

Clinical Feasibility of Low-Magnification Fluorescence Imaging for Breast Cancer Margin Detection Using Texture Analysis and Deep Learning

TIANLING NIU ¹⁺, POUYA AFSHIN ²⁺, TONGTONG LU ³, DAVID HELMINIAK ⁴, JULIE JORNS ⁵, MOLLIE PATTON ⁵, TINA YEN ⁶, DONG HYE YE ², AND BING YU ^{1*}

¹Joint Department of Biomedical Engineering, Marquette University and Medical College of Wisconsin, Milwaukee, WI, USA

²Department of Computer Science, Georgia State University, Atlanta, GA, USA

³Department of Engineering and Engineering Technology, University of Wisconsin–Oshkosh, Oshkosh, WI, USA

⁴Department of Computer Engineering, Marquette University, Wisconsin, WI, USA

⁵Department of Pathology, Medical College of Wisconsin, Milwaukee, WI, USA

⁶Department of Surgery, Medical College of Wisconsin, Milwaukee, WI, USA

*Corresponding author: bingyu@marquette.edu

⁺These authors contributed equally to this work.

Abstract:

High-resolution images of unprocessed surgical breast tissue can be obtained using microscopy with ultraviolet surface excitation (MUSE). This technique is considered a promising method for checking surgical margins during breast cancer surgery. In this study, MUSE images at 4× and 10× magnifications were compared using patch-level classification methods. Texture analysis (TA) based on local binary patterns (LBP) and deep learning (DL) with a base Vision Transformer (ViT) model were used. Both methods achieved similar performance at both magnifications. Using DL method, both 4× and 10× magnifications achieved 96.30% sensitivity, 100% specificity and 98.18% accuracy. Using TA method, 4× achieved better specificity (100% vs 93.33%) and 10× yielded higher sensitivity (100% vs 93.33%), but both had the same accuracy (96.67%). No clear improvement in performance was observed with 10× magnification. These results show that 4× imaging achieves the same diagnostic accuracy as 10× imaging. At the same time, 4× offers a larger field of view and faster image capture. Therefore, lower magnification can be effectively used in MUSE systems for accurate and efficient intraoperative margin assessment.

1. Introduction

Breast cancer has had the highest incidence among cancers. It is the second leading cause of cancer-related death among women in the United States, even with improvements in screening and treatment [1, 2]. Surgical treatment is the most prevalent option for breast cancer patients, and breast-conserving surgery is recommended for the earliest-stage patients [3]. However, approximately 20% of patients require secondary surgery due to incomplete excision during the primary procedure [4]. This indicates that achieving complete tumor removal remains a challenge in breast-conserving surgery. The surgical margin, the portion of the patient’s removed tissue, is examined using hematoxylin and eosin (H&E) staining. This process usually takes several days because the tissue must first be fixed in formalin and embedded in paraffin (FFPE) before analysis [5]. Margins are classified as either negative or positive; the latter indicates that cancer cells remain at the surgical margin, thereby increasing the risk of recurrence. Therefore, additional surgery is often required to obtain a negative margin [6]. Secondary surgery can place both emotional and financial burdens on patients, highlighting the need for methods that allow

real-time, intraoperative margin assessment during breast-conserving surgery [7].

Microscopy with ultraviolet surface excitation (MUSE) applies ultraviolet (UV) light to image the tissue surface at high resolution rapidly and without sectioning, making it a promising approach for real-time margin evaluation [8–12]. Unlike traditional H&E, which requires thin tissue sections and relies heavily on pathologist expertise, MUSE offers rapid surface imaging with distinctive contrast patterns that enable clinicians to distinguish healthy from cancerous tissue and to clearly identify tumor boundaries, both of which are essential during surgery. The shallow penetration of UV light limits imaging to the tissue surface without the need for sectioning, which greatly shortens image acquisition time and makes MUSE suitable for intraoperative assessment [8–12]. Moreover, the MUSE system has a simple optical setup and does not require complex components, making it a cost-effective method for margin assessment that is easy to use in clinical settings [13].

Beyond intraoperative margin assessment, MUSE has been applied to rapid histology, surface topography, and nerve imaging [9–15]. Its objective magnification can be adjusted: higher magnifications (10× or 20×) capture finer details when speed is less critical. This allows MUSE to complement conventional histopathology and visualize nerve microanatomy [9–15]. For example, Levenson et al. used 10× for surface tomography and histopathology [9, 13, 15], while others used 10× for nerve morphology [14] and breast margins [10], or 10–20× for skin and breast tissue imaging [12]. Higher magnification reveals finer details but captures a smaller field of view (FOV), slowing scanning of the entire tissue, an important factor for rapid imaging during surgery. To address this limitation, a MUSE system with a 4× magnification was developed, achieving classification accuracies of 90.3% with texture analysis (TA) and 96.47% with deep learning (DL) [16, 17], demonstrating that computational approaches can compensate for reduced visual detail.

Building on this, TA provides a simple, computationally efficient method for quantifying tissue patterns. Local binary pattern (LBP), a common TA method, compares each pixel with its neighbors and encodes these relationships to capture micro-texture. LBP performs well for classifying MUSE fluorescence breast images with minimal computational cost [16, 18, 19]. Although TA is fast and straightforward, DL can automatically capture more complex patterns from image patches, including fine cellular details and overall tissue structure. When applied to MUSE images, DL improves margin assessment, reduces the need for follow-up surgeries, and supports more accurate detection and treatment decisions for breast cancer [17, 24].

Because MUSE images are very large, analyzing them at high resolution with DL is computationally challenging, even at 4× magnification. Downsampling can also remove fine details needed for accurate classification. To solve this, researchers divide MUSE images into smaller patches, analyze each patch separately, and then combine the results to classify tissue [10, 16, 21]. Many studies have used Convolutional Neural Networks (CNNs) [22], such as ResNet-50 with XGBoost [21], or have improved training with synthetic patches generated by diffusion models [23]. CNNs work well but may miss overall tissue context due to the limited field of view [24]. Vision transformers (ViTs) capture both fine details and the full tissue structure, using self-attention to capture relationships among subpatches within each patch. Patch-level ViTs, when combined with Grad-CAM++, can highlight important regions and facilitate interpretation of predictions [24]. In contrast, others used self-supervised synthetic MUSE images to enhance training and classification accuracy further [17]. Inspired by these results, the current study uses a patch-level ViT framework to analyze MUSE images.

From an optical perspective, higher magnification provides better resolution but also increases acquisition time. For example, acquiring an image of all scanned breast tissues generally takes a few minutes with a 4× objective but more than 15 minutes with a 10× objective. Additionally, higher-magnification objectives have a shallower depth of field and a narrower field of view, which require more precise focusing and longer exposure times. Therefore,

