Publishing Excellence
Inspiring Progress
ARISTON PUBLICATIONS

CompSci & AI Advances

ISSN: xxxx-xxxx https://aristonpubs.com/computers-ai-advances https://doi.org/10.69626/cai.2024.0140

RESEARCH ARTICLE

Enhanced Breast Cancer Classification from Histopathological Images Using Integrated SVM and CNN Frameworks

Ashish Thapa [1], Sami Azam [2], Nur Alam MD [1], Zubaer Ibna Mannan [1, *]

ABSTRACT: Breast cancer, particularly Invasive Ductal Carcinoma (IDC), remains one of the most prevalent and life-threatening cancers affecting women globally. Early and accurate detection is critical for improving treatment outcomes, making automated classification systems increasingly vital in healthcare. This study proposes an automated breast cancer classification system that integrates Support Vector Machine (SVM) and Convolutional Neural Network (CNN) classifiers, utilizing features such as data patches and color variations extracted from histopathological images. The proposed system focuses on accurately classifying breast cancer histopathological images into benign or malignant categories. The SVM classifier demonstrates superior performance, outperforming CNN and traditional classifiers such as Decision Tree, Naïve Bayes, and Nearest Neighbors in terms of accuracy, sensitivity, and specificity. It also surpasses CNN in key metrics, including precision, recall, and F1-score, establishing its robustness and reliability in image classification tasks. The study evaluates the proposed SVM architecture on the IDC dataset, achieving an accuracy of 94%. This result underscores the potential of the SVM classifier for developing computer-assisted diagnostic (CAD) systems, which can enhance the efficiency and accuracy of breast cancer diagnosis. Furthermore, the system demonstrates the ability to localize cancerous tissues within whole histopathological images, offering a comprehensive diagnostic tool for pathologists. This study highlights the effectiveness of integrating machine learning models for breast cancer classification and emphasizes the importance of automated systems in augmenting diagnostic accuracy, reducing workload, and advancing personalized cancer care.

Keywords: Plant Leaf Disease Detection, Image Processing, Grayscale Conversion, Thresholding Techniques, Segmentation.

Received: 05 April 2024; Revised: 30 May 2024; Accepted: 05 July 2024; Published Online: 10 July 2024

1. INTRODUCTION

Breast cancer is the most prevalent cancer among females in the whole world according to WHO, with a ratio of 1:8 in the USA [1]. Detection of breast cancer involves three procedures: screening, biopsy, and core needle biopsy. Mammography screening is the most widely used screening method for cancer detection, consisting of an X-ray device used to analyze the breast tissues [2-3]. If the screening procedure shows that the patient has the risk of developing malignant tissue, a breast biopsy is suggested. Histopathological analysis on the tissue acquire from biopsy helps to differentiate between normal (benign) and abnormal (malignant) lesions [4]. Histopathological staining is employed to visualize the structures in a slide with the help of a microscope. The most famous staining system is hematoxylin and eosin (H&E) [5]. However, the time-consuming manual assessment of stained histopathology slides are prone to intra- and inter-observer variability and suffers from low throughput [6-7]. Therefore, a growing interest is emerging for digital pathology (DP) where the computer-assisted diagnosis (CAD) relieves the workload of pathologists and swiftly performs a digital assessment [8-9].

Digital pathology refers to machine learning-based classification, detection, and segmentation of histopathology

[1] Department of Smart Computing, Kyungdong University (Global Campus), Gangwon-do, Republic of Korea.
[2] College of Engineering, IT and Environment, Charles Darwin University, NT, Australia.

* Author to whom correspondence should be addressed: zimannan@kduniv.ac.kr (Z. I. Mannan)

images [10]. Current developments in artificial intelligence and computer vision have enriched the modules of CAD systems that can competently detect the breast cancer using histopathological images [11]. The complex tissue textures, overlapping of cell nuclei, abnormal color variations caused by stain, and light normalization of the histopathological images are the main challenges for detection and classification of breast cancer.

For classification, SVM is used to classify the multiscale color texture features [12-13]. Thus, the classifier employs the benefits of the multiresolution structure of complex wavelet for multiscale feature extraction. The use of color texture analysis techniques for diagnosing histopathological sample images of breast has not been reported in the literature to our best knowledgee. This paper attempts to use SVM or multiple SVMs to classify the multiscale color texture features and eventually classify the cancerous region. For example, a raw image and breast tissue slice are shown in Figure 1(a) and (b) whereas the SVM is going employed on the possible cancerous patch (colored in red, shown in Figure 1(c)) to predict whether the patches are malignant or benign.

In this paper we proposed a novel SVM based classifier that classify the multiscale color texture features. We also employed a CNN classifier and compare between the proposed SVM and CNN classifier where SVM outperforms CNN classifier. We compare the SVM classifier with other classifiers such as Decision Tree, Naïve Bayes, and Nearest Neighbors, where SVM outperforms all of these classifier in accuracy, sensitivity, and specificity aspects.

