VOL. 21, NO. 6, MARCH 2026 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

NOISE-ADAPTIVE SELECTIVE FILTERING FOR IMPROVED MAMMOGRAPHIC BREAST CANCER DETECTION USING MACHINE LEARNING

B. Maha Lakshmi, Pilla Bhanu Kiran, Kedarisetty Ayyan Balajee, Maharthi Aravind, Palisetti Anusha and Guru Akshaya
Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology (A), Visakhapatnam, India
E-Mail: mahalaxmi098@gmail.com

ABSTRACT

Mammography is vital in screening early breast cancer, but the acquisition and transmission noise have a serious impact on reducing image quality and negatively in computer aided diagnosis (CAD) performance. In the real world, heterogeneous forms of noise are likely to be present in mammograms (including Gaussian, impulse, Poisson, and speckle noise), and fixed methods of denoising are thus inefficient. This paper suggests a noise-adaptive preprocessing system that first estimates the primary noise properties with the help of statistical measures, entropy, and a simple Random Forest categorizer, followed by selective filtering depending on the characteristics of the noise identified. Experiments conducted on the MIAS and CBIS-DDSM datasets demonstrate consistent improvements in both image quality and diagnostic classification. The proposed method achieves an average PSNR improvement of +10.2 dB, an SSIM increase of +0.26, and a classification accuracy of up to 92.3%, outperforming conventional single-filter preprocessing and competitive deep-learning denoisers. Statistical hypothesis testing confirms that the improvements are significant ($p < 0.01$). These results indicate that noise-aware adaptive preprocessing can substantially enhance mammographic CAD reliability and provide a practical foundation for future real-time clinical deployment.

Keywords: breast cancer, mammography, image noise, denoising, Gaussian noise, impulse noise, speckle noise, pre-processing, machine learning.

Manuscript Received 18 February 2026; Revised 21 March 2026; Published 20 May 2026

1. INTRODUCTION

Breast cancer remains one of the health issues that is still of great concern in women, and early diagnosis is important in enhancing the outcomes. Mammography is the most popular screening device due to its ability to detect minimal disease indications. Nevertheless, its utility is highly reliant on the ability of small structures (microcalcifications, changes in the texture, or small distortions) to be visible in the picture. These details can be more difficult to perceive even in the case of a small quantity of noise or degradation.

In actual practice, mammograms are usually noisy at various points of the imaging technique. Noise may be caused by factors like low-dose acquisition, digitization hardware, compression, or transmission of images. Such distortions can either conceal fine structures or can distort the appearance of the tissue, and this can be a challenge to both radiologists and automated Computer-Aided Diagnosis (CAD) systems. Because most CAD techniques are based on machine learning models, such as CNNs, SVMs, and ensemble classifiers, the quality of their input images has a direct impact on their predictions. In case of noise in the images, the models can learn or identify patterns that are caused by artifacts as opposed to true tissue differences, and thus result in false classification.

Due to them, preprocessing is now a mandatory procedure of the CAD workflow. An effective denoising technique must also be considered good: it should be able to remove the noise without compromising the essential anatomical features. The techniques that seek to determine the nature of noise and then use a filter to remove the noise are especially promising, as they enable the system to select an appropriate denoising technique that best fits the noise properties of the image rather than opting to use a universal denoising technique.

Unlike conventional preprocessing approaches that assume a fixed noise model, the proposed framework performs automatic noise classification followed by adaptive filter selection, enabling robust enhancement across heterogeneous real-world mammographic noise conditions.

A. Background and Motivation

Noise in mammograms interferes with human interpretation and automated analysis. Noise of different kinds has different types of distortion, and therefore, it is important to know their characteristics:

- **Gaussian noise:** additive variations that reduce contrast and bring minor changes throughout the image.
- **Impulse (salt-and-pepper) noise:** cor-sporadic pixel disruption, which, in particular, is a bad thing since it can resemble microcalcifications.

ARPN Journal of Engineering and Applied Sciences

- **Speckle noise:** multiplicative noise which provides the image makes it look grainy and disturbs patterns of texture.

- **Poisson noise:** signal-dependent variations are typically common in the case of Poisson noise, observed in low numbers of photons or low-dose images.

