Applied Data Science and Analysis
Vol. **(2023)**, 2023, **pp**. 78–87
DOI: https://doi.org/10.58496/ADSA/2023/006; ISSN: xxx-xxxx
https://mesopotamian.press/journals/index.php/ADSA

Research Article

Automated Grading System for Breast Cancer Histopathological Images Using Histogram of Oriented Gradients (HOG) Algorithm

Mohammed Saher[1, *], Muneera Alsaedi [1,2], Ahmed Al Ibraheemi[3, 4],

[1]*Biomedical Informatics College, University of Information Technology and Communication, Baghdad, Iraq.*
[2]*Concordia University, Montreal, Canada.*
[3]*Al Ameed University/College of Medicine, Karbala, Iraq.*
[4]*Al Kafeel Super Specialist Hospital, Karbala, Iraq.*

ARTICLEINFO

Article History

Received 05 Jun 2023
Accepted 11 Aug 2023
Published 29 Aug 2023

Keywords

Breast cancer grading

Automated malignancy grading

Pattern analysis

Histopathological image

ABSTRACT

Breast cancer is the most common type of cancer in the world, affecting both men and women. In 2023, the American Cancer Society's reported that there will be approximately 297,800 new cases of invasive breast cancer in women and 2,850 in men, along with 55,750 cases of ductal carcinoma in situ (DCIS) in women. Further, an estimated 43,750 deaths are expected from breast cancer, of which approximately 43,180 are among women and 570 are among men. In this paper, we propose an automated grading system for breast cancer based on tumor's histopathological images using a combination of the Histogram of Oriented Gradients (HOG) for feature extraction and machine learning algorithms. The proposed system has four main phases: image preprocessing and segmentation, feature extraction, classification, and integration with a website. Grayscale conversion, enhancement, noise and artifact removal methods are used during the image preprocessing stage. Then the image is segment during the segmentation phase to extract regions of interest. And then, features are extracted from the obtained region of interest using the Histogram of Oriented Gradients (HOG) algorithm. The next, the images are classified into three distinct breast cancer grades based on the extracted features using machine learning algorithms. Moreover, the effectiveness of the proposed system was evaluated and reported using vary evaluation methods and the results showed a remarkable accuracy of up to 97% by the SVM classifier. Finally, the machine learning model is integrated into a website to improve the detection and diagnosis of breast cancer disease and facilitate the access and use of patient data. This will make the work easier for physicians to enhance breast cancer detection and treatment.

1. INTRODUCTION

Breast cancer is a leading cause of death in women and is of concern to the global medical community. Nearly 2.3 million women were diagnosed with breast cancer in 2020 alone, resulting in 685,000 deaths [1][2]. One of the main factors for successful and effective treatment and bettering patient outcomes is early detection and accurate diagnosis [3]. Breast cancer detection and grading rely heavily on histopathology images acquired via medical imaging techniques. Pathologists can investigate the cellular and tissue structures of breast cancer tumors using these images, which can provide valuable details regarding the severity and progress of the disease [4].

The analysis of histopathology images of cancerous tissue is utilized to grade breast cancer. Using grading systems such as the Nottingham Histologic Grade (NHG), also referred to as the Elston-Ellis grading system, pathologists assess tissue samples obtained through biopsy or surgical resection. Tubule formation, nuclear pleomorphism, and mitotic count are the three main features of cancer cells that the NHG system takes into consideration [4]. Tubule formation assesses the extent to which cancer cells form tubular structures resembling normal breast tissue. Nuclear pleomorphism evaluates the variability in size, shape, and staining intensity of cancer cell nuclei. Mitotic count indicates the number of cells undergoing

Corresponding author. Email: Mohammed.sahir.IMS.2019@uoitc.edu.iq

cell division, reflecting the rate of cancer cell growth [5]. Based on the scoring system defined for tubule formation, nuclear pleomorphism, and mitotic count, each feature is assigned a score ranging from 1 to 3, as presented in Table 1. The scores are then summed to calculate a total score ranging from 3 to 9, as depicted in Table 2. A total score of 3-5 corresponds to grade 1 (low-grade), 6-7 to grade 2 (intermediate-grade), and 8-9 to grade 3 (high-grade) [5]. In summary, breast cancer grading involves analyzing histopathological images and applying the NHG system, as outlined in detail in Tables 1 and 2 [6]. The growing interest in computer-aided analysis of histopathological or cytological images has emerged due to the increasing demand for improved accuracy in determining the malignancy grade of tumors [7]. Accurate grading significantly impacts the selection of appropriate treatment courses for cancer patients. Computer-aided diagnosis provides automated and unbiased assessments, offering support to pathologists who may face challenges such as inexperience, heavy workloads, or fatigue, and helping to prevent grading errors. This technique provides an immediate second expert opinion, particularly in situations in which a second specialist is required to conduct an additional manual examination [8] In this research study, we propose an automated system based on machine learning algorithms and web technologies for breast cancer grade classification. This technology aims to help pathologists determine the grade of breast cancer accurately, quickly, and with reliable results.

