Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

Classification of Breast Cancer Detection Using K-Nearest Neighbor Algorithm Trained with Wisconsin Dataset

Dr. Jayanthi N
Assistant Professor(Sr.G)
Dept. of Computer Engineering
KPR Institute of Engineering and Technology,
Coimbatore, Tamil Nadu, India.
jayanthi.n@kpriet.ac.in

Gitanjali Wadhwa
Assistant Professor
Dept. of Computer Engineering
KPR Institute of Engineering and Technology,
Coimbatore, Tamil Nadu, India.
Gitanjali.wadhwa1994@gmail.com

Abstract: Cancer is a type of diseases, which is caused by changes in body cells and a huge increase in number of their control and growth. Prediction of breast cancer is necessary to increase the mortality rate of those patients who are suffering from this breast cancer. Breast cancer diagnosis required analysis of medical images. There are many breast cancer detection algorithms has been proposed like SVM, KNN,K-MEAN, RAMDON FOREST, NAVIE BAYES.The most invasive cancer in females: breast cancer is affecting 12% for population worldwide. Breast cancer has no physical symptoms unless there is sign of painless lump. Breast cancer pathology images are classified into: normal, benign and malignant. Breast cancer cell have sub classes interrelated to cells' variability, organization and density with the structure of tissues and morphology.Eventually various studies were systematically reviewed related articles published from 2010 to 2020 were considered.Result is calculated based on performance metric elements: Accuracy, sensitivity. In this we are using a dataset that obtains 569 samples with 30 features, this dataset is majorly called the Wisconsin dataset. Many techniques are implemented on this dataset we are using Machine learning methodology (KNN) for the diagnosis and training purpose, it is supervised machine learning technique; the initial step of this classifier is to calculate distance of the input test sample with all the dataset entries of training samples. The best accuracy achieved from KNN is 96.49%.

Keywords— Breast cancer, machine learning, deep learning, KNN

I. Introduction:

In recent decades, a huge advancement has been performed in the field of cancer research. Breast cancer rate is increasing day by day and many researchers are using different techniques in this research area. The techniques such as image processing, machine learning, screening just to identify the stage of that particular cancer in the particular patient before the symptoms arrives. Experts are developing new approaches for the early stage detection and prediction for cancer treatments. With the introduction of new detection techniques, the accurate prediction of a cancer is one of the most challenging tasks for researchers.Breast cancer patients' rate in India is increasing uncontrollably means it needs to be handled. About 262468 breast cancer patients were diagnosed in India in the year 2019, out of 262468 approximately 84909 were died because they were diagnosed in late stages. This is somehow the Factors for late marriage, having no children i.e. problem in conceiving a baby. This cancer starts with uncontrolled and fast multiplication of breast tissues which are classified as benign and malignant. Benign are just abnormality present in breast tissues but may not cause death to the patient on the other side malignant is type of tumor tissues their timely diagnosis can increase the mortality rate. The accuracy is calculated from the classification techniques is the rate of test sets classified accurately. Many researches shows that various techniques of machine learning have been used for breast cancer detection then among all others KNN is most useful and efficient for classification. Machine learning has huge application in medical domain and it is growing rapidly because of its effective approaches in prediction and classification. The dataset used in this approach is WDBC (Wisconsin Diagnostic Breast Cancer)[3]. Different MLTs were used in detection process of breast cancer generally at stage III or stage IV. Classification of breast cancer found in

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

females using ML algorithm is done by using dataset and that dataset is divided into two parts: training dataset and testing dataset.
Table 1 represents the complete overview of machine learning techniques and their related solutions in breast cancer detection.

S.NO.	TECHNIQUE	SOLUTION
1.	Support Vector Machine (SVM)	Breast Cancer prognosis
2.	KNN	Breast cancer risk classification
3.	K-mean	Predict breast cancer susceptibility and recurrences.
4.	Neural network	Medical diagnosis
5.	Artificial Neural Network	Breast cancer risk estimation
6.	Decision tree	Prediction of breast cancer
7.	Logistic regression	Breast cancer recurrence prediction
8.	Random forest	Big data analysis in spark cloud computing

Table 1: Machine learning techniques and solution

According to the related table study some machine learning algorithms are mainly involved in prediction and diagnosis.Breast cancer is most eminent disease in the area of medical detection which is rapidly increasing day by day.
These techniques are performed on Wisconsin Breast Cancer Dataset (WBCD) to prognosis the breast cancer recurrence.The oncologist can diagnose this cancer by using various modalities to detect cancer as well as its region and feature. Various modalities can be mammography, MRI, biopsy, etc. After the diagnosis, of breast cancer is confirmed then a biopsy is done which help the oncologist to detect the cancerous tissues of lymph [2].

These techniques are also used for classification purposes either the cells are benign or malignant. Early diagnosis of cancer will help the patient to take proper treatment so that to do the right things for their survival and fight against this disease. These learning techniques are used to classify tumors.This research is done for the Wisconsin dataset [3].For pre-training, we use sequential model ResNet. For testing and training, we are using WBCD (Wisconsin dataset) [4].

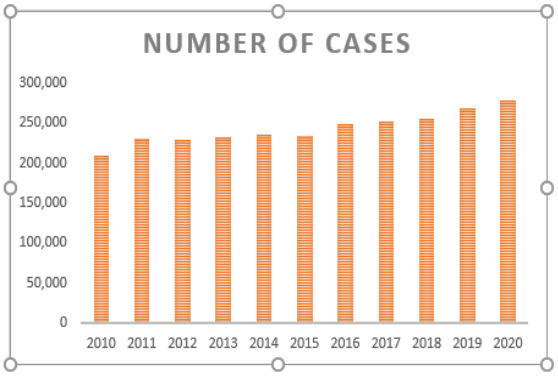
Different classifiers are designed for WBCD for classification. The main aim of this paper is to classify benign and malignant.

