Breast Cancer Detection and Classification using Deep Learning Xception Algorithm

Basem S. Abunasser[1], Mohammed Rasheed J. AL-Hiealy[2], Ihab S. Zaqout[3], Samy S. Abu-Naser[4]
University Malaysia of Computer Science & Engineering (UNIMY), Cyberjaya, Malaysia[1, 2]
Faculty of Engineering and Information Technology, Al-Azhar University, Gaza, Palestine[3, 4]

Abstract*—Breast Cancer (BC) is one of the leading cause of deaths worldwide. Approximately 10 million people pass away internationally from breast cancer in the year 2020. Breast Cancer is a fatal disease and very popular among women globally. It is ranked fourth among the fatal diseases of different cancers, for example colorectal cancer, cervical cancer, and brain tumors. Furthermore, the number of new cases of breast cancer is anticipated to upsurge by 70% in the next twenty years. Consequently, early detection and precise diagnosis of breast cancer plays an essential part in enhancing the diagnosis and improving the breast cancer survival rate of patients from 30 to 50%. Through the advances of technology in healthcare, deep learning takes a significant role in handling and inspecting a great number of X-ray, Magnetic Resonance Imaging (MRI), computed tomography (CT) images. The aim of this study is to propose a deep learning model to detect and classify breast cancers. Breast cancers has eight classes of cancers: benign adenosis, benign fibroadenoma, benign phyllodes tumor, benign tubular adenoma, malignant ductal carcinoma, malignant lobular carcinoma, malignant mucinous carcinoma, and malignant papillary carcinoma. The dataset was collected from Kaggle depository for breast cancer detection and classification. The measurement that was used in the evaluation of the proposed model includes: F1-score, recall, precision, accuracy. The proposed model was trained, validated and tested using the preprocessed dataset. The results showed that Precision was (97.60%), Recall (97.60%) and F1-Score (97.58%). This indicates that deep learning models are suitable for detecting and classifying breast cancers precisely.

Keywords*—*Breast cancer; deep learning; xception

I. Introduction

Breast Cancer is considered one of the most common type of cancers and is taken as the second, to cause death in women worldwide. In 2002, it was the second most common cancer globally, by means of exceeding of one million new cases. Despite the enhancements in early detection and knowing of the molecular foundations of the biology of the breast cancer, nearly 30% of the patients with "early-stage" breast cancer have disease recurred.

Determining the most real and least toxic therapy, molecular and clinical features of the tumor in certain need of inspection. General treatment of breast cancer comprises hormonal agents, immunotherapy and cytotoxic. These medications are used in adjuvant, transitional, and neoadjuvant modes. Overall, systemic agents are vigorous at the start of management in 90% of main breast cancers and 50% of metastases. Though, after a flexible period of time, progress does follow. At the present, opposition to treatment is not only popular but anticipated [7, 9].

Diagnosing BC early is the most important features of its treatment. Amongst the diverse types of diagnosing technique of BC, imaging is the main diagnostic method that can make offer important data on the patients of BC. It was revealed that a few techniques of imaging e.g.: "Magnetic Resonance Imaging (MRI), Single-Photon Emission Computed Tomography (SPECT), mammography, Computerized Tomography (CT), and Positron Emission Tomography (PET)" can be put into use for the identifying and observing of the patients of BC in diverse stages. Moreover, techniques of imaging, and the use of biomarkers of biochemical e.g.: "proteins, mRNAs, DNA, and microRNAs" can be used as an innovative analytic and therapy gears for BC patients [12, 13 and 14].

Deep Learning (DL) as a type of Machine Learning method has its working mechanism inspiration from the way human brain neurons process information. The most basic element of the DL networks are small nodes known as artificial neurons , which are usually arranged in layers wherein each neuron has connections to every neuron in the subsequent layer via weighted connections. Recently, the rise of DL technique has encouraged different areas of study to solve complex problem or enhance performances of existing study using the new technology. Example of application of DL includes machine translation, speech recognition, sentiment analysis, image recognition, face recognition, signal processing, etc. [17].

