

FREQTRANS: ADAPTIVE FREQUENCY FEATURE FUSION FOR ROBUST BREAST ULTRASOUND IMAGE ANALYSIS

Hongyang Zhao¹ Dan Lu^{1,3} Yingnan Zhao^{1,3,*} Yanchen Xu¹ Yi Lin^{2,*}

¹ College of Computer Science and Technology, Harbin Engineering University, China

² School of Interdisciplinary Medicine and Engineering, Harbin Medical University, China

³ National Engineering Laboratory for Modeling and Emulation in E-Government
Harbin Engineering University, Harbin, 150001, China

ABSTRACT

Breast ultrasound imaging is widely used for early detection of breast cancer but is challenged by low contrast, blurred lesion boundaries, and speckle noise, which complicate the identification of subtle tumors. Here, we present FreqTrans, a frequency-domain model that decomposes images into low- and high-frequency components and applies attention mechanisms to enhance diagnostically relevant features. By integrating local and global cues and fusing multi-level representations, FreqTrans improves edge delineation and suppresses noise, enabling robust lesion detection. On public breast ultrasound datasets, FreqTrans outperforms standard convolutional and vision Transformer models, achieving higher sensitivity and overall detection accuracy, highlighting its potential for clinical application.

Index Terms— Breast ultrasound, Tumor detection, Frequency-domain analysis, Adaptive fusion, Deep learning.

1. INTRODUCTION

Accurately detecting breast cancer in medical images is crucial for early diagnosis and effective treatment[1]. However, breast ultrasound images present multiple challenges that complicate automated analysis. Firstly, low contrast and blurred lesion boundaries make tumors difficult to distinguish from surrounding tissue, especially for small or early-stage lesions [2]. Secondly, speckle noise and redundant frequency information can obscure clinically relevant structures, reducing the effectiveness of conventional feature extraction [3]. Thirdly, the insufficient integration of local and global features in standard convolutional neural networks (CNNs) or vision transformers (ViTs) limits their ability to simultaneously capture fine lesion edges and overall tissue architecture [4][5]. Additionally, variations in imaging modalities, acquisition protocols, and patient anatomy, coupled with limited availability of well-annotated datasets, further hinder deep learning performance [6].

Existing deep learning architectures have shown promising results but still exhibit critical limitations in breast ultrasound analysis. For example, VGG16 [7] and similar CNN-based backbones focus mainly on spatial domain convolutions, which excel at capturing local patterns but struggle with long-range dependencies, and often fail to generalize under noisy ultrasound conditions. Their reliance on deep stacking also leads to redundant features, making them less efficient for subtle lesion detection[8]. More recently, transformer-based models such as HoVerTrans attempt to incorporate global context by leveraging self-attention mechanisms[9]. While they improve

long-range feature modeling, they still suffer from weak representation of fine-grained lesion boundaries[10], as spatial-domain attention alone cannot fully disentangle noise and frequency redundancies inherent in ultrasound imaging[11]. Moreover, both CNNs and vanilla transformers often underperform when applied directly to highly heterogeneous medical datasets without explicit mechanisms to handle frequency-specific cues[12].

To address these challenges, we propose FreqTrans, a hybrid model that integrates spatial and frequency-domain representations [13]. The model extracts low-frequency features to capture global tissue architecture and overall lesion distribution [14], while high-frequency features highlight fine-grained details such as lesion margins and micro-calcifications [15]. An Adaptive Frequency Fusion (AFF) module combines these complementary features, adaptively weighting low- and high-frequency information to enhance subtle structures, reduce noise interference, and improve contrast in low-visibility regions [16]. A Multi-Scale Feature Extraction and Abstraction further refines pixel embeddings augmented with frequency information [17], and a merge module reconstructs multi-level feature maps for final classification [18].

By jointly leveraging global context and fine details, FreqTrans effectively improves sensitivity to small or low-contrast lesions, suppresses noise and redundant information, and balances local and global feature integration, outperforming conventional CNNs and standard vision transformers in complex breast cancer image analysis.