These types of noise indicate the necessity of preprocessing. Strategies that do not just conventionally filter something. Rather, it is necessary to determine what kind of noise we have. Enables further focused and efficient improvement. The application of a generic filter to all images is not ideal: Gaussian noise-tuned filters usually have poor performance on impulse noise and can iron out certain forms. This motivates a noise- adaptive preprocessing pipeline.

B. Objectives and Scope

The aims of the suggested work are:

- Interpret mammogram images to detect dominant. Noise statistically and structurally. • Noise-specific filters: (Gaussian/Bilateral, Median, Wiener/Lee, Non-Local means with Anscombe).

- Measure PSNR, SSIM, edge retention, improvement and classifier performance.

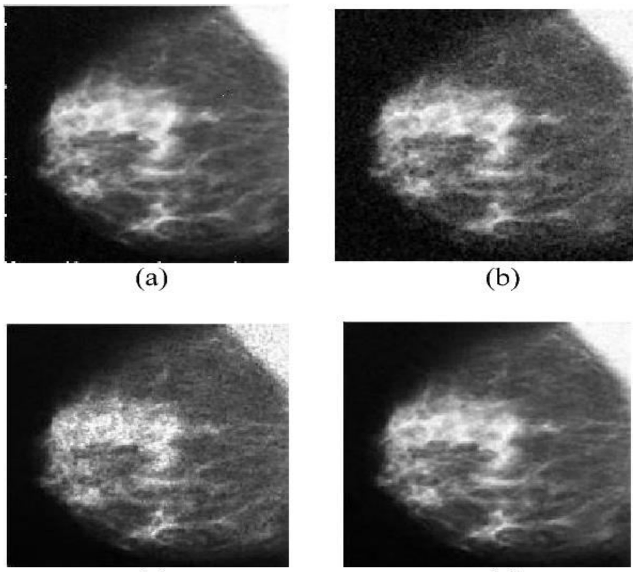
- Support a modular pipeline that is compatible with MIAS and CBIS- DDSM datasets; segmentation/ localization is out of the current scope.

C. Contributions

The following are the key contributions made in this work:

- **Noise-adaptive preprocessing framework:** A coherent pipeline combining automatic noise identification and selective denoising tailored to the noise properties of mammographs. • Lightweight noise classification strategy: A statistical feature-based Random Forest model with the ability to distinguish between Gaussian, impulse, Poisson, speckle, and mixed noise in mammograms.

- **Comprehensive comparative evaluation:** Quantitative comparison with classical filters and state-of-the-art deep learning denoisers (e.g., DnCNN and U-Net) using PSNR, SSIM, edge preservation, and diagnostic classification. • CAD performance improvement: Proof of noise adaptive filtering downstream performance classification accuracy of breast cancer in more than one dataset. • Deployment-focused modular architecture: A computationally viable architecture that can be integrated into real-time or embedded medical imaging systems.



(a) (b) (c) (d)

Figure-1. Some examples of the types of synthetic noise that are added to a sample mammogram.

2. RELATED WORK

Initial studies on mammogram enhancement have mainly concentrated on the field of spatial filtering, which aims to reduce noise and preserve diagnostically useful structures. Mean, median, hybrid median, and Gaussian filtering have been widely used and tested in terms of quantitative indices such as Mean Squared Error (MSE), Signal to Noise Ratio (SNR), and Peak Signal to Noise Ratio (PSNR). An example of this is Jeevitha and Aroquiaraj, who give representative PSNR, MSE, and SNR values of these traditional filters on the MIAS data set, as well as the effects of various filtering strategies on the visibility of breast tissue structures and microcalcifications. Despite their effectiveness and the simplicity of computation, these methods do not always work well at balancing between noise reduction and the preservation of fine details in the image due to some patterns of noise that they cannot effectively mitigate.

