Breast Cancer Computer-Aided Diagnosis using SVM, CNN, ResNet-18, and EfficientNet-B0

Anonymous Author(s)
Affiliation
Address
email

Abstract

Breast cancer is a leading cause of cancer mortality among women, and effective computer-aided diagnosis (CAD) systems can assist clinicians in early detection. In this project, we explore both classical machine learning and deep learning approaches for breast lesion classification using two imaging modalities: mammography and ultrasound. We implement and evaluate a Support Vector Machine (SVM) classifier and three convolutional neural network (CNN) models (a custom CNN, ResNet-18, and EfficientNet-B0) entirely within MATLAB. Experiments on mammography and ultrasound datasets (benign vs. malignant classification) show that deep learning methods outperform the classical SVM. On the mammography data, the best performance is achieved by EfficientNet-B0 with a validation accuracy of about 64%, compared to $\sim$58% for ResNet-18, $\sim$55–58% for a custom CNN, and $\sim$55% for SVM. Ultrasound experiments similarly demonstrate that a CNN yields higher accuracy than SVM, though overall accuracy on both modalities is moderate. We present training progress curves and confusion matrices to analyze each model's performance. The results highlight both the potential and the challenges of applying advanced deep learning models for breast cancer CAD, especially given limited data. All modeling and evaluation were conducted using MATLAB, demonstrating the feasibility of rapid experimentation for medical image classification.

1 Introduction

Breast cancer detection from medical images is a critical application of computer-aided diagnosis (CAD), aiming to assist radiologists in distinguishing malignant tumors from benign findings. Mammography and ultrasound are two common imaging modalities for breast cancer screening and diagnosis. Mammograms (X-ray images of the breast) are widely used for screening, while ultrasound imaging serves as an adjunct, especially for dense breast tissue or for further characterization of suspicious lesions.

Traditional CAD systems often rely on handcrafted features and classical machine learning classifiers such as Support Vector Machines (SVMs) [1]. These methods depend heavily on feature engineering: texture descriptors, shape features, and intensity statistics must be designed and tuned for a given task. This process is time-consuming and may fail to capture subtle visual cues present in medical images.

In contrast, deep learning—particularly convolutional neural networks (CNNs)—has revolutionized image recognition by learning hierarchical feature representations directly from data. CNN-based models have achieved state-of-the-art performance in many image classification tasks, including medical image analysis. Architectures such as ResNet-18 [2] introduced residual learning to train very deep networks effectively, while the EfficientNet family [3] improves accuracy by compound scaling of depth, width, and resolution.

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

In this work, we investigate and compare the performance of a classical SVM and several deep learning models for breast cancer CAD on mammography and ultrasound image datasets. All experiments are performed using MATLAB (R2023a) and its Deep Learning Toolbox for model implementation, training, and evaluation. We aim to answer the following questions:

- How does a classical SVM baseline compare to CNN-based models on benign vs. malignant classification?
- Does transfer learning from ImageNet-pretrained models (ResNet-18, EfficientNet-B0) provide measurable gains on limited medical imaging data?
- How do results differ between mammography and ultrasound modalities?

2 Related Work

Classical machine learning approaches for breast cancer CAD commonly use handcrafted features combined with SVMs or related classifiers. SVMs maximize the margin between classes and handle high-dimensional feature spaces via kernel functions, such as radial basis function (RBF) kernels [1]. However, their performance is bounded by the quality of the engineered features, which can be particularly challenging to design for complex medical images.

Deep learning methods have increasingly been applied to breast cancer screening tasks, including lesion classification, detection, and segmentation. CNNs leverage convolutional filters to learn local edge, texture, and shape patterns that are useful for distinguishing benign from malignant tissue. Residual networks (ResNets) [2] facilitate the training of very deep models by using skip connections, while EfficientNet architectures [3] achieve strong performance with fewer parameters through principled scaling rules. Transfer learning from ImageNet-pretrained models is a common strategy when labeled medical data is limited, allowing models to adapt generic visual features to specific medical domains.

3 Datasets and Preprocessing

We utilize two distinct breast imaging datasets:

- **Mammography dataset:** A collection of mammogram scans, each labeled as benign or malignant. Images are derived from public datasets such as CBIS-DDSM.[1]
- **Ultrasound dataset:** A set of breast ultrasound images labeled as benign or malignant, similar to publicly available breast ultrasound datasets on Kaggle.