To sum up, the contributions of this work are as follows:

- Enhancing low-contrast lesion visibility: By decomposing images into low- and high-frequency components, FreqTrans improves the representation of blurred tumor boundaries and subtle lesion regions.
- Suppressing noise and reducing frequency redundancy: The proposed frequency-domain attention selectively emphasizes informative channels while suppressing speckle noise and irrelevant spectral components.
- Balancing local and global feature integration: Through adaptive frequency fusion and a Multi-Scale Feature Extraction and Abstraction method, FreqTrans effectively combines fine edge details with global tissue context for more robust classification.

2. METHODS

2.1. Overview

The proposed FreqTrans framework, illustrated in Fig.1. The pipeline consists of four major components: (1) Fourier Preprocessing, which decomposes the input into low- and high-frequency sig-

*Corresponding authors

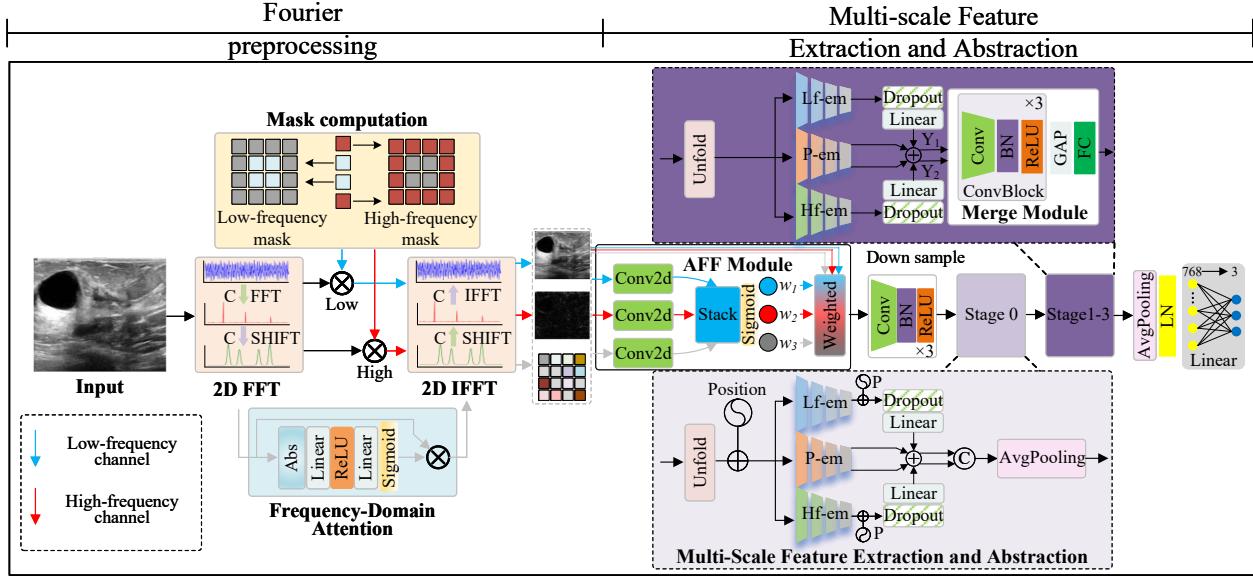


Fig. 1. Overall architecture of the proposed FreqTrans framework. It consists of four main components: (1) Fourier Preprocessing, which decomposes the input image into low- and high-frequency representations; (2) Frequency-Domain Attention, which selectively enhances informative spectral channels and suppresses noise; (3) Adaptive Frequency Fusion (AFF), which dynamically balances complementary frequency cues; (4) Multi-Scale Feature Extraction and Abstraction with Merge Module, which integrates local lesion edges and global tissue structures for robust classification.

nals; (2) Frequency-Domain Attention, which adaptively enhances spectral channels most relevant for tumor detection; (3) Adaptive Frequency Fusion, which dynamically balances complementary frequency cues; (4) Multi-Scale Feature Extraction and Abstraction with Merge Module, which integrates local lesion edges and global tissue structures for robust classification.