TABLE I. THE MEDICAL SCHEMA OF GRADING BREAST CANCER HISTOLOGICALLY [9]

Score	Glandular / Tubular Differentiation	Nuclear Pleomorphism	Mitotic Count
1	>75% of tumor forms glands	Uniform cells with small nuclei similar in size to normal breast epithelial cells	< 7 mitoses per 10 high power fields
2	10% to 75% of tumor forms glands	Cells larger than normal with open vesicular nuclei, visible nucleoli, and moderate variability in size and shape	8-15 mitoses per 10 high power fields
3	<10% of tumor forms glands	Cells with vesicular nuclei, prominent nucleoli, marked variation in size and shape	> 16 mitoses per 10 high power fields

TABLE II. BREAST CANCER GRADING TYPES [9]

Scores	Grade
Scores of 3, 4, or 5	Grade I
Scores of 6 or 7	Grade II
Scores of 8 or 9	Grade III

The grade of breast cancer is an essential factor in determining the appropriate treatment plan for the patient, as well as predicting his probability of recovery and response to treatment. High-grade breast cancers are more aggressive and grow and spread rapidly, and as a result, these tumors have a greater chance of the disease returning after treatment. Thus, a multidisciplinary approach combining several treatment modalities is used to successfully target and eliminate cancer cells to control high-grade breast cancers. As for low-grade tumors, they are generally treated with surgical interventions, such as lumpectomy and mastectomy, in which the tumor is removed along with a portion of the surrounding healthy tissue. In contrast, patients with higher-grade breast cancer need additional treatments to completely eliminate cancer cells and prevent the development of the disease. Chemotherapy, radiation therapy, targeted therapy, and hormonal therapy are used to combat the spread of the disease and eliminate the risk of high-grade tumors. physicians use chemotherapy to kill cancer cells throughout the body, while radiotherapy uses high-energy radiation to terminate cancer cells and their spread in a specific area, while targeted therapy prevents specific molecules involved in the growth and spread of cancer cells, as for hormonal therapy for breast cancers with hormone receptor-positive, which inhibits hormones that fuel tumor growth [10]. By determining the grade of breast cancer, physicians can develop treatment plans that suit each patient's condition separately, which helps improve treatment results. Early detection, accurate determination of the grade of breast cancer, and the development of a correct treatment plan are among the most important factors for the success of breast cancer management and the improvement of the patient's condition. Moreover, the grade of breast cancer greatly affects the possibility of the disease returning. Recent research has proven that high-grade breast cancers have a higher risk of cancer recurrence and the success rate of treatment is lower than low-grade tumors. Therefore, it is necessary to accurately determine the grade of breast cancer to develop a correct treatment plan for the patient.

2. METHODOLOGY

As shown in Figure 1, the system proposed in this paper includes four main components: preprocessing, feature extraction, classification, and website building. The preprocessing of the data is initiated which includes several steps to enhance the histopathological images. Firstly, the images must be converted to grayscale, so that the data processing process in the other stages is easier. Then, image enhancement techniques were used to improve image visibility and contrast, after that denoising algorithms were applied to histopathological images to remove unwanted noise and artifacts that might affect the accuracy of subsequent algorithms. The segmentation process is then used for the histopathology images in order to isolate the region of interest. The segmented region is used to extract the necessary features during the process of feature extraction. The Histogram of Oriented Gradients (HOG) algorithm is applied to extract distinct patterns and textures from histopathological images. To effectively describe the special characteristics of breast cancer tumors, this technique extracts features by analyzing the gradient distribution inside image blocks. Machine learning algorithms are then used to classify the features extracted with the Histogram of Oriented Gradients (HOG) algorithm into three different grades of breast cancer. The labeled data is used to train these algorithms to learn patterns and relationships between the extracted features, increasing classification accuracy for breast cancer grades. In the final stage of the system proposed in this research paper, a website is created that provides an easy-to-use interface, as it enables the user to upload a histopathological image or enter the URL of the image. The uploaded image is then processed and the breast cancer grade determined based on the output of machine learning algorithms. Additionally, the website provides important breast cancer information to increase user understanding and awareness. Figure 1 illustrates the overall structure and logic of the suggested web-based breast cancer grading system, highlighting the interactions between each component and their roles to provide accurate and efficient grading.

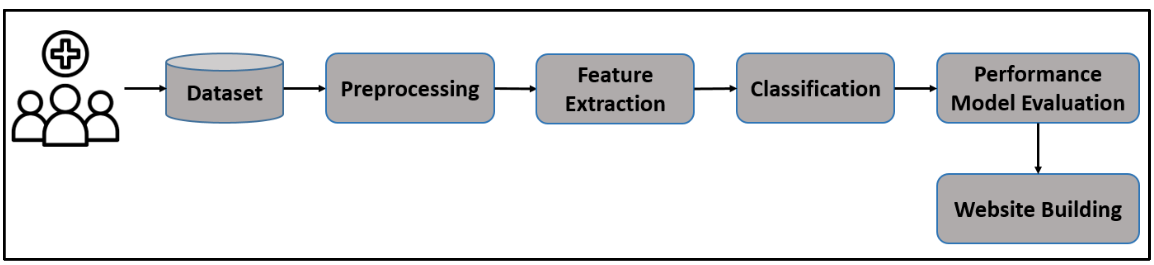

Fig. 1. The proposed Automated Grading System for breast cancer

The proposed system provides the advantage of access from any location provided that the user is connected to the Internet. This is one of the most important advantages for pathologists who work in remote areas and need to consult other specialists in some difficult cases. The significance of this paper lies in developing a system that takes advantage of medical image processing techniques and relies on machine learning algorithms and feature extraction techniques to provide accurate and effective classification results and combines this with web technologies to provide a system that reduces the workload on pathologists and enhances the accuracy of diagnosing the grade of breast cancer. The development of such a system that integrates all these advanced technologies is a significant advance in this field, providing a valuable tool for pathologists to diagnose and treat breast cancer.