The current trend of applying DL technique in medical applications has reaped great success with the potential of DL technology to perform faster analysis with higher accuracy when compared with human practitioners. To give an example, a notable study by Google on the diagnostic classification of diabetic retinopathy has shown remarkable performance that exceeds the capabilities of domain experts [22]. In addition, the application of transfer learning techniques can be seen in several studies. As opposed to training a model from scratch, transfer learning method allowed the use of weight s trained previously on a specific task to be reused as the starting point for a model on another task.

It comes with the benefits of shorter training time and is possible to deliver better results. CNN models such as Alex Net, VGG, ResNet, and Xception are the most dominant models applied in transfer learning approach. Generally, there are two ways of applying transfer learning technique; First, pre-trained model with weights trained on ImageNet dataset can be used as feature extractor; Second, fine-tuning of pre-trained model on a new problem. Study utilized transfer learning technology with Xception neural network to speed up the training of model for distinguishing subjects [6].

II. Literature Review

Many researchers have employed artificial intelligence, expert systems, and neural networks in the diagnosis of BC to increase the screening accuracy. Usually hospitals use x-rays for diagnosis of BC, but lately, hospitals have been using mammography images as a substitute of x-rays, due to their easiness in analyzing and studying through intelligent models, which have increased the efficiency and accuracy of BC diagnosis.

A number of models and methods were proposed for increasing the efficacy of the diagnosis of BC. These methods include: "Linear Regression (LR), Artificial Neural Network (ANN), K-Nearest Neighbors (KNN) search, Softmax Regression, and Support Vector Machine (SVM), and Convolutional Neural Network (CNN)".

The authors in [3] and [18] collected their datasets from Kaggle depository. They proposed a prediction model to predict whether or not a person has breast cancer and to provide awareness or diagnosis about it. The authors of these studies made a comparisons using accuracy of each results of the SVM, random forest (RF), Naive Bayes classifier, and logistic regression on the dataset to deliver a precise model for breast cancer prediction. The outcome of their experiments indicated that the techniques of machine learning models that were applied in their studies predicted breast cancer disease with an accuracy between 52.63% and 98.24%.

In these studies, the authors propose new KNN models for the need for early diagnosis and precise diagnostic procedures that clinicians can use to classify whether cancer is benign or malignant. The main objective of their study was to compare the results of supervised learning classification algorithms and to combine these algorithms using a classification technique called voting. Voting was a grouping method because they could combine multiple models to achieve higher classification accuracy. The datasets were collected from the University of Wisconsin. [8] achieved 98.90%, [16] achieved an accuracy of 97.60%, [17] achieved 97.13%, [10] achieved 99.9%, [27] achieved 98.10% [28] achieved 98.23% and [6] achieved 83.45%.

In the following studies like [15, 21, 29, 30, 31] the authors argued that early detection and prevention can significantly reduce the chances of passing away. An important fact about breast cancer prognosis was to improve the likelihood of cancer recurrence. Their studies aimed to find the probability of breast cancer recurrence using different machine learning techniques like SVM. The authors presented new approaches in order to improve the accuracy of these models. Cancer patient data were collected from the Wisconsin Dataset of the UCI Machine Learning Repository. The dataset contains a total of 35 attributes in which they applied the Naive Bayes, C4.5 Decision Tree and SVM algorithms and measured their prediction accuracy. The efficient feature selection algorithm helped them improve the accuracy of each model by reducing some lower-order features. Not only are the contributions of these traits much lower, but their addition also misleads the classification algorithms. After careful selection of higher-order attributes, they significantly improved accuracy rate for all algorithms.