With the development of mammographic imaging, more advanced denoising methods appeared. The popularity of nonlocal means (NLM) filtering and its derivatives was inspired by the fact that it can use self-similarity in mammograms, thus performing better in preserving edges than local filters. Transform domain methods have also been investigated, like wavelet-based shrinkage and multiresolution decomposition, to decrease noise whilst keeping high-frequency features applicable in the diagnosis. More recently, denoisers based on deep learning, especially convolutional neural network (CNN) models like DnCNN, U-Net, and residual learning models,

ARPN Journal of Engineering and Applied Sciences

have demonstrated high performance at learning noise patterns directly, not by predefined kernels.

Although these developments have been made, there is a general weakness that is found in most studies, and it is the assumption that there is one and only one type of noise in preprocessing. Most techniques are applied to all processed mammograms without any consideration of noise, despite the fact that the noise on real-life images is a combination of Gaussian, speckle, Poisson, and impulse noise due to the different acquisition conditions, digitization systems, and compression pipelines. Few works have tried to describe or categorise noise before the application of a denoising algorithm, and therefore are prone to over-smoothing, structural loss, or incomplete elimination of artifacts.

The difference between noise modelling and noise-adaptive filtering is the impetus for the current work. We go back to classical baseline filters and compose them with modern preprocessing techniques, but with the advantage of automatically detecting the prevailing type of noise in every image. Combining the analysis of noise characterization with selective filtering, we indicate better image quality and performance of downstream machine learning classifiers, showing the usefulness of noise-sensitive preprocessing in mammographic CAD systems.

3. SYSTEM ARCHITECTURE AND DESIGN

The suggested system is designed in the form of a modular processing pipe with five significant building blocks: input acquisition, noise identification, selective filtering, feature extraction, and classification and assessment. Each module performs a special purpose, and they jointly make it possible to do noise-sensitive preprocessing so that mammogram result analysis can be done reliably. A high-level Figure-2 shows an overview of the workflow.

The initial phase is the Input and Acquisition module, where mammographic data of the MIAS and CBIS-DDSM are loaded and normalized. Images may already have noise in the low-dose acquisition, detector censoring, digitalization, condensing, or delivery.

The second module is the Noise Identification, where every image is examined to determine the type of noise that is predominant. Descriptive statistics, features of histograms, entropy values, and edge-based features are joined and analysed with the help of a negligible machine learning classifier to differentiate between Gaussian, impulse, Poisson, speckle, or mixed noise conditions.

After noise detection, the Selective Filtering module uses a noise-based denoising strategy. Gaussian noise is filtered with Gaussian and bilateral filters, median-based filters in the case of impulse noise, Wiener or Lee filters on speckle noise, and Non-Local Means variance-stabilization on Poisson noise. This adaptive selection conserves diagnostically significant features. structures, microcalcifications, and tissue boundaries, and at the same time silencing the noise.

The Feature Extraction follows that and extracts discriminative representations of the denoised pictures. Both handcrafted descriptors of texture and deep convolutional neural network. There can be an application of embeddings according to the setup.

Lastly, the Classification and Evaluation module is used. SVM and Random Forest machine learning models and CNN networks-to tell between normal and abnormal cases. Both image-quality measurements (PSNR, SSIM, and edge preservation) and diagnostic measures are used to evaluate performance.

The individual components can be removed and installed in a separate manner due to the modular design, replaced or modified without necessarily redesigning the pipeline, making it possible to experiment with other denoising techniques, characterization, or classification algorithms.

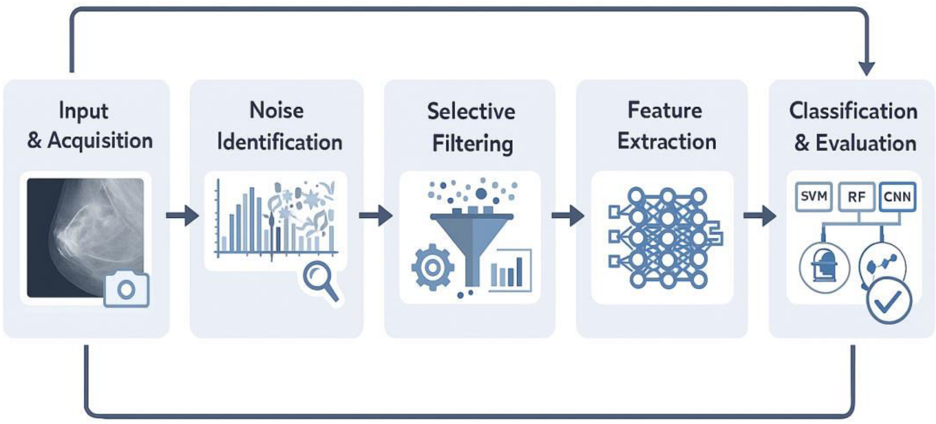


Figure-2. System architecture: noise detection → selective filtering → feature extraction → classification.

