Classification Of Invasive Ductal Carcinoma And Invasive Lobular Carcinoma Of Breast Cancer Using The Artificial Neural Network Algorithm

ABSTRACT

Aims: The purpose of this study is to classify invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC) of breast cancer using the artificial neural network (ANN) algorithm.
Study design: The method employed is a cross-sectional design
Place and Duration of Study: The research was conducted in the Computer Laboratory Department of Informatics, Faculty of Mathematics and Natural Sciences, Udayana University, Bali Indonesia.
Methodology: Utilizing physical parameters from mammographic images as input variables for the artificial neural network algorithm.
Results: For Invasive Ductal Carcinoma, the accuracy is 77.5%, sensitivity (recall) is 55%, precision is 100%, F1-Score is 60.97%, specificity is 100%, FPR is 0, and TPR is 0.55. For Invasive Lobular Carcinoma, the accuracy is 77.5%, sensitivity (recall) is 100%, precision is 68.97%, F1-Score is 81.63%, specificity is 55%, FPR is 0.45, and TPR is 1.
Conclusion: The artificial neural network algorithm is capable of classifying Invasive Ductal Carcinoma and Invasive Lobular Carcinoma effectively.

Keywords:invasive ductal carcinoma; invasive lobular carcinoma; breast cancer; artificial neural networks; IDC; ILC, mammography.

1. INTRODUCTION

Breast cancer is a type of cancer that occurs when abnormal cells grow and develop in the breast tissue.It is a major public health issue due to the increasing incidence and mortality rates [1]. Breast cancer is the most common cancer among women [2]. Although more common in women, breast cancer can also affect men [3]. Symptoms may include the presence of a lump in the breast, changes in skin texture, or the shape of the breast [4]. Early detection is crucial for successful treatment, and women are advised to perform routine breast self-examinations and consult a doctor if there are concerns [5]. Several breast cancer screening tools are available for early detection, such as mammography, which is the current standard method. However, mammography is less effective for patients under the age of 40 and has limited sensitivity for detecting tumors smaller than 1 mm and for dense breast tissue [6]. Ultrasonography (US) is used for screening dense breast tissue [7]. Magnetic Resonance Imaging (MRI) has the ability to screen for small tumors that cannot be detected by mammography [8]. Positron Emission Tomography (PET) is a highly accurate method for visualizing the spread of breast cancer [9]. Computed Tomography (CT) scan is the most accurate method for observing the metastasis of breast cancer [10]. The presence of microcalcifications also plays a significant role in the early detection of breast cancer. The presence of microcalcifications in a mammogram is an indicator of breast cancer. Many early detection methods have been developed by scientists to detect microcalcifications, including: Texture Analysis [11], Neural Networks [12], Edge Detection [13], Adaptive k-means Clustering [14], Self-Similar Fractals [15], Orthogonal Polynomials Model [16], Gabor Features [17], and Vector Quantization Technique [18]. Once microcalcifications are identified, it indicates the presence of malignancy, prompting researchers to investigate whether the malignancy is classified as Invasive Ductal Carcinoma (IDC) or Invasive Lobular Carcinoma (ILC). Several methods have been developed to classify IDC and ILC, including the SqueezeNet Architecture [19], Bloom Richardson Grading [20], Weakly-Supervised and Transfer Learning [21],

Convolutional Neural Networks [22], [23], Gene Expression Profiling [24], Adaptive Mask Region-based Convolutional Network [25]. However, none of these methods have utilized physical parameters for detecting IDC and ILC. Therefore, in this study, we will use physical parameters to detect IDC and ILC using the Artificial Neural Network algorithm.

