



Breast Cancer Histopathology Image Classification Using a Set of Deep Learning Models and VGG16 and VGG19 Architectures

M. Shirin Manesh^{1,*}, V. Razmavar²

¹ Master of Science in Biomedical Engineering, Islamic Azad University, Garmsar Branch, Iran.

² Assistant Professor, Department of Biomedical Engineering, Faculty of Engineering, Islamic Azad University, Garmsar Branch, Iran.

ARTICLE INFO	ABSTRACT
<p>Article History: Received 6 June 2024 Received in revised form 23 September 2024 Accepted 13 November 2024 Available online 1 December 2024</p>	<p>Breast cancer is one of the most prevalent and serious public health challenges worldwide, being the leading cause of cancer-related deaths among women. Early detection is a critical factor in improving survival rates, as it allows for timely intervention and treatment. The complexity of diagnosing breast cancer from histopathology images has led to the development of advanced techniques using artificial intelligence (AI) and machine learning. This study introduces a novel deep learning ensemble approach to classify breast cancer histopathology images using publicly available datasets. The primary objective of this research is to improve the classification accuracy of breast cancer images by leveraging a combination of two deep learning models. The proposed approach utilizes the VGG16 and VGG19 models, which were both fine-tuned to enhance their performance. The results demonstrate that the ensemble method, which averages the predicted probabilities from both models, leads to a more robust classifier. Specifically, the fine-tuning of the VGG16 and VGG19 models contributes significantly to improving the model's performance. The ensemble model exhibits competitive results in classifying complex histopathology images of breast cancer, with a recall value of 97.73% for the cancer class in both the full training and fine-tuning approaches. This research highlights the effectiveness of ensemble learning in medical image classification, paving the way for more accurate and reliable tools in the diagnosis of breast cancer.</p>
<p>Keywords: Image Classification, Breast Cancer Histopathology, Deep Learning, VGG16 Architecture, VGG19 Architecture.</p>	

1. INTRODUCTION

Breast cancer is a leading cause of high mortality rates among women, commonly recognized as the primary cause of cancer-related deaths in women worldwide [1]. Typically, this cancer is diagnosed through a type of surgery [2], which is the most accurate method among available options; however, this surgery is very costly and highly invasive [3]. Additionally, it causes anxiety and psychological and emotional issues for the patient. These concerns have led to the gradual adoption of predictive systems in various medical fields [4]. Cancer is one of the most significant public health issues globally [5]. According to the Global Burden of Disease (GBD) study, in 2017, there

* Corresponding Author: shahdadshirinmanesh9@gmail.com

Master of Science in Biomedical Engineering, Islamic Azad University, Garmsar Branch, Iran



were 24.5 million cases of cancer [6] and 9.6 million cancer-related deaths worldwide [7]. These statistics reveal that cancer incidence globally increased by 33% from 2007 to 2017 [8]. Specifically, breast cancer is the most common malignancy and the leading cause of cancer-related deaths among women worldwide [9]. Therefore, early detection of this pathology is crucial to prevent its progression and reduce its complications in women [10]. Breast cancer is a heterogeneous disease consisting of multiple entities with distinct biological, histological, and clinical features [11].

This malignancy arises from the growth of abnormal cells in the breast and may invade nearby healthy tissues [12]. Its clinical screening is initially performed using radiology images, such as mammography, ultrasound imaging, and magnetic resonance imaging (MRI) [13]. However, these non-invasive imaging techniques may not always effectively identify cancerous regions [14]. In such cases, biopsy is commonly used for more comprehensive analysis of breast cancer malignancy [15]. The biopsy process involves collecting tissue samples, placing them on microscope slides, and staining these slides to enhance the visualization of nuclei and cytoplasm [16]. Pathologists then perform microscopic analysis of these slides to confirm the breast cancer diagnosis [17].