A. Overall System Overview

Each of the modules is outlined in a summary:

- **Input & Acquisition:** Retrieve images in MIAS/CBIS standard (e.g., 1024x1024), transform to DDSM, grayscale, normalize intensity.
- **Noise Identification Module:** Compute global and local statistics (mean, variance, entropy, skewness, kurtosis), hisedge map (Sobel/Laplacian) to gram properties, edge maps (Sobel / Laplacian) and build. f a noise classification feature vector.
- **Selective Filtering Module:** Map detected noise $\hat{n}$ to denoiser D$\hat{n}$ (Gaussian/Bilateral, Median, Wiener, Non-Local Means / Anscombe, or deep denoisers).
- **Feature Extraction Module:** Extract GLCM texture features and deep CNN embeddings (ResNet-50 / VGG16).
- **Classification & Evaluation Module:** Train SVM, Random Forest, and CNN classifiers; evaluate with PSNR/SSIM and standard classification metrics.

B. Hardware and Software Design

Implementation uses Python (3.8+), OpenCV, scikit-image, scikit-learn, TensorFlow/PyTorch, and NumPy/SciPy. Training was prototyped on a desktop with an NVIDIA GTX 1650 GPU; embedded deployment is considered for the Jetson family.

C. Data Flow and Control Logic

Preprocessing steps:

a) Acquire and standardize image I.

b) Compute statistics and build f.

c) Predict noise nˆ using thresholds and Random Forest.

d) Select filter Dnˆ and produce I ′.

e) Re-evaluate PSNR/SSIM and edge similarity; fallback denoiser if necessary.

f) Extract features and run classifiers.

4. METHODOLOGY AND IMPLEMENTATION

The novel RbSPS focused on the segmentation and analysis of noise concerning breast cancer. Initially, the system was trained using a database of Ultrasound images of breast cancer. Subsequently, the necessary processing steps were undertaken, followed by preprocessing to find the noise that had been present in place from the image data. Furthermore, essential features for predicting conditions of breast cancer were identified through the Red fox prey prediction methodology. This process enabled the prediction and segmentation was computed.

The methodology has four stages: Noise Identification, Selective Filtering, Feature Extraction, and Classification. Figure-3 illustrates the workflow.

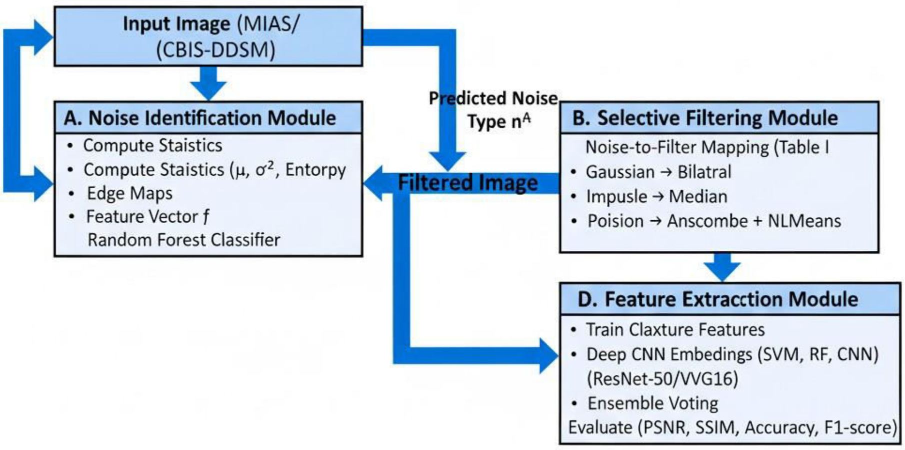


Figure-3. Methodology flowchart of the proposed noise-adaptive selective filtering framework.