2. MATERIAL AND METHODS

This study used a cross-sectional design method. The cross-sectional design is defined as comparing the original data (which has been labeled) with the predicted data, and then calculating the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The sample population used in this study consists of 200 samples, with 100 IDC and 100 ILC cases. 80% of the total samples were used for training, and 20% were used for testing. The instruments used in this study are medical record documents and mammographic images that have been examined by pathologists, with a confirmed diagnosis. The research procedure is as follows: First, the patient records are obtained from the anatomical pathology department, and then the mammographic images are retrieved according to the patient's record number from the radiology database. The suspicious areas in the images are cropped to a size of 2 x 2 cm. Subsequently, 10 physical parameters from the mammographic images are calculated using the formulas provided in the journal [26]. The calculated parameter data are transformed into binary using the formula: Benary = 0.8 * (X – Min) / (Max – Min) + 0.1. The data are then divided into 80% for training and 20% for testing. These physical parameter data are input into the ANN algorithm, and the TP, TN, FP, FN values are calculated using the formulas from the journal [26]. Finally, the values for Accuracy, Recall (Sensitivity), Precision, F1-Score, and Specificity are computed, as shown in the block diagram in Figure 1.

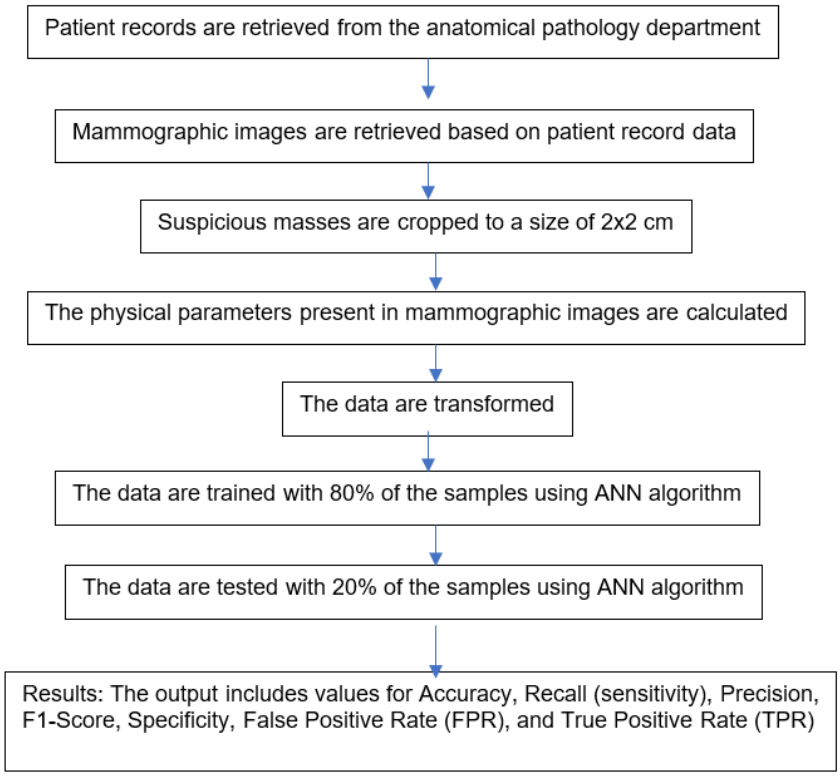

Fig. 1. Research block diagram

3. RESULTS AND DISCUSSION

3.1 Results

There is a significant difference between the data visualizations of IDC and ILC, as shown in Figure 2.