2. RELATED WORKS

Pathologists grade cancer by observing the micro and macro-structures present in the histopathology slides. Conventional approaches in CAD involve feature-extraction based on the texture and appearances of the nuclei and its micro-environment [11-13]. These features are fed to analysis methods like fuzzy-C means, Gaussian mixture models, SVM, MLP and clustering algorithms to decide the class for the histopathology entities like nuclei or patch. These methods were popular with small datasets, however with large dataset they fail to generalize.

Deep learning models are well suited to the cases with large amount of data, as they can learn intricate features from the histopathology slides and generalize well across patients, disease conditions, and hospitals and are even robust to human-induced errors in the slide preparation. Deep learning-based approaches are often based on CNNs [14], and SVM [15], but little work has been done on visualization or localization on histopathology datasets. Gradient-based methods such as guided-backpropagation and deconvolution [16] provide good visualization on natural images, but fail to provide reasonable localization on histopathology images due to challenges of large size, variation across disease states, and human-induced errors in slide preparation.

Attention-based multiple instances learning (SVM) is an effective approach for classification and localization of relevant areas in histopathology images. SVM is efficient in high dimensional spaces, and it is memory efficiency and has ability to specify different kernel functions to make a proper decision boundary [17]. However, Artificial Neural Networks (ANNs) suffer with hardware dependence, determination of proper network structure and unexplained behaviors of ANNs itself [18]. Recurrent Neural Network (RNN) are complex and suffers with gradient vanishing and exploding problems. Training an RNN is very difficult task, and it cannot process very long sequences if using tanh or relu as an activation function [19].

3. PROPOSED ARCHITECTURE

3.1. Support Vector Machine (SVM)

Support Vector Machines (SVMs) are the most widely used machine learning-based pattern classification technique. The primary aim of this statistical technique is to project nonlinearly separable samples onto another higher dimensional space by using different types of kernel functions [20].

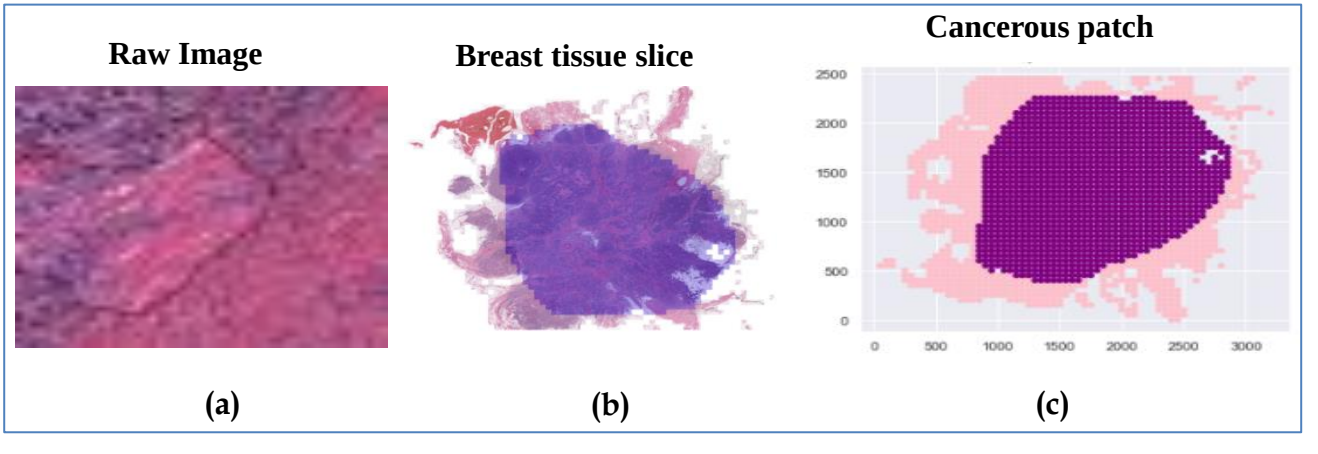

Fig. 1. Breast cancer tissue. (a) Raw image, (b) breast tissue slice, and (c) cancerous patch.