Unlike conventional CNNs or Vision Transformers that process only spatial features, FreqTrans explicitly models frequency-domain information, making it particularly effective for capturing both fine-grained boundaries (from high-frequency channels) and global context (from low-frequency channels).

2.2. Fourier Preprocessing

Breast ultrasound images often contain subtle lesions that are difficult to detect in the raw spatial domain. To alleviate this, we employ a Fourier transform to convert the image into the frequency domain:

$$F^s = \text{fftshift}(\mathcal{F}(x)), \quad (1)$$

where the shift operation centers the low-frequency spectrum. We then generate two complementary masks: a circular low-frequency mask and a high-frequency mask. These masks separate global structural patterns (low-frequency) from edges and microcalcifications (high-frequency). Applying the masks yields:

$$F_{low} = F^s \odot M_{low}, \quad F_{high} = F^s \odot M_{high}. \quad (2)$$

Inverse FFT recovers spatial maps x_{low} and x_{high} , which explicitly highlight global tissue distribution and fine lesion boundaries, as shown in Fig.1. The above preprocessing operation ensures that

the subsequent modules receive disentangled frequency cues instead of raw images. This step directly mitigates low contrast and blurred boundaries by separating global intensity distributions (low-frequency) from sharp structural edges (high-frequency), enabling clearer tumor boundary representation.

2.3. Frequency-Domain Attention

Although frequency decomposition provides rich information, not all channels contribute equally to lesion classification. Some frequency bands may contain irrelevant noise, while others emphasize diagnostically important structures. Inspired by prior studies on channel attention, we extend this idea to the frequency domain[19]. To address this, we compute per-channel frequency statistics:

$$s_c = \frac{1}{HW} \sum_{i,j} |F^s(c, i, j)|, \quad (3)$$

which measure the overall energy of each channel. These statistics are passed through a small MLP to produce attention weights $\alpha_c \in [0, 1]$.

The frequency attention mechanism enhances discriminative channels (e.g., strong edge responses around tumor boundaries) while suppressing uninformative ones. The enhanced spectrum is:

$$F^{attn} = F^s \odot \alpha^{attn}. \quad (4)$$

This module tackles noise interference and frequency redundancy by selectively amplifying informative channels (e.g., edges or calcification signals) while suppressing channels dominated by noise, thus reducing false activations.

2.4. Adaptive Frequency Fusion (AFF)

The low-, high-, and attention-enhanced frequency features provide complementary views of the image. However, their importance varies depending on the lesion type and imaging conditions. For example, micro-calcifications require high-frequency emphasis[20][21], while diffuse low-contrast lesions benefit from low-frequency dominance.

To achieve dynamic balance, we concatenate the three feature streams and compute adaptive weights via a lightweight attention mechanism:

$$F^{fused} = \alpha^{low} \cdot F_{low} + \beta^{high} \cdot F_{high} + F^{attn}. \quad (5)$$

In this method, α and β are adaptively learned for each image, allowing the network to flexibly emphasize edge cues or global structures, allowing the network to flexibly emphasize edge cues or global structures. This adaptivity is crucial for handling diverse ultrasound cases. This dynamic weighting mechanism alleviates the imbalance between local and global features, as the network adaptively emphasizes high-frequency detail for sharp boundaries or low-frequency smoothness for diffuse lesions.

2.5. Multi-Scale Feature Extraction and Abstraction

After adaptive frequency fusion, the network leverages multiple embeddings rather than a single fused feature. Specifically, three types of embeddings are constructed: low-frequency embeddings (X^{lf}), high-frequency embeddings (X^{hf}), and position-aware embeddings (X^p). These streams provide complementary perspectives: X^{lf} encodes global tissue structure, X^{hf} highlights sharp lesion boundaries, and X^p introduces positional context.