However, manual analysis of histopathological images is a time-consuming and tedious process, prone to errors due to the complex nature of the images [18]. Furthermore, the morphological criteria used to classify these images are somewhat subjective, resulting in an average diagnostic agreement among pathologists of approximately 75% [19]. Therefore, computer-aided diagnosis plays a significant role in assisting pathologists in analyzing histopathology images [20]. Specifically, it enhances the diagnostic accuracy of breast cancer by reducing variability in decisions among pathologists [21]. However, conventional computer diagnostic methods, ranging from rule-based systems to machine learning techniques, may not effectively handle intra-class diversity and inter-class consistency in breast cancer histopathology images [22]. Additionally, these methods largely rely on feature extraction techniques such as scale-invariant feature transform, robust speed features, and local binary patterns, all of which are based on supervised information and may yield biased results during classification [23].

Hence, the demand for efficient diagnosis has led to the development of advanced computational models based on multiple layers of nonlinear processing units, known as deep learning [24]. Recently, deep learning models have made significant advancements in computer vision, especially in biomedical image processing, due to their ability to learn complex and advanced features from images. They have been widely used in cancer detection [-27-25], prompting various researchers to employ these models, particularly Convolutional Neural Networks (CNN), because of their ability to effectively share parameters across different layers in a deep learning model. CNNs are extensively used in image-related tasks [28]. Several CNN-based architectures have been proposed in recent years. However, AlexNet is considered the first deep CNN to achieve notable accuracy in the large-scale ImageNet classification challenge (ILSVRC) in 2012. Subsequently, the VGG architecture introduced the concept of applying deeper networks with smaller convolutional filters and secured second place in ILSVRC 2014 [29-30]. The intuition behind using multiple smaller stacked filters is that it can provide an effective receptive field. This idea is also employed in recent pre-trained models, including the Inception Network and Residual Neural Network (ResNet) [31]. Based on the information presented, this study aims to utilize two different VGG architecture approaches for efficient breast cancer histopathology classification using the created datasets.

2. RESEARCH METHODOLOGY

The present research is descriptive-analytical in nature and is applied in terms of purpose. The primary objective of this study is to classify breast cancer histopathology images using a combination of deep learning models, employing two distinct VGG architecture approaches for efficient classification of breast cancer histopathology images based on the created datasets.

Initially, a private dataset of full slide images from breast cancer patients, previously created, was selected. Image patches from these slides, consisting of non-cancerous and carcinomatous classes, were then extracted. Various combinations of pre-trained VGG16 and VGG19 deep learning architectures were selected and trained. Finally, the models' accuracy was evaluated and compared with other models.

In total, 544 full slide images (WSI) from 80 breast cancer patients were collected from the pathology department at the University of Cundinamarca, Bogotá, Colombia (pre-recorded data). 845 regions were extracted from the

WSIs, 408 of which were non-cancerous and 437 cancerous. The carcinomas class includes images of malignant tumors, while the non-carcinoma class includes images of healthy tissues and benign tissue formations. These regions were captured at a resolution of 200× (50-micron resolution) and were converted to png format using the Qupath 0.1.2 software. The dimensions of these images are 1278 × 760 pixels. This dataset is considered balanced, and its statistics are shown in Table (1).

Table 1. Features of the Proposed Dataset

Staining	Color Model	Number	Image Type
H & E	RGB	437	Cancerous
H & E	RGB	408	Non-Cancerous
H & E	RGB	845	Total

For both individual and ensemble models, 80% of the images were selected for training and 20% for testing, with an equal percentage of cancerous and non-cancerous images in both sets. Thus, 675 images were used for training, and 170 remaining images were kept for model testing. In this study, five-fold cross-validation was used on the training images, meaning that 540 images were used for training and 135 images for validation, maintaining the same percentage of cancerous and non-cancerous images in both training and validation sets. These statistics for training, validation, and testing the models are shown in Table (2).

Table 2. Criteria for Selecting Training, Validation, and Test Images

Percentage	Number of Images	-
64%	540	Training
16%	135	Validation
20%	170	Testing
100%	845	Total

3. MODELING AND ANALYSIS

The overall performance of our proposed model relies on the elements of the confusion matrix, which is also referred to as the error matrix or probability table. This evaluation matrix includes four terms: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). In the present case, TP refers to images that are correctly classified as carcinoma, while FP represents non-cancerous images that are incorrectly classified as carcinoma. FN refers to images from the carcinoma class that are incorrectly classified as non-cancerous, and TN refers to non-cancerous images that are correctly classified.