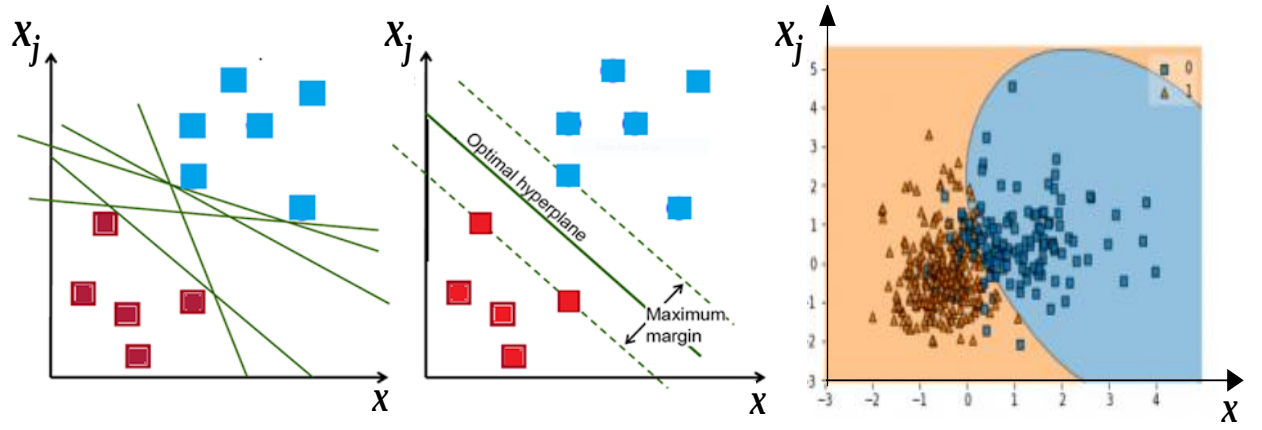

Fig. 2. Finding technique of hyper-planes in the SVM.

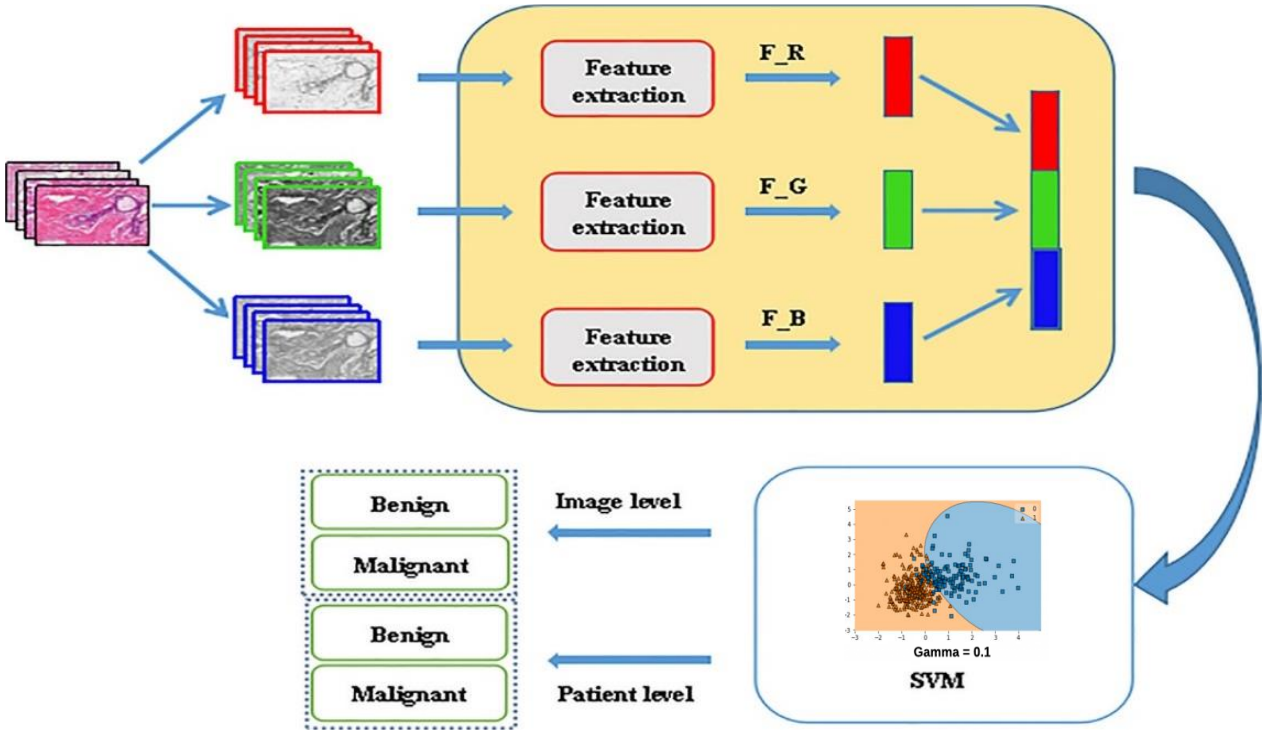

Fig. 3. Finding technique of hyper-planes in the SVM.

