Machine Learning-based Classification and Analysis of Breast Cancer Pathological Images

Prabha Ravi[1], Neha Sara John[2], Prithvi B. S.[3], Supriya B. S.[4], Elavaar Kuzhali S.[5], Varsha M. S.[6]

Abstract—Histopathology is the microscopic examination of diseased cells and tissues present in the body. Determining the degree of severity of illness is very essential in the diagnosis process. Assessing the variations that take place during the illness requires an in-depth understanding of the fundamental structures and functions of various tissues. This study aimed to develop a diagnostic tool to aid clinicians in the preliminary digital examination of histopathological breast tissueslides obtained after surgical biopsy. The designed study was carried out on the open-source Breakhis dataset. The breast tumor microscopic tissue images were divided into 5 categories, including benign breast tumors and 4 different types of malignant breast tumors (lobular carcinoma, ductal carcinoma, papillary carcinoma, and mucinous carcinoma). The images were prepared for classification by performing preprocessing steps such as color normalization and nuclei segmentation using the K-means algorithm. A comparative analysis was performed using four image classification algorithms: AlexNet, Convolutional Neural Networks(CNN), Support Vector Machine (SVM), and K-nearest Neighbors (KNN). The AlexNet performed at a significantlyhigher caliber with an accuracy of 98%. A GUI, or graphic user interface, was created as an assist platform for physicians to utilize the designed breast tumor tissue image analysis pipeline.

Index Terms—breast cancer, histopathology, nuclear segmentation, machine learning, CNN

1. Introduction

The breast area is composed of many different types of tissues, from extremely hard tissues to significantly fatty tissues. Breast cancer occurs when healthy cells undergo mutations causing them to divide and proliferate out of control, resulting in mass of cells referred to as a tumor. There are two types of tumors: cancerous and abnormal noncancerous benign tumors. A malignant tumor is one that has the potential to invade nearby tissues and has ability to spread to other regions of the body. The term "benign tumor" refers to a tumor that are closely enclosed and they grow slowly but does not spread to surrounding tissues. Cancerous cells have the ability to break away from the primary tumor and enter into neighboring organs or go through lymphatic vessels and/or blood vessels to other body parts[12]. This process is referred as "metastasis". The cancer cells in breast spreads to nearby lymph nodes, as well as it can also invade organs suchasthe bones, brain, lungs, liver etc. Such cancer is referred as metastatic breast cancer. The stage of breast cancer is a significant aspect and are determined by various factors. The aspects such as extent of the cancer's growth, size of the tumor and where it has spread are dictates the stage of breast cancer.

A. Histopathology

Histopathology is the study of diseased or unusual tissues present in the body. It can be very useful in disease diagnosis, level of damage of tissues and spread of the disease. It is essential to comprehend the fundamental structure and function of various tissues to assess changes caused by diseases. Tissue sections are prepared at histopathology laboratories in order to organize the histopathologic diagnosis.Histotechnology is one of the most effective techniques for diagnosing tissue abnormalities and malignant illnesses due to its

[1]Department of Medical Electronics
Ramaiah Institute of Technology
Bengaluru, Karnataka
prabharavi@msrit.edu
[2]Department of Medical Electronics
Ramaiah Institute of Technology
Bengaluru, Karnataka
[3]Department of CSE
Jyothy Institute of Technology
Bengaluru, Karnataka
bsprithvi1992@gmail.com
[4]Department of Medical Electronics
Ramaiah Institute of Technology
Bengaluru, Karnataka
[5]Department of Electronics & Instrumentation Engineering
Ramaiah Institute of Technology
Bengaluru, Karnataka
kuzhalisubbnaiah@msrit.edu
[6]Department of Medical Electronics
Ramaiah Institute of Technology
Bengaluru, Karnataka

persuasive results and evidence. The majority of the specimens submitted to the histopathology laboratory come from the operating room in the form of small bits of tissue called biopsies, as well as organs. Fresh (unfixed) or submerged in fixative fluid specimens are provided. Fixation, embedding, microtomy, staining, mounting, and microscopic inspection are the basic procedures in specimen processing.