The classification performance of the proposed model on the test set was evaluated using four performance metrics based on the confusion matrix: accuracy, sensitivity (recall), overall precision, and F1 score, using the scikit-learn Python modules. These performance metrics can be calculated as follows:

- **Accuracy:** The accuracy value determines how well the model performs and shows the proportion of cancerous images that are correctly classified from the predicted class images.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

Sensitivity: Also known as "Recall," sensitivity measures the completeness of a model. It represents the ratio of images correctly classified as carcinoma to the total number of cancer images.

$$Sensitivity = \frac{TP}{TP+FN} \tag{2}$$

Accuracy: Accuracy evaluates a model and represents the ratio of the number of correctly classified images to the total number of test images.

$$\text{Overall Accuracy} = \frac{TP+FN}{TP+TN+FN+FP} \tag{3}$$

F1 Score: The harmonic mean of precision and recall, which is commonly used to optimize a model towards either precision or recall.

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

3.1. Neural Networks and Hyperparameter Optimization

Neural networks have a powerful capability for learning complex relationships between their inputs and outputs. However, some of these connections may result from sampling noise, making them prominent during the training process, but they might not exist in the actual test dataset. This issue can lead to overfitting problems, which may consequently reduce the predictive performance of a deep learning model. Therefore, hyperparameter tuning is essential to generalize the performance of the proposed model. The method used for selecting optimal hyperparameters is as follows: Initially, binary cross-entropy is chosen as the loss function for the binary classification problem. Then, the Adam optimizer (adaptive moment estimation) is applied during the training process to optimize over 200 epochs. At this stage, three different learning rates (0.001, 0.0001, and 0.00001) and three different batch sizes (16, 32, and 64) were tested.

Throughout the model training, our primary goal was to minimize the generalization gap between training loss and validation loss. It was found that a batch size of 64 worked well with a learning rate of 0.0001. Additionally, to prevent overfitting during training, a dropout rate of 0.3 was used. In the next step, the weights of the five best models, based on the minimum validation loss reduction using five-fold cross-validation, were saved. Finally, these weights were used to predict the class in the test dataset. Notably, in this study, convolutional filters, composite filters, strides, and padding with default values mentioned in the original VGG16 and VGG19 architectures were used. All optimal hyperparameter values used in this study are provided in Table (3).

Table 3. Hyperparameters Used in Individual and Ensemble Models

Hyperparameters	VGG16	VGG19
Training Approach	5-fold cross-validation	5-fold cross-validation
Optimizer	Adam	Adam
Loss Function	Binary Cross-Entropy	Binary Cross-Entropy
Learning Rate	0.0001	0.0001
Batch Size	64	64
Convolution Specifications	3×3 with stride 1	3×3 with stride 1
Padding	Same	Same
Pooling	2×2	2×2
Epoch Size	250	250
Dropout Rate	0.3	0.3
Architecture	Full Training and Fine-tuning	Full Training and Fine-tuning

3.2. Performance Evaluation of the Proposed Deep Learning Models

In this section, the performance of the proposed deep learning models is evaluated based on the average predicted probabilities. Initially, the performance metrics of the individual models are highlighted, and then the competitiveness of the proposed models is discussed, especially in terms of cancer classification.

3.3. VGG16 Architecture Results

The performance metrics of the full training VGG16 architecture on the present dataset are shown in Table (4). It can be observed that these metrics vary across different layers, even though the same test samples are used. Notably, the average recall (sensitivity) value for the malignant class is 98.18%. Also, the highest accuracy and F1 score are observed in the first iteration, compared to the lowest values in the second iteration. The overall accuracy of the VGG16 model after full training is 95.88%, with an average F1 score of 95.86%. The accuracy curves of this model are shown in Figure (1), while the loss curves are displayed in Figure (2).