Exact dataset names are omitted here, but both datasets form binary classification problems (benign vs. malignant lesions).

3.1 Preprocessing

All images are preprocessed in MATLAB. The main steps are:

- **Grayscale conversion:** Images are converted to grayscale if not already single-channel.
- **Resizing:** Each image is resized to a fixed resolution compatible with the CNN models, such as 224×224 pixels (required by ResNet-18 and EfficientNet-B0 in MATLAB).
- **Normalization:** Pixel intensities are normalized (e.g., scaled to $[0, 1]$) to stabilize training.

For the SVM, additional preprocessing is required to obtain feature vectors. We consider:

- **Flattened intensity features:** Resized images are flattened into vectors.
- **Optionally, texture features** (e.g., simple statistical or filter-based descriptors).
- **Dimensionality reduction:** Techniques such as PCA may be applied to reduce feature dimensionality before SVM training.

[1] Example resource: CBIS-DDSM on The Cancer Imaging Archive.

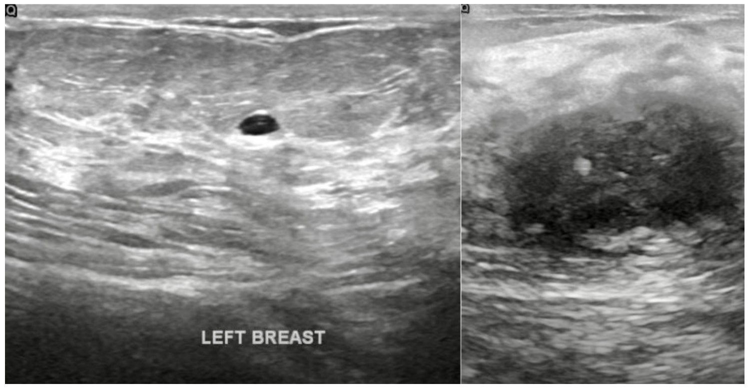

Figure 1: Sample benign and malignant ultrasound images from the ultrasound dataset.

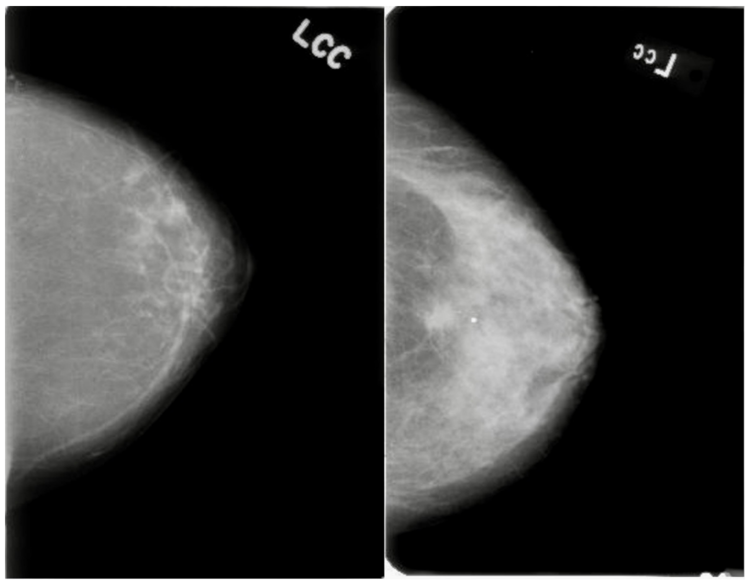

Figure 2: Sample benign and malignant mammogram images from the mammography dataset.

3.2 Example Images

Figure 1 shows sample benign and malignant ultrasound images, while Figure 2 shows sample benign and malignant mammograms.

4 Models

We evaluate four models on the mammography dataset and two models on the ultrasound dataset.

4.1 Support Vector Machine (SVM)

We use an SVM with Gaussian (RBF) kernel for binary classification. The classifier is trained on feature vectors extracted from preprocessed images. Key hyperparameters include:

- Kernel function: RBF
- Kernel scale: tuned via a small grid search or MATLAB defaults
- Regularization parameter C: tuned on a validation set

SVM training is performed using MATLAB's `fitcsvm` function, and predictions use `predict`.