2.1. DATA DESCRIPTION

This paper made use of a dataset made available by "Hamidreza Bolhasani" via Kaggle. The dataset was made up of three different classes of histological breast cancer images. Each class represent a certain grade of breast cancer. The dataset included four folders for each grade, corresponding to different image magnifications: 4x, 10x, 20x, and 40x [11]. However, this paper focused exclusively on images with a magnification of 40x, as they provided the most detailed information for accurate grading. Within the subset of images at 40x magnification, Grade 1 comprised 132 images, Grade 2 consisted of 174 images, and Grade 3 contained 137 images. These images were used for the determination of breast cancer grade, which relied on accurately classifying the malignancy level of the breast cancer images. Figure 2 displays examples of the breast cancer images belonging to these three different grades.

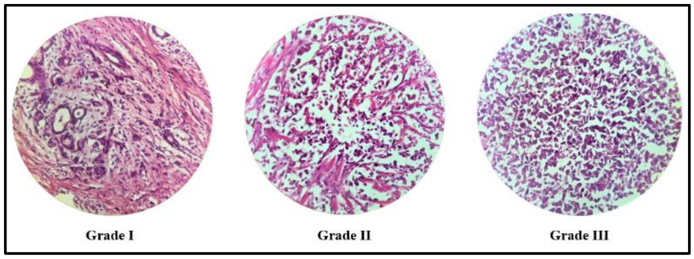

Fig. 2. Histopathological images example of breast cancer with three different classes

2.2. IMAGE PREPROCESSING

Preprocessing is a critical step in the computer-aided grading system as it transforms raw data into a format that is suitable for analysis by machine learning algorithms [12]–[14]. The quality of the preprocessing stage significantly impacts the accuracy and performance of the classification model [15][16]. Effective preprocessing enhances the input data quality and facilitates efficient segmentation, allowing the machine learning algorithm to extract more meaningful insights [17][18]. The preprocessing stage in this study encompasses several steps. Firstly, the original images are converted into grayscale images to streamline the data by reducing color information and emphasizing brightness details [19]. This conversion simplifies the subsequent processing steps and focuses on the essential features of the images [20][21]. Secondly, a normalization process is applied to the grayscale images to adjust the range of pixel values to a standardized scale. Normalization ensures that all images have a consistent pixel intensity distribution, which aids in accurate feature extraction and classification [22]. Next, image enhancement techniques are employed to improve the visual appeal and highlight crucial details that may have been obscured due to poor lighting or uneven illumination. Subtle details are crucial for accurate diagnosis and analysis in medical and scientific fields, which is why image enhancement is of particular importance in these fields [23]. In the last steps of the preprocessing stage, a denoise process is applied to reduce any noise present in the histopathological images. The denoising process is performed within the preprocessing stage to mitigate the effects of any noise present in the histopathological images, since the process of extracting meaningful information may be affected by the noise present in the images. Through this process, the accuracy of analysis and prediction is improved with higher accuracy. Figure 3 shows the effect of each preprocessing step on the histopathological images, and provides an example from the dataset used in this paper. The preprocessing stage plays an important role in the process of feature extraction and then classification, and thus it makes an effective contribution to the accuracy and overall performance of the automated grading system.

2.2.1. IMAGE SEGMENTATION

An image can be divided into useful regions or objects using image segmentation. Figure 3 shows the results of this process in a more practical way. The goal of image segmentation is to identify and extract areas that can be further analyzed and processed [24].

2.2.2. DATA AUGMENTATION

Data augmentation is a commonly used technique to artificially expand a data set by creating new forms of existing data [25]. In this paper, data augmentation was used to generate additional images through rotation transformations. Specifically, 8 rotations were applied to each image, with a range of degrees starting from -90° and ending with 90°. Figure 4 shows the resulting augmented images generated through this rotation process.

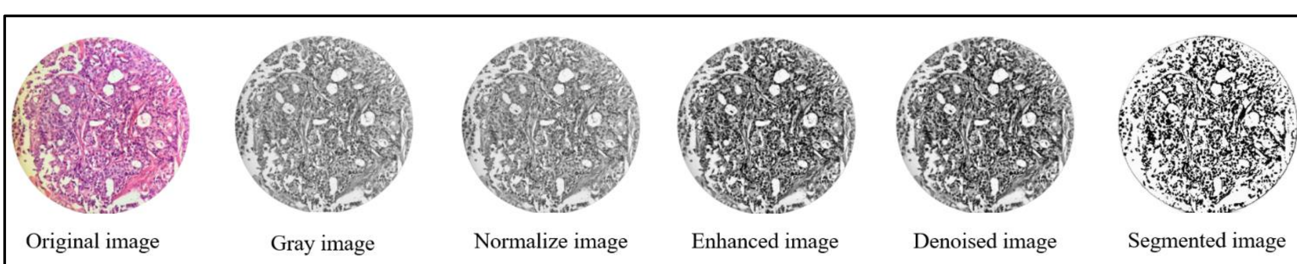

Fig. 3. An example of steps of preprocessing stage used in this paper.