Table 4. Performance Metrics of VGG16 Architecture on the Dataset

Architecture	Fold	Class	Accuracy	Recall	F1 Score	Overall Accuracy	Average F1 Score
Full Training VGG16	Fold 1	Non-cancer	98.70%	92.68%	95.60%	95.88%	95.87%
		Cancer	93.55%	98.86%	96.13%		
	Fold 2	Non-cancer	95.89%	85.37%	90.32%	91.18%	91.11%
		Cancer	87.63%	96.59%	91.89%		
	Fold 3	Non-cancer	98.72%	93.90%	96.25%	96.47%	96.46%
		Cancer	94.57%	98.86%	96.67%		
	Fold 4	Non-cancer	98.75%	96.34%	97.53%	97.65%	97.64%
		Cancer	96.67%	98.86%	97.75%		
	Fold 5	Non-cancer	97.59%	98.78%	98.18%	98.24%	98.23%
		Cancer	98.85%	97.73%	98.29%		
Average		Non-cancer	97.93%	93.41%	95.58%	95.88%	95.86%
		Cancer	94.25%	97.18%	96.15%		

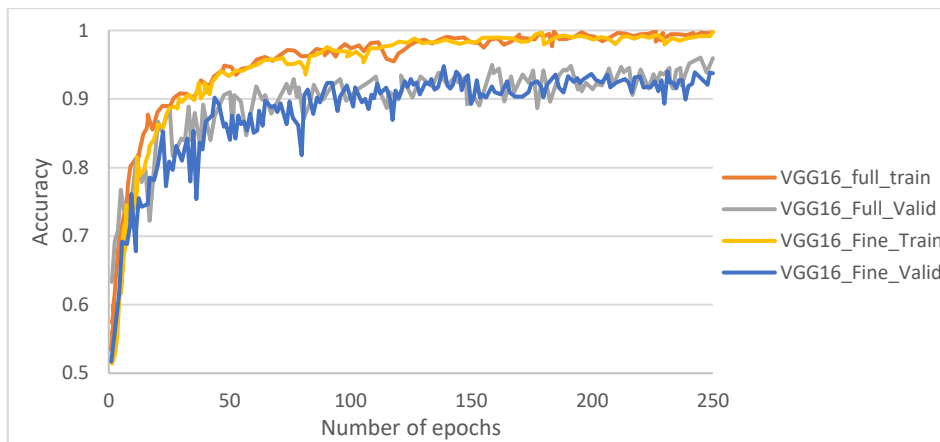


Fig. 1. Training and Validation Accuracy Curves for Fully Trained and Fine-tuned VGG16 Models

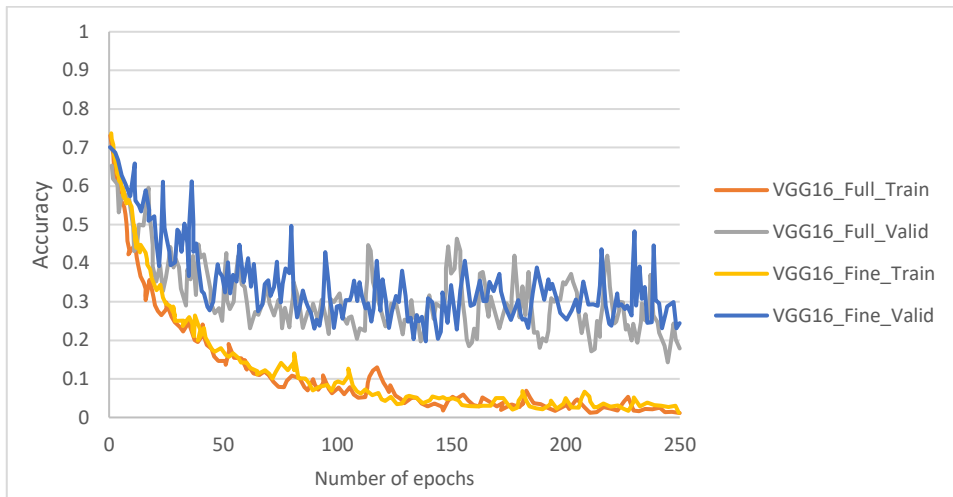


Fig. 2. Training and Validation Loss Curves for Fully Trained and Fine-tuned VGG16 Models

Similar to the fully trained VGG16 architecture, the performance metrics for the fine-tuned VGG16 framework are provided in Table (5). Again, the same test set was used across all folds. Here, the average recall value for the malignant class is observed as 97.96%. Additionally, the highest accuracy and F1 scores are found in fold 5, while the lowest values appear in fold 1. Overall, the fine-tuned VGG16 models delivered an average accuracy of 96.84% and an average F1 score of 96.83%. The accuracy curves of this model are also shown in Figure (1), while its loss curves are presented in Figure 4-2. Finally, the training and prediction times for the fully trained and fine-tuned VGG16 models are provided in Table (5).