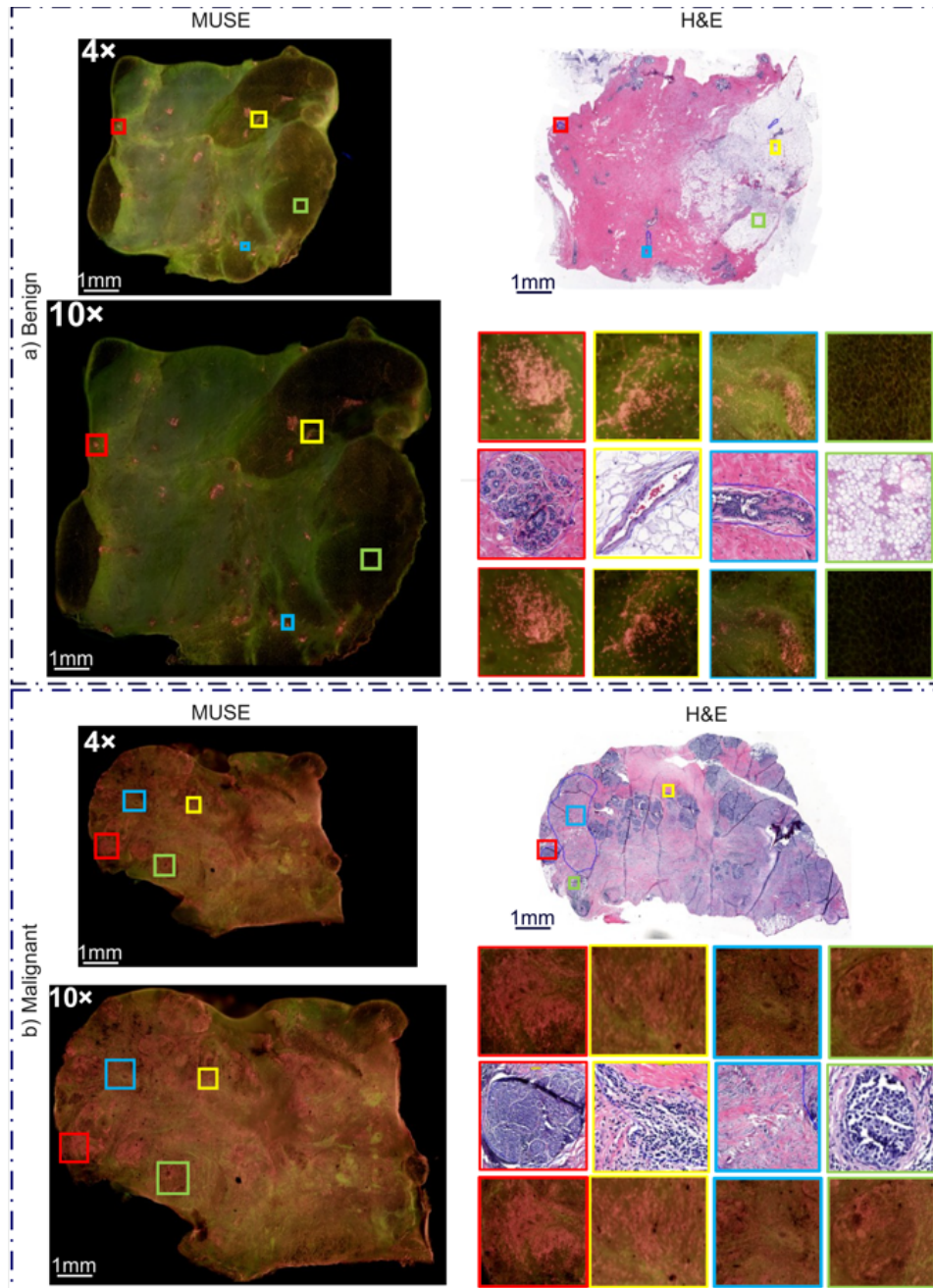


Fig. 1. Representative MUSE and H&E images of (a) benign tissue and (b) invasive lobular carcinoma (ILC). For each case, 4x MUSE (top left), H&E (top right), and 10x MUSE (bottom left) images are shown, with corresponding zoomed-in regions (bottom right) that highlight key morphological structures between tissue types. In (a), glands (red), vessels (yellow), ducts (blue), and adipocytes (green) are indicated; in (b), lobular neoplasia (red/green), chronic inflammation (yellow), and high tumor cellularity (blue) are highlighted.

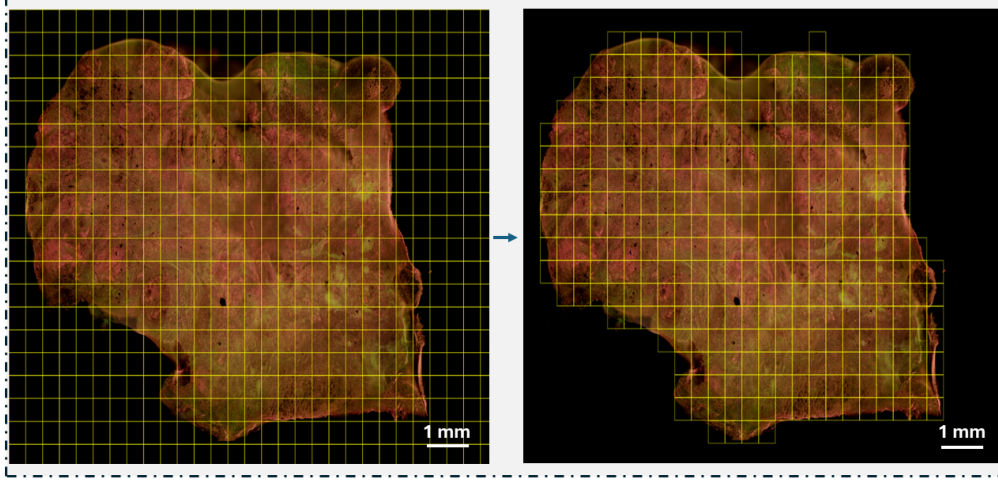


Fig. 2. Illustration of patch extraction from a 10 $\times$ MUSE whole-slide image of a grade 2 invasive lobular carcinoma (ILC) sample. Patches are selected using a 20% threshold, retaining only those with more than 20% of pixels above intensity 5 (background). Extracted patches are square, sized 1000 $\times$ 1000 pixels for 10 $\times$.

there is a clear trade-off between image resolution and imaging speed in the evaluation of intraoperative margins [20]. Although previous studies have reported promising results with MUSE in various applications, most have evaluated it at a single magnification. As a result, it remains unclear how different magnifications affect classification performance, particularly for binary classification of tumor versus normal tissue [8–15]. To address this gap, this study compares the classification performance of 4 $\times$ and 10 $\times$ magnifications using texture analysis and deep learning models [16, 21, 24]. By directly comparing them, we provide clear results to inform the selection of the most suitable objective for intraoperative assessment with MUSE. In particular, this study investigates whether 10 $\times$ yields a significant improvement in classification accuracy or whether 4 $\times$ suffices while allowing faster image acquisition.

2. Materials and Methods

2.1. Pre-processing

Before classification, all fluorescence images are prepared to ensure consistent and informative input data. The fluorescence dyes used in this study are propidium iodide (PPID) and eosin Y (EY). PPID stains the nuclei and produces red fluorescence, while EY stains the cytoplasm and produces green fluorescence. All fluorescence images are stitched MUSE images, as shown in Fig. 1 a) presents a representative example of benign breast tissue, including MUSE images acquired at 4 $\times$ and 10 $\times$ magnifications, along with the corresponding H&E images for comparison. Zoomed-in views are provided to highlight benign glands, blood vessels, benign ducts, and adipocytes, each indicated by a distinct color. Fig. 1 b) shows a representative example of malignant breast tissue, including MUSE images acquired at 4 $\times$ and 10 $\times$ magnifications, along with the corresponding H&E images for comparison. Zoomed-in regions highlight lobular neoplasia, chronic inflammation, and areas of high cellular density. The fluorescence images were divided into patches for analysis because the original stitched images were very large, often exceeding 10,000 $\times$ 10,000 pixels.

During the preprocessing, the stitched MUSE images $\{x_i\}_{i=1}^M \in \mathcal{X}$ are divided into non-overlapping patches $\mathbf{p}_i^j$ of size 400 $\times$ 400 pixels for 4 $\times$ magnification and 1000 $\times$ 1000 pixels

for 10 $\times$ magnification. The patch sizes were scaled according to the magnification difference ($10\times / 4\times = 2.5$) to maintain an approximately consistent physical tissue area across different imaging magnifications. The patch extraction process is indicated in Fig. 2. The red channel is extracted, and a threshold of 5 is applied to identify informative pixels, based on empirical observations. Pixels with intensity above the threshold are considered informative, and only patches $\mathbf{p}_i^j$ containing more than 20% informative pixels are retained for classification; all other patches are discarded.