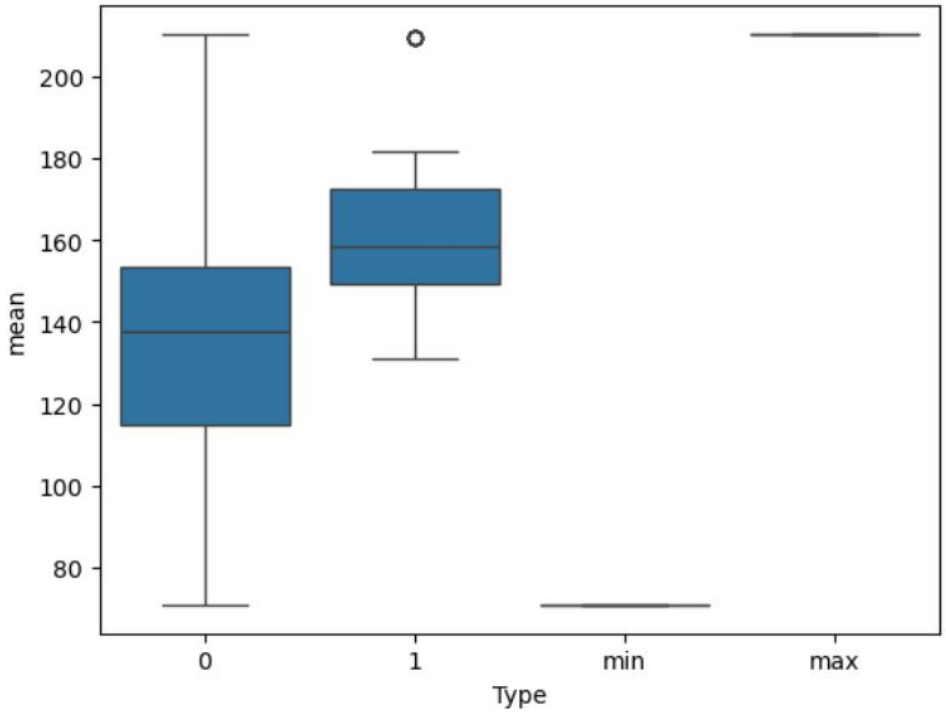

Fig. 2. The data visualizations of IDC (0) and ILC (1).

The results for TP, TN, FP, and FN are presented in Table 1.

Table 1 Confusion Matrix

		ORIGINAL	
		IDC	ILC
PREDICTIONS	IDC	TP=10	FP=10
	ILC	FN=0	TN=20

The testing results yielded values for Accuracy, Recall, Precision, Specificity, F1-Score, FPR, and TPR, as shown in Table 2.

Table 2. The results of the calculations for Accuracy, Recall, Precision, Specificity, F1-Score, FPR, and TPR.

	IDC	**ILC**
Accuracy	77.5 %	77.5 %
Recall (Sensitivity)	55 %	100%
Precision	100%	68.97 %
Specificity	100%	55 %
F1-Score	70.97 %	81.63 %
False Positive Rate (FPR)	0	0.45
True Positive Rate (TPR)	0.55	1