There are three top kernels used in the field such Linear, Polynomial, and Radial Basis Function (RBF). At its simplest, an SVM is looking to find a way to separate the variables into classes based on the target value (as shown in Eq. 1). The goal of the SVM is to find the optimal hyper-plane between the data as shown in Figure 2. The graph on the right shows the optimal hyperplane that creates the largest margin between the classes.

$$K(\bar{x}) = \begin{cases} 1 & if \ \ \|\bar{x}\| \leq 1 \\ 0 & otherwise \end{cases} \quad (1)$$

The proposed SVM based classifier framework, shown in Figure 3, used Radial Basis Function (RBF) kernel functions for the classification between benign and malignant classes as RBF is used to nonlinearly separate the classes. Specifically, RBF performs proper separation between the classes of the unseen data. The RBF kernel function computes the similarity or the distance between two points x and xj, which can be mathematically represented as follows:

$$F\ (x, x_j) \ = \ exp\ \left(-\frac{\|x- \ x_j\|^2}{2\sigma^2}\right) \quad (2)$$

Where, 'σ' is the variance and hyperparameter and ||x - xj|| is the Euclidean distance between the two points. We used σ = 0.1 in our SVM classifier, as one of the best outcomes for separating the classes can be achieved with this hyper parameter value [21].

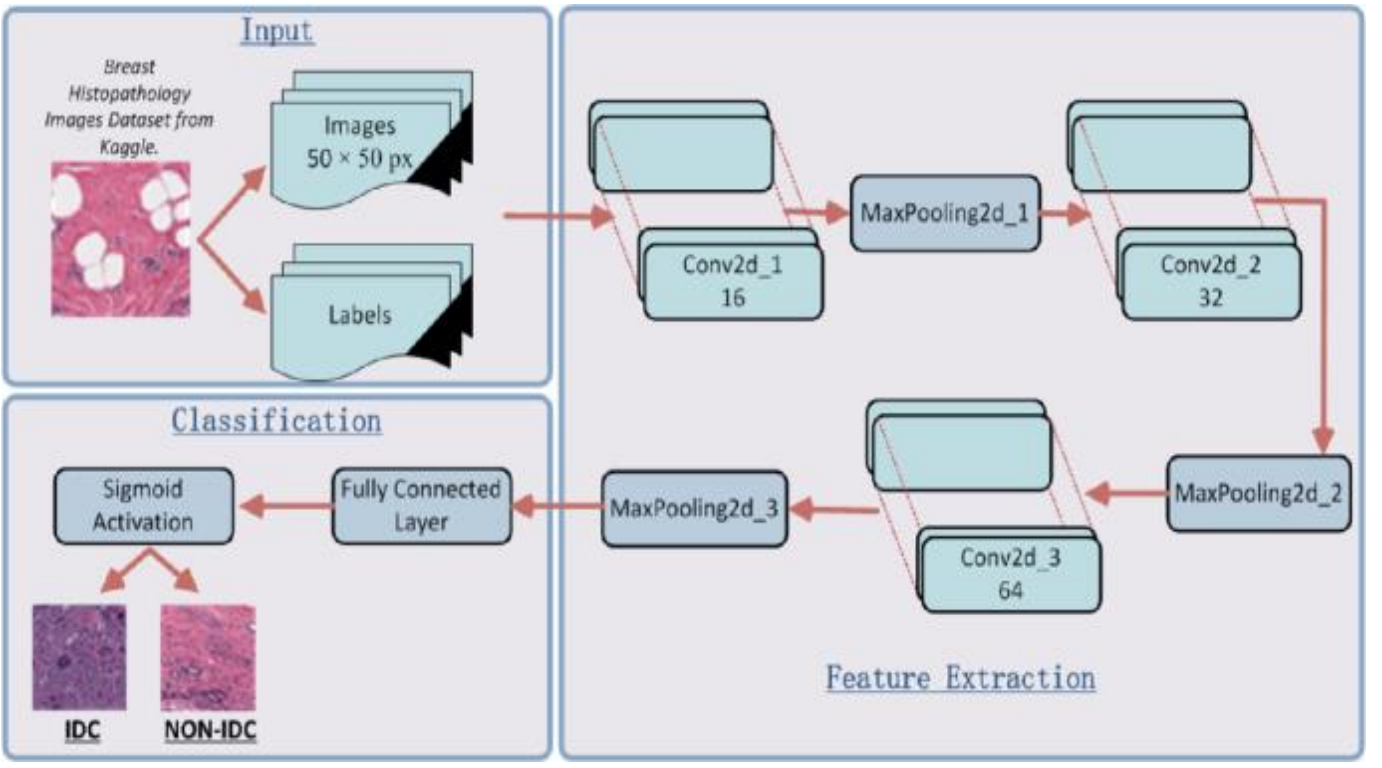

Fig. 4. IDC classification of histopathology images using CNN.

3.2. Convolution Neural Network (CNN)

For CNN based classifier, same IDC image dataset is applied to classify the cancer. Same data preprocessing and augmentation techniques are employed for detection and color normalization. For feature extraction from the CNN-based classifier, shown in Figure 4, we employed three convolution layers followed by max-pooling layers. In the convolution layers, we employed 3×3 convolution with 16, 32, and 64 filters along with Relu as an activation function in each CNN layers. We used sigmoid activation for the classifier along with the binary cross entropy as loss function.