Patch labels were generated based on the corresponding H&E images using semi-automated annotation transfer method and verified by our pathologist [37]. Patches from benign regions were labeled as normal, while patches from malignant regions were labeled as tumor. To minimize the risk of excluding tumor tissue due to boundary uncertainty, patches extracted from the transitional borders between benign and malignant zones were labeled as tumor. Let $J_i = \{j \mid \mathbf{p}_i^j \in x_i\}$ denote the set of indices of all valid patches extracted from MUSE image x_i . Each patch $\mathbf{p}_i^j$ is labeled according to the corresponding image of H&E, which is the gold standard clinically. Patches along the edge of cancerous tissue were also labeled as malignant.

In this study, two approaches are used for margin classification: a traditional TA method and a DL method. In the TA approach, texture features are extracted from image patches and are used for classification with a conventional machine learning model. In the DL approach, discriminative features are automatically learned from image patches using a fine-tuned Vision Transformer model. In both methods, patch-level predictions are aggregated to obtain the final margin-level classification.

2.2. Texture Analysis

Texture analysis has been an effective method for capturing local patterns and structural information in medical images, helping to differentiate normal from cancerous tissue based on textural differences [25]. In this study, texture features from the red channel of each MUSE patch $\mathbf{p}_i^j$ are extracted using the local binary pattern (LBP) method [16, 26, 29, 30]. Red channel is used because it represents the nuclei information and provides the highest contrast. Before extracting LBP features, the patches are denoised using a Wiener filter and histogram equalized to ensure uniform pixel intensities. These features are then used for classification. Specifically, from each patch $\mathbf{p}_i^j$, $j \in J_i$, 14 LBP features are extracted using the number of neighbors $N = 12$ and a circular pattern radius $R = 3$. This configuration samples 12 evenly spaced points on a circle of radius 3 centered at each pixel and compares the intensity values at these locations with that of the center pixel to encode local texture patterns.

Linear interpolation and rotation-invariant uniform patterns are used in feature computation to ensure robust LBP feature extraction when the sampled neighbor points do not align exactly with pixel grid locations and to reduce sensitivity to image rotation. The extracted LBP features are stored in a feature matrix $\mathbf{F}_i \in \mathbb{R}^{J_i \times 14}$. The nuclei-to-cytoplasm (NC) ratio from each patch $\mathbf{p}_i^j$ is also calculated using a dark threshold of 115 (set empirically) and stored in a separate matrix $\mathbf{R}_i \in \mathbb{R}^{J_i \times 1}$. The feature matrices $\mathbf{F}_i$ and $\mathbf{R}_i$, together with the healthy/malignant labels y_i^j for $j \in J_i$, are used as inputs to a support vector machine (SVM) to perform classification and predict labels for the test set, which works by drawing a boundary that best separates different tissue types based on extracted image features. The polynomial kernel helps the model draw a curved boundary rather than a straight line, which is useful when the data are not easily separable. The regularization parameter C was set to 20 in MATLAB and controls how strictly the boundary follows the training data: a higher value places more emphasis on correctly classifying the training samples, while still limiting overly complex boundaries.

Then five-fold cross-validation is used to evaluate patch-level classification performance. Min-max normalization is applied to all features to prevent data leakage prior to margin-level classification. The trained patch classifier produces class predictions $\hat{y}_i^j$ for all valid patches

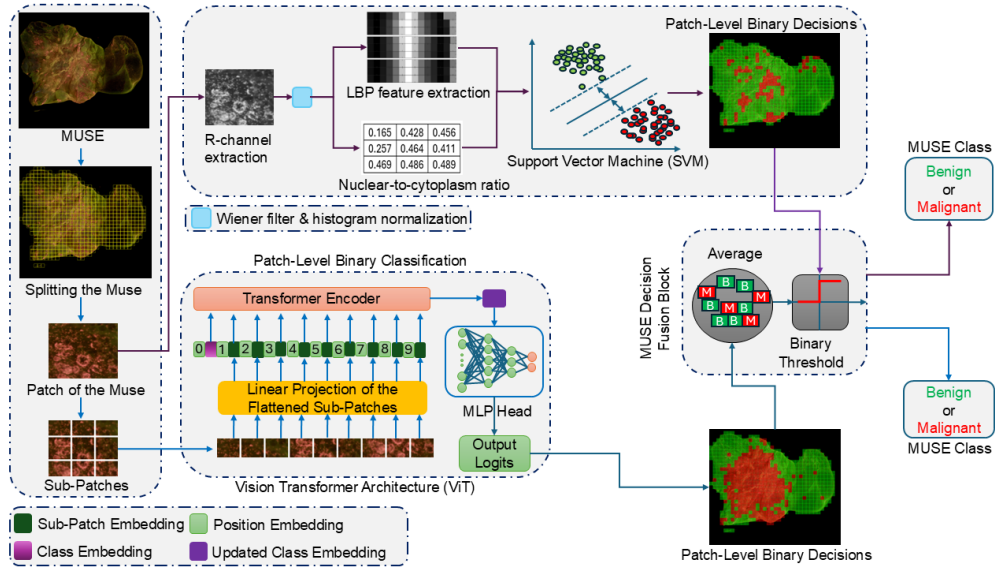


Fig. 3. Workflow for classifying MUSE images using TA (top branch) and DL (bottom branch). MUSE images are divided into patches at two magnifications. For TA, the red channel of each patch is first preprocessed. Then, LBP features are extracted and an SVM model predicts the class of each patch. For DL, each patch is converted into embeddings and passed through a ViT model with an MLP head to generate patch-level predictions. In both methods, the patch predictions are combined and compared with a threshold to determine the final margin-level classification.

$\mathbf{p}_i^j \in J_i$. The tumor and normal predictive scores for the MUSE image x_i were computed as

$$S_i^{\text{tumor}} = \sum_{j \in \mathcal{J}_i} \hat{y}_i^j, \quad (1)$$

and the normal predictive score was computed as

$$S_i^{\text{normal}} = \sum_{j \in \mathcal{J}_i} (1 - \hat{y}_i^j). \quad (2)$$

The final margin-level label $\hat{y}_i$ was determined by comparing the tumor and normal predictive scores:

$$\hat{y}_i = \begin{cases} 1, & \text{if } S_i^{\text{tumor}} > \theta \times S_i^{\text{normal}}, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

A margin was classified as malignant when the tumor prediction score exceeded the normal prediction score scaled by a threshold; otherwise, it was classified as benign. The complete workflow is illustrated in Fig. 3.