All embeddings are projected into patch tokens and processed by S parallel streams, each corresponding to a different receptive field or scale. For scale s , the representation is refined through a Transformer block:

$$X'_s = X_s + \text{MSA}(\text{LN}(X_s)) + \text{FFN}(\text{LN}(X_s)), \quad s = 1, 2, \dots, S. \quad (6)$$

Here, MSA denotes multi-head self-attention, FFN is the feed-forward network, and LN is layer normalization. This enables each stream to capture long-range dependencies while preserving scale-specific details.

To encourage cross-scale integration, outputs from all streams are aggregated via weighted fusion:

$$Z = \sum_{s=1}^S \gamma_s \cdot X'_s, \quad \sum_{s=1}^S \gamma_s = 1, \quad (7)$$

where γ_s are learnable weights adjusting the relative importance of each scale. The aggregated representation Z is then reshaped into spatial maps and refined with convolution:

$$F_{ms} = \text{ConvBlock}(Z). \quad (8)$$

Through this design, the module explicitly disentangles frequency-aware embeddings into low- and high-frequency branches while introducing positional context, before re-integrating them at multiple scales. This ensures that both global structural information and fine-grained boundaries are preserved, yielding a unified abstraction F_{ms} that is robust for lesion classification.

2.6. Merge Module

Finally, the outputs from the two main branches are merged to generate the final prediction. The frequency branch provides $Y_1 \equiv F^{fused}$

(adaptive frequency-enhanced representation), while the multi-scale spatial branch provides $Y_2 \equiv F_{ms}$ (multi-scale abstraction from low-, high-frequency, and position-aware embeddings). Both Y_1 and Y_2 are reshaped into spatial feature maps, concatenated, and passed through a convolutional fusion block:

$$F_{out} = \text{ConvBlock}([Y_1, Y_2]). \quad (9)$$

Global average pooling then aggregates the fused representation, and a fully connected layer outputs the classification logits:

$$\hat{y} = \text{FC}(\text{GAP}(F_{out})). \quad (10)$$

This hierarchical merging ensures that frequency-based cues (captured in F^{fused}) and spatial multi-scale abstractions (captured in F_{ms}) complement each other. The final representation thus leverages both enhanced spectral information and refined spatial context, leading to robust lesion classification even under low-contrast or noisy imaging conditions.

2.7. Summary

In summary, FreqTrans combines Fourier decomposition, adaptive spectral enhancement, and Multi-Scale Feature Extraction and Abstraction into a unified framework. Each module is tightly connected: Fourier preprocessing disentangles signals, frequency attention enhances them, AFF balances them, and the transformer learns complex interactions. This synergy enables the model to detect subtle and low-contrast lesions more effectively than conventional CNN or ViT baselines.

3. EXPERIMENTS AND ANALYSIS

3.1. Dataset

We evaluate our method on two widely used breast ultrasound datasets: UDIAT[22] and BUSI[23]. UDIAT contains 163 images, and BUSI consists of 647 images, including benign, malignant, and normal cases. For fair evaluation, we adopt a 4-fold cross-validation strategy on both datasets, with all images resized to 256×256 . Data augmentation, such as horizontal flipping and random rotation, is applied to improve generalization.

3.2. Implementation Details

All experiments are conducted on an NVIDIA RTX 4070. The proposed FreqTrans framework is trained from scratch without pre-training. We use the Adam optimizer with $\beta = (0.9, 0.999)$. The batch size is set to 32 and the learning rate is 1×10^{-6} . Training proceeds for 250 epochs, and the best model is selected based on validation F1-score. For fair comparison, all baseline CNN and Transformer models are trained under the same settings. The best results are highlighted in bold in the tables.