Table 5. Training and Prediction Times of Fully Trained and Fine-tuned Models

Model	Single Training Time	5-Fold Training Time	Prediction Time
VGG16 Full Training	16 min 20 sec	67 min	25 sec
VGG16 Fine-tuning	16 min	61 min	27 sec
VGG19 Full Training	18 min 50 sec	96 min	28 sec
VGG19 Fine-tuning	18 min 10 sec	92 min	31 sec

3.4. VGG19 Architecture Results

The performance metrics of the fully trained VGG19 architecture on the present dataset are presented in Table (6). In this case, the average recall (sensitivity) for cancer images is 97.95%, which is 0.77% higher than the fully trained VGG16 model. Furthermore, the maximum accuracy and F1 scores were observed in fold 3, while the minimum values were seen in fold 4. Finally, the overall accuracy of the fully trained VGG19 model is 94.59%, with an average F1 score of 94.56%. The accuracy curves of this model are shown in Figure (3), while its loss curves are displayed in Figure (4).

Table 6. Performance metrics of the VGG19 architecture on the dataset.

Architecture	Fold	Performance Metric	Average
		Accuracy	Recall
Fully Trained VGG19	Fold 1	Non-cancer: 98.63%	87.80%
		Cancer: 89.69%	98.86%
	Fold 2	Non-cancer: 97.44%	92.68%
		Cancer: 93.48%	97.73%
	Fold 3	Non-cancer: 98.68%	91.46%
		Cancer: 92.55%	98.86%
	Fold 4	Non-cancer: 96.10%	90.24%
		Cancer: 91.40%	96.59%
	Fold 5	Non-cancer: 97.44%	92.68%
		Cancer: 93.48%	97.73%
Mean	Non-cancer: 97.66%	90.98%	
	Cancer: 92.12%	97.95%	
Fine-tuned VGG19	Fold 1	Non-cancer: 98.59%	85.37%
		Cancer: 87.88%	98.86%
	Fold 2	Non-cancer: 96.25%	93.90%
		Cancer: 94.44%	96.59%
	Fold 3	Non-cancer: 97.50%	95.12%
		Cancer: 95.56%	97.73%
	Fold 4	Non-cancer: 97.56%	95.24%
		Cancer: 95.56%	97.73%
	Fold 5	Non-cancer: 96.25%	93.90%
		Cancer: 94.44%	96.59%
Mean	Non-cancer: 97.23%	92.71%	
	Cancer: 93.58%	97.50%	

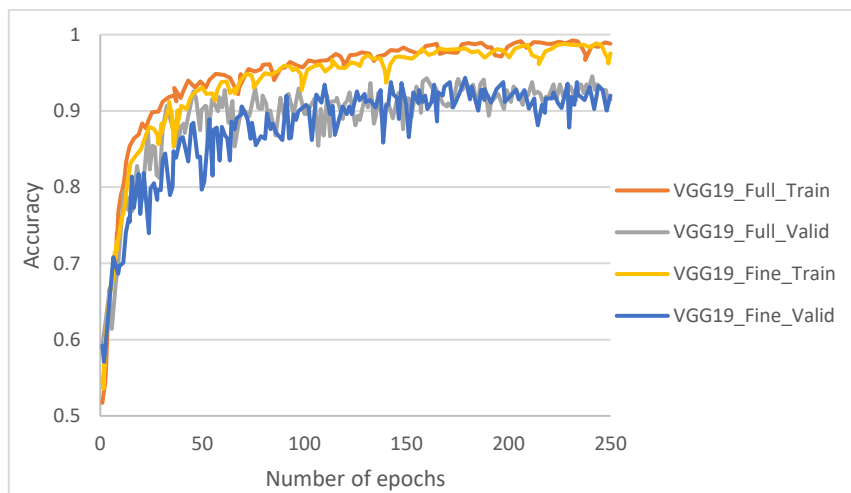


Fig. 3. Training and Validation Accuracy Curves for Fully Trained and Fine-tuned VGG19 Models