2.3. Deep Learning with Patch-Level ViT

A pre-trained ViT [27] is used to classify patches from MUSE Images and aggregate the results for margin-level predictions. Using transfer learning, the model recognizes common image features, thereby reducing overfitting and accelerating training [28]. After fine-tuning the MUSE

patch-level dataset, the model extracts detailed features and classifies each patch as benign or malignant. To that end, each extracted patch $\mathbf{p}_i^j$ is divided into N non-overlapping sub-patches $\mathbf{s}_i^{jk}$ (see Fig. 3), which are reshaped and linearly projected into a D -dimensional embedding space $\hat{\mathbf{s}}_i^{jk} = \text{vec}(\mathbf{s}_i^{jk})\mathbf{E} \in \mathbb{R}^{1 \times D}$ using a trainable projection matrix $\mathbf{E} \in \mathbb{R}^{(P^2 \cdot C) \times D}$. A learnable class token $\mathbf{p}_i^{\text{class}} \in \mathbb{R}^{1 \times D}$ is prepended to represent global patch information, and positional embeddings $\mathbf{E}_i^{\text{pos}} \in \mathbb{R}^{(N+1) \times D}$ are added to encode spatial relationships. The resulting input sequence is given by [17]

$$\mathbf{h}_0 = [\mathbf{p}_i^{\text{class}}, \hat{\mathbf{s}}_i^{j1}, \dots, \hat{\mathbf{s}}_i^{jN}] + \mathbf{E}_i^{\text{pos}}. \quad (4)$$

The sequence $\mathbf{h}_0$ is processed by a vision transformer encoder composed of L layers with normalization, self-attention, and multi-layer perceptron (MLP) blocks. The final class token is normalized and passed to a classification head to predict the patch label [17]:

$$\hat{y}_i^j = \text{MLP}_{\text{head}}(\text{LN}(\mathbf{p}_i^{\text{class}})) \in \{0, 1\}, \quad (5)$$

where 0 and 1 denote benign and malignant classes, respectively. The margin-level label $\hat{y}_i$ is obtained by averaging the patch predictions over the valid patch indices $\mathcal{J}_i$ and comparing to a threshold θ [17]:

$$\hat{y}_i = \begin{cases} 1, & \text{if } \frac{1}{|\mathcal{J}_i|} \sum_{j \in \mathcal{J}_i} \hat{y}_i^j > \theta, \\ 0, & \text{otherwise.} \end{cases}$$

For the 4 $\times$ and 10 $\times$ system setups, the 4 $\times$ model was optimized with the AdamW optimizer at a learning rate of 5×10^{-5} and a weight decay of 1×10^{-4} . In contrast, for the 10 $\times$ setting, a higher learning rate of 4×10^{-4} was employed. For the 10 $\times$ configuration, the model was trained with stochastic gradient descent (SGD) with a momentum of 0.9 and the same weight decay parameter. Preliminary ablation experiments using weighted cross-entropy loss to address patch imbalance did not improve performance and yielded no consistent gains in sensitivity or specificity; therefore, the final models were trained with the standard unweighted loss formulation. All models were trained for 60 epochs. Patch-level classification was performed using the ViT-B/16 architecture with 5-fold cross-validation. The data split was performed at the margin-level to prevent data leakage. In each fold, a set of MUSE images was held for testing, while patches extracted from the remaining MUSE images were further divided into 80% for training and 20% for validation.

Table 1. Dataset summary at 4 $\times$ and 10 $\times$ magnifications including margin-level and patch-level distributions.

	4 $\times$ magnification				10 $\times$ magnification			
	Samples	Margin (%)	Patches	Patch (%)	Samples	Margin (%)	Patches	Patch (%)
Healthy	28	50.9	29,897	76.4	28	50.9	27,419	75.1
Malignant	27	49.1	9,248	23.6	27	49.1	9,078	24.9
Total	55	100	39,145	100	55	100	36,497	100

3. Results

3.1. Dataset and Experimental Setup

The dataset in this study is obtained from the Medical College of Wisconsin. It consists of 55 breast tissue samples, including 28 benign and 27 malignant cases. From these samples, a total of 39,145 patches are extracted at a magnification of 4 $\times$ and 36,497 patches at a magnification

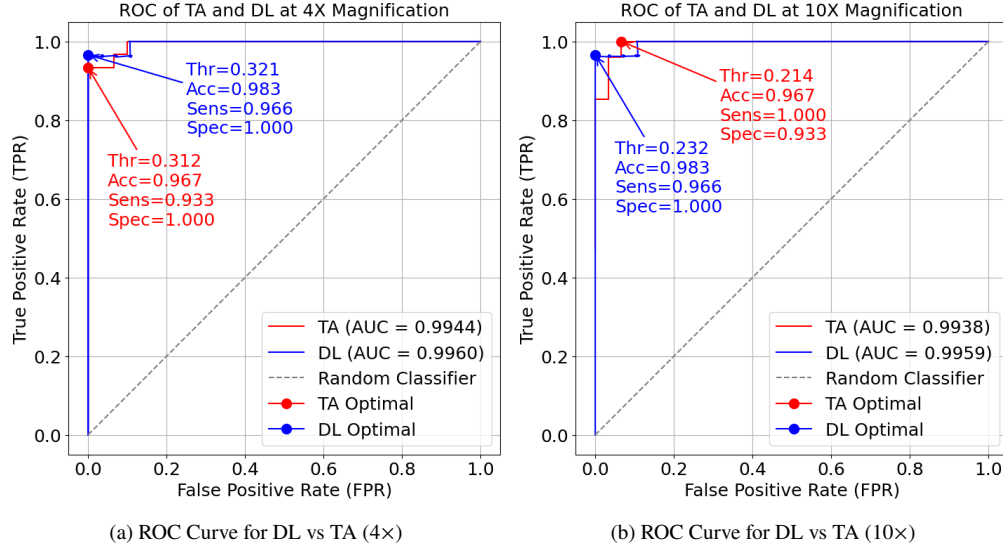


Fig. 4. ROC curves comparing margin-level classification performance of DL and TA at both magnifications.

of 10 $\times$. At 4 $\times$, 29,897 patches are from benign tissue and 9,248 patches from malignant tissue, whereas at 10 $\times$, there are 27,419 benign patches and 9,078 malignant patches. Although the margin-level dataset is nearly balanced (50.9% benign vs. 49.1% malignant), the patch-level dataset is imbalanced across both magnifications, with a higher proportion of benign patches. This imbalance reflects the larger spatial extent of benign tissue within surgical margins. A summary of the patch-level and margin-level datasets is presented in Table 1.

3.2. Quantitative Results

To evaluate classification performance fairly, the best binary threshold was selected for each magnification using grid search along the ROC curves as shown in Figure 4. Instead of using a random cutoff value, all possible thresholds were tested, and the threshold that gave the best margin-level performance was chosen. This ensures that both TA and DL were evaluated under their best operating conditions and allows a reliable comparison between 4 $\times$ and 10 $\times$ magnifications. To compare the two methods, both TA and DL were evaluated using the same margin-level classification framework.

Table 2 summarizes the margin-level results at 4 $\times$ and 10 $\times$ magnifications. At 4 $\times$, TA achieved 96.67% accuracy, 93.33% sensitivity, and 100.00% specificity at the optimal threshold of 0.312 (see Fig. 4a). Two malignant margins were classified as negative. At 10 $\times$, the accuracy remained 96.67%. Sensitivity increased to 100.00%, while specificity decreased to 93.33% at the optimal threshold of 0.214 (see Fig. 4b), resulting in two false positives. Although different errors were observed, the overall accuracy was the same at both magnifications. This means that using 10 $\times$ does not significantly improve LBP-based margin classification. With an appropriate threshold, 4 $\times$ magnification already provides sufficient texture information for reliable classification. In addition, 4 $\times$ imaging gives a larger field of view and faster image capture. This makes it more practical and efficient for real-time margin assessment without reducing performance.

The DL framework also showed consistent margin-level performance at both magnifications, as shown in Table 2. At 4 $\times$ magnification, the model achieved 98.18% accuracy, 96.30% sensitivity, and 100.00% specificity at the optimal threshold of 0.321 (see Fig. 4a). It correctly identified 26 of 27 malignant margins and all 27 normal margins, with only one false negative and no

Table 2. Margin-level classification results for TA and DL at 4× and 10× magnifications.