The ROC graph for ILC and ILC is shown in Figure 3

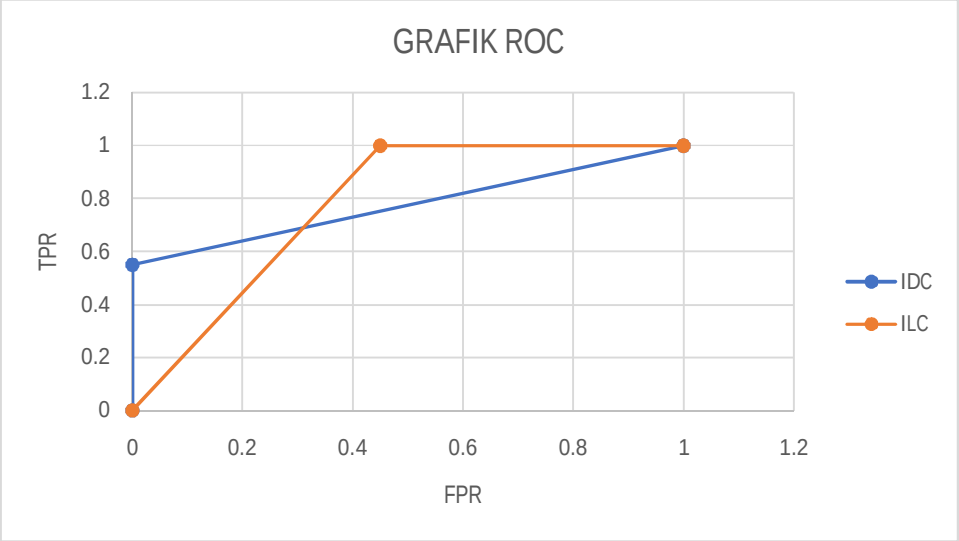

Fig. 3. IDC and ILC ROC graph

3.2 Discussions

The artificial neural network is able to classify Invasive Ductal Carcinoma with an accuracy of 77.5%, sensitivity (recall) of 55%, precision of 100%, F1-Score of 60.97%, specificity of 100%, FPR of 0, and TPR of 0.55. For Invasive Lobular Carcinoma, the accuracy is 77.5%, sensitivity (recall) is 100%, precision is 68.97%, F1-Score is 81.63%, specificity is 55%, FPR is 0.45, and TPR is 1. These results are comparable to those of previous studies, which used various methods, including Whole Slide Image (WSI) processing, a commonly used technique in the analysis performed by pathologists. The procedure begins with the collection of tissue from the human body, followed by fixation, dehydration, clearing, infiltration, embedding, sectioning, and staining. The tissue is then placed on a glass slide and examined using a light microscope [7]. Prior to visual analysis, the tissue is typically stained with hematoxylin and eosin (H&E) [8]. Previous researchers have extensively utilized the results of Whole Slide Image (WSI) to classify IDC-ILC using image processing methods, including machine learning techniques. For example, Chaudhury et al. [9] used a fast AI technique and the Sqeezenet architecture to classify IDC. Chaudhury's research focused on the color similarity in histological images and provided a solution based on Grad-CAM for feature extraction. The results yielded an accuracy of 90.3%, sensitivity of 78.70%, specificity of 93.78%, and precision of 83.46%. Talpur et al. [10] classified IDC based on Bloom Richardson Grading to determine the grade of breast cancer, achieving an accuracy of 92.81%. Kanavati et al. [11] classified IDC and benign cases using Whole Slide Image (WSI) with a Transfer Learning method, achieving ROC AUC values ranging from 0.95 to 0.98. Kanavati et al. [12] also classified ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC) using Whole Slide Image (WSI) with a deep learning CNN model, obtaining ROC AUC values between 0.93 and 0.97. In classifying DCIS and IDC, Kanavati focused on the presence or absence of micro-invasion at the lesion margins and whether it extended beyond the myoepithelium and basement membrane. Kawattikul et al. [7] employed the graph-cut technique to classify IDC and non-IDC from histological images. Kawattikul divided the entire slide into two areas: IDC and non-IDC, and achieved a precision value of 85% and recall (sensitivity) of 84%. Cruz-Roa et al. [13] automatically detected IDC in the WSI regions of breast cancer histopathology using a CNN method, achieving an accuracy of 84.23%, sensitivity of 79.60%, specificity of 88.86%, and precision of 65.40%. Hannemann et al. [14] classified ductal carcinoma in situ (DCIS) using gene expression profiling. Hannemann was able to differentiate between DCIS and invasive breast carcinoma with an average performance of 91%. Suvarna Vani et al. [15] detected and classified IDC using Artificial Intelligence Adaptive Mask Region Convolutional Network (Mask R-CNN) and Deep CNN, achieving an accuracy of 99.02%, precision of 98.93%, sensitivity (recall) of 98.5%, and F1-Score of 99.23%. Ganesh et al. [16] employed various deep learning methods to classify IDC and non-IDC. For the Logistic Regression method, an accuracy of 66.39%, precision of 66%, sensitivity (recall) of 67%, and F1-Score of 67% were obtained. The Random Forest method achieved an accuracy of 71.8%, precision