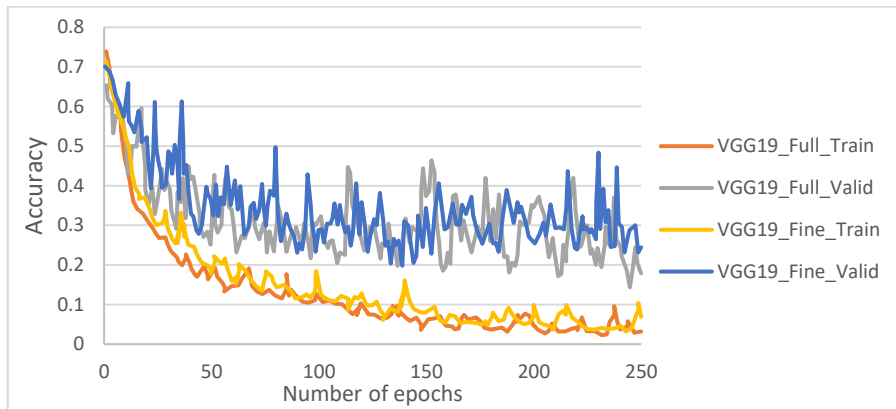


Fig. 4. presents the training and validation loss curves for the fully trained and fine-tuned VGG19 models.

Similar to the fully trained VGG19 model, the performance metrics of the fine-tuned VGG19 architecture are summarized in Table 6. The average recall for cancerous cases reaches 97.50%, which is 0.54% higher than that of the fine-tuned VGG16 model. Notably, the highest precision and F1-score are achieved in folds 3 and 4, while the lowest values occur in fold 1. The average precision and F1-score across folds are 95.18% and 95.16%, respectively. Additionally, as with the VGG16 models, the training and prediction times for the fully trained and fine-tuned VGG19 frameworks are also reported in Table 6.

3.5. Performance of Combined VGG16 and VGG19 Models

The performance metrics for the combination of VGG16 and VGG19 architectures are reported in Table 7. In this approach, the fully trained and fine-tuned VGG16 and VGG19 frameworks are combined by averaging the output probabilities across all folds. Notably, the recall for the cancer class reaches the same value of 97.73% under both the fully trained and fine-tuned strategies. However, the fine-tuned approach delivers superior overall precision and F1-score compared to the fully trained approach, as presented in Table 7.

Table 7. Performance metrics of the combined VGG16 and VGG19 architectures.

Combined Approach	Class	Accuracy	Recall	F1-score	Precision	F1
Fully Trained VGG16 + VGG19	Non-cancer	97.47%	93.90%	95.65%	95.88%	95.87%
	Cancer	94.51%	97.73%	96.09%		
Fine-tuned VGG16 + VGG19	Non-cancer	97.50%	95.12%	96.30%	96.47%	96.46%
	Cancer	95.56%	97.73%	96.63%		

4. CONCLUSION

In this study, an ensemble deep learning approach was proposed for classifying breast cancer histopathology images using a curated dataset. The primary objective was to achieve effective cancer image classification. The results demonstrate that averaging the predicted probabilities of two distinct models leads to improved performance. To this end, an ensemble of fine-tuned VGG16 and VGG19 models was employed, yielding a relatively stronger predictive framework. The proposed ensemble approach offers competitive performance in classifying complex breast cancer histopathology images. According to the findings, the recall for the cancer class consistently reached 97.73% under both the fully trained and fine-tuned strategies.