Method	4×						10×				
		P	N	Total	Se (%)	Sp (%)	P	N	Total	Se (%)	Sp (%)
TA	P	25	0	25	93.33	–	27	2	29	100.00	–
	N	2	28	30	–	100.00	0	26	26	–	93.33
	Total	27	28	55	–	–	27	28	55	–	–
	Accuracy (%)	96.67					96.67				
DL	P	26	0	27	96.30	–	26	0	27	96.30	–
	N	1	28	29	–	100.00	1	28	29	–	100.00
	Total	27	28	55	–	–	27	28	55	–	–
	Accuracy (%)	98.18					98.18				

false positives. Similar results were seen at 10× magnification, with 98.18% accuracy, 96.30% sensitivity, and 100.00% specificity at the optimal threshold of 0.232 (see Fig. 4b). At this magnification, most malignant samples were correctly detected, with only a small number of normal samples misclassified as malignant.

Although the optimal binary thresholds were different across magnifications, the final classification performance was the same for both TA and DL. The ROC curves showed strong separation between malignant and benign margins. For DL, the AUC values were 0.9960 at 4× and 0.9959 at 10×. For TA, the AUC values were 0.9944 at 4× and 0.9938 at 10× (see Fig. 4). Based on the results, magnification primarily affects the distribution of prediction scores, but does not significantly alter classification performance. The best binary thresholds were selected via grid search along the ROC curve to maximize margin-level accuracy. Overall, TA and DL showed similar margin-level performance at both magnifications. Using a higher-resolution 10× objective did not affect classification performance at the margin-level. This shows that a lower magnification of 4× already provides sufficient tissue structure and texture information for reliable classification. It also reduces computational cost and scanning time without degrading margin-level performance in real-time use.

3.3. Qualitative Results

Fig. 5 presents examples of correctly classified samples at both 4× and 10× magnifications using TA and DL methods. The figure shows patch-level prediction results for representative tissue samples, where (a) shows ground-truth labels overlaid on MUSE images, (b) shows TA prediction maps, and (c) shows DL prediction maps. The results demonstrate that the sample was correctly classified at both magnifications using both methods. As shown in the figure, malignant and benign regions are accurately distinguished, with most patches correctly predicted by both TA and DL models. Although a few patches exhibit minor misclassifications, the overall patch fusion results lead to correct image-level classification. These findings indicate that both TA and DL methods can effectively capture discriminative histological features across magnifications, supporting robust performance for surgical margin assessment.

To better understand the structure of the high-dimensional MUSE image patches, t-distributed Stochastic Neighbor Embedding (t-SNE) [32] was used. In this method, high-dimensional data are projected into two- or three-dimensional spaces while preserving local relationships

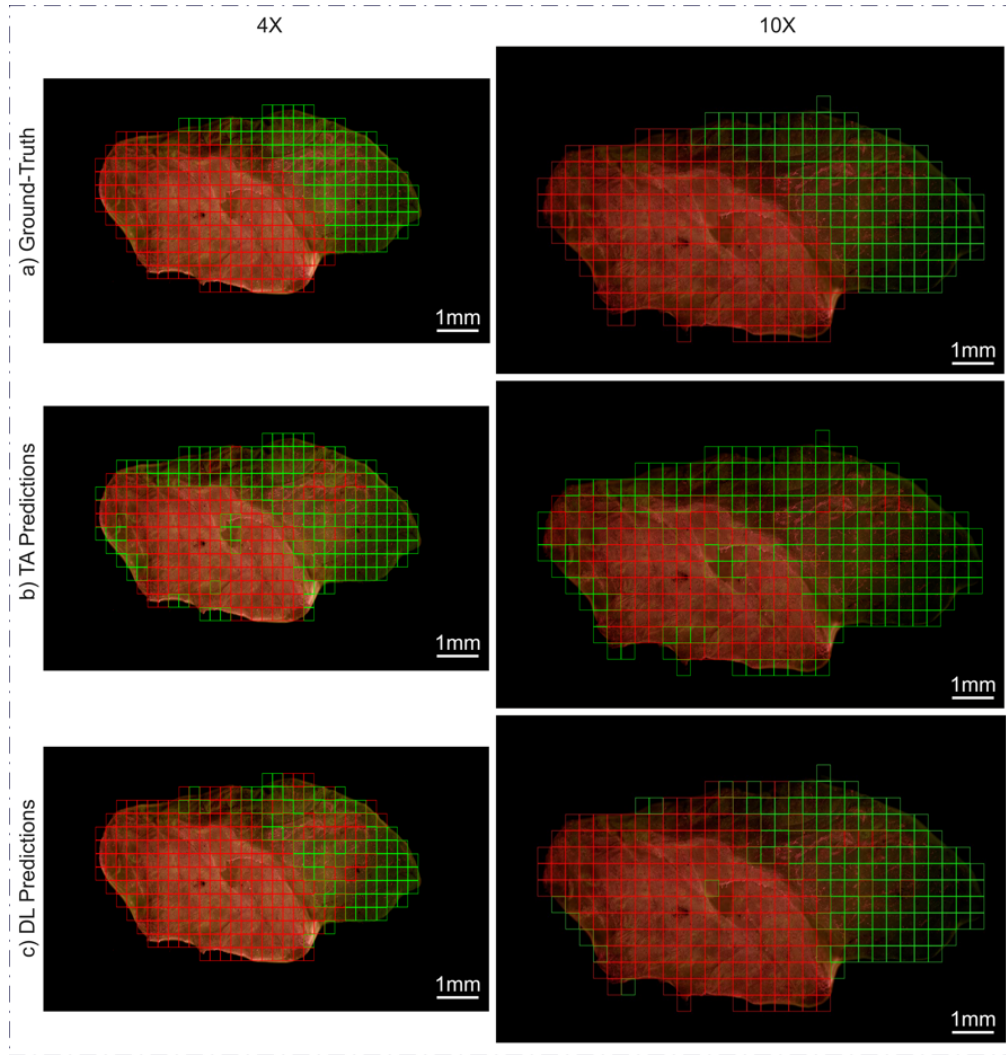


Fig. 5. Visualization of classification results at 4 $\times$ (left column) and 10 $\times$ (right column) magnifications. (a) Ground truth labels overlaid on MUSE images, (b) TA predictions, and (c) DL predictions. Ground truth annotations are derived from corresponding H&E images.

among points. t-SNE also preserves important global structures, making it effective for exploring patterns in high-dimensional data.

As depicted in Figure 6, the t-SNE plots show feature representation for both magnifications from the ViT class token, which summarizes the global structure of each patch. Both 4 $\times$ and 10 $\times$ patches form well-separated clusters for benign and malignant tissue, showing that the ViT learns useful features even before the final classification layer. The 4 $\times$ clusters the features slightly better, whereas the 10 $\times$ feature representations are more mixed at the edges, which might be due to local variation in fine cellular details and downsampling. Some overlap is observed across both magnifications, likely reflecting the model's difficulty in distinguishing features or where benign and malignant patterns appear similar or ambiguous. Overall, most patches form clear clusters, which support the use of the patch-level framework for MUSE-level classification. Even