The study of [4, 10, 11, 24, 25, 26, 32] proposed method that uses CNNs which is a particular type of deep learning, feedforward network, and some preliminary experiments used the deep learning approach to classify breast cancer mammography images from BreaKHis, which is a public dataset. Method based on the extraction of image patches for training the CNN and the combination of these patches for final classification. All their convolutional network technique for categorizing screening mammograms reached good accuracies. The finest performance on an independent test set of digitized film mammograms from the Digital Database for Screening Mammography was 0.88%, (the sensitivity: 86.2%, the specificity: 80.2%).

The aim of the studies [1, 2, 5] was to develop methods for classifying cancers into specific prognostic categories based on gene expression signatures using artificial neural networks. ANN were trained to use small round blue cell tumors (SRBCTs) as a model. These cancers belong to four distinct diagnostic classes and often present diagnostic dilemmas in clinical practice. ANN properly classified all samples and recognized the genes most related to the classification. The experimental results suggested that the new strategies were able to improve the stability of the selection results as well as the sample classification accuracy. The new algorithms achieved accuracy of classification about 99%.

A. Previous Studies Summary

The previous studies aim was to detect if an image has breast cancer or not. The current study's aim is to classify eight types of breast cancers. A summary was made of the studies discussed in the previous section in terms of the following criteria: Machine Learning methods used, programming language used, best result attained, best method, and source of the dataset used in Table I.

There are numerous studies and suggestions that put on to deal with breast cancer. Each suggestion or study has a different way of detecting or dealing with it. Some of the studies use methods that rely on image processing and the use of systems for this purpose and other studies that address the form of the breast and observe any changes in it. Some studies have enhanced the quality of algorithms that exist for detecting the disease. All these studies concentrate on the detection of whether breast cancer exists or not. None of these studies classified the eight different classes of the breast cancer. In the current study, the concentrate will be on the classification of the eight classes of breast cancers: "benign adenosis, benign fibroadenoma, benign phyllodes tumor, benign tubular adenoma, malignant ductal carcinoma, malignant papillary carcinoma, malignant mucinous carcinoma, and malignant lobular carcinoma".

TABLE I. A SUMMARY OF THE PREVIOUS STUDIES BY METHODS, BEST METHOD, LANGUAGE, RESULT, DATASET USED

Reference	Methods Used	Best Methods	Language	Best Result	Data Provider
[1]	ANN	ANN	-	96.00	Private
[2]	ANN	ANN	Python	90.30	Kaggle
[3]	LR	LR	-	97.00	Wisconsin
[4]	CNN	CNN	Python	89.50	BreaKHis
[5]	ANN	ANN	-	99.00	Wisconsin
[6]	KNN	KNN	-	83.45	Wisconsin
[8]	KNN	KNN	-	98.90	FMC
[10]	CNN	CNN	Python	90.00	BreaKHis
[11]	CNN	CNN	Python	88.00	DDSM
[15]	SVM,KNN	SVM	-	99.00	Wisconsin
[16]	SVM,KNN,ANN	SVM	-	97.60	Wisconsin
[17]	SVM,KNN	KNN	-	97.13	Wisconsin
[18]	SVM,LR	LR	Python	98.24	Kaggle
[21]	SVM	SVM	-	92.30	Wisconsin
[22]	SVM,KNN	KNN	-	99.90	Wisconsin
[23]	SVM,ANN	SVM	-	98.82	Wisconsin
[24]	SVM,KNN,CNN	CNN	-	97.00	Wisconsin
[25]	SVM,KNN,CNN	CNN	Python	99.30	Wisconsin
[26]	SVM,ANN	ANN	Python	94.00	DDSM
[27]	SVM,KNN	SVM	Minitab	98.10	Wisconsin
[28]	SVM,KNN	SVM	WEKA	97.13	Wisconsin
[29]	SVM,KNN,LR,ANN	ANN	-	99.30	Wisconsin
[30]	SVM,KNN	SVM	SAE	98.90	Wisconsin
[31]	SVM,ANN	ANN	-	95.70	Private
[32]	KNN,LR,SVM,CNN	CNN	Python	97.20	Wisconsin

III. METHODOLOGY

The methodology that was used in the study consists of dataset collection, dataset preparation, dataset splitting, create the proposed model, model training, validating and testing as can be seen in Fig. 1.