of 70%, sensitivity (recall) of 76%, and F1-Score of 73%. The Gradient Boosting method yielded an accuracy of 77%, precision of 78%, sensitivity (recall) of 76%, and F1-Score of 77%. The Extra Trees method resulted in an accuracy of 73%, precision of 72%, sensitivity (recall) of 77%, and F1-Score of 74%. The Convolutional Neural Networks (CNNs) method achieved an accuracy of 80.36%. Araujo et al. [8] used Convolutional Neural Networks to classify two classes: carcinoma and non-carcinoma, achieving an accuracy of 83.3%. When classifying four classes (normal tissue, benign lesion, in situ carcinoma, and invasive carcinoma), an accuracy of 77.8% and sensitivity of 95.6% were obtained.
In this study, the physical parameters present in mammographic images were used as input variables for the ANN method to classify IDC and ILC. This approach represents a novel method that has not been previously employed by other researchers, who mostly rely on biopsy results. The outcomes obtained are not significantly different from those of the methods explored in earlier studies.

Related Research

S. Chaudhury et al. (2023) utilized the SqueezeNet architecture method to identify IDC. The SqueezeNet architecture is a variant of CNN that yields results and accuracy nearly identical to ImageNet, while utilizing fewer parameters by replacing 3x3 filters with 1x1 filters. Prior to being input into the SqueezeNet architecture, the data was first normalized to a minimum of 0 and a maximum of 1 by dividing the pixel intensity values by 255. The SqueezeNet architecture performed convolution ten times. The obtained result achieved an accuracy of 90.3%.
S. Talpur et al. (2022) employed Bloom Richardson Grading to identify IDC and determine the grade value of patients by evaluating three factors: tubular formation, nuclear pleomorphism, and mitotic rate. The accuracy values obtained for each grade were 95.58%, 92.81%, and 96.41% for grade 1, grade 2, and grade 3, respectively.
F. Kanavati et al. (2021) employed Weakly-Supervised and Transfer Learning to classify IDC and benign cases. Four test sets were used: one biopsy test set and three surgical test sets. The obtained ROC values ranged from 0.95 to 0.98.
A. Cruz-Roa et al. (2014) employed Convolutional Neural Networks to detect IDC and Normal cases. Whole slide images (WSI) were divided into 100×100 pixel image sections. The sampling process for WSI images by pathologists involved dividing the images into non-overlapping sections, with tissue regions marked in red (IDC) and green (Normal) being selected. The CNN architecture was constructed with two convolutional layers, two pooling layers, and one fully connected layer. An accuracy of 84.23% was obtained.
J. Hannemann et al. (2016) employed gene expression profiling to classify invasive breast carcinoma and ductal carcinoma in situ (DCIS). The results indicated that the overall gene expression profile of in situ samples with moderate differentiation components was more similar to well-differentiated DCIS than to poorly differentiated DCIS
K. Suvarna Vani et al. (2022) employed Artificial Intelligence to detect and classify IDC. Suvarna utilized the trained LeNet-5 network on the ImageNet database. LeNet-5 consists of a total of seven layers: an input layer, two pooling layers, two fully connected layers, two convolutional layers, and an output layer. Convolution was performed using a 5x5 kernel, and pooling was done using a 2x2 size. The proposed method achieved an accuracy of 99.02%, precision of 98.84%, recall of 97.26%, and F1-score of 98.54%.

4. CONCLUSIONS

The artificial neural network (ANN) was able to classify Invasive Ductal Carcinoma (IDC) and Invasive Lobular Carcinoma (ILC) in breast cancer. The values for Accuracy, Recall (sensitivity), Precision, F1-Score, Specificity, False Positive Rate (FPR), and True Positive Rate (TPR) using the artificial neural network method were 77.5%, 50%, 100%, 70.97%, 100%, 0, and 0.55 for IDC, respectively. For ILC, the values for Accuracy, Recall (sensitivity), Precision, F1-Score, Specificity, False Positive Rate (FPR), and True Positive Rate (TPR) using the artificial neural network method were 77.5%, 100%, 68.97%, 81.63%, 55%, 0.45, and 1, respectively

ACKNOWLEDGEMENTS

The authors would like to express the sincere gratitude to the Academic Advisor, the Supervisors and

Examiners of the thesis, the Head of the Department of Informatics Engineering, and the Dean of the Faculty of Mathematics and Natural Sciences at Udayana University, whose support and assistance have contributed significantly to the successful completion of this research.