Declaration

We acknowledge that we used ChatGPT to enhance the academic writing of our manuscript while ensuring the originality and integrity of our work.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

- [1] Emami, Z., & Shamsi, M. (2017). A review of breast cancer tumor detection methods using machine learning techniques. 2018 World Congress on Intelligent Technologies, Tehran, Conference Secretariat.
- [2] Tavakoli, B., & Nourafza, N. (2016). A review of several breast cancer prediction methods. International Conference on Engineering and Computer Science, Najafabad, Islamic Azad University, Najafabad Branch.
- [3] Lahoura, V., Singh, H., Aggarwal, A., Sharma, B., Mohammed, M. A., Damaševičius, R., ... & Cengiz, K. (2021). Cloud computing-based framework for breast cancer diagnosis using extreme learning machine. *Diagnostics*, 11(2), 241. <https://doi.org/10.3390/diagnostics11020241>
- [4] Lei, Y., Chen, X., Min, M., & Xie, Y. (2020). A semi-supervised Laplacian extreme learning machine and feature fusion with CNN for industrial superheat identification. *Neurocomputing*, 381, 186-195. <https://doi.org/10.1016/j.neucom.2019.11.012>
- [5] Li, Y., Wu, J., & Wu, Q. (2019). Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning. *IEEE Access*, 7, 21400-21408. <https://doi.org/10.1109/ACCESS.2019.2898044>
- [6] Rozehbahani, M., Ghasemzadeh, M.R., & Maghouli, K. (2015). Prediction of breast cancer types using data mining techniques and their comparison. National Conference on Technology, Energy, and Data with an Electrical and Computer Engineering Approach, Kermanshah, Iranian Association of Electrical and Electronics Engineers - Western Branch.
- [7] Shamrat, F. J. M., Raihan, M. A., Rahman, A. S., Mahmud, I., & Akter, R. (2020). An analysis on breast disease prediction using machine learning approaches. *International Journal of Scientific & Technology Research*, 9(02), 2450-2455.
- [8] Gayathri, B. M., & Sumathi, C. P. (2016, December). Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer. In 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICIC) (pp. 1-5). IEEE. <https://doi.org/10.1109/ICIC.2016.7919576>
- [9] Yue, W., Wang, Z., Chen, H., Payne, A., & Liu, X. (2018). Machine learning with applications in breast cancer

diagnosis and prognosis. *Designs*, 2(2), 13. <https://doi.org/10.3390/designs2020013>

- [10] Shukla, R., Yadav, V., Pal, P. R., & Pathak, P. (2019). Machine learning techniques for detecting and predicting breast cancer. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(7), 2658-2662.
- [11] Lou, S. J., Hou, M. F., Chang, H. T., Chiu, C. C., Lee, H. H., Yeh, S. C. J., & Shi, H. Y. (2020). Machine learning algorithms to predict recurrence within 10 years after breast cancer surgery: A prospective cohort study. *Cancers*, 12(12), 3817. <https://doi.org/10.3390/cancers12123817>
- [12] Kumar, M., Khatri, S. K., & Mohammadian, M. (2020). Breast cancer identification and prognosis with machine learning techniques-An elucidative review. *Journal of Interdisciplinary Mathematics*, 23(2), 503-521. <https://doi.org/10.1080/09720502.2020.1731963>
- [13] Sarhadi, R., Akbari, R., Tahmasebi, S., & Zangouri, V. (2018). Breast cancer detection using SVM. Third International Conference on Electrical Engineering, Computer Science, and Information Technology, Hamedan, Permanent Secretariat of the Conference.
- [14] Morshedi, F., Ali Esfahani, S., & Ghasemi, L. (2015). Prediction and recurrence of breast cancer after treatment using data mining. Second National Conference on Computer Engineering and Information Technology Management, Tehran, Bouali Research Group.
- [15] Mahtabi, A., & Parvin Nia, E. (2018). Breast cancer prediction using a hybrid algorithm of imperialist competitive algorithm and random forest. Fourth National Conference on Soft Computing in Electrical and Computer Engineering, Ardabil, Islamic Azad University, Ardabil Branch.
- [16] Antropova, N., Huynh, B. Q., & Giger, M. L. (2017). A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. *Medical physics*, 44(10), 5162-5171. <https://doi.org/10.1002/mp.12453>
- [17] Ghosh, P., Azam, S., Hasib, K. M., Karim, A., Jonkman, M., & Anwar, A. (2021, July). A performance-based study on deep learning algorithms in the effective prediction of breast cancer. In 2021 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE. <https://doi.org/10.1109/IJCNN52387.2021.9534293>
- [18] Mahvash, P., & Ghaemi, R. (2016). A review of breast cancer detection methods. First National Conference on Computer Engineering, Computer Science, and Information Technology, Qom, University of Applied Science and Technology, Qom Governorate.
- [19] Zarrabi Baboldasht, Sh., & Behzadfar, N. (2021). Breast cancer detection from mammographic microcalcifications using fuzzy clustering and neural networks. *Advanced Signal Processing Semiannual Journal*, Vol. 5, No. 1.
- [20] Ranjouri, R., Haji Alilou, H., & Ranjouri, Y. (2022). Automatic image classification for breast cancer detection in women using artificial neural networks. Second International Conference on Mechanics, Electrical Engineering, Aerospace Engineering, and Engineering Sciences.
- [21] Khorang, H. (2022). Breast cancer detection using a fuzzy neural system optimized with the ICA algorithm. Fourteenth National Conference on Electrical, Computer, and Mechanical Engineering, Shirvan.
- [22] Nasiri, H., Gharacian, H., & Gharacian, K. (2021). A novel approach based on deep learning and the XGBoost algorithm for benign and malignant breast cancer detection. Eighth National Congress on New Advances in Electrical and Computer Engineering of Iran, Tehran.