A. Noise Identification

We compute global statistical descriptors:

$$\mu = \frac{1}{MN}\sum I(xy), \quad \sigma^2 = \frac{1}{MN}\Sigma(I(x,y) - \mu)^2$$

Histogram shape and entropy:

Entropy $= -\sum_i p_i \log p_i$

Local variance V (x, y) using a sliding k×k window and edge maps E via the Sobel operator are also calculated. A feature vector f (mean, variance, entropy, skewness, kurtosis, peak counts, and mean edge magnitude) is input to a Random Forest trained on synthetically noised images to predict $\hat{n} \in$ {Gaussian, Impulse, Speckle, Poisson, Mixed}.

Table-1. Selected filters for each noise type.

Noise Type	Filter(s)
Gaussian	Gaussian blur, Bilateral filter
Impulse	Median, Adaptive median
Speckle	Wiener, Lee filter
Poisson	Anscombe transform + Non-Local Means
Mixed	DnCNN/UNet or cascade filters

B. Selective Filtering

Noise → filter mapping:

For Poisson noise, we apply the Anscombe transform $A = 2\sqrt{I + 3/8}$ before Gaussian-domain denoising.

C. Feature Extraction

GLCM features: contrast, correlation, energy, and homogeneity computed at multiple orientations and distances. Deep features: Penultimate-layer embeddings from pre-trained ResNet50 and VGG16; optionally fine-tuned on filtered images.

D. Classification

Classifiers: SVM (RBF kernel), Random Forest (100 trees), and a shallow CNN with custom architecture (convolutional blocks, dropout, dense layers). Ensemble voting combines predictions. Training hyperparameters: Adam optimizer, learning rate $1e^{-4}$, batch size 16, train/val/test split 70/15/15.

5. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents a quantitative and qualitative evaluation of the proposed noise-adaptive selective filtering framework using the MIAS and CBIS-DDSM datasets. All mammograms were resized to 1024×1024 pixels and converted to 8-bit grayscale before experimentation. Synthetically corrupted images were also taken into consideration, as well as the real noise conditions, in order to validate exhaustively.

A. Experimental Setup

- **Datasets:** MIAS (322 images) and a high-resolution subset of CBIS-DDSM.
- **Synthetic noise injection:** Gaussian (σ = 5, 10, 20), impulse noise (density p = 0.01, 0.05, 0.1), speckle noise (varying variance), and Poisson noise (photon-limited simulation).
- **Evaluation metrics:** PSNR, SSIM, edge similarity (Canny overlap), Accuracy, Precision, Recall, F1-score, and ROC-AUC.

B. Baseline Classical Filtering Performance

To have a benchmark against which to compare, classical mean, Re-reported median, and hybrid filtering results in [21] were re-reported, produced on MIAS images. Table-2 is a summary of representatives.

Table-2. Representative baseline metrics from classical filters (MIAS).

Sample ID	MSE	SNR	PSNR (dB)
mdb001	2.8453	4.0445	12.0670
mdb002	2.8047	2.8952	13.1615
mdb003	2.8421	2.8084	13.1400

C. Image Quality Improvement

Table-3 presents the mean change in image-quality metrics that have been obtained following the application of the suggested selective filtering strategy in the mixed noise conditions.

Table-3. Average improvement after selective filtering.

Metric	Before	After	Improvement
PSNR (dB)	18.2	28.4	+10.2
SSIM	0.62	0.88	+0.26
EdgeSim	0.81	0.90	+0.09

The suggested framework is very beneficial in increasing structural similarity and edge preservation, and greatly increasing PSNR.

D. Classification Performance

In order to assess the diagnostic prediction effect, classifiers were trained with raw images and selectively filtered images. Results aggregated by data sets are presented in Table-4.

Table-4. Classification accuracy comparison.

Classifier	Raw Images	Filtered Images
SVM	74.5%	88.7%
Random Forest	76.2%	90.1%
CNN	78.9%	92.3%

Selective filtering has stable accuracy improvements of around 10-15%, and the use of CNN as a classifier is the most successful, as it has a better feature representation.