A. Dataset

The dataset has been collected from Kaggle depository. The Breast "Cancer Histopathological Image Classification" (BreakHis) consists of 7909 microscopic images of breast cancer tissue gathered from 82 patients using various magnifying factors (40X, 100X, 200X, and 400X). It contains 2,480 benign and 5,429 malignant samples (700X460 pixels, three-channel RGB, eight-bit depth in each channel, PNG image format). This dataset was built in collaboration with the P&D Laboratory – Pathological Anatomy and Cytopathology, Parana, Brazil"

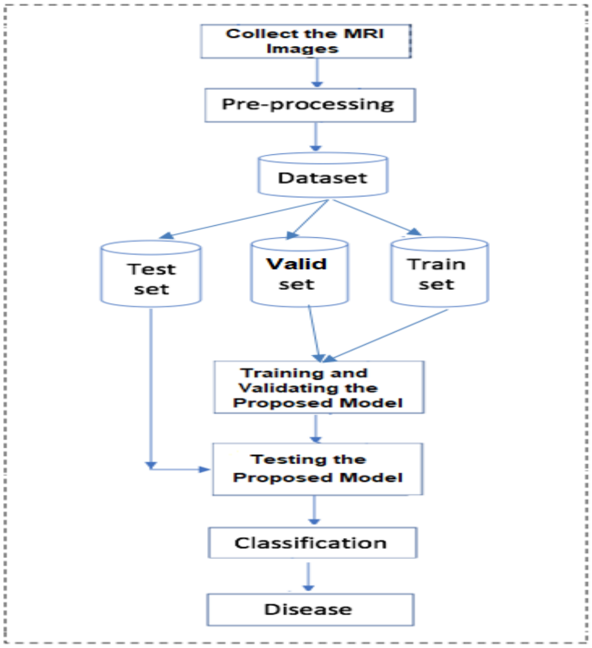

Fig. 1. Methodology Flowchart.

B. Data Perpetration

The number of images in the original BreakHis dataset has 7,909 images. This number of images is considered low and thus one can use new methods to generate more images to boost the dataset called Generative Adversarial Networks" (GANs). GAN is an algorithmic constructions that utilizes two NNs, fighting one in contradiction of the other (consequently the 'adversarial') to be able to produce new, fake instances of images that can pass as real images. They are very popular in image generation, voice generation, and video generation [19, 20, 23]. By using GAN the dataset was increased to 10,000 images. Each class of the eight classes has 1250 images. Sample of the eight classes are shown in Fig. 2.

C. Data Splitting

The dataset has been split into three datasets: training, validating, and testing datasets. The ratio of splitting is 60%, 20%, and 20% respectively.

D. Performance Measures

The most popular measurements were used in the performance of the proposed model: Accuracy is represented in (1), Precision is represented in (2), Recall is represented in (3) and F1-score is represented in (4).

$$\text{Accuracy} = (TP + TN) / (TP+TN+FP+FN) \quad (1)$$

$$\text{Precision} = (TP / (TP+FP) \quad (2)$$

$$\text{Recall} = (TP / (TP+FN) \quad (3)$$

$$\text{F1-score}= 2x(\text{Precision x Recall})/(\text{Precision}+\text{Recall}) \quad (4)$$

Where TP = True Positive, TN = True Negative

FP = False Positive, FN = False Negative

E. Proposed Model

A pre-trained Deep Learning model called Xception was proposed to be fine-tuned for the detection and classification of breast cancer. Xception model is considered to be among the beast pre-trained model of all classical Deep learning models due to its high accuracy in classifying the1000 natural images of ImageNet [17].

To be able to use a pre-trained model in the current dataset, it must be fine-tuned. In the current dataset, there are eight classes only. That means the top layer has to be removed from the Xception model which is called the classifier and to be replaced with the current classifier of the eight classes as illustrated in Fig. 3.