1) Benefits of Histological Image Analysis: Computer aided diagnosis to histological scan images of has reliable and consistent patient prognoses. On the other hand, conventional feature extraction techniques can only extract a small number of low-level picture elements, and selecting valuable aspects that are significantly influenced by individual's needs prior knowledge. Automatically extracting high-level abstract characteristics from images is possible with deep learning techniques. As a result, we use supervised and unsupervised deep convolutional neural networks to assess histopathology images of breast cancer. This work provides a thorough examination of the computational stages involved in automated breast cancer detection. The following steps are followed for the computer aided diagnosis: 1) image preprocessing to identify focus areas, 2) feature extraction to measure the characteristics of these focal areas, and 3) categorizing the focused areas as malignant or non-malignant, or finding out their level of malignancy. Following the design, this stage calculates the system's accuracy. In this study, we aim to describe these computational procedures in detail, examine the obstacles they pose, and explore solutions to these problems, indicating the requirement for developing standard data sets. By comparing different features and system designs using benchmark data sets, it is possible to avoid estimating system accuracy incorrectly. This enables the identification of subsets of distinguishing features, the creation of new features, and an improvement in the accuracy of automated cancer detection[13].

B. Previous Works

Numerous classification and pre-processing techniques have been used in previous research on breast cancer[14,17]. Along with the application of conventional Neural Networks (NN), logic-based classifiers like the Support Vector Machines (SVM), Random Forest (RF) algorithm, Bayesian approaches, and unsupervised and a few semi-supervised techniques for breast image categorization. (b). The SVM-LDA and NN-LDA models outperformed the other ML classifier models under simulations dimensionality reduction techniques and three common machine learning algorithms to classify malignant and benign tumors in the WDBC dataset. (c) For cluster fusion, the two of the most popular clustering techniques with excellent performance, the K-means (with HCA) and PAM clustering algorithms were used. (d)

Other papers focused on the usage of Hough transform as a faster approach for nuclei segmentation and also gives a decent nuclei representation. (e) Several studies tested the effectiveness of machine-learning methods like Decision Trees, K-nearest Neighbors, and Support Vector Machines to identify which classifier categorizes breast cancer the best. The quadratic support vector machine has the highest performance measure with an accuracy of 98.1% and the lowest false discovery rates, according to the results of one of these studies. (h) These techniques have played a crucial (often bottleneck) part in cancer research by minimizing human intervention, gathering pertinent second opinions, and providing verifiable clinical data. This study also enables us to evaluate the challenges that still need to be surpassed in order to achieve prognosis support in digital pathology, high-content imaging with diagnostic biomarkers, and robust whole-slide image processing[16].

2. Materials And Methods

A. Data Description

The open-source breast cancer database known as Breast Cancer Histopathological Image Classification (Breakhis) was used in this study[1]. 7909 histopathology images from 82 individuals made up the dataset. The dataset included both malignant and benign images. Each sample was collected upon surgical open biopsy and was prepared for a histological study by the pathologists at the P& D Laboratory, Parana, Brazil. The images obtained after the complete preparation procedure was completed were raw, without normalization or color standardization. Using various magnifying factors, 82 patients' microscopic pictures were obtained. (40x, 100x, 200x and 400x). Table I displays a break of the dataset according to the level of magnification. For this study, 200x magnified images were used. The dataset included four histologically distinct kinds of malignant tumors, including papillary carcinoma, ductal carcinoma, and lobular and mucinous carcinoma. Adenosis, phyllodes tumor, fibroadenoma, and tubular adenoma are just a few of the four different benign tumor kinds that were present.

Table I
BREAKHIS DATASET DESCRIPTION

Magnification	Benign	Malignant	Total
40x	625	1370	1995
100x	644	1437	2081
200x	623	1390	2013
400x	588	1232	1820
Total	2480	5429	7909
# Patients	24	58	82

B. Preprocessing of the Data

1) Color Normalization: The application of different scanners and stain colorings from different manufacturers can result in color variations across histopathology images obtained from different sources. A reference image is used as the source of the mean color for the target image during the image preprocessing method known as "color normalization." This step enables the separating of the stains present in the histopathology images. Color normalization has been provento provide enhanced qualitative and quantitative results when compared with other existing histopathology image preprocessing methods. Image analysis of histopathology images with color variation can result in incorrect diagnosis. Color normalization is performed to prevent such occurrences. Past studies converted the RGB images to grayscale to avoid color variation errors. This process, however, led to the loss of information from the histopathology images as well as the addition of noise artifacts that give the cell nuclei a shadowy appearance in some of the images. Numerous color normalization techniques such as Stain Color Descriptor (SCD), histogram specification, Macenko Method, Reinhard method, etc., have been used in past studies to perform color transformation without converting the images to grayscale. In this study, the histogram of the 3-dimensional image was adjusted to matchthe histogram of the reference image[2].