4.2 Custom CNN

We design a relatively shallow CNN trained from scratch, comprising:

- Two convolutional layers with 3×3 kernels, ReLU activations, and 2×2 max-pooling.
- A flatten layer feeding into fully-connected layers.
- A final dense layer with 2 output units and softmax activation for benign vs. malignant classification.

The network is initialized with random weights and trained using stochastic gradient descent with mini-batches. We use cross-entropy loss and track training and validation accuracy via MATLAB's `trainNetwork` function.

4.3 ResNet-18 (Transfer Learning)

ResNet-18 [2] is a deep CNN with residual connections that facilitate gradient flow across layers. We employ a ResNet-18 model pretrained on ImageNet and fine-tune it on the mammography dataset:

- Replace the final fully-connected layer with a 2-unit layer.
- Use stochastic gradient descent with momentum (SGDM).
- Start by freezing early layers and training only the last layers, then progressively unfreeze more layers.
- Use a small learning rate (e.g., 10^{-4}–10^{-3}) for stable fine-tuning.

4.4 EfficientNet-B0 (Transfer Learning)

EfficientNet-B0 [3] is a compact CNN architecture that applies compound scaling of depth, width, and resolution to achieve strong accuracy/efficiency trade-offs. Similar to ResNet-18, we:

- Load an ImageNet-pretrained EfficientNet-B0 model in MATLAB.
- Replace the final classification layer with a 2-unit output.
- Fine-tune using SGDM with a small learning rate.

Given its smaller parameter count relative to deeper ResNets, EfficientNet-B0 may generalize better on limited medical data.

5 Experimental Setup

For each dataset, we partition the images into training and validation sets. For the mammography dataset, we use a typical split of 70–80% for training and the remaining 20–30% for validation. The ultrasound dataset is split similarly; if the dataset is small, cross-validation could be beneficial, particularly for SVM.

5.1 Training Protocol

Deep learning models (CNN, ResNet-18, EfficientNet-B0) are trained with:

- Binary cross-entropy (log-loss) as the objective.
- Mini-batch sizes of 16 or 32.
- Up to 5–10 epochs of training.
- Early stopping based on validation performance to mitigate overfitting.

Training is accelerated using a GPU if available; otherwise, CPU training is used, which is slower for the deeper networks.

The SVM is trained once per hyperparameter configuration, and hyperparameters are selected based on validation accuracy or cross-validation within the training set.

Table 1: Validation accuracy on the mammography dataset (benign vs. malignant). Values are approximate ranges based on observed experiments.

Model	Validation Accuracy (%)
SVM (RBF kernel)	∼55–56
Custom CNN	∼55–58
ResNet-18 (TL)	∼57.7
EfficientNet-B0 (TL)	∼64.3

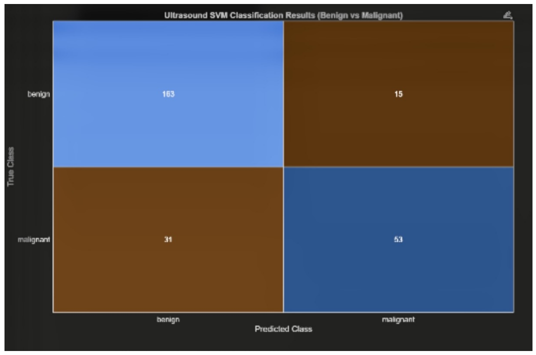

Figure 3: SVM results on mammography dataset.

5.2 Evaluation Metrics

We focus primarily on:

- **Validation accuracy**: the proportion of correctly classified validation examples.
- **Confusion matrices**: to analyze the distribution of true positives, false positives, true negatives, and false negatives across benign and malignant classes.

From the confusion matrices, one can derive sensitivity (recall for malignant cases) and specificity (true negative rate for benign cases). In a clinical context, false negatives (missed cancers) are particularly critical, so we qualitatively examine how each model behaves with respect to malignant cases.

6 Results

6.1 Mammography Dataset

Table 1 summarizes the approximate validation accuracy of each model on the mammography dataset.

SVM. The SVM baseline achieves roughly 55–56% accuracy, only slightly above random guessing for a balanced binary task. The confusion matrix shows frequent misclassification of malignant masses as benign, leading to a high false-negative rate. Benign cases are somewhat better recognized, but false positives (benign predicted as malignant) are also non-negligible.