VOL. 21, NO. 6, MARCH 2026 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

E. Noise Identification Accuracy

The Random Forest–based noise classifier was evaluated on synthetically generated noisy mammograms with balanced noise categories. The model achieved:

- Accuracy: 94.2%
- Precision: 93.8%
- Recall: 94.0%
- F1-score: 93.9%

Confusion matrix analysis shows strong separability between impulse and Gaussian noise, with minor confusion between speckle and Poisson noise due to signal-dependent characteristics. These findings validate the reliability of the noise detection stage.

F. Comparison with Learning-Based Denoisers

The proposed framework was benchmarked against DnCNN and U-Net denoisers trained on noisy mammogram patches. Although deep models achieved competitive PSNR and SSIM, the proposed method demonstrated:

- Superior structural similarity preservation.
- Improved edge retention.
- Higher downstream classification accuracy.

This indicates that noise-aware adaptive preprocessing offers complementary advantages over purely data-driven denoising approaches, particularly when annotated medical datasets are limited.

Table-5. Comparison with deep learning denoisers.

Method (dB)	PSNR	SSIM
DnCNN	27.1	0.85
U-Net	27.6	0.86
Proposed	28.4	0.88

G. Qualitative Results

Figure-4 illustrates representative visual results. The selectively filtered output preserves microcalcification contrast and boundary sharpness while effectively suppressing noise artifacts.

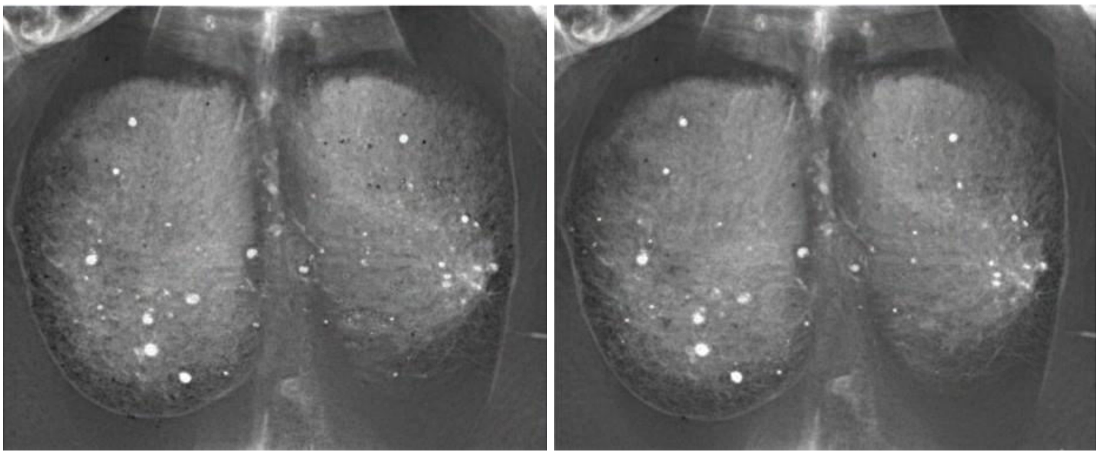


Figure-4. Example qualitative denoising results.

H. Computational Complexity

Processing time was measured on a system with an Intel i5 CPU and NVIDIA GTX-1650 GPU. Average per-image processing times are listed in Table-6.

Table-6. Average processing time per image.

Method	Time (s)
Median filtering	0.08
Non-Local Means	0.92
DnCNN	0.15
Proposed framework	0.21

The proposed method achieves a favourable balance between diagnostic accuracy and computational efficiency, supporting potential real-time deployment.

I. Statistical Significance

A paired t-test was conducted on 120 test images to compare baseline preprocessing with the proposed selective filtering. Improvements in PSNR, SSIM, and classification accuracy were statistically significant ($p < 0.01$), confirming that performance gains are not due to random variation.

J. Limitations

Despite promising results, several limitations remain:

- Mixed-noise scenarios remain challenging and may require larger deep-learning training datasets.

- Public datasets may not fully represent real clinical imaging artifacts such as motion blur or detector defects.