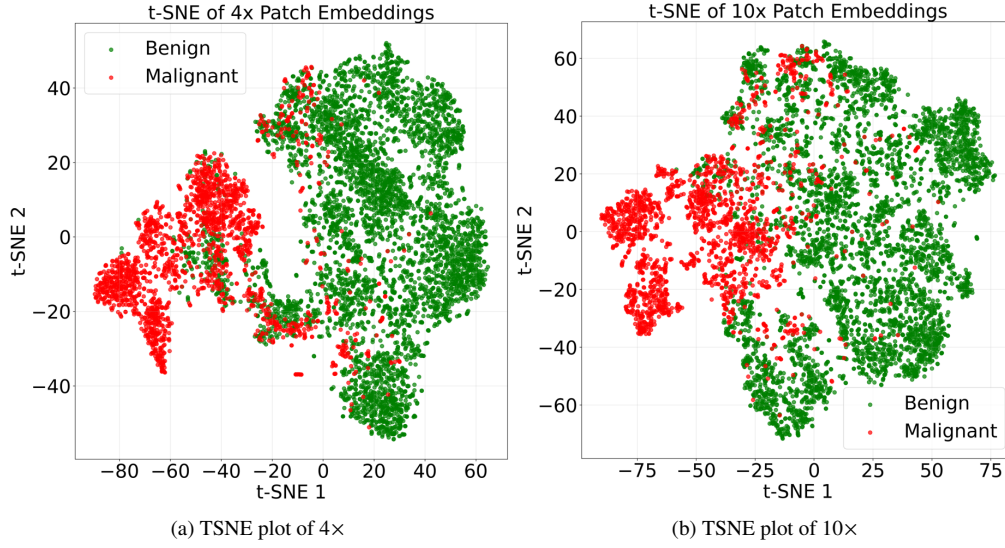


Fig. 6. TSNE plots of visualization of features across both magnifications.

after downsampling, the ViT still captures meaningful features.

4. Discussion

4.1. Main Findings

In this study, the effects of magnifications 4 $\times$ and 10 $\times$ on the classification at the margin-level of MUSE images were investigated using TA and DL. Both methods demonstrated consistent performance, indicating that automated approaches can distinguish benign from malignant breast tissue. Specifically, TA achieves a similar general accuracy at the optimal threshold across magnifications; however, the magnification 4 $\times$ shows better performance in detecting benign samples, while the magnification 10 $\times$ demonstrates a greater sensitivity to identify malignant tissues. In contrast, the DL method exhibited stable, consistent performance at both magnifications. Overall, both methods achieved high margin-level accuracy, with more than 96% accuracy at 4 $\times$ and 10 $\times$.

4.2. Analysis of Misclassified Cases

Fig. 7 shows examples of TA-misclassified samples at two different magnifications. On the left, (a) shows the H&E image of a malignant sample, (b) shows the corresponding ground-truth labels overlaid on the MUSE images, (c) shows the TA prediction map overlaid on the 4 $\times$ MUSE images. In the right column, (a) shows a benign sample H&E that was misclassified as positive (malignant), (b) shows the corresponding ground-truth labels, (c) shows the TA prediction map overlaid on the MUSE image.

According to the results in Fig. 7, misclassifications were primarily due to local texture similarities between benign and malignant tissues rather than to limitations in spatial resolution. At 4 $\times$ magnification, some dark malignant regions were misinterpreted as healthy tissue at the patch level, thereby reducing the number of malignant patches and leading to an overall false-negative classification. At 10 $\times$ magnification, more detailed cellular structures were visible, but some benign lobules and glandular regions were misclassified as malignant because their nuclei appeared clustered, resulting in false-positive predictions. In general, these errors were due mainly to textural similarities between benign and malignant regions rather than magnification

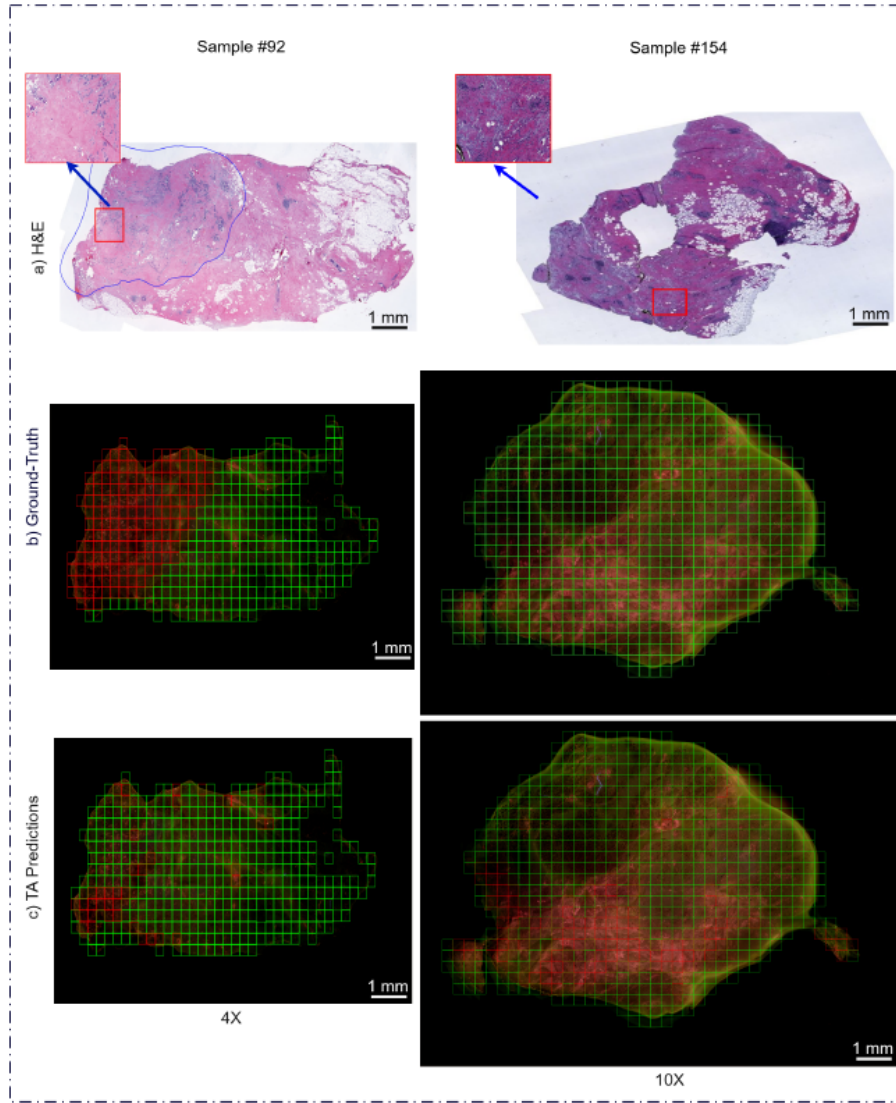


Fig. 7. Examples of misclassified samples using the TA method at 4 $\times$ (left column) and 10 $\times$ (right column). The left sample is malignant but misclassified as negative, whereas the right sample is benign but misclassified as malignant. (a) Corresponding H&E images for each sample, including zoomed-in patches highlighting the misclassified regions. (b) Ground-truth labels overlaid on the MUSE images. (c) TA predictions overlaid on the MUSE images.

itself. This indicates that both 4 $\times$ and 10 $\times$ yield a similar diagnostic performance, with most errors arising from challenging tissue patterns rather than from resolution.

Fig. 8 displays the same misclassified malignant sample at both 4 $\times$ and 10 $\times$ magnifications by DL. (a) corresponds to the H&E image of a malignant sample with highlighting the malignant patch in the H&E image. (b) shows the ground truth patches overlaid on 4 $\times$ (left) and 10 $\times$ (right) magnifications. (c) displays the DL predictions overlaid for these magnifications.

As shown in the DL predictions (Fig. 8c), for both the 4 $\times$ and 10 $\times$ magnifications, models correctly classified most benign patches and successfully detected the main malignant core.