F. Model Training and Validating and Testing

The newly customized Xception model was trained with the prepared training dataset and cross-validating it with the validation dataset for 120 epochs. During the training, Learning Rate (0.0001), Batch Size (128), and the Optimizer is Adam were used. Furthermore, data augmentation was used during training to overcome the problem of overfitting. Fig. 4 and Fig. 5 shows the loss and accuracy of the training and validation of the xception model.

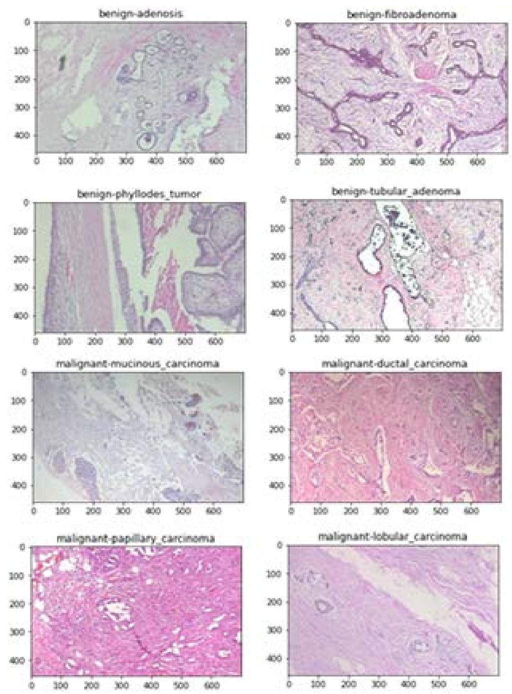

Fig. 2. Sample of the Dataset with its 8 Different Classes.

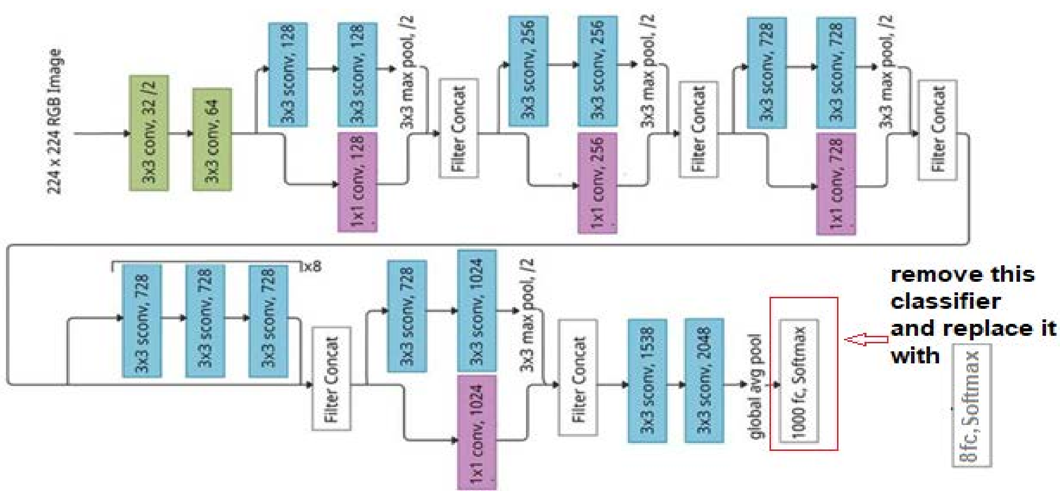

Fig. 3. Architecture of the Customized Xception Model.

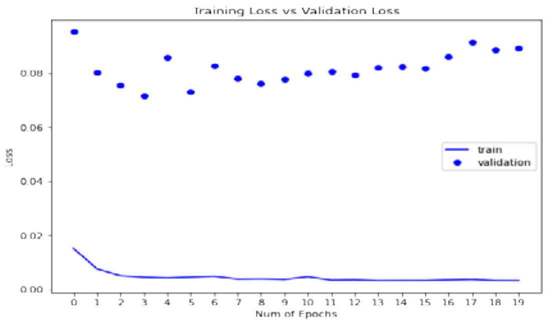

Fig. 4. Training and Validation Loss.