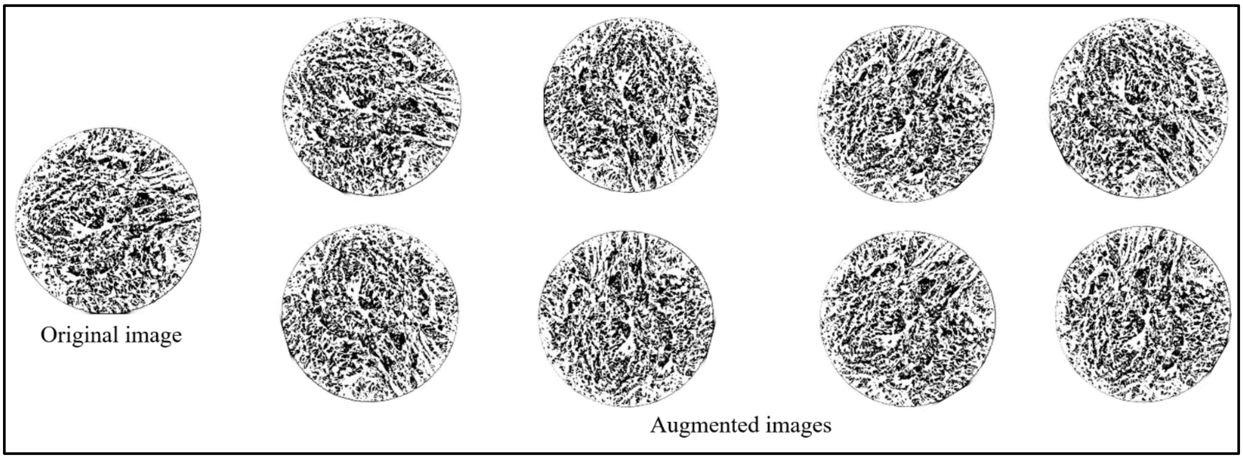

Fig. 4. An example results of data augmentation step.

2.2.3. FEATURE EXTRACTION

Feature extraction is one of the essential techniques in computer vision and image processing. This technique is vital to identifying relevant patterns and structures within the image [26]–[28]. So that the extracted features are meaningful in order to reduce the complexity of the data and thus be more amenable to analysis and decision-making tasks. [29] In this paper, the Histogram of Oriented Gradients (HOG) algorithm was used for feature extraction HOG is a technique commonly used in computer vision that captures local shape and texture information. This technique works by dividing images into small, overlapping cells and then calculating the gradient direction for each cell. Then, a histogram of cell orientations is generated that represents the local gradient distribution of the image. This histogram helps provide a concise representation of important edge and texture information. The HOG algorithm can be expressed mathematically as follows[30]:

1. Convert the color image to a grayscale image.
2. Gradients are calculated in both the horizontal (G_x) and vertical (G_y) directions, usually using operators such as the Sobel operator.
3. The magnitude (M) and orientation (θ) of the gradient vectors are calculated.
4. The image is divided into small cells and then the gradient orientations are collected within each cell.
5. A histogram of gradient orientations is generated within each cell.
6. The histograms in each block are normalized, taking into account other neighboring cells.
7. The normalized histograms are concatenated to form the final feature vector.

This paper benefited from the HOG algorithm from the power extraction force to represent the essential features of the images that were later used in the machine learning algorithms. The HOG has provided a higher capacity for the proposed system in this paper in discovering and identifying objects. The results obtained in this study show the high performance of the HOG algorithm and its facilitation in the advanced analysis of the histopathological images of breast cancer.

2.3. MODEL DEVELOPMENT

At this stage, the classification process is carried out (predicting the class label for a given input data) using machine learning algorithms, in this study various algorithms were used to classify the grade of breast cancer with high accuracy, as the algorithms used are Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbors (KNN), and Adaptive Boosting (AB) [31]–[37]. Figure 3 shows the results obtained using these algorithms in detail. To evaluate the performance of the results of the machine learning models used in this research, several evaluation metrics were used. These metrics include Accuracy, Precision, Recall, and F-Score, Confusion Matrix. Accuracy represents the overall correctness of the model's predictions by calculating the ratio of correctly classified data instances to the total number of data instances [38]. Precision measures how well the model can accurately spot true positive cases among all the cases it predicted as positive [39][40]. Recall, which is also called sensitivity, gauges how effectively the model can correctly detect positive cases among all the real positive cases. [39][41].

The F-Score combines both Precision and Recall to give us a balanced measure. It calculates the harmonic mean of these two factors. [39][42]. Finally, the Confusion Matrix offers a table that shows how well the model is doing. It includes values for true positives, true negatives, false positives, and false negatives. [43][44]. By using these evaluation measures, we can learn a lot about where our models excel and where they fall short. This helps us make smart choices when it comes to grading breast cancer effectively. Our goal is to create a strong and precise model for breast cancer grading by trying out various classification methods and assessing how well they work with the right metrics.