2) Thresholding and Active Contouring: Image thresholding is a processing technique with the goal of creating a binarycharacterization of the target image. This is done by classifyingeach pixel into two categories on account of their variations in illumination. Since adaptive thresholding enables the estimation of various threshold value that corresponds to every pixel in the image, it was used in this research. It is highly flexible in taking into account the variations in illumination across different images. The Bradley method was utilized for adaptive thresholding[3]. The pixel's surrounding mean intensity is used as a locally adaptive threshold to binarize the histological image.
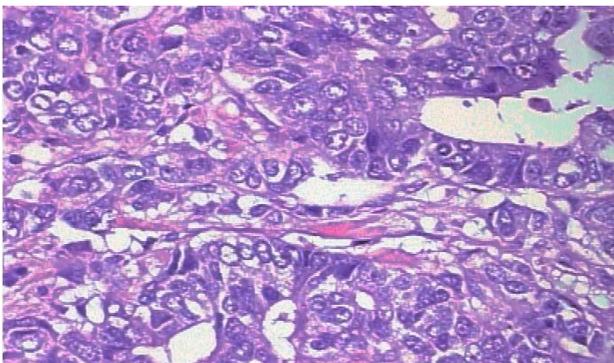
Fig. 1. Color Normalized Image.

The images were then passed through an active contouring model. Also known as the snake's algorithm, this model aids inthe detection of objects with boundaries that may not be linearor defined by a gradient. Through the Chase-Vess technique, a curve is evolved around the target object and it moves inwardsuntil it stops at the boundary of the object[4].
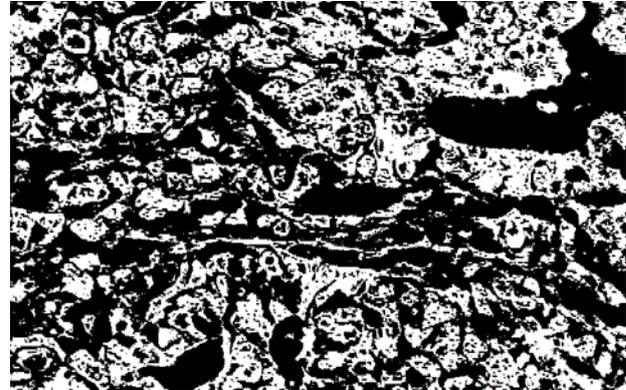
Fig. 2. Histopathological Image after Thresholding and Active Contouring

C. Nuclei Segmentation

The primary function of segmentation is to distinguish foreground objects from the background. Nuclei segmentation aids clinicians in the isolation of nuclei clusters from the cytoplasm to identify distinct nuclei patterns that may indicate different carcinoma types. Computer- aided diagnosis of breast canceris strengthened by modern techniques such as the automated segmentation of the nucleus of breast tissue cells. Due to intensity differences brought on by the uneven absorption of the hematoxylin and eosin (HE) stain as well as noise, researchers have encountered difficulties in the segmentation of densely clustered nuclei.In this way, under and overestimating the boundaries of the nuclei remains a current challenge. In this work, nuclei segmentation was carried out using the K-Means algorithm.
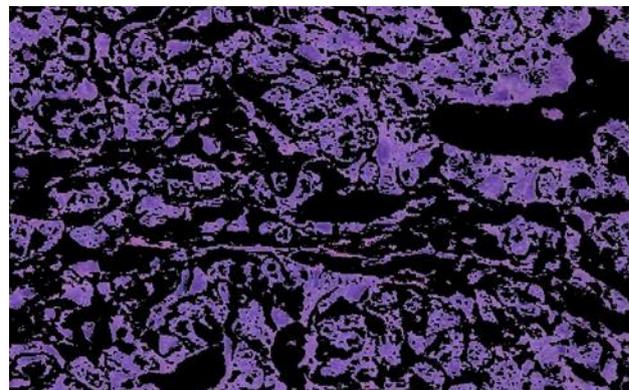
Fig. 3. Segmented Nuclei obtained on using the K-Means algorithm