4. EXPERIMENTAL SETUP

4.1. Dataset

The dataset used in this paper contains images from automatic detection of IDC in whole slide images [22- 24]. The data itself contains 162 whole mount slide images of Breast Cancer specimens scanned at 40x. Each of the slides is resized into 50×50 patches, equating to 277,524 total images. Each of the patches is labeled as malignant or benign.

4.2. Preprocessing and Data Augmentation

Pre-processing is the most important step in histopathology image analysis due to poor captured image quality, noises, changes in resolution, and weak contrast [25]. To avoid aforementioned limitation, we employed HSV color model on the microscopic RGB images. HSV remaps the RGB primary colors into 3-dimensions where Hue specifies the angle of the color on the RGB color circle, Saturation controls the amount of color used, and Value controls the brightness of the color. To avoid false classification and complexities, we also used gray microscopic images to extract features for SVM and CNN classifiers.

To avoid overfitting, data augmentation is often performed for the training process after dataset splitting. The strategies used include random rotation, flipping transformation, and shearing transformation. Random rotation involves rotating the images randomly, while flipping transformation involves rotating the images with fixed angles. Shearing transformation involves zooming in or zooming out images in different directions [26].

4.3. Implementation

The implementation of the proposed work involved four main stages: loading the dataset, analyzing the data, splitting the dataset into training and testing sets, and creating a simple SVM to train and test on the dataset, as shown in Figure 5. The first step is to employ the image preprocessing algorithms and resize the microscopic images of the IDC dataset. Open CV (a common computer vision package in Python) is used to read the images and fatten them into a Pandas data frame.

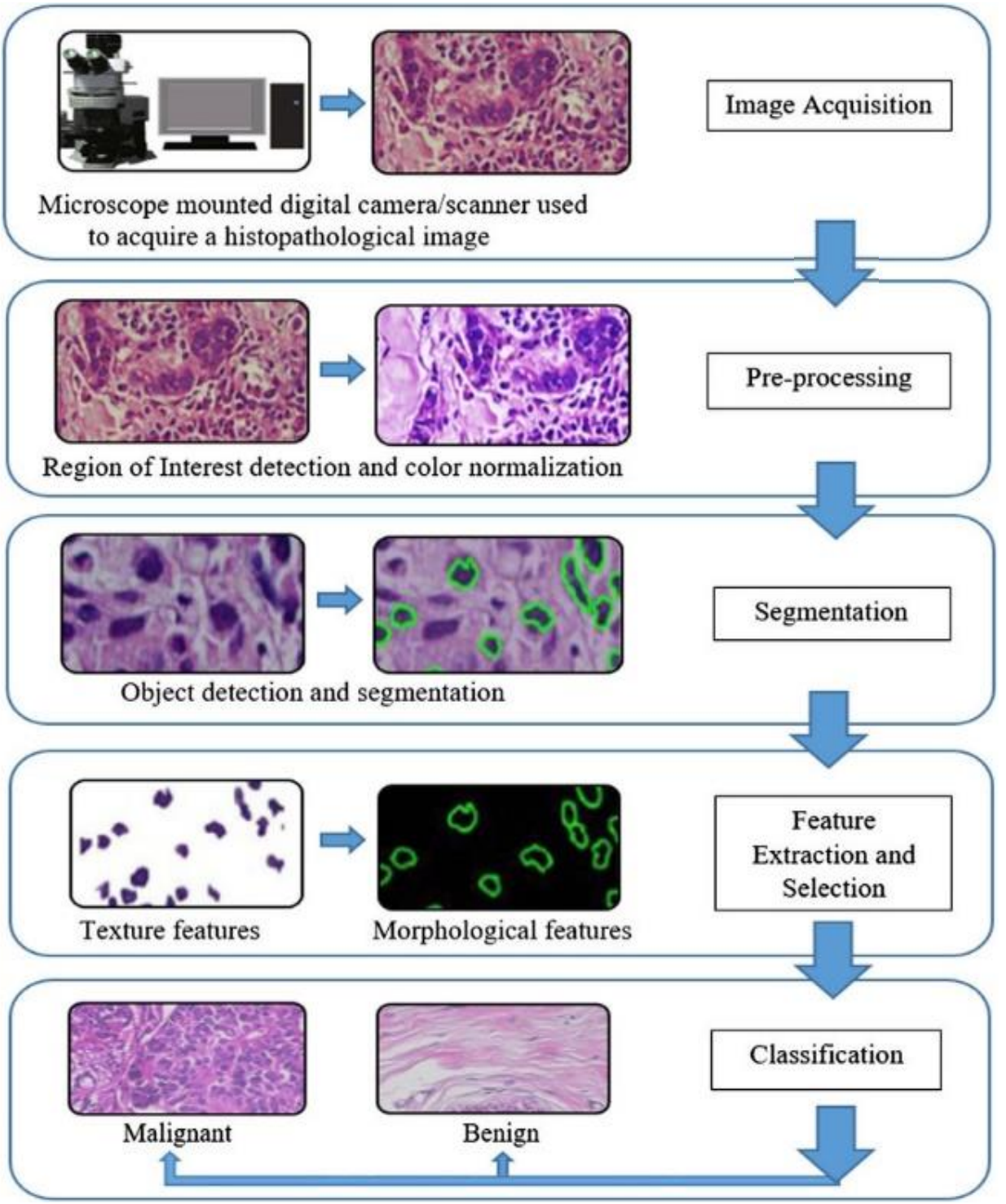

Fig. 5. Comprehensive view of the proposed SVM based CAD system for breast cancer using histopathology.