COMPETING INTERESTS

For all authors, there are no conflicts of interest

AUTHORS' CONTRIBUTIONS

Frady: project design; data analysis; script writing. Suwiprabayanti: Manuscript editing, data collection; data analysis. Arta: Manuscript editing, data collection; data analysis. sanjaya: program maker, manuscript editor..

REFERENCES

[1] E. Mirbagheri, M. Shafiee, M. Shanbezadeh, and H. Kazemi-Arpanahi, "Developing the required data set for the integration of breast cancer registry systems in Iran,"Informatics Med. Unlocked, vol. 32, Jan. 2022, doi: 10.1016/j.imu.2022.101011.

[2] HA Geuzinge, EAM Heijnsdijk, IM Obdeijn, HJ de Koning, and MMA Tilanus-Linthorst, "Experiences, expectations and preferences regarding MRI and mammography as breast cancer screening tools in women at familial risk,"Breast, vol. 56, pp. 1–6, Apr. 2021, doi: 10.1016/j.breast.2021.01.002.

[3] I. Bin Khalid, MA Parvaiz, A. Sarwar, OS Shaikh, N. Javed, and A. Fatima, "A leafy surprise: Case report of male breast malignant phyllodes,"Int. J. Surg. Case Rep., vol. 88, no. October, p. 106536, 2021, doi: 10.1016/j.ijscr.2021.106536.

[4] RA Waldman, J. Finch, J.M. Grant-Kels, C. Stevenson, and D. Whitaker-Worth, "Skin diseases of the breast and nipple: Benign and malignant tumors,"Journal of the American Academy of Dermatology, vol. 80, no. 6. Mosby Inc., pp. 1467–1481, June. 01, 2019. doi: 10.1016/j.jaad.2018.08.066.

[5] S. Bayardet al., "Screening mammography mitigates breast cancer disparities through early detection of triple negative breast cancer," Clin. Imaging, vol. 80, pp. 430–437, 2021, doi: 10.1016/j.clinimag.2021.08.013.

[6] B. Malik, E. Iuanow, and J. Klock, "An Exploratory Multi-reader, Multi-case Study Comparing Transmission Ultrasound to Mammography on Recall Rates and Detection Rates for Breast Cancer Lesions,"Acad. Radiol., vol. 29, pp. S10–S18, Jan. 2022, doi: 10.1016/j.acra.2020.11.011.

[7] T. Uematsu, A. Izumori, and WK Moon, "Overcoming the limitations of screening mammography in Japan and Korea: a paradigm shift to personalized breast cancer screening based on ultrasonography,"Ultrasonography, vol. 42, no. 4, pp. 508–517, 2023, doi: 10.14366/usg.23047.

[8] J. Wang, L. Wang, Z. Yang, W. Tan, and M. Luo, "Multifractal analysis of MRI . images from breast cancer patients,"Multimed.Tools Appl., no. 0123456789, 2023, doi: 10.1007/s11042-023-17380-9.

[9] CE Edmonds, SR O'Brien, DA Mankoff, and AR Pantel, "Novel applications of molecular imaging to guide breast cancer therapy,"Cancer Imaging, vol. 22, no. 1, pp. 1–17, 2022, doi: 10.1186/s40644-022-00468-0.

[10] M. Shawky, Z. Abd, E. Ali, and DH Hashem, "Role of positron-emission tomography_computed tomography (PET_CT) in breast cancer _ Enhanced Reader.pdf," vol. 9, 2020.

[11] ZM Sarvestani, J. Jamali, M. Taghizadeh, and MHF Dindarloo, "A novel machine learning approach on texture analysis for automatic breast microcalcification diagnosis classification of mammogram images,"J. Cancer Res. Clin. Oncol., vol. 149, no. 9, pp. 6151–6170, 2023, doi: 10.1007/s00432-023-04571-y.