2.4. APPLICATION DEVELOPMENT

To develop the website and integrate it to the machine learning model, HTML, CSS, and JavaScript were used to develop the front-end, and Flask, a framework in Python, was used to develop the back-end. These technologies were used to provide a user interface that provides information about breast cancer and can predict the grade of breast cancer through a machine learning model.

3. RESULTS AND DISCUSSION

The classification models were trained using the features extracted by the Histogram of Oriented Gradients (HOG) algorithm in conjunction with seven different classification algorithms: SVM, Logistic Regression, Random Forest, Decision Tree, Naive Bayes, KNN, and Adaptive Boosting. Table 3 presents the output results of the performance parameters (accuracy, precision, recall, and f-score) used to evaluate the performance of these classification algorithms in determining the grade of breast cancer.

TABLE III. THE SUMMARY RESULTS OF THE USED CLASSIFIERS

ML Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM	97	97	95	96
Logistic regression (LR)	94	95	92	93
Random Forest (RF)	84	90	76	81
Decision Tree (DT)	64	58	59	58
Naïve Bayes (NB)	34	50	45	32
k-nearest neighbors (KNN)	62	46	50	45
Adaptive Boosting (AB)	68	66	61	63

Furthermore, the performance of each model was thoroughly evaluated using the confusion matrix. Examples of the confusion matrix can be seen in Figures 5 and 6. In Figure 5, the confusion matrix represents the predicted classes in the rows and the actual classes in the columns. The elements of the confusion matrix can be described as follows:

1. True Positives (TP): The SVM algorithm accurately predicts the positive class for each grade. The true positives are indicated by the values on the diagonal starting from the top-left (973, 1009, and 1282), representing the SVM's accurate predictions. For example, Grade I have 973 correctly predicted instances, Grade II has 1009 instances, and Grade III has 1282 instances.
2. True Negatives (TN): The SVM algorithm correctly predicts the negative class for each grade. To calculate the true negatives, we consider the values outside the corresponding row and column for each grade. For example, for Grade I, the sum of values outside the first row (30 + 54 + 40 + 1282) is 1406. Similarly, for Grade II, the sum of values outside the second row (30 + 68 + 40 + 1282) is 1420. And for Grade III, the sum of values outside the third row (30 + 54 + 1009 + 54) is 1147.
3. False Positives (FP): The SVM algorithm incorrectly predicts a positive class when the actual class is negative. To calculate the false positives, we sum the values in each column for a specific grade, excluding the true positives for that grade. For example, for Grade I, the sum of values in the first column, excluding the true positive (30 + 54 + 68), is 152. Similarly, for Grade II, the sum of values in the second column, excluding the true positive (30 + 40 + 30), is 100. And for Grade III, the sum of values in the third column, excluding the true positive (32 + 40 + 1282), is 1354.
4. False Negatives (FN): The SVM algorithm incorrectly predicts a negative class when the actual class is positive. To calculate the false negatives, we sum the values in each row for a specific grade, excluding the true positives for that grade. For example, for Grade I, the sum of values in the first row, excluding the true positive (30 + 30), is 60.

5. Similarly, for Grade II, the sum of values in the second row, excluding the true positive (54 + 32), is 86. And for Grade III, the sum of values in the third row, excluding the true positive (68 + 40), is 108.