The data is skewed towards non-cancerous, so we alter the data set and test data, so that the number of non-cancerous patches is equal to the number of cancerous patches. We feed-forward non-cancerous and cancerous data-augmented patches to the proposed SVM and CNN classifier models, mentioned in Section 3, for training and testing on the IDC dataset using the sklearn library. In this work, we split the dataset into 70% for training and 30% for validation, and testing.

4.4. Evaluation Matrices

For fair comparison of the classifications obtained by SVM and CNN, we calculate the precision (which defines the ability of the model to locate relevant objects only), recall

(which evaluates true positive detections relative to all ground truths), and F1 score (the harmonic-mean between precision and recall) for the classification tasks with the following expressions:

$$Precision, P = \frac{TP}{TP+FP}, \quad (3)$$

$$Recall, P = \frac{TP}{TP+FN}\ , \quad (4)$$

$$F_1\ \ Score, F_1 = \frac{2\times P\times R}{(P+R)}\ , \quad (5)$$

Where TP, FP, and FN are the true positive, false positive, and false negative detections of the model, respectively.

To compare the proposed SVM and CNN classifiers with Decision Tree, Naïve Bayes, and Nearest neighbors classifiers, we compute accuracy (which evaluates number of correct predictions in total number of predictions), sensitivity (provide the opportunity to assess the input variables in terms of the importance of their impact on the output variable and indicate insignificant variables), and specificity (which evaluate model's ability to predict a true negative of each category available), for the classification tasks with the following expressions:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}\ , \quad (6)$$

$$Sensitivity = \frac{TP}{TP+FN}\ , \quad (7)$$

and

$$Specificity = \frac{TN}{TN+FP}\ . \quad (8)$$

5. RESULTS AND DISCUSSION

This section discusses the comparative performance of the proposed SVM classifier against various models, including CNN, Decision Tree, Naïve Bayes, and Nearest Neighbors. The performance evaluation is based on key metrics, such as precision, recall, F1-score, accuracy, sensitivity, and specificity. These metrics are essential for determining the effectiveness of classification models in distinguishing between benign and malignant breast cancer histopathological images.

The first part of the analysis compares the performance of the SVM and CNN classifiers in terms of precision, recall, and F1-score, as detailed in Table 1. For benign cases, the SVM classifier achieves a precision of 0.91, recall of 0.99, and F1-score of 0.95, which are slightly higher than the CNN classifier's corresponding values of 0.90, 0.99, and 0.94. For malignant cases, the SVM classifier records a precision of 0.98, recall of 0.84, and F1-score of 0.91, outperforming the CNN classifier's F1-score of 0.89 and recall of 0.81.

The macro and weighted averages also highlight the superiority of the SVM classifier, which achieves a macro average precision, recall, and F1-score of 0.95, 0.92, and 0.93, respectively, compared to CNN's averages of 0.94, 0.90, and 0.92. These results clearly indicate that the SVM classifier has better overall performance, particularly for malignant classifications where accurate identification is critical for timely intervention.

The better performance of the SVM classifier can be attributed to its ability to handle high-dimensional data effectively by finding the optimal hyperplane for separation. Conversely, the CNN classifier struggles when the dataset contains overlapping target classes, occluded boundaries, and noise in microscopic images. The noise hampers the CNN classifier's ability to distinguish between subtle differences in image features, which is crucial in histopathological image analysis. This limitation necessitates additional post-processing steps, such as voting methods, to mitigate noise-related misclassifications, thereby increasing computational overhead.

In the second part of the analysis, the SVM and CNN classifiers are compared with other traditional machine learning models, including Decision Tree, Naïve Bayes, and Nearest Neighbors, as summarized in Table 2. The SVM classifier achieves the highest accuracy (94%), sensitivity (95%), and specificity (94%), significantly outperforming the other models. In contrast, the Decision Tree classifier shows the poorest performance, with an accuracy of 65.19%, sensitivity of 66.41%, and specificity of 63.96%.

Table 1. IDC classification with SVM and CNN classifiers.

