfer learning using convolutional neural networks (CNNs) has shown promise in distinguishing between benign and malignant breast lesions, potentially reducing the rate of unwarranted biopsies.

Understanding Transfer Learning in Breast Ultrasound

Transfer learning utilises pre-trained CNNs to perform specific classification tasks with less data and computational power than would be needed to train a model from scratch. This approach is particularly beneficial in medical imaging, where obtaining large datasets can be challenging. CNNs such as InceptionV3, Xception, ResNet50, VGG16 and DenseNet121 have been employed in various applications for object detection and classification. These sophisticated architectures are particularly effective at identifying complex features in images.

In transfer learning, the foundational layers of these pre-trained models, which capture basic image characteristics like edges and textures, remain intact. Only the top layers are retrained or fine-tuned to specialise in new tasks, such as differentiating between types of breast lesions. This approach allows for efficient training, mitigates overfitting and is adaptable even when training data is limited. The "feature extractor" method, which keeps all but the last layers frozen and the "fine-tuning" method, which updates more layers, are common strategies employed to customise CNNs for breast ultrasound analysis.

Evaluating CNN Performance on Ultrasound Data

Studies utilising both public and institutional datasets have demonstrated the efficacy of transfer learning in breast ultrasound imaging. For instance, when CNNs trained on the Breast Ultrasound Image (BUSI) dataset were tested, models such as DenseNet121 and VGG16 achieved high diagnostic accuracy with AUROC scores ranging from 0.938 to 0.996. DenseNet121, in particular, demonstrated a sensitivity of 96.2% and specificity of 99%, showcasing its robustness in initial tests.

However, when these models were validated against an institutional dataset containing more diverse and complex data, performance dropped, especially for intermediate BI-RADS 4 lesions, which pose a classification challenge due to their varied malignancy risk. While DenseNet121 still performed reliably with an AUROC of 0.814 on the institutional dataset, the difference highlighted the limitations of applying models trained on public datasets directly to real-world data. The variability among datasets, influenced by factors such as image acquisition techniques and patient demographics, necessitates tailored fine-tuning to achieve optimal results across different clinical settings.

Optimising Clinical Use with Threshold Adjustments

One significant finding in applying CNNs to breast ultrasound is the role of classification thresholds in balancing sensitivity and specificity. The default threshold for classification is typically set at 50%, where a lesion is considered malignant if its predicted probability exceeds this value. However, lowering this threshold to values such as 2% has increased sensitivity substantially. For example, at a 2% threshold, networks like Xception and DenseNet121 achieved a sensitivity of 98.3%, ensuring that nearly all malignant cases were identified, which is vital in preventing false negatives.

The trade-off for increased sensitivity is reduced specificity, leading to more false positives. At a 2% threshold, specificity dropped to around 15-

18%, meaning that while more benign cases might be flagged as suspicious, the reduction in missed malignancies is a critical advantage in clinical practice. This balance is particularly useful for BI-RADS 3 lesions, where current practice often leads to excessive biopsies. By implementing CNNs with adjusted thresholds, unnecessary biopsies could be reduced by up to 15-18%, offering a safer and more cost-effective approach to patient management.

Moreover, combining predictions from multiple top-performing models, such as Xception, InceptionV3 and DenseNet121, has further enhanced performance. This ensemble method can achieve a sensitivity of 100%, ensuring no cancer is missed while maintaining a specificity that, although lower than desired, still supports clinical decision-making.

Challenges and Future Considerations

Despite promising results, there are challenges to the widespread implementation of transfer learning for breast ultrasound. One limitation is the need for extensive validation using images from different ultrasound machines and patient populations to ensure generalisability. Although substantial, the current study's institutional dataset comprised 392 images, which may not fully represent the variability found in broader clinical practice. Additionally, while transfer learning mitigates the need for large datasets, manually contouring lesions for training remains time-consuming. Future research should focus on automated segmentation techniques that can pre-process images efficiently.

Another consideration is the variability in performance across different BI-RADS categories. BI-RADS 4 lesions, with their diverse risk levels, often lead to reduced model accuracy compared to simpler cases like BI-RADS 3 or 5. Customised strategies, such as further fine-tuning of model layers or incorporating more clinical parameters, could improve results in these challenging categories.

The integration of transfer learning with CNNs in breast ultrasound analysis represents a significant advancement in improving diagnostic accuracy and reducing unnecessary biopsies. By adjusting classification thresholds, these models can achieve high sensitivity, which is crucial for ensuring early cancer detection. While specificity may decrease, the potential to reduce biopsy rates by up to 18% without missing malignancies makes this an attractive clinical tool. Continued refinement, broader dataset validation and prospective trials are essential to establish CNNs as reliable decision support systems, enhancing the efficacy and efficiency of breast cancer diagnosis and patient care.

Source: [European Radiology Experimental](#)

Image Credit: [iStock](#)

Published on : Sun, 10 Nov 2024