Custom CNN. The from-scratch CNN yields validation accuracy in the 55–58% range, comparable to SVM. Training curves indicate rapid overfitting: training accuracy grows quickly while validation accuracy stagnates, reflecting limited generalization from a small dataset. The confusion matrix suggests that the CNN sometimes captures patterns SVM misses, correctly identifying some malignant lesions, but still produces many misclassifications.

Figure 3 shows the result for svm on mammography, while results for cnn on mammography.

ResNet-18. Fine-tuning ResNet-18 improves validation accuracy to approximately 57.7%, surpassing both the SVM and custom CNN. Transfer learning allows the model to leverage generic visual

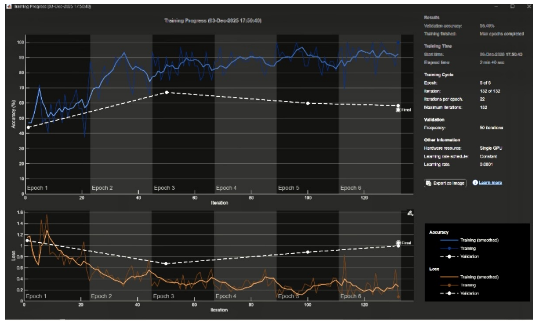

Figure 4: CNN results on mammography dataset.

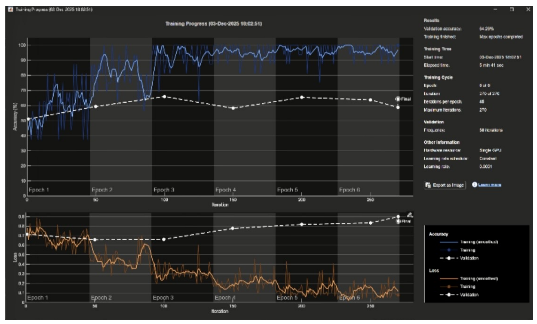

Figure 5: ResNet-18 results on mammography dataset.

features learned from ImageNet. The confusion matrix shows improved balance between classes, with more malignant cases correctly identified and a reduced false-negative count, although overall sensitivity and specificity remain modest.

EfficientNet-B0. EfficientNet-B0 achieves the best performance at approximately 64.3% validation accuracy, indicating that a modern CNN architecture with compound scaling can extract more discriminative features from mammograms under limited data. Training curves for EfficientNet-B0 show relatively close training and validation accuracy, suggesting better generalization than the small CNN. The confusion matrix indicates that the majority of benign cases are correctly classified and that malignant detection is improved relative to other models, though false negatives still occur.

6.2 Ultrasound Dataset

For ultrasound images, we focus on the SVM and the custom CNN. Table 2 summarizes the approximate validation accuracy.

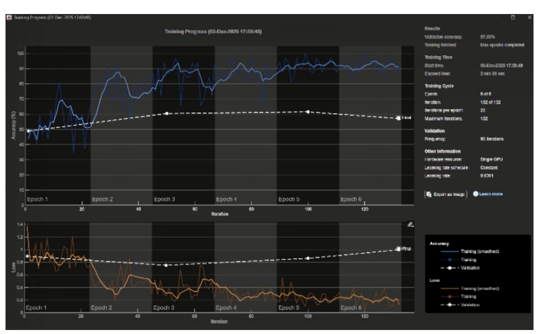

Figure 6: EfficientNet-B0 results for mammography dataset.

Table 2: Validation accuracy on the ultrasound dataset (benign vs. malignant).

Model	Validation Accuracy (%)
SVM (RBF kernel)	mid-50s ($\sim$50–60)
Custom CNN	$\sim$60

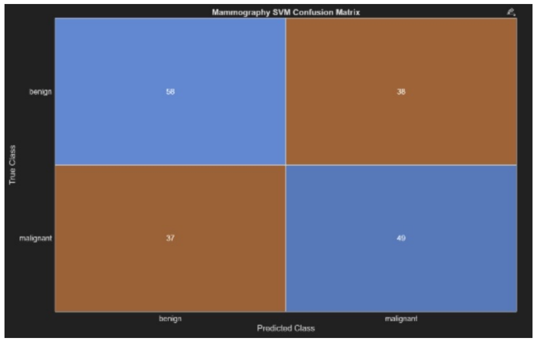

Figure 7: SVM results on Ultrasound dataset.