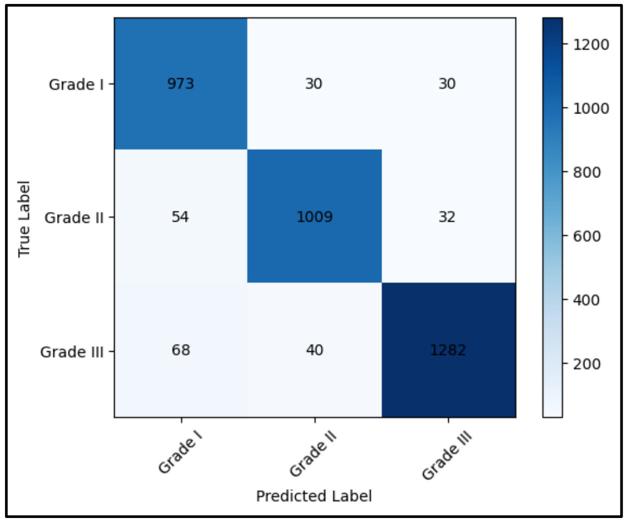

Fig. 5. Confusion matrix of SVM

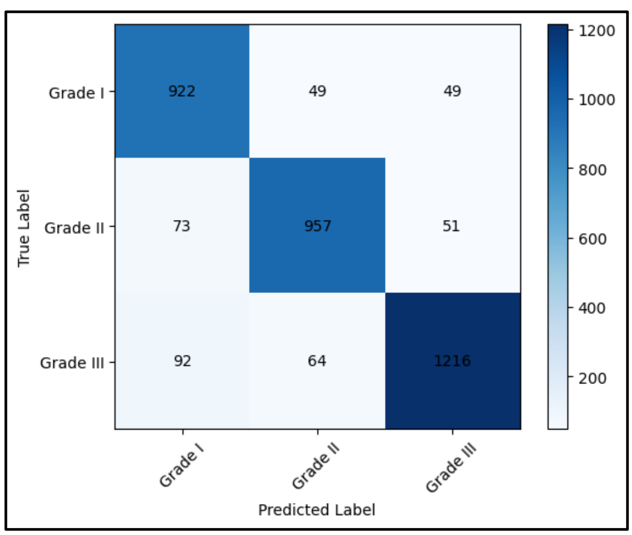

Fig. 6. Confusion matrix of Logistic regression

In addition, the user-friendly website interface is depicted in Figures 7 and 8. Users can upload images or enter the URL of an image to determine the grade of breast cancer. Upon uploading an image and clicking the "Upload" button, the grading results are promptly displayed for the user to view and interpret. The straightforward process ensures a seamless user experience and facilitates quick access to the breast cancer grading results.

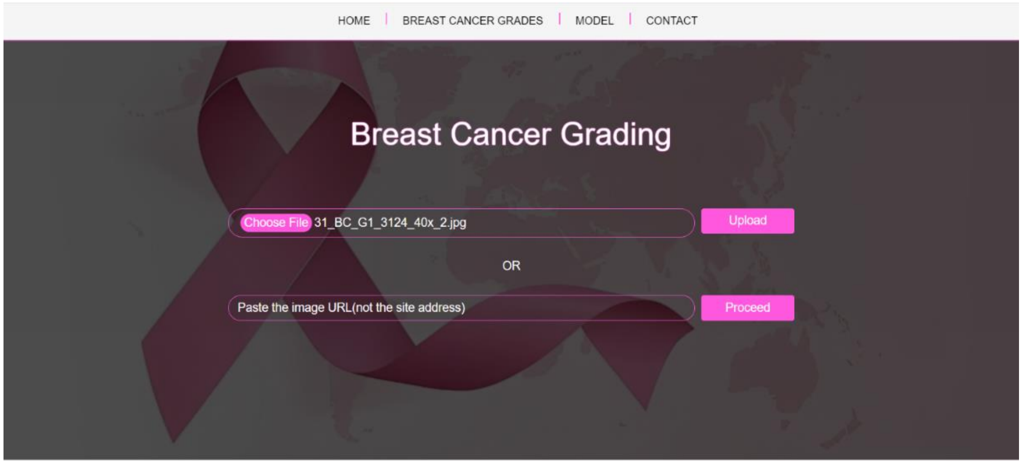

Fig. 7. Breast Cancer Grading Page

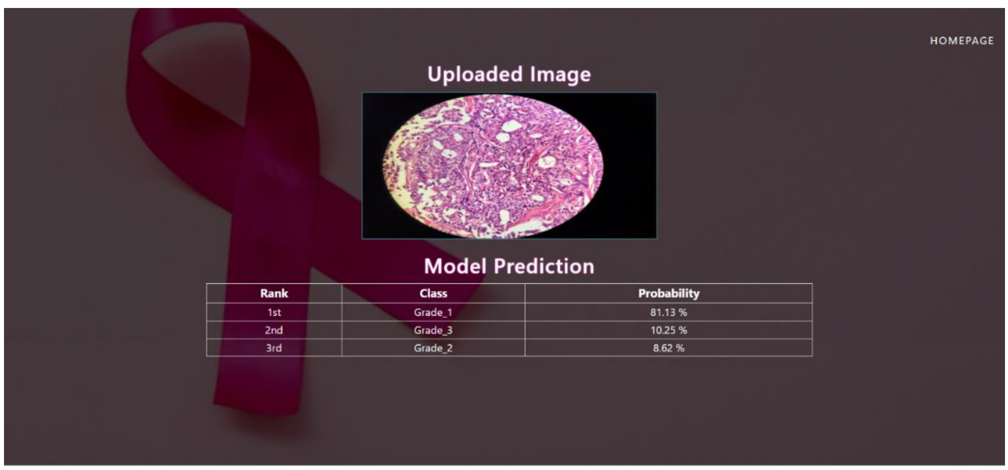

Fig. 8. Result Page

4. CONCLUSION

The proposed web-based computer-aided grading system for breast cancer, which integrates the HOG algorithm and various machine-learning techniques, has demonstrated promising results. The SVM algorithm achieved the highest accuracy of 97%, followed by LR with an accuracy of 94%. These results illustrate the ability of these algorithms to accurately classify the grade of breast cancer. RF, this algorithm shows less accuracy (84%) but is considered reasonable. However, the DT, NB, KNN, and AB algorithms showed lower accuracy scores than the other algorithms and had limitations in accurately classifying breast cancer grades. This paper provides an effective contribution to providing a reliable and effective approach for grading breast cancer. The proposed system provides an easy-to-use interface to provide ease of use by healthcare providers and reduce time consumption at work. The proposed system also provides easy online access and quick interpretation of results. By integrating machine learning and image processing techniques, the system provides accurate and effective classification, making it easier for pathologists to develop an appropriate treatment plan for breast cancer patients. Future studies should focus on improving the performance of classification models, experimenting with other techniques such as deep learning and other modern techniques, and exploring other techniques for extracting features and collecting larger and more diverse data. In addition, the system should be improved by adding other diagnostic features and expanding its capabilities to support other types of cancer grading. In general, the system proposed in this study has the ability to help healthcare professionals classify breast cancer with high accuracy, which helps in making appropriate decisions, developing appropriate treatment plans for patients, and improving patient outcomes.