1) K-Means Algorithm: In this study, nuclei segmentation of the breast cancer microscopic images was performed by applying the K-means clustering algorithm. The methodology followed is depicted in the

figure below. The raw histopathology image obtained from the pathologist is an RGB image, that is, it is a 3- dimensional image with color intensities of 3 channels: red, green and blue respectively[5]. The first step involves the application of a median filter on the RGB images to safeguard the edges of the image while eliminating noise. Following this, the histological image is changed from an RGB to a L*a*b* color space. This is done since the correlation between the red, green, and blue channels of an RGB image make it difficult for researchers to execute further image processing techniques. L*a*b* is a manifestation of the lightness (L*) as well as colors where red minus blue is represented by a* and green minus blue is represented by b*. These images are device independent, and it makes it a feasible method of analyzing images obtained from different equipment and location sites. The colors in the a* and b* spaces are sorted into 3 clusters using the Euclidean distance as a measurement metric. This sorting is performed by using the K-means clustering algorithm. This segments the L*a*b* image such that one cluster will contain the nuclei whereas the remaining two will consist of the background objects of the image[6]. The final choice of the cluster is decided manually by the clinician.

D. Feature Extraction

For the feature extraction, mainly binary features were taken into consideration. The main aim was to understand the impact of binary feature extraction on the accuracy and classification of histopathological cancer images.

1) Eccentricity: Eccentricity can be estimated as the ratio of the distance between the foci of the ellipse and its major axis length. The calculated values range between 0 to 1, where the values 0 and 1 are the exceptional cases. The eccentricity value 0 of a ellipse is a circle and an ellipse with eccentricity value as 1 is a line segment.

$$Eccentricity = \frac{Distance\ between\ foci}{Length\ of\ the\ Major\ Axis} \qquad (1)$$

2) Area: The details on the exact count of pixels within the Region of Interest will be provided by Area.

$$Area(A) = \sum_{i,j=0}^{M,N} b(i,j) \qquad (2)$$

where $b(i,j)$ represents the pixel value for the coordinates i,j within the region of interest.

3) Perimeter: The distance that exists between each adjacent pair of pixels along the edge of the area of interest is used for calculating the perimeter. A scalar value representing the distance around the region's border is returned.

$$Perimeter(P) = \sum_{i,j=0}^{M,N} E_d(i,j) \qquad (3)$$

where $E_d(i,j)$ represents the distance between adjoining boundary pixels of the connected regions of the image. i,j represents the coordinates within the region of interest.

4) Orientation: Orientation can be defined as a direction without a sign. Orientation is how the relative pieces of an object would be arranged. The value of orientation varies from -90 degrees to 90 degrees.

$$Orientation\ Angle\ (\alpha) = tan^{-1} tan\vartheta(\frac{I_x}{I_y}) \qquad (4)$$

where I_x and I_y represents the moments of inertia along the x-axis and y-axis respectively

5) Solidity: A particle's total concavity, or solidity, is measured. The image area and convex hull area, therefore, approach one another as the particle solidifies, resulting in the solidity value being equal to one.

6) Mean: Mean is the most basic of all the statistical measures. Mean is often used for analysis and in geometry. The arithmetic mean filter also referred to as the averaging filter, works with a sliding "m*n" window by determining the average of all the pixel values inside the window and substituting that value for the center pixel in the destination image.

$$f(x,y) = \frac{1}{mn} \sum_{s,t \in S_{xy}} g_r(s,t) \qquad (5)$$

where $g_r(s,t)$ is the sliding window region of the image S_{xy}. s,t represents the coordinates of the window region.

7) Standard Deviation: The most frequently used statisticians' parameter for measuring variability or variation is the standard deviation. It is apparent from the image processing that there is variation or dispersion from the mean. A low standard deviation means that the data points typically fall within a small range of the mean, whereas a high standard deviation means that the data points are dispersed across a wide range of values.

$$f(x,y) = \frac{1}{mn-1} \sum_{s,t \in S_{xy}} g_r(s,t) - \frac{1}{mn-1} \sum_{s,t \in S_{xy}} g_r(s,t)^2 \qquad (6)$$

where

8) Convex Area: It gives the number of pixels in the convex image, which is the picture that defines the convex hull, with all of the pixels in the hull filled in.