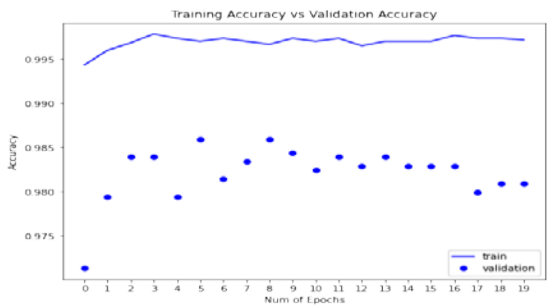

Fig. 5. Training and Validation Accuracy.

After the training was finished and validating the customized Xception model, it was tested with the testing dataset and the different performance measures were recorded.

IV. Result and Discussion

The model achieved Training Accuracy (99.78%), Validating Accuracy (98.59%) and Testing Accuracy (97.60%). In the customized model Training Loss was (0.00315), Validating Loss (0.07326), Testing Loss (0.09518). In terms of the time required for training and testing, the Xception model needed 2944 seconds for training and 5.32 seconds for testing.

Table II shows the precision, Recall, F1-Score of each class in the dataset in terms of the eight classes that the models are used for classifying: benign adenosis (BA), benign fibroadenoma (BF), benign phyllodes tumor (BPT), and benign tubular adenoma (BTA), malignant lobular carcinoma (MLC), malignant mucinous carcinoma (MMC), and malignant papillary carcinoma (MPC), malignant ductal carcinoma (MDC). The customized model achieved average Precision (97.60%), Recall (97.60%) and F1-Score (97.58%). Furthermore, the ROC Curve measure for each class in the dataset reached 100% as shown in Fig. 6.

TABLE II. XCEPTION PRECISION, RECALL, F1-SCORE OF EACH CLASS IN THE DATASET

Measure	**BA**	**BF**	**BPT**	**BTA**	**MLC**	**MMC**	**MPC**	**MDC**	**Model AVG**
Precision	99.20%	97.63%	100.00%	98.01%	95.34%	94.16%	97.62%	98.81%	97.60%
Recall	99.60%	98.80%	98.80%	98.40%	90.00%	96.80%	98.40%	100.00%	97.60%
F1-Score	99.40%	98.21%	99.40%	98.20%	92.59%	95.46%	98.01%	99.40%	97.58%

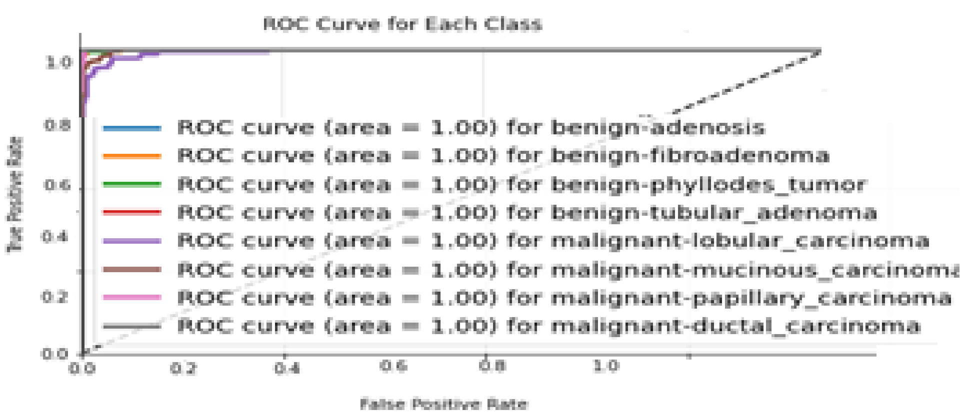

Fig. 6. ROC Curve for each Class in the Dataset.