Conflicts of Interest
The authors confirm no conflicts of interest.
Funding
The author's paper explicitly states that no funding was received from any institution or sponsor.
Acknowledgment
The authors convey their thankfulness to the reviewers for their diligent contributions.

References

[1] "Cancer Facts & Figures 2023," *Am. Cancer Soc.*, 2023.
[2] "Breast cancer," *World Health Organization*, 2021.
[3] S. S. Joudar *et al.*, "Artificial intelligence-based approaches for improving the diagnosis, triage, and prioritization of autism spectrum disorder: a systematic review of current trends and open issues," *Artif. Intell. Rev.*, pp. 1–65, Jun. 2023, doi: 10.1007/s10462-023-10536-x.
[4] E. A. Rakha *et al.*, "Breast cancer prognostic classification in the molecular era: the role of histological grade," *Breast cancer Res.*, vol. 12, no. 4, pp. 1–12, 2010.
[5] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up," *Histopathology*, vol. 19, no. 5, pp. 403–410, 1991.
[6] M. Santos *et al.*, "Value of the Nottingham histological grading parameters and Nottingham prognostic index in canine mammary carcinoma," *Anticancer Res.*, vol. 35, no. 7, pp. 4219–4227, 2015.
[7] A. S. Albahri *et al.*, "A Systematic Review of Trustworthy and Explainable Artificial Intelligence in Healthcare: Assessment of Quality, Bias Risk, and Data Fusion," *Inf. Fusion*, Mar. 2023, doi: 10.1016/j.inffus.2023.03.008.
[8] M. Alsaedi, T. Fevens, A. Krzyżak, and Ł. Jeleń, "Cytological malignancy grading systems for fine needle aspiration biopsies of breast cancer," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 705–709.
[9] "Staging & Grade," *Johns Hopkins University*.
[10] F. Cardoso *et al.*, "Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up," *Ann. Oncol.*, vol. 30, no. 8, pp. 1194–1220, 2019.
[11] H. BOLHASANI, "Breast Cancer Grading Data." Kaggle, 2021.
[12] D. C. Prakash, R. C. Narayanan, N. Ganesh, M. Ramachandran, S. Chinnasami, and R. Rajeshwari, "A study on image processing with data analysis," in *AIP Conference Proceedings*, 2022, vol. 2393, no. 1, p. 20225.
[13] M. E. Alqaysi, A. S. Albahri, and R. A. Hamid, "Hybrid Diagnosis Models for Autism Patients Based on Medical and Sociodemographic Features Using Machine Learning and Multicriteria Decision-Making (MCDM) Techniques: An Evaluation and Benchmarking Framework," *Comput. Math. Methods Med.*, vol. 2022, p. 9410222, 2022, doi: 10.1155/2022/9410222.
[14] Samar Hazim Hammed and A.S. Albahri, "Unlocking the Potential of Autism Detection: Integrating Traditional Feature Selection and Machine Learning Techniques," *Appl. Data Sci. Anal.*, vol. 2023, no. SE-Articles, pp. 42–58, May 2023, doi: 10.58496/ADSA/2023/003.
[15] F. Merchant and K. Castleman, *Microscope image processing*. Academic press, 2022.
[16] M. E. Alqaysi, A. S. Albahri, and R. A. Hamid, "Diagnosis-Based Hybridization of Multimedical Tests and Sociodemographic Characteristics of Autism Spectrum Disorder Using Artificial Intelligence and Machine Learning Techniques: A Systematic Review," *Int. J. Telemed. Appl.*, vol. 2022, 2022, doi: 10.1155/2022/3551528.
[17] T. N. Ataiwe, "Using Image Processing for Automatic Detection of Pavement Surface Distress," *Al-Salam J. Eng. Technol.*, vol. 2, no. 1, pp. 46–52, 2023.
[18] S. S. Joudar, A. S. Albahri, and R. A. Hamid, "Intelligent triage method for early diagnosis autism spectrum disorder (ASD) based on integrated fuzzy multi-criteria decision-making methods," *Informatics Med. Unlocked*, vol. 36, p. 101131, 2023, doi: 10.1016/j.imu.2022.101131.
[19] C. Kanan and G. W. Cottrell, "Color-to-grayscale: does the method matter in image recognition?," *PLoS One*, vol. 7, no. 1, p. e29740, 2012.
[20] R. T. Shinohara *et al.*, "Statistical normalization techniques for magnetic resonance imaging," *NeuroImage Clin.*, vol. 6, pp. 9–19, 2014.
[21] S. S. Joudar, A. S. Albahri, and R. A. Hamid, "Triage and priority-based healthcare diagnosis using artificial intelligence for autism spectrum disorder and gene contribution: A systematic review," *Comput. Biol. Med.*, vol. 146, p. 105553, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105553.
[22] H. S. S. Ahmed and M. J. Nordin, "Improving diagnostic viewing of medical images using enhancement

algorithms," *J. Comput. Sci.*, vol. 7, no. 12, p. 1831, 2011.