3.3. Overall Comparison

We first compare our proposed FreqTrans with representative CNN- and Transformer-based baselines on two public breast ultrasound datasets: UDIAT and BUSI. As illustrated in Fig.2, the original image, its low-frequency representation, and its high-frequency representation are presented from left to right. This visualization highlights how frequency decomposition separates global tissue structures from fine lesion boundaries, providing intuitive evidence for the effectiveness of our frequency-domain design and demonstrating why both components are essential for accurate lesion detection. Tables 1 and 2 present the performance metrics of FreqTrans

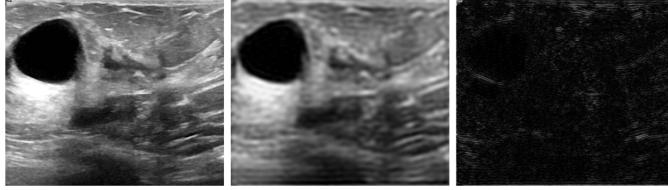


Fig. 2. From left to right: original image, low-frequency representation, and high-frequency representation.

and baseline methods. Our model consistently achieves higher recall and F1-score, indicating its strength in detecting subtle lesions and maintaining balanced sensitivity and precision.

Table 1. Comparison of different methods on UDIAT.

Method	Acc	Auc	Rec	F1-score
Vgg16	0.7632	0.8710	0.7400	0.7614
ResNet50	0.7885	0.9063	0.7269	0.7539
Vit	0.7838	0.8730	0.7460	0.7625
Efficient3	0.7568	0.8946	0.6333	0.7227
Gsm	0.8108	0.8612	0.7667	0.7729
FreqTrans (Ours)	0.8205	0.8545	0.8274	0.8183

Table 2. Comparison of different methods on BUSI.

Method	Acc	Auc	Rec	F1-score
Vgg16	0.7949	0.8619	0.7677	0.7775
ResNet50	0.7895	0.8972	0.7575	0.7920
Vit	0.8000	0.9306	0.8181	0.7944
Efficient3	0.8129	0.9690	0.6901	0.7681
Gsm	0.8258	0.9335	0.7435	0.7845
FreqTrans (Ours)	0.8462	0.9156	0.8274	0.8372

In both datasets, FreqTrans achieves the best recall, which is particularly important for clinical applications, as missing malignant lesions can lead to severe consequences. While the AUC does not always increase proportionally, this outcome results from the design of our model. Specifically, FreqTrans emphasizes enhancing high-frequency components that capture subtle lesion boundaries and low-frequency components that highlight global tissue context. This design inherently prioritizes sensitivity to small or low-contrast lesions, leading the model to predict more positive cases. Although this improves recall, it may also introduce additional false positives, which in turn reduces specificity and slightly lowers the overall AUC.

Such a trade-off between sensitivity and AUC has also been widely discussed in prior medical imaging studies. In highly imbalanced datasets, or when the clinical priority is to avoid missing any positive cases, models often deliberately bias toward high recall at the expense of precision[24]. As several works have noted, ROC-based AUC can remain relatively stable even when false positives increase, whereas precision-recall metrics provide a more sensitive reflection of performance under class imbalance[25]. Consequently, our results are consistent with this phenomenon: FreqTrans achieves higher sensitivity by capturing subtle lesion cues, while a moderate decrease in specificity slightly lowers AUC but aligns with the overarching clinical objective of minimizing false negatives.

3.4. Ablation Study

To evaluate the contribution of each proposed module in our framework, we conduct ablation experiments on the UDIAT dataset. The tested variants include: (1) baseline backbone only, (2) + Fourier Preprocessing (FP)[26][27], (3) + FP and Adaptive Frequency Fusion (AFF), (4) + FP+AFF with Frequency Attention (FA)[28], and (5) the full FreqTrans model. The results demonstrate that each module progressively enhances performance: FP improves feature representation by transforming data into the frequency domain, AFF balances low- and high-frequency information to further boost recall, and FA adaptively emphasizes discriminative spectral regions. Finally, the complete FreqTrans achieves the best results, confirming the complementary benefits of all proposed modules.

Table 3. Ablation study on Dataset BUSI.