IDC Classification	SVM Classifier			CNN Classifier		
	Precision	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
Benign	0.91	0.99	0.95	0.90	0.99	0.94
Malignant	0.98	0.84	0.91	0.98	0.81	0.89
Accuracy		0.94			0.92	
Macro avg.	0.95	0.92	0.93	0.94	0.90	0.92
Weighted avg.	0.94	0.94	0.93	0.93	0.92	0.92

Table 2. Comparison with other classifiers.

Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)
CNN	92	92	93
Decision Tree	65.19	66.41	63.96
Naïve Bayes	65.95	64.65	67.25
Nearest neighbors	65.72	69.47	61.97
SVM	94	95	94

The Naïve Bayes and Nearest Neighbors classifiers also exhibit relatively lower performance, with accuracies of 65.95% and 65.72%, respectively. The sensitivity and specificity values for these classifiers remain in the range of 61% to 69%, making them less suitable for clinical applications where high sensitivity and specificity are paramount. The performance superiority of the SVM classifier stems from its ability to integrate clinically important features, such as color, geometrical shape, and texture, for robust image classification. This integration ensures that the classifier effectively distinguishes between benign and malignant tissues based on subtle histopathological differences. Additionally, the SVM model demonstrates better generalization capabilities, maintaining consistent performance across multiple classification trials.

The results highlight the potential of the SVM classifier for practical applications in computer-assisted diagnosis (CAD) systems. Its high sensitivity ensures accurate identification of malignant cases, reducing the risk of false negatives, which are critical in breast cancer diagnosis. Similarly, its high specificity minimizes false positives, alleviating unnecessary psychological stress for patients and avoiding unnecessary medical interventions. The ability of the SVM model to outperform CNN and traditional classifiers underscores its robustness and reliability in handling complex datasets with noise and overlapping class distributions. By leveraging the SVM classifier, CAD systems can provide pathologists with a powerful tool to enhance diagnostic accuracy, reduce workload, and improve efficiency. Furthermore, the SVM's scalability and consistency make it a viable candidate for real-time clinical applications.

While the SVM classifier demonstrates remarkable performance, further work is needed to address some limitations. For instance, the current dataset's noise and overlapping boundaries highlight the importance of data preprocessing techniques to further improve classifier performance. Additionally, incorporating hybrid models that combine SVM with feature extraction methods from deep learning models, such as CNN, could enhance classification accuracy and robustness. Future studies could also focus on expanding the dataset with more diverse histopathological images to improve generalizability. Incorporating explainability techniques could make the SVM model's decisions more interpretable for clinicians, fostering trust in AI-based diagnostic tools. The results demonstrate the efficacy of the proposed SVM classifier in breast cancer classification and its potential to revolutionize histopathological image analysis for improved patient outcomes.

6. CONCLUSION

This study introduces an automated classification system for breast cancer histopathological images, utilizing an integrated Support Vector Machine (SVM) classifier that combines multiple feature sets, such as data patches and color variations. By focusing on accurate and reliable classification of breast cancer into benign and malignant categories, the proposed model provides a significant contribution to the field of computer-assisted diagnosis (CAD). The performance of the SVM model was rigorously evaluated against other machine learning models, including Convolutional Neural Networks (CNN), Decision Trees, Naïve Bayes, and Nearest Neighbors. Notably, the SVM classifier outperformed CNN in key metrics, achieving higher precision, recall, and F1-score values. It also exhibited superior performance in terms of accuracy, sensitivity, and specificity when compared to traditional classifiers. On the IDC dataset, the SVM model achieved an impressive accuracy of 94%, underscoring its effectiveness and robustness. A major advantage of the proposed SVM-based approach is its ability to localize cancerous tissues within whole histopathological images, enabling a more targeted diagnostic process. This capability enhances its potential as a valuable tool in CAD systems, which are increasingly recognized for their ability to reduce pathologists' workloads while improving diagnostic speed and consistency. While the results are promising, future work could focus on integrating deep learning techniques for feature extraction and leveraging larger, more diverse datasets to enhance the system's generalizability. Additionally, incorporating explainability modules could provide clinicians with better insights into classification decisions, fostering trust in AI-based systems. This study demonstrates the potential of the

proposed SVM model for automated breast cancer classification, paving the way for efficient, accurate, and scalable diagnostic solutions in modern healthcare.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests.

ACKNOWLEDGEMENT

This study is supported by Kyungdong University Research Fund, 2021.

REFERENCES

[1] Alqudah, A. and Alqudah, A.M., **2022.** Sliding window based support vector machine system for classification of breast cancer using histopathological microscopic images. *IETE Journal of Research*, *68*(1), pp.59-67.

[2] Elmore, J.G., Armstrong, K., Lehman, C.D. and Fletcher, S.W., **2005.** Screening for breast cancer. *Jama*, *293*(10), pp.1245-1256.

[3] Roque, A.C. and André, T.C., **2002.** Mammography and computerized decision systems: a review. *Annals of the New York Academy of Sciences*, *980*(1), pp.83-94.