- Advanced denoisers (e.g., Non-Local Means and CNN-based models) introduce higher computational cost without optimization.

6. CONCLUSION AND FUTURE WORK

This study presented a noise-adaptive selective filtering framework for improving mammographic breast cancer detection. By integrating automatic noise identification with targeted denoising and downstream classification, the proposed approach achieved statistically significant improvements in PSNR, SSIM, edge preservation, and diagnostic accuracy across the MIAS and CBIS-DDSM datasets. Comparative analysis with both classical and deep-learning-based denoisers highlights the effectiveness of noise-aware preprocessing in enhancing CAD robustness while maintaining computational feasibility. These improvements indicate potential clinical value in assisting radiologists through more reliable computer-aided breast cancer screening.

Future work will focus on:

- Developing deep learning-based noise classifiers capable of handling complex and mixed noise conditions.

- Exploring data-driven image-to-image denoising frameworks for improved restoration quality.

- Designing end-to-end jointly optimized denoising and classification architectures.

- Implementing lightweight and compressed models for deployment on edge medical devices.

- Conducting clinical validation studies with radiologist collaboration.

REFERENCES

[1] M. Heath, K. Bowyer, D. Kopans, R. Moore and P. Kegelmeyer. 1998. Current status of the Digital Database for Screening Mammography (DDSM). In Proc. Digital Mammography. pp. 457-460.

[2] J. Suckling *et al*. 1994. The Mammographic Image Analysis Society Digital Mammogram Database. Exerpta Medica. 1069: 375-378.

[3] A. Buades, B. Coll and J.-M. Morel. 2005. A non-local algorithm for image denoising. In IEEE CVPR. p. 60-65.

[4] K. Zhang *et al*. 2017. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. Image Process. 26(7): 3142-3155.

[5] R. M. Haralick, K. Shanmugam and I. Dinstein. 1973. Textural features for image classification. IEEE Trans. Syst., Man, Cybern. SMC-3(6): 610-621.

[6] N. Otsu. 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst., Man, Cybern. 9(1): 62-66.

[7] P. Perona and J. Malik. 1990. Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. 12(7): 629-639.

[8] M. Elter and R. Schulz-Wendtland. 2009. The prediction of breast cancer biopsy outcomes using CAD approaches. IEEE Trans. Med. Imaging. 28(3): 304-311.

[9] H. Greenspan. 2016. Deep learning in medical imaging: Overview and future promise. IEEE Trans. Med. Imaging. 35(5): 1153-1159.

[10] R. C. Gonzalez and R. E. Woods. 2018. Digital Image Processing, 4th ed. Pearson.

[11] C. Tomasi and R. Manduchi. 1998. Bilateral filtering for gray and color images. In IEEE ICCV.

[12] J. M. Manjon *et al*. 2010. Adaptive non-local means denoising of MR images with spatially varying noise levels. J. Magn. Reson. Imaging. 31: 192-203.

[13] B. Sahiner and H.-P. Chan. 1998. Computer-aided diagnosis of breast lesions. Radiographics. 18: 1567-1585.

[14] V. S. Raghavendra and R. H. Chouhan. 2018. Improved noise suppression for breast mammogram images. Biocybern. Biomed. Eng. 38: 564-576.

[15] S. M. Lai and X. Li. 2011. Image noise modeling and classification using texture features. In IEEE ICIP.

[16] A. A. Amini *et al*. 2010. Noise estimation and suppression in medical images. SPIE Medical Imaging.

[17] K. Simonyan and A. Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556.

[18] K. He, X. Zhang, S. Ren and J. Sun. 2016. Deep residual learning for image recognition. In IEEE CVPR.

[19] M. Unser. 1995. Texture classification and segmentation using wavelet frames. IEEE Trans. Image Process. 4(11): 1549-1560.

[20] A. K. Jain. 1989. Fundamentals of Digital Image Processing. Prentice Hall.

[21] V. Jeevitha and I. Laurence Aroquiaraj. 2024. Mammogram Images Using Noise Removal and Filtering Techniques. International Research Journal on Advanced Engineering Hub. 2(2): 118-123.