Method	Acc	Rec	F1-score
Baseline (Backbone only)	0.7483	0.6944	0.6944
+ FP	0.7631	0.7231	0.7588
+ FP + AFF	0.8000	0.7534	0.7682
+ FP+ AFF +FA	0.8102	0.7765	0.7873
+ Full Model (Ours)	0.8205	0.8274	0.8183

3.5. Module-Level Insights

To quantitatively assess the contribution of each module in FreqTrans, we performed ablation experiments on the BUSI dataset (Table 3). Compared with the baseline backbone, adding Fourier Preprocessing (FP) improves accuracy from 0.7483 to 0.7631 and recall from 0.6944 to 0.7231, showing that transforming inputs into the frequency domain enhances representation. Adaptive Frequency Fusion (AFF) further raises accuracy to 0.8000 and recall to 0.7534, indicating that balancing low- and high-frequency information helps capture both global structures and fine lesion details. With Frequency Attention (FA), accuracy reaches 0.8102 and recall 0.7765, highlighting the benefit of emphasizing discriminative spectral components while suppressing noise. Finally, the full FreqTrans achieves the best results (Acc: 0.8205, Rec: 0.8274, F1: 0.8183), confirming that FP, AFF, and FA provide complementary benefits. The progressive gains demonstrate the value of each module for lesion detection.

4. CONCLUSION

In this paper, we proposed FreqTrans, a Multi-Scale Feature Extraction and Abstraction model with frequency-domain embedding for breast ultrasound classification. By decomposing images into low- and high-frequency components, our method enhances edge details and global context simultaneously. The Adaptive Frequency Fusion (AFF) and Frequency Attention (FA) modules further balance spectral information and highlight discriminative regions.

Extensive experiments on two public datasets (UDIAT and BUSI) show that FreqTrans outperforms CNNs and vision transformers in recall and F1-score, which are critical for detecting subtle and low-contrast lesions. While some baselines achieve higher AUC in specific cases, our method offers the best trade-off between sensitivity and precision, demonstrating robustness under diverse imaging conditions.

In summary, the frequency-domain design integrates global spectral context with local spatial cues, offering a new perspective for medical image analysis. Future work will extend FreqTrans to larger datasets and explore its application to other imaging modalities.