V. Conclusion and Future Work

Breast Cancer is the top cause of deaths worldwide. About ten million people died globally from cancer in the year 2020. Breast Cancer is a fatal disease among women worldwide. With technical advances in healthcare, machine learning and deep learning play an important role in processing and analyzing a large number of medical images. The objective of this study is to propose a deep learning model for detecting classifying Breast Cancer. Xception model was used and customized to fit our current breast cancer eight classes dataset. The Dataset was collected from Kaggle and boosted using GAN. The dataset was split into three datasets: training, validating and testing. The customized Xception model was trained, validated, and tested. The mode achieved Precision (97.60%), Recall (97.60%) and F1-Score (97.58%).

In future work, other techniques of generating images will be used and the Xception model will be tested with newly generated images. Furthermore, other pre-trained model will be tested and compared with the current results.

REFERENCES

[1] Westermann et al., "Classification and diagnostic prediction of pediatric cancers using gene expression profiling and artificial neural networks", *GBM Annual Fall meeting Halle 2020*, vol. 2020.

[2] S. Joo, Y. Yang, W. Moon and H. Kim, "Computer-Aided Diagnosis of Solid Breast Nodules: Use of an Artificial Neural Network Based on Multiple Sonographic Features", IEEE Transactions on Medical Imaging, vol. 23, no. 10, pp. 1292-1300, 2021.

[3] T. Kiyan and t. Yildirim, "breast cancer diagnosis using statistical neural networks", Istanbul university journal of electrical & electronics engineering, vol. 4, no. 2, 2021.

[4] F. Spanhol, L. Oliveira, C. Petitjean and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks", 2017.

[5] H. Salem, G. Attiya, N. El-Fishawy. Early diagnosis of breast cancer by gene expression profiles. Pattern Analysis and Applications. Vol. 20, no. 2, pp. 567–578, 2017.

[6] U. K. Kumar, M. B. S. Nikhil and K. Sumangali, "Prediction of breast cancer using voting classifier technique," 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2017, pp. 108-114.

[7] Albatish, I.M., Abu-Naser, S.S. Modeling and controlling smart traffic light system using a rule based system. Proceedings - 2019 International Conference on Promising Electronic Technologies, ICPET 2019, 2019, pp. 55–60.

[8] N. Ponraj, E. Jenifer., P. Poongodi, S. Manoharan. "Morphological operations for the mammogram image to increase the contrast for the efficient detection of breast cancer", European Journal of Scientific Research, (ISSN) 1450-216X, vol. 68, no.4, pp. 494-505, 2021.

[9] Naser, S. S. A. JEE-Tutor: An intelligent tutoring system for java expressions evaluation. Information Technology Journal, 2008, Vol. 7, No. 3, 528-532.

[10] B. Gayathri, C. Sumathi, T. Santhanam, "Breast cancer diagnosis using machine learning algorithm a survey". International Journal of Distributed and Parallel Systems, vol. 4, no. 3, pp. 39-50, 2017.

[11] L. Shen, L. Margolies, J. Rothstein. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. Sci Rep vol. 9, no.1, pp. 24-35, 2019.

[12] World health organization https://www.who.int/.

[13] Abu-Naser, S.S., El-Hissi H., Abu-Rass, M., & El-khozondar, N. An expert system for endocrine diagnosis and treatments using JESS. Journal of Artificial Intelligence, 2010, Vol. 3, No. 4, pp. 239-251.

[14] Abu Naser, S.S. Evaluating the effectiveness of the CPP-Tutor, an intelligent tutoring system for students learning to program in C++. Journal of Applied Sciences Research, 2009, Vol. 5, No. 1, pp. 109-114.

[15] A. F. Agarap. On breast cancer detection: an application of machine learning algorithms on the wisconsin diagnostic dataset. ArXiv, abs/1711.07831, 2018.