[23] A. E. Ilesanmi and T. O. Ilesanmi, "Methods for image denoising using convolutional neural network: a review," *Complex Intell. Syst.*, vol. 7, no. 5, pp. 2179–2198, 2021.

[24] W. Esam Noori and A. S. Albahri, "Towards Trustworthy Myopia Detection: Integration Methodology of Deep Learning Approach, XAI Visualization, and User Interface System," *Appl. Data Sci. Anal.*, vol. 2023, no. SE-Articles, pp. 1–15, Feb. 2023, doi: 10.58496/ADSA/2023/001.

[25] L. Alzubaidi *et al.*, "A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications," *J. Big Data*, vol. 10, no. 1, p. 46, Apr. 2023, doi: 10.1186/s40537-023-00727-2.

[26] Hiba Mohammed Talib, A.S. Albahri, and Thierry O. C. EDOH, "Fuzzy Decision-Making Framework for Sensitively Prioritizing Autism Patients with Moderate Emergency Level," *Appl. Data Sci. Anal.*, vol. 2023, no. SE-Articles, pp. 16–41, Mar. 2023, doi: 10.58496/ADSA/2023/002.

[27] N. Basil, M. E. Alqaysi, M. Deveci, A. S. Albahri, O. S. Albahri, and A. H. Alamoodi, "Evaluation of autonomous underwater vehicle motion trajectory optimization algorithms," *Knowledge-Based Syst.*, p. 110722, 2023, doi: https://doi.org/10.1016/j.knosys.2023.110722.

[28] M. Talal, A. H. Alamoodi, O. S. Albahri, A. S. Albahri, and D. Pamucar, "Evaluation of remote sensing techniques-based water quality monitoring for sustainable hydrological applications: an integrated FWZIC-VIKOR modelling approach," *Environ. Dev. Sustain.*, pp. 1–45, 2023, doi: 10.1007/s10668-023-03432-5.

[29] C. Tomasi, "Histograms of oriented gradients," *Comput. Vis. Sampl.*, pp. 1–6, 2012.

[30] M. Tyagi, "HOG (Histogram of Oriented Gradients): An Overview," *Towards Data Science*, 2021.

[31] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, 2021.

[32] F. Pedregosa *et al.*, "Scikit-learn: Machine Learn-ing in Python, Journal of Machine Learning Re-search, 12," 2011.

[33] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC Med. Inform. Decis. Mak.*, vol. 19, no. 1, p. 281, Dec. 2019, doi: 10.1186/s12911-019-1004-8.

[34] M. Rathi, A. Malik, D. Varshney, R. Sharma, and S. Mendiratta, "Sentiment analysis of tweets using machine learning approach," in *2018 Eleventh international conference on contemporary computing (IC3)*, 2018, pp. 1–3.

[35] D. T. Larose and C. D. Larose, "k-nearest neighbor algorithm," 2014.

[36] S. M. Sherif *et al.*, "Lexicon annotation in sentiment analysis for dialectal Arabic: Systematic review of current trends and future directions," *Inf. Process. Manag.*, vol. 60, no. 5, p. 103449, 2023, doi: 10.1016/j.ipm.2023.103449.

[37] Z. T. Al-Qaysi *et al.*, "A systematic rank of smart training environment applications with motor imagery brain-computer interface," *Multimed. Tools Appl.*, vol. 82, no. 12, pp. 17905–17927, May 2023, doi: 10.1007/s11042-022-14118-x.

[38] D. Müller, I. Soto-Rey, and F. Kramer, "Towards a guideline for evaluation metrics in medical image segmentation," *BMC Res. Notes*, vol. 15, no. 1, pp. 1–8, 2022.

[39] S. A. Hicks *et al.*, "On evaluation metrics for medical applications of artificial intelligence," *Sci. Rep.*, vol. 12, no. 1, p. 5979, 2022.

[40] M. A. Ahmed *et al.*, "Intelligent Decision-Making Framework for Evaluating and Benchmarking Hybridized Multi-Deep Transfer Learning Models: Managing COVID-19 and Beyond," *Int. J. Inf. Technol. Decis. Mak.*, Apr. 2023, doi: 10.1142/S0219622023500463.

[41] A. S. Albahri *et al.*, "A Systematic Review of Using Deep Learning Technology in the Steady-State Visually Evoked Potential-Based Brain-Computer Interface Applications: Current Trends and Future Trust Methodology," *Int. J. Telemed. Appl.*, vol. 2023, 2023.

[42] Z. T. Al-Qaysi *et al.*, "Systematic review of training environments with motor imagery brain–computer interface: Coherent taxonomy, open issues and recommendation pathway solution," *Health Technol. (Berl).*, vol. 11, no. 4, pp. 783–801, Jul. 2021, doi: 10.1007/s12553-021-00560-8.

[43] D. Valero-Carreras, J. Alcaraz, and M. Landete, "Comparing two SVM models through different metrics based on the confusion matrix," *Comput. Oper. Res.*, vol. 152, p. 106131, 2023.

[44] N. S. Baqer, H. A. Mohammed, and A. S. Albahri, "Development of a real-time monitoring and detection indoor air quality system for intensive care unit and emergency department," *Signa Vitae*, vol. 1, p. 16, 2022, doi: 10.22514/sv.2022.013.