9) Number of Nuclei: Once the segmentation has been performed, the object of interest is the segmented

nucleus.It is the textural feature that is widely used for the further classification of malignant and benign tissues[7,8].

E. Classification

The classification model adopted for this study was a modification of a convolutional neural network (CNN). The neural network is a deep learning model that automatically extracts image features and uses these features—which are self-evaluated, to classify an image into the anticipated classes. CNNs are a type of feed-forward artificial neural network that has achieved great success in recent years when used to analyze computer vision issues. In this study, the model was provided with nuclei-segmented image patches to train the network. The performance of the neural networks superseded the results of the KNN[9] and SVM[10] algorithms designed as seen in Table IV. The decision to proceed with the CNN for the automated analysis of histopathological breast images was finalized upon performing the comparative analysis[11,18,19].

AlexNet is an eight-layer convolutional neural network, with five convolutional layers and three completely connected layers. Rectified Linear Units (ReLU) are used by AlexNet, which significantly reduces training time. By placing half of the model's synapses on one GPU and the other half on a different GPU, the Net enables multi-GPU training. This shortens the training period and permits the training of a larger model. It offers overlapping pooling, resulting in a 0.5% decrease in error. The dropout technique, used in AlexNet, involves "turning off" neurons with a preset probability. Since each iteration employs a different sample of the model's parameter, each neuron is compelled to have more robust features that can be applied to other random neurons.

In our work, the pre-trained AlexNet had the input image Size of 227*227*3, the final three output layers were finely tuned to fit for the classification of BreakHis database which is the classification of a total 5 classes (1 class of benign and 4 classes of malignant). The bias learning factor considered was 20. The solver considered for the training is stochastic gradient descent with momentum optimizer (sgdm) with a maximum number of epochs as four with a learning rate alpha as 0.01 with Validation frequency as 30.

F. Graphical User Interface

A graphical user interface (GUI) was developed using MAT LAB. GUI was designed to easily perform nuclei segmentation and/or classification of the required image to help them determine the category that the tumor belongs to, with just a few clicks. The designed interactive user interface for our system is shown in Figure 4. To view the test histopathological images, click the "Load Image" button. Once the button is clicked, the selected image is displayed. The RGB image is displayed on the main viewing screen. The "Perform Nuclei Segmentation" button is then pressed to initiate the nuclei segmentation using the K-Means algorithm. The 3 clusters are displayed in the column of viewing screens on the right side of the screen. The user can then select the cluster which displays the segmented nuclei. Then on performing thresholding and active contouring, the grayscale image is displayed on the main viewing screen. Then, on clicking the "Perform Classification" button, the AlexNet model is executed. The test image is thus classified, and the result of the classification is displayed in a pop-up dialog box.
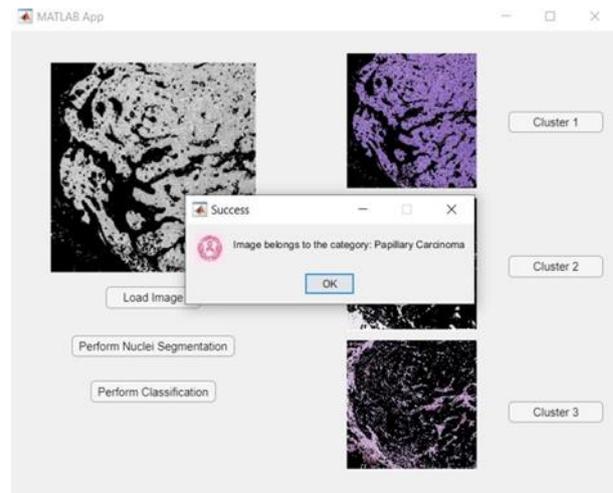
Fig. 4. GUI platform designed

3. Results

A. Preprocessing Performance Metrics

Quality performances of different color normalization methods are evaluated and compared in terms of PSNR and NMSE.The results are presented in Table II.

1) Power Signal to Noise Ratio: The ratio between a signal's maximum possible value and the amount of noise that distorts an image is expressed using the term PSNR.

$$PSNR = 20\log(MAX) - 10\log(MSE) \qquad (7)$$

2) Normalized Mean Square Error: The NMSE is a measure of the mean relative scatter and reflects the random error. Normalized error ranges between 0 to 1. The lower the value of NMSE, the better the preprocessing algorithm applied.

$$NMSE = \frac{\sum_{I=1}^{L} |Y_{desired}(I) - Y_{estimated}(I)|^2}{\sum_{I=1}^{L} |Y_{desired}(I)|^2} \qquad (8)$$

Where $Y_{desired}(I)$ is the true value or the expected value $Y_{estimated}(I)$ is the obtained value.