5. REFERENCES

- [1] J. G. Elmore, S. M. Reisch, L. Carney, K. Abraham, A. B. Fosse, B. M. Longton, and C. D. Nelson, “Diagnostic Concordance Among Pathologists Interpreting Breast Biopsy Specimens,” *JAMA*, vol. 313, no. 11, pp. 1122–1132, 2015.
- [2] Q. Huang, Y. Luo, and C. Zhang, “Breast Ultrasound Image Segmentation: A Survey,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, pp. 197–208, 2020.
- [3] H. R. Koo, Y. Cho, J. Song, and H. Park, “Noise Reduction in Breast Ultrasound Imaging: A Review,” *Ultrasonography*, vol. 38, no. 3, pp. 244–252, 2019.
- [4] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, et al., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” in *International Conference on Learning Representations (ICLR)*, 2021.
- [5] X. Chen, S. Xie, and K. He, “An Empirical Study of Training Self-Supervised Vision Transformers,” in *International Conference on Computer Vision (ICCV)*, pp. 9640–9649, 2021.
- [6] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, “A Survey on Deep Learning in Medical Image Analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [7] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” in *International Conference on Learning Representations (ICLR)*, 2015.
- [8] S. Wang, Y. Zhang, and X. Li, “Noise-Robustness Test for Ultrasound Breast Nodule Neural Network,” *IEEE Transactions on Medical Imaging*, vol. 42, no. 5, pp. 1301–1312, 2023.
- [9] Yuhao Mo, Chu Han, Yu Liu, Min Liu, Zhenwei Shi, Jiatai Lin, Bingchao Zhao, Chunwang Huang, Bingjiang Qiu, Yanfen Cui, Lei Wu, Xipeng Pan, Zeyan Xu, Xiaomei Huang, Zaiyi Liu, Ying Wang, Changhong Liang, “HoVer-Trans: Anatomy-aware HoVer-Transformer for ROI-free Breast Cancer Diagnosis in Ultrasound Images,” *IEEE Transactions on Medical Imaging*, vol. 42, no. 5, pp. 1696–1706, 2023.
- [10] Y. Feng, H. Zhang, X. Li, and J. Wang, “EH-former: Regional easy-hard-aware transformer for breast lesion segmentation,” *Expert Systems with Applications*, vol. 202, p. 117254, 2024.
- [11] F. Liu, M. Tan, and J. Zhou, “Ultrasound Signal Processing: From Models to Deep Learning,” *Ultrasonics*, vol. 122, p. 106768, 2022.
- [12] Y. Li, X. Zhang, and J. Wang, “A survey on deep learning in medical ultrasound imaging,” *Frontiers in Physics*, vol. 12, p. 1398393, 2024.
- [13] R. Acharya, O. Faust, S. Molinari, L. S. Sree, and J. S. Suri, “Nonlinear Analysis of Ultrasound Images for Breast Cancer Diagnosis,” *Medical and Biological Engineering and Computing*, vol. 46, no. 12, pp. 999–1009, 2008.
- [14] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I.-C. Chang, “Deep Learning of Feature Representation with Multiple Instance Learning for Medical Image Analysis,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 12, pp. 2794–2805, 2017.
- [15] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, “Perceptual Blur and Ringing Metrics: Application to JPEG2000,” *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 163–172, 2004.
- [16] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19, 2018.
- [17] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10012–10022, 2021.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.
- [19] J. Hu, L. Shen, and G. Sun, “Squeeze-and-Excitation Networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7132–7141, 2018.
- [20] Y. Ouyang, Z. Zhou, W. Wu, J. Tian, F. Xu, S. Wu, and P.-H. Tsui, “A review of ultrasound detection methods for breast microcalcification,” *Mathematical Biosciences and Engineering*, vol. 16, no. 4, pp. 1761–1785, 2019.
- [21] K. Frick, T. M. L. van der, A. J. M. Nooijen-Parker, et al., “Pre-operative localization of breast microcalcification using high-frequency ultrasound,” *Clinical Radiology*, vol. 52, no. 12, pp. 924–926, 1997.
- [22] A. Oliver, J. Freixenet, J. Martí, E. Pérez, J. Pont, E. R. Denton, and R. Zwiggelaar, “A Novel Breast Tissue Density Classification Methodology,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 55–65, 2008.
- [23] M. Al-Dhabayani, M. Gomaa, H. Khaled, and A. Fahmy, “Dataset of Breast Ultrasound Images,” *Data in Brief*, vol. 28, p. 104863, 2020.
- [24] Saito, T. and Rehmsmeier, M., “The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets,” *PLOS ONE*, vol. 10, no. 3, e0118432, 2015.
- [25] Lever, J., Krzywinski, M., and Altman, N., “Classification Evaluation: Part 2 — ROC, Precision-Recall and Imbalanced Data,” *Nature Methods*, vol. 13, pp. 132–133, 2016.
- [26] Y. Chen, H. Li, Z. Wang, and X. Liu, “Syn-Net: A Synchronous Frequency-Perception Fusion Network for Breast Tumor Segmentation in Ultrasound Images,” *Computers in Biology and Medicine*, vol. 165, p. 107393, 2023.
- [27] S. Ying, X. Song, and H. Wang, “High-frequency-based Multispectral Attention for Domain Generalization,” *Artificial Intelligence Review*, pp. 1–23, 2025.
- [28] M. Zhang, L. Zhao, and Q. Huang, “FA-Unet: A Deep Learning Method with Fusion of Frequency Domain Features for Fruit Leaf Disease Identification,” *Horticulturae*, vol. 11, no. 7, p. 783, 2023.