[16] D. R. Pavithra, S. Preethi, & A. K. SriRakshitha. Breast Cancer Classification using the Supervised Learning Algorithms. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1492-1498, 2021.

[17] H. Asri, H. Mousannif, H. Moatassime, & T. Noël. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. ANT/SEIT, 2017.

[18] A. M. Gonzalez-Angulo, F. Morales-Vasquez, G. N. Hortobagyi. Overview of resistance to systemic therapy in patients with breast cancer. Adv Exp Med Biol. 2017; vol. 608, pp. 1-22.

[19] Saleh, A., Sukaik, R., Abu-Naser, S.S. Brain tumor classification using deep learning. Proceedings - 2020 International Conference on Assistive and Rehabilitation Technologies, iCareTech 2020, 2020, pp. 131–136, 9328072.

[20] Arqawi, S., Atieh, K.A.F.T., Shobaki, M.J.A.L., Abu-Naser, S.S., Abu Abdulla, A.A.M. Integration of the dimensions of computerized health information systems and their role in improving administrative performance in Al-Shifa medical complex, Journal of Theoretical and Applied Information Technologythis link is disabled, 2020, Vol. 98, No. 6, pp. 1087–1119.

[21] A. I. Pritom, M. A. R. Munshi, S. A. Sabab and S. Shihab, "Predicting breast cancer recurrence using effective classification and feature selection technique," 2016 19th International Conference on Computer and Information Technology (ICCIT), 2017, pp. 310-314.

[22] A. Joshi, D. A. Mehta. Comparative Analysis of Various Machine Learning Techniques for Diagnosis of Breast Cancer. International Journal on Emerging Technologies, vol. 8, (NCETST-2017), pp. 522–526, 2017.

[23] A. D. Omondiagbe, V. Shanmugam, S. S. Amandeep. Machine Learning Classification Techniques for Breast Cancer Diagnosis, vol. 5, no. 3, 2019.

[24] K.Wadkar, P. Pathak, N. Wagh. Breast cancer detection using Ann network and performance analysis with SVM. International journal of computer engineering and technology, vol. 10, no. 3, 2019, https://doi.org/10.34218/ijcet.10.3.2019.009.

[25] M. TIWARI, R. Bharuka, P. Shah, R. Lokare. Breast Cancer Prediction Using Deep Learning and Machine Learning Techniques. SSRN Electronic Journal. Published. https://doi.org/10.2139/ssrn.3558786, 2020.

[26] D. A. Ragab, M. Sharkas, S. Marshall, J. Ren. Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ, vol. 7, no. 6, pp. 20-31, 2019. https://doi.org/10.7717/peerj.6201.

[27] M. F. Ak. A Comparative Analysis of Breast Cancer Detection and Diagnosis Using Data Visualization and Machine Learning Applications. Healthcare, vol. 8, no. 2, pp.1-11, 2020.

[28] H. Mousannif, H. Asri, H. A. Moatassime, T. Noel. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. Procedia Computer Science, vol. 83, pp. 1064–1069, 2017.

[29] R. MurtiRawat, S. Panchal, V. K. Singh and Y. Panchal. Breast Cancer Detection Using K-Nearest Neighbors, Logistic Regression and Ensemble Learning. International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 534-540, doi: 10.1109/ICESC48915.2020.9155783.

[30] D. Selvathi and A. A. Poornila. Performance Analysis of Various Classifiers on Deep Learning Network for Breast Cancer Detection. International Conference on Signal Processing and Communication (ICSPC), pp. 359-363, IEEE 2017.

[31] A. LG, E. AT. Using Three Machine Learning Techniques for Predicting Breast Cancer Recurrence. Journal of Health & Medical Informatics, vol. 04, no. 02.2018.

[32] M. A. Naji, S. E. Filali, K. Aarika, E. H. Benlahmar, R. A. Abdelouhahid, O. Debauche. Machine Learning Algorithms for Breast Cancer Prediction and Diagnosis. Procedia Computer Science, vol. 191, pp. 487–492, 2021.