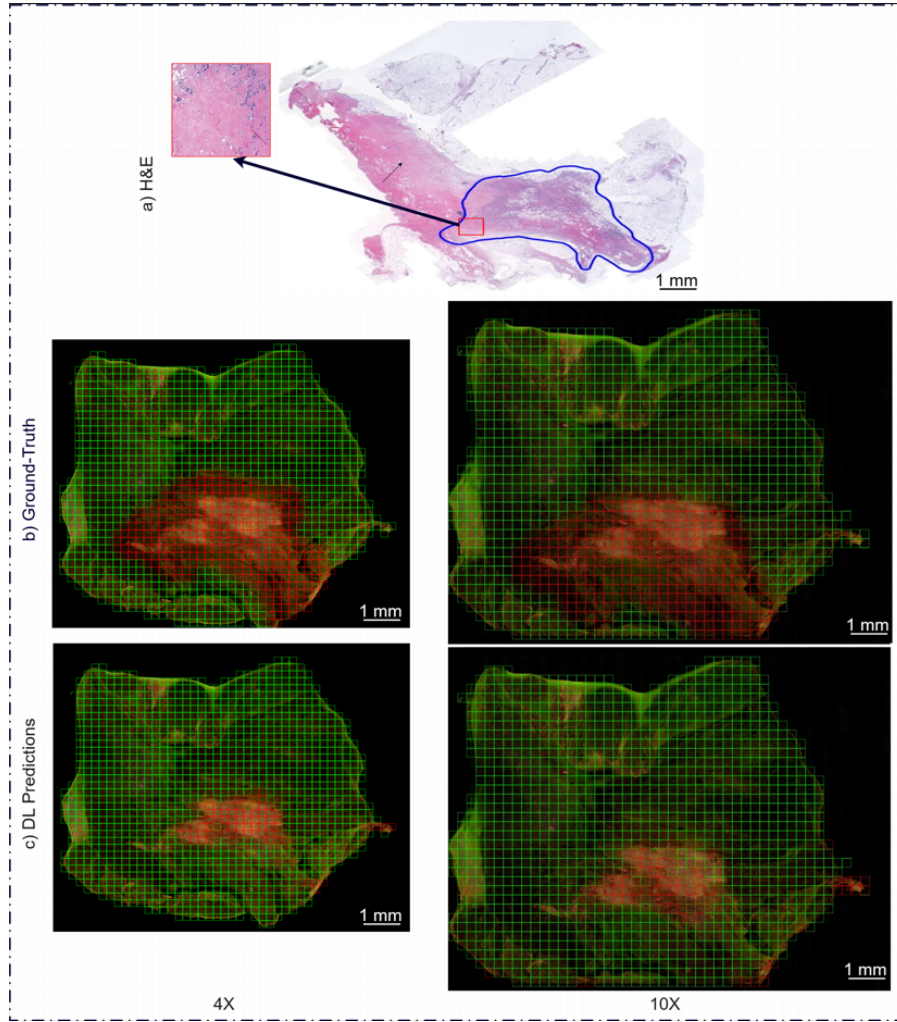


Fig. 8. Examples of a misclassified malignant sample using the DL method at 4 $\times$ (left column) and 10 $\times$ (right column). The same case was incorrectly predicted as benign at both magnifications. (a) H&E images for each sample, including zoomed-in patches highlighting the misclassified regions. (b) Ground-truth labels overlaid on the MUSE images. (c) DL predictions overlaid on the MUSE images.

However, at both cases, models failed to classify several malignant patches near the tumor boundary, which resulted in false-negative regions compared with the ground truth annotations (Fig. 8b). Importantly, the misclassified regions were very similar at both magnifications, suggesting that the errors were not primarily due to the imaging magnification. These results show that the misclassified areas contain mixed or unclear features, which the model failed to classify correctly. Because the final decision is based on majority voting, samples with too few correctly identified malignant patches were ultimately classified as false negatives.

4.3. Limitations and Future Directions

Objective magnification is an important design parameter for the MUSE system [33,34]. Although higher magnification provides finer tissue details, image resolution alone should not determine

system optimization. For classification tasks, the algorithm's ability to handle different resolutions must be considered, as increased magnification does not always improve performance. In this study, the performance of a 10× objective lens was compared with that of a 4× lens using TA and DL methods. While the balance between sensitivity and specificity was affected at 10×, overall accuracy was not improved, and the longer acquisition time makes it less suitable for time-sensitive intraoperative use. These findings can also inform other optical imaging systems that apply LBP for TA and ViT for DL in classification tasks [31, 35].

Several limitations should be acknowledged in this study. First, because our comparison is conducted at the margin level on identical specimens rather than patch-by-patch, the 4× and 10× images necessarily capture different contextual and structural information, patch sizes are scaled by magnification (2.5×) to provide comparable tissue context rather than to enforce exact physical-area equivalence, and this difference reflects the magnification effect under study rather than a confound [36]. While analyzing the whole-slide images without patch-based processing would be ideal, the stitched MUSE images are exceptionally large (often exceeding 10,000 × 10,000 pixels). Passing images of this dimensionality directly through a ViT or extracting LBP features without aggressive downsampling exceeds current GPU memory limits and destroys the fine micro-textural details required for diagnosis. Thus, patch-based processing is a computational necessity.

Second, the 10× objective lens has a limited depth of field, and some regions of the MUSE images may be slightly out of focus, potentially affecting classification accuracy. Third, the dataset was imbalanced at the patch level, with more than 75% benign patches and less than 25% malignant patches at both magnifications. This scenario might affect the margin-level decision. In the case of small malignant areas, where majority voting may allow the larger number of benign patches to dominate, this can lead to false-negative margin predictions.

Future work should consider adaptive patch-sizing methods that better capture tissue context across different magnifications. In addition, the models were not externally validated on an independent dataset, which will be necessary to confirm generalizability. Automated focusing techniques might be worth applying to improve image quality, especially at higher magnifications. In addition, a weighted majority voting method can be explored to improve MUSE's margin-level classification by assigning greater weight to suspicious patches. Collecting more malignant samples will help balance the dataset and improve the model's robustness and generalizability.

5. Conclusion

This study investigated how 4× and 10× magnifications affect margin-level classification of MUSE images using both TA and DL methods, employing a patch-level framework. For TA, the red channel of each patch was preprocessed, LBP features were extracted, and an SVM predicted the patch-level class. For DL, patch embeddings were processed through a ViT with an MLP head to get patch-level predictions. In both methods, patch predictions were combined and compared to a binary threshold to predict the final MUSE image prediction. Quantitative results showed that DL and TA performed consistently across both magnifications, though TA was slightly less accurate. Overall, 10× magnification did not improve margin-level accuracy compared to 4×. Since 10× takes longer to acquire and requires more computation, 4× is a practical and efficient choice for intraoperative margin assessment. Overall, MUSE has the potential to require less time than traditional histopathology methods, such as frozen-section analysis, and to provide diagnostic results using DL, thereby reducing the need for manual image interpretation by pathologists. These findings can help develop more accurate diagnostic models, support careful margin inspection during breast-conserving surgery, and improve patient outcomes.

Funding. National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number R01EB033806. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgment. The authors thank the Medical College of Wisconsin Tissue Bank for providing de-identified human breast tissue specimens.

Disclosures. The authors have no conflicts of interest to claim.