- [23] Agarwal, R., & Sharma, H. (2021). A New Enhanced Recurrent Extreme Learning Machine Based on Feature Fusion with CNN Deep Features for Breast Cancer Detection. In *Advances in Computer, Communication and Computational Sciences* (pp. 461-471). Springer, Singapore. https://doi.org/10.1007/978-981-15-4409-5_42
- [24] Mahmood, T., Arsalan, M., Owais, M., Lee, M. B., & Park, K. R. (2020). Artificial intelligence-based mitosis detection in breast cancer histopathology images using faster R-CNN and deep CNNs. *Journal of clinical medicine*, 9(3), 749. <https://doi.org/10.3390/jcm9030749>
- [25] Seryasat, O. R., & Haddadnia, J. (2018). Evaluation of a new ensemble learning framework for mass classification in mammograms. *Clinical breast cancer*, 18(3), e407-e420. <https://doi.org/10.1016/j.clbc.2017.05.009>
- [26] Zadeh, H. G., Haddadnia, J., Seryasat, O. R., & Isfahani, S. M. M. (2016). Segmenting breast cancerous regions in thermal images using fuzzy active contours. *EXCLI journal*, 15, 532.
- [27] Rahmani-Seryasat, O., Haddadnia, J., & Ghayoumi-Zadeh, H. (2015). A new method to classify breast cancer tumors and their fractionation. *Ciência e Natura*, 37(4), 51-57. <https://doi.org/10.5902/2179460X19428>
- [28] Touahri, R., Azizi, N., Hammami, N. E., Aldwairi, M., & Benaida, F. (2019, April). Automated breast tumor diagnosis using local binary patterns (LBP) based on deep learning classification. In *2019 International Conference on Computer and Information Sciences (ICCIS)* (pp. 1-5). IEEE. <https://doi.org/10.1109/ICCISci.2019.8716428>
- [29] Wang, P., Song, Q., Li, Y., Lv, S., Wang, J., Li, L., & Zhang, H. (2020). Cross-task extreme learning machine for breast cancer image classification with deep convolutional features. *Biomedical Signal Processing and Control*, 57, 101789. <https://doi.org/10.1016/j.bspc.2019.101789>
- [30] Tumuluru, P., Lakshmi, C. P., Sahaja, T., & Prazna, R. (2019, December). A review of Machine Learning techniques for breast cancer diagnosis in medical applications. In *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 618-623). IEEE. <https://doi.org/10.1109/I-SMAC47947.2019.9032427>
- [31] Nallamala, S. H., Mishra, P., & Koneru, S. V. (2019). Breast cancer detection using machine learning way. *Int J Recent Technol Eng*, 8(2-3), 1402-1405. <https://doi.org/10.35940/ijrte.B1260.0782S319>