Table II

PERFORMANCE METRICS OF THE PREPROCESSING TECHNIQUES

Performance parameter	Benign	Malignant
Normalized MSE (NMSE)	0.01718	0.0180
Peak Signal to Noise Ratio (PSNR)	20.1841	21.8507

B. Feature Extraction Performance Metrics

Sequential feature selection using custom criteria before sending the data through the KNN and SVM models was performed. The feature ranking obtained, as a result, is displayed in Table III.

Table III

FEATURES RANKED IN ORDER OF IMPORTANCE

Features	Ranking
Eccentricity	1
Perimeter	3
Area	2
Orientation	5
Solidity	6
Number of nuclei	4
Convex Area	7
Mean	8
Standard Deviation	9

C. Classification Performance Metrics

The accuracy, sensitivity, and specificity were calculated for all four models and are displayed in Table IV.

Table IV

PERFORMANCE METRICS OF THE CLASSIFICATION ALGORITHMS

Classifier model used	Accuracy	Specificity	Sensitivity
KNN	80.21	0.7770	0.8286
SVM	64.93	0.6284	0.6714
CNN	92.2	0.92	0.833
AlexNet	**96.53**	**0.72**	**0.992**

4. Discussion

The histopathological images were preprocessed, which mainly included color normalization of the images mainly to handle the artifacts that arise due to insufficient lighting during the imaging procedure. The following preprocessing techniques followed the color normalization, that is enhancement of the images followed by global thresholding and active contouring. Mean Squared Error, Peak Signal to Noise Ratio, and Normalized Mean Squared Error were used to assess the image preprocessing effectiveness. K means segmentation followed the preprocessing. In this work, the main focus was provided on extracting the textural and color histogram features. The features were selected through custom criteria before passing through KNN and SVM classification models[15]. A comparison of four classifications that are KNN, SVM, CNN, and AlexNet was performed with the pre-trained AlexNet providing the effective classification with the highest accuracy of 96.53%.

5. Conclusion

The breast cancer histopathology images were obtained from the open-source BreakHis database and analyzed in detail, along with a thorough examination of available histopathology image processing techniques. The features extracted were used as inputs to develop a unique classification model of breast cancer tissue samples into normal and abnormal classes after preprocessing the data, which included color normalization and noise removal. This algorithm assisted in fabricating an image-processing tool to automate pathological studies under the development of a classification model to identify further sub-abnormalities within the abnormal tissue samples. Previously applied and currently automated techniques such as SVM, KNN, and CNN were compared and analyzed in terms of their performance which provided results of the CNN supervised learning model as the most successful at classifying the images with high accuracy. The GUI designed on the same model (using MATLAB) aids in facilitating the pathologist with a guided automated image processing tool to easily perform nuclei segmentation and/or classification of the required image to help them determine the category that the tumor belongs to, with just a few clicks.

References

[1] Spanhol, Fabio A., Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte. "A dataset for breast cancer histopathological image classifi- cation." *Ieee transactions on biomedical engineering* 63, no. 7 (2015): 1455-1462.

[2] Roy, Santanu, Alok kumar Jain, Shyam Lal, and Jyoti Kini. "A study about color normalization methods for histopathology images." *Micron* 114 (2018): 42-61.

[3] Bradley, Derek, and Gerhard Roth. "Adaptive thresholding using the integral image." *Journal of graphics tools* 12, no. 2 (2007): 13-21.

[4] Chan, Tony F., and Luminita A. Vese. "Active contours without edges." *IEEE Transactions on image processing* 10, no. 2 (2001): 266-277.

[5] Whitaker, Ross T. "A level-set approach to 3D reconstruction from range data." *International journal of computer vision* 29, no. 3 (1998): 203- 231.

[6] Abdolhoseini, Mahmoud, Murielle G. Kluge, Frederick R. Walker, and Sarah J. Johnson. "Segmentation of heavily clustered nuclei from histopathological images." *Scientific reports* 9, no. 1 (2019): 1-13.