Data Availability Statement. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. R. L. Siegel, T. B. Kratzer, A. N. Giaquinto, *et al.*, “Cancer statistics, 2025,” *CA: A Cancer Journal for Clinicians*, vol. 75, no. 1, p. 10, 2025.
2. E. A. Davis and B. N. Joe, “Rising Rates of Distant-Stage Breast Cancer: Trends Among US Women,” *Radiology: Imaging Cancer*, vol. 7, 2025, e259012.
3. M. L. Czajka and C. Pfeifer, “Breast cancer surgery,” in *StatPearls [Internet]*, StatPearls Publishing, 2023.
4. K. Kaczmarek, P. Wang, R. Gilmore, *et al.*, “Surgeon re-excision rates after breast-conserving surgery: a measure of low-value care,” *Journal of the American College of Surgeons*, vol. 228, pp. 504–512.e502, 2019.
5. L.-J. Tseng, A. Matsuyama, and V. MacDonald-Dickinson, “Histology: The gold standard for diagnosis?,” *The Canadian Veterinary Journal*, vol. 64, p. 389, 2023.
6. A. Nayyar, K. K. Gallagher, and K. P. McGuire, “Definition and management of positive margins for invasive breast cancer,” *Surgical Clinics*, vol. 98, pp. 761–771, 2018.
7. P. S. Mullenix, D. G. Cuadrado, S. R. Steele, *et al.*, “Secondary operations are frequently required to complete the surgical phase of therapy in the era of breast conservation and sentinel lymph node biopsy,” *The American Journal of Surgery*, vol. 187, pp. 643–646, 2004.
8. F. Fereidouni, Z. T. Harmany, M. Tian, *et al.*, “Microscopy with ultraviolet surface excitation for rapid slide-free histology,” *Nature Biomedical Engineering*, vol. 1, pp. 957–966, 2017.
9. W. Xie, Y. Chen, Y. Wang, *et al.*, “Microscopy with ultraviolet surface excitation for wide-area pathology of breast surgical margins,” *Journal of Biomedical Optics*, vol. 24, 026501, 2019.
10. T. Lu, J. M. Jorns, M. Patton, *et al.*, “Rapid assessment of breast tumor margins using deep ultraviolet fluorescence scanning microscopy,” *Journal of Biomedical Optics*, vol. 25, 126501, 2020.
11. R. C. Niemeier, G. N. McKay, M. T. Chen, *et al.*, “Sequential staining technique for acquiring paired images with H&E and microscopy with ultraviolet surface excitation (MUSE),” in *Photonics in Dermatology and Plastic Surgery 2022*, SPIE, 2022, pp. 20–25.
12. T. Yoshitake, M. G. Giacomelli, L. M. Quintana, *et al.*, “Rapid histopathological imaging of skin and breast cancer surgical specimens using immersion microscopy with ultraviolet surface excitation,” *Scientific Reports*, vol. 8, p. 4476, 2018.
13. F. Fereidouni, A. D. Mitra, S. Demos, *et al.*, “Microscopy with UV Surface Excitation (MUSE) for slide-free histology and pathology imaging,” in *Optical Biopsy XIII: Toward Real-Time Spectroscopic Imaging and Diagnosis*, SPIE, 2015, pp. 46–51.
14. C. Kolluru, A. Todd, A. R. Upadhye, *et al.*, “Imaging peripheral nerve micro-anatomy with MUSE, 2D and 3D approaches,” *Scientific Reports*, vol. 12, p. 10205, 2022.
15. R. M. Levenson and F. Fereidouni, “MUSE: a new, fast, simple microscopy method for slide-free histology and surface topography,” *The FASEB Journal*, vol. 30, pp. 51.53–51.53, 2016.
16. T. Lu, J. M. Jorns, D. H. Ye, *et al.*, “Automated assessment of breast margins in deep ultraviolet fluorescence images using texture analysis,” *Biomedical Optics Express*, vol. 13, pp. 5015–5034, 2022.
17. P. Afshin, D. Helminiak, T. Niu, *et al.*, “Self-learned representation-guided latent diffusion model for breast cancer classification in deep ultraviolet whole surface images,” in *2026 IEEE 23rd International Symposium on Biomedical Imaging (ISBI)*, London, United Kingdom, 2026, pp. 1–5, doi: 10.1109/ISBI61048.2026.11516002.
18. K.-C. Song, Y.-H. Yan, W.-H. Chen, X. Zhang, “Research and perspective on local binary pattern,” *Acta Automatica Sinica*, vol. 39, no. 6, pp. 730–744, 2013.
19. M. Heikkilä, M. Pietikäinen, C. Schmid, “Description of interest regions with local binary patterns,” *Pattern Recognition*, vol. 42, no. 3, pp. 425–436, 2009.
20. T. Niu, T. Lu, J. M. Jorns, *et al.*, “Texture Analysis of MUSE Images of Different Resolutions for Breast Tumor Margin Detection,” in *Bio-Optics: Design and Application*, Optica Publishing Group, 2025, p. DTu3A.4.
21. T. To, T. Lu, J. M. Jorns, *et al.*, “Deep learning classification of deep ultraviolet fluorescence images toward intra-operative margin assessment in breast cancer,” *Frontiers in Oncology*, vol. 13, p. 1179025, 2023.
22. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Backpropagation applied to handwritten zip code recognition,” in *Neural Computation*, vol. 1, no. 4, 1989, pp. 541–551.
23. S. S. Ghahfarokhi, T. To, J. Jorns, T. Yen, B. Yu, and D. H. Ye, “Deep learning for automated detection of breast cancer in deep ultraviolet fluorescence images with diffusion probabilistic model,” in *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, IEEE, May 2024, pp. 1–5.
24. P. Afshin, D. Helminiak, T. Lu, T. Yen, J. M. Jorns, M. Patton, B. Yu, and D. H. Ye, “Breast cancer classification in deep ultraviolet fluorescence images using a patch-level vision transformer framework,” in *2025 47th Annual*

- International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2025, pp. 1–6, doi:10.1109/EMBC58623.2025.11253275.
25. S. Wan, H. C. Lee, X. Huang, T. Xu, T. Xu, X. Zeng, *et al.*, “Integrated local binary pattern texture features for classification of breast tissue imaged by optical coherence microscopy,” *Medical Image Analysis*, vol. 38, pp. 104–116, 2017.
 26. P. Král and L. Lenc, “LBP features for breast cancer detection,” in *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2016, pp. 2643–2647.
 27. A. Dosovitskiy, “An image is worth 16×16 words: Transformers for image recognition at scale,” arXiv preprint arXiv:2010.11929, 2020.
 28. K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *Journal of Big Data*, vol. 3, p. 9, 2016.
 - Pietikainen2010, Liao2009.
 29. M. Pietikäinen, “Local binary patterns,” *Scholarpedia*, vol. 5, p. 9775, 2010.
 30. S. Liao, M. W. Law, and A. C. Chung, “Dominant local binary patterns for texture classification,” *IEEE Transactions on Image Processing*, vol. 18, pp. 1107–1118, 2009.
 31. B. Gheflati and H. Rivaz, “Vision transformers for classification of breast ultrasound images,” in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2022, pp. 480–483.
 32. L. van der Maaten and G. Hinton, “Visualizing data using t-SNE,” *Journal of Machine Learning Research*, vol. 9, no. 11, pp. 2579–2605, 2008.
 33. R. Kingslake, “Optical system design,” Academic Press, 2012.
 34. D. N. Frolov, O. A. Vinogradova, V. N. Frolov, *et al.*, “Optical design and unification of optical systems of objectives for microscopes,” in *Optical Design and Engineering VII*, SPIE, 2018, pp. 428–439.
 35. S. Naresh and S. V. Kumari, “Breast cancer detection using local binary patterns,” *International Journal of Computer Applications*, vol. 123, 2015.
 36. O. Akbarzadeh, M. R. Khosravi, B. Khosravi, *et al.*, “Medical image magnification based on original and estimated pixel selection models,” *Journal of Biomedical Physics & Engineering*, vol. 10, p. 357, 2020.
 37. T. Niu, E. Ampo, J. M. Jorns, M. Patton, T. Lu, D. H. Ye, T. W. F. Yen, and B. Yu, “Effective transfer of tumor annotations from hematoxylin and eosin to fluorescence images of breast and lung tissues,” *Journal of Biomedical Optics*, vol. 31, no. 1, pp. 016501–016501, 2026.