[12] YS Leong, K. Hasikin, KW Lai, N. Mohd Zain, and MM Azizan, "Microcalcification Discrimination in Mammography Using Deep Convolutional Neural Network: Towards Rapid and Early Breast Cancer Diagnosis,"Front. Public Heal., vol. 10, no. April, 2022, doi: 10.3389/fpubh.2022.875305.

[13] SS Yasiranet al., “Microcalcifications segmentation using three edge detection techniques,” Int. Conf. Electrons. Devices, Syst. Appl., pp. 207–211, 2012, doi: 10.1109/ICEDSA.2012.6507798.
[14] SH Abdulla, AM Sagheer, and H. Veisi, “Breast cancer segmentation using K-means clustering and optimized region-growing technique,”Bull. Electr. Eng. Informatics, vol. 11, no. 1, pp. 158–167, 2022, doi: 10.11591/eei.v11i1.3458.
[15] B. Charan Patel, "Early Detection of Breast Cancer using Self Similar Fractal Method," 2010.
[16] R. Krishnamoorthy, N. Amudhavalli, and MK Sivakolundu, “Identification of Microcalcifications with Orthogonal Polynomials Model,” 2010.
[17] Y. Zheng, “Breast cancer detection with gabor features from digital Mammograms,”Algorithms, vol. 3, no. 1, pp. 44–62, March. 2010, doi: 10.3390/a3010044.
[18] HB Kekre, TK Sarode, and SM Gharge, “Tumor Detection in Mammography Images using Vector Quantization Technique,” 2009.
[19] S. Chaudhury, K. Sau, M.A. Khan, and M. Shabaz, “Deep transfer learning for IDC breast cancer detection using fast AI technique and Sqeezenet architecture,”Mathematical Biosciences and Engineering, vol. 20, no. 6. American Institute of Mathematical Sciences, pp. 10404–10427, 2023. doi: 10.3934/mbe.2023457.
[20] S. Talpur, M. Rashid, S. Jawaid Khan, and S. Abid Syed, “Automatic Detection System to Identify Invasive Ductal Carcinoma by Predicting Bloom Richardson Grading from Histopathological Images,”J. Indep. Stud. Res. Comput., vol. 20, p. 2022, doi: 10.31645/JISRC.22.20.1.6.
[21] F. Kanavati and M. Tsuneki, “Breast Invasive Ductal Carcinoma Classification on Whole Slide Images with Weakly-Supervised and Transfer Learning,” 2021, doi: 10.3390/cancers.
[22] A. Cruz-Roaet al., “Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks,” in Medical Imaging 2014: Digital Pathology, SPIE, Mar. 2014, p. 904103. doi: 10.1117/12.2043872.
[23] T. Araujoet al., “Classification of breast cancer histology images using convolutional neural networks,” PLoS One, vol. 12, no. 6, June. 2017, doi: 10.1371/journal.pone.0177544.
[24] J. Hannemann, A. Velds, JBG Halfwerk, B. Kreike, JL Peterse, and MJ Van de Vijver, “Classification of ductal carcinoma in situ by gene expression profiling,”Breast Cancer Res., vol. 8, no. 5, Oct. 2006, doi: 10.1186/bcr1613.
[25] K. Suvarna Vaniet al., “Detection and Classification of Invasive Ductal Carcinoma using Artificial Intelligence Detection and Classification of Invasive Ductal Carcinoma using Artificial Intelligence,” 2022, doi: 10.21203/rs.3.rs-2069384/v1.
[26] English: A.A.N Gunawanet al., “Digital Mammography Data Transformation Model to Improve the Accuracy of Benign and Malignant Detection in Breast Cancer,” Int. J. Intell. Syst. Appl. Eng., vol. 12, no. 1, pp. 647–665, 2024.