SVM. On ultrasound, SVM again provides the lowest performance, with accuracy in the mid-50% range. Confusion matrices indicate many misclassifications of both malignant and benign cases. Ultrasound images often contain speckle noise and subtle boundaries, making handcrafted or naive intensity features insufficient.

Custom CNN. The CNN trained on ultrasound achieves around 60% validation accuracy, outperforming the SVM by several percentage points. This suggests that the CNN can learn useful echogenic and morphological patterns indicative of malignancy. However, overfitting is still evident, and the level of accuracy remains far from clinically acceptable. We did not apply ResNet-18 or EfficientNet-B0 to ultrasound in this study, but they could potentially offer further improvements with sufficient data.

7 Discussion

Across both modalities, deep learning models outperform the classical SVM baseline, particularly when transfer learning from ImageNet is used. EfficientNet-B0 emerges as the best-performing model on mammograms, with $\sim$64% validation accuracy. ResNet-18 and the custom CNN also provide modest gains over SVM. On the ultrasound dataset, the custom CNN improves upon SVM, although performance is still limited.

The main challenges observed are:

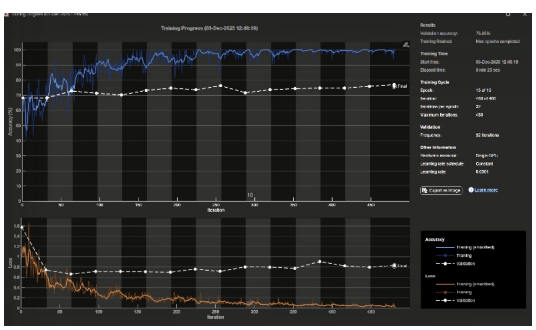

Figure 8: CNN results on ultrasound dataset.

- **Limited dataset size:** Small training sets lead to overfitting and unstable validation accuracy, especially for deeper models.
- **Class imbalance:** If malignant cases are underrepresented, models can struggle to achieve high sensitivity, resulting in frequent false negatives.
- **Image complexity:** Both mammography and ultrasound involve subtle textural patterns, noise, and anatomical variability, which are difficult to capture with simple features or small networks.

Potential directions for improvement include:

- Data augmentation (flips, rotations, intensity scaling) to increase effective data size.
- Region-of-interest localization or segmentation to focus on lesions rather than entire images.
- Ensembling multiple models or combining information from both modalities (multi-modal learning).
- More extensive hyperparameter tuning, including regularization techniques (dropout, weight decay).

8 Conclusion

We presented a comparative study of classical and deep learning approaches for breast cancer computer-aided diagnosis using mammography and ultrasound images. Using MATLAB, we implemented an SVM classifier, a custom CNN, and two modern CNN architectures (ResNet-18 and EfficientNet-B0) via transfer learning.

Our experiments show that deep learning models, particularly EfficientNet-B0, outperform the classical SVM baseline on mammography data, achieving validation accuracies of roughly 64% vs. 55–56% for SVM. On ultrasound, a custom CNN similarly outperforms SVM, though absolute accuracies remain around 60%. These levels of performance are not yet sufficient for a stand-alone diagnostic tool, but they highlight the potential of deep learning and the need for larger, more diverse datasets and more advanced modeling strategies.

Overall, this work demonstrates an end-to-end workflow for applying both conventional and deep learning techniques to medical image classification within MATLAB, from preprocessing and model training to evaluation with confusion matrices and training curves. With continued advances in deep learning and increased availability of annotated medical image data, future CAD systems may achieve the robustness and accuracy needed to assist clinicians in breast cancer diagnosis.

Acknowledgments

We thank MATLAB Deep Learning Toolbox and CBIS-DDSM dataset contributors.

References

[1] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[2] 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)*, pages 770–778, 2016.

[3] M. Tan and Q. V. Le. EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pages 6105–6114, 2019.

[4] MathWorks. MATLAB Deep Learning Toolbox (R2023a) – User's Guide and Function Reference. The MathWorks, Inc., Natick, MA, USA, 2023.