[7] Sarrafzadeh, Omid, and Alireza Mehri Dehnavi. "Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing." Advanced biomedical research 4 (2015).

[8] Win, Khin Yadanar, Somsak Choomchuay, and Kazuhiko Hamamoto. "K mean clustering based automated segmentation of overlapping cell nuclei in pleural effusion cytology images." In 2017 International Conference on Advanced Technologies for Communications (ATC), pp. 265-269. IEEE, 2017.

[9] Bakthavachalam, M. D., and Dr S. Raj. "A Study Of Breast Cancer Analysis Using K-Nearest Neighbor With Different Distance Measures And Classification Rules Using Machine Learning." European Journal of Molecular & Clinical Medicine 7, no. 3 (2020): 4842-4851.

[10] Sewak, Mihir, Priyanka Vaidya, Chien-Chung Chan, and Zhong-Hui Duan. "SVM approach to breast cancer classification." In Second Inter- national Multi-Symposiums on Computer and Computational Sciences (IMSCCS 2007), pp. 32-37. IEEE, 2007.

[11] Kumar, Kundan, and Annavarapu Chandra Sekhara Rao. "Breast cancer classification of image using convolutional neural network." In 2018 4th International Conference on Recent Advances in Information Technology (RAIT), pp. 1-6. IEEE, 2018.

[12] Feng, Yixiao, Mia Spezia, Shifeng Huang, Chengfu Yuan, Zongyue Zeng, Linghuan Zhang, Xiaojuan Ji et al. "Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis." *Genes* & diseases 5, no. 2 (2018): 77-106.

[13] Nahid, Abdullah-Al, and Yinan Kong. "Involvement of machine learning for breast cancer image classification: a survey." Computational and mathematical methods in medicine 2017 (2017).

[14] Omondiagbe, David A., Shanmugam Veeramani, and Amandeep S. Sidhu. "Machine learning classification techniques for breast cancer di- agnosis." In IOP Conference Series: Materials Science and Engineering, vol. 495, no. 1, p. 012033. IOP Publishing, 2019.

[15] Agrawal, Utkarsh, Daniele Soria, Christian Wagner, Jonathan Garibaldi, Ian O. Ellis, John MS Bartlett, David Cameron, Emad A. Rakha, and Andrew R. Green. "Combining clustering and classification ensembles: A novel pipeline to identify breast cancer profiles." Artificial intelligence in medicine 97 (2019): 27-37.

[16] Jeleń, Łukasz, Thomas Fevens, and Adam Krzyżak. "Classification of breast cancer malignancy using cytological images of fine needle aspiration biopsies." International Journal of Applied Mathematics and Computer Science 18, no. 1 (2008): 75-83.

[17] Avramov, Todor K., and Dong Si. "Comparison of Feature Reduction Methods and Machine Learning Models for Breast Cancer Diagnosis." In Proceedings of the International Conference on Compute and Data Analysis, pp. 69-74. 2017.

[18] Obaid, O. Ibrahim, Mazin Abed Mohammed, M. K. A. Ghani, A. Mostafa, and Fahad Taha. "Evaluating the performance of machine learning techniques in the classification of Wisconsin Breast Cancer." International Journal of Engineering & Technology 7, no. 4.36 (2018): 160-166.

[19] Irshad, Humayun, Antoine Veillard, Ludovic Roux, and Daniel Raco- ceanu. "Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential." IEEE reviews in biomedical engineering 7 (2013): 97-114.

[20] Christopher Thomas, Stephen Wright, Juan Martinez, Alejandro Perez, Luis Rodriguez. Integrating Machine Learning and Decision Science for Intelligent Automation. Kuwait Journal of Machine Learning, 2(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/213

[21] Uppal, A. ., Naruka, M. S. ., & Tewari, G. . (2023). Image Processing based Plant Disease Detection and Classification . International Journal on Recent and Innovation Trends in Computing and Communication, 11(1s), 52–56. https://doi.org/10.17762/ijritcc.v11i1s.5993

[22] Reddy, B.R.S., Saxena, A.K., Pandey, B.K., Gupta, S., Gurpur, S., Dari, S.S., Dhabliya, D. Machine learning application for evidence image enhancement (2023) Handbook of Research on Thrust Technologies? Effect on Image Processing, pp. 25-38.