[4] Freedman, R.A., Keating, N.L., Pace, L.E., Lii, J., McCarthy, E.P. and Schonberg, M.A., **2017.** Use of surveillance mammography among older breast cancer survivors by life expectancy. *Journal of Clinical Oncology*, *35*(27), pp.3123-3130.

[5] Contourlet, M.U., **2013.** A robust approach to classify microcalcification in digital mammograms using contourlet transform and support vector machine. *American Journal of Engineering and Applied Sciences*, *6*(1), pp.57-68.

[6] Spanhol, F.A., Oliveira, L.S., Petitjean, C. and Heutte, L., **2015.** A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, *63*(7), pp.1455-1462.

[7] Elmore, J.G., Longton, G.M., Carney, P.A., Geller, B.M., Onega, T., Tosteson, A.N., Nelson, H.D., Pepe, M.S., Allison, K.H., Schnitt, S.J. and O'Malley, F.P., **2015.** Diagnostic concordance among pathologists interpreting breast biopsy specimens. *Jama*, *313*(11), pp.1122-1132.

[8] Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M. and Yener, B., **2009.** Histopathological image analysis: A review. *IEEE reviews in biomedical engineering*, *2*, pp.147-171.

[9] Ilyas, T., Mannan, Z.I., Khan, A., Azam, S., Kim, H. and De Boer, F., **2022**. TSFD-Net: Tissue specific feature distillation network for nuclei segmentation and classification. *Neural Networks*, *151*, pp.1-15.

[10] Madabhushi, A. and Lee, G., **2016.** Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical image analysis*, *33*, pp.170-175.

[11] Filipczuk, P., Fevens, T., Krzyżak, A. and Monczak, R., **2013.** Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. *IEEE Transactions on Medical Imaging*, *32*(12), pp.2169-2178.

[12] Zhang, Y., Zhang, B., Coenen, F. and Lu, W., **2013.** Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles. *Machine Vision and Applications*, *24*(7), pp.1405-1420.

[13] George, Y.M., Zayed, H.H., Roushdy, M.I. and Elbagoury, B.M., **2013**. Remote computer-aided breast cancer detection and diagnosis system based on cytological images. *IEEE Systems Journal*, *8*(3), pp.949-964.

[14] Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., Polónia, A. and Campilho, A., **2017**. Classification of breast cancer histology images using convolutional neural networks. *PloS one*, *12*(6), p.e0177544.

[15] Spanhol, F.A., Oliveira, L.S., Petitjean, C. and Heutte, L., **2016**, July. Breast cancer histopathological image classification using convolutional neural networks. In *2016 international joint conference on neural networks (IJCNN)* (pp. 2560-2567). IEEE.

[16] Rawat, W. and Wang, Z., **2017**. Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*, *29*(9), pp.2352-2449.

[17] Bo'riboyevna, A.G.Z. and Bo'riboyevna, A.D., **2023.** Support vector machine algorithm. *Central asian journal of mathematical theory and computer sciences*, *4*(5), pp.164-167.

[18] Kathirgamanathan, A., Patel, A., Khwaja, A.S., Venkatesh, B. and Anpalagan, A., **2022**. Performance comparison of single and ensemble CNN, LSTM and traditional ANN models for short-term electricity load forecasting. *The Journal of Engineering*, *2022*(5), pp.550-565.

[19] Marhon, S.A., Cameron, C.J. and Kremer, S.C., **2013**. Recurrent neural networks. *Handbook on Neural Information Processing*, pp.29-65.

[20] Laskov, P., **2002**. Feasible direction decomposition algorithms for training support vector machines. *Machine Learning*, *46*, pp.315-349.

[21] Liu, Q., Chen, C., Zhang, Y. and Hu, Z., **2011**. Feature selection for support vector machines with RBF kernel. *Artificial Intelligence Review*, *36*(2), pp.99-115.

[22] Deepa, B.G. and Senthil, S., **2022**. Predicting invasive ductal carcinoma tissues in whole slide images of breast Cancer by using convolutional neural network model and multiple classifiers. *Multimedia Tools and Applications*, *81*(6), pp.8575-8596.

[23] Janowczyk, A. and Madabhushi, A., **2016**. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, *7*(1), p.29.

[24] Cruz-Roa, A., Gilmore, H., Basavanhally, A., Feldman, M., Ganesan, S., Shih, N.N., Tomaszewski, J., González, F.A. and Madabhushi, A., **2017**. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. *Scientific Reports*, *7*(1), p.46450.

[25] Gupta, V., Chaurasia, V. and Shandilya, M., 2015. Random-valued impulse noise removal using adaptive dual threshold median filter. *Journal of Visual Communication and Image Representation*, *26*, pp.296-304.

[26] Zhu, C., Song, F., Wang, Y., Dong, H., Guo, Y. and Liu, J., 2019. Breast cancer histopathology image classification through assembling multiple compact CNNs. *BMC Medical Informatics and Decision Making*, *19*, pp.1-17.