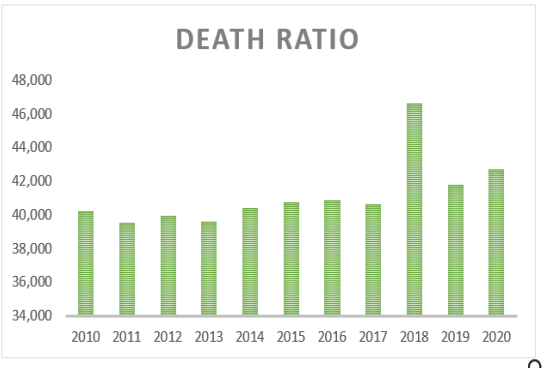
Table 2. Table shows number of new cases and death ratio from 2010 to 2020

Year	Number of cases	Death ratio
2010	209,060	40,230
2011	230,480	39,520
2012	229,060	39,920
2013	232,340	39,620

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

2014	235,030	40,430
2015	234,190	40,730
2016	249,260	40,890
2017	252,710	40,610
2018	255,350	46,640
2019	268,600	41,760
2020	279,100	42,690





These graphsreveals the overall statistic of US that shows the worst situation of breast cancer increasing year by year. Figure 1 illustrates breast cancer circumstances in US for the most recent 10 year how the graph is growing with cancer patients. Graph also illustrates number of new breast cancer cases in males and females. In 2010 the new cancer patients was 209060 and in 2020 it increases and reaches to an alarming rate i.e. 279100 number of new cancer patients are growing in number year by year at high rate. Figure 2 shows number of deaths cases of males and females those were suffering from breast cancer. As number of registered cases of breast cancer are increasing similarly death rates is also increasing with years by year. In 2010 the death that are due to this breast cancer were in number 40230 that include males as well as females as years passed number increases up to 42690.

In short, we proposed a technique to diagnose breast cancer using the Wisconsin dataset which is a collection of 569 samples and 30 features computerized from FNA of breast mass [6] by using a CNN classifier as well as KNN classifier. The main features of the proposed work are as follow:

1) We calculate the performance by using the KNN algorithm.
2) Then we calculate the performance by using CNN classifier.
3) Then we compare their performance based on features.

This Paper is managed in a certain pattern as followed. In the Next Section I.e. Related work,explanation of brief reviews related to our effortoftheanalysisofbreast cancer. Then in the next Section III, i.e. Methodology, we define our projected technique in specifics. Experimental outcomes are specifiedin SectionIVtrackedbydiscussionsinSectionV.Finally, theconclusions are accessible in SectionVI.

II. RELATED WORK

Researcher have found that the employments of machine learning for analysis of diseases by framing the examples of medical dataset.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

Kavita mittal and Gaurav aggarwal [1] have experimented two major techniques one is of supervised learning and one is of unsupervised learning KNN and k-mean respectively. Later on the performance of both the algorithms have been compared on the same dataset i.e. WBCD.
Moh'dRasoul Al-hadidi [2] perform image processing and use different machine learning algorithm. He proposed a new technique for detection of breast cancer which have two parts, in first part image processing is done to prepare the mammography images which are used for feature extraction and pattern matching for they are using two different algorithms Back Propagation Neural Network (BPNN) and Logistic Regression (LR). Their accuracy will also compare.
RikuTurkki, DmitriiByckhov [3] their findings shows that the feasibility and reliability of their learning prediction signals in tumor cells images without their exact domain knowledge. Their study shows that machine learning techniques can extract diagnostic related information for tumor patients. Prediction is done by classifying the samples on the basis of Digital risk score (DRS) groups.

Dina A. Ragab, MahaSharkas[4]they detected the masses form patients just to classify cancerous and non cancerous tissues using mammograms. Proposed Computer Aided Detection system was designed in those two segmentation techniques was included. Firstly, the ROI (Region of Interest) was decided manually from the actual image using round contour. Secondly, threshold was set by using region used method, which is equal to 76 and calculate the huge area that includes this threshold.

David A. Omondiagbe, Shanmugam Veeramani [5] they uses machine learning techniques for detection breast cancer by using the dataset WDBC (Wisconsin Diagnostic Breast Cancer). This paper has proposed a hybrid technique for breast cancer detection by decreasing the high dimensionality of features extracted by using linear discriminate analysis (LDA). This proposed method achieved an accuracy 98.82%, specificity 99.07% and curve of 0.9994.

Kandhasomy et al. [6] in this paper they looked at a wide range of machine learning called as KNN, SVM (support vector machine), decision tree, and random forest for distinguishing patients having diabetes mellitus. Their efficiency and performance was tested in terms of accuracy, specificity and sensitivity. These experiments are performed with the K = 1 for KNN approach and it is noticed that random forest and KNN provides the best performance.

Manjisha et al. [7] works on KNN classifiers for classification and k-mean for clustering of epilepsy within EEG signals and concluded that K-means performed better from KNN in different terms of performance factors.
Abien Fred M. Agarap [8] Implementing SVM (support vector machine) with Gaussian radial basis function (RBF) for classification. Proposed GRU-SVM model, cross validation technique such as f- fold validation is used. We train SVM model to predict the data and 83 features from unstructured clinical narratives and 18 features from structured clinical data to identify distant recurrences in breast cancer.

Anusha Bharat, Pooja N, R Anishka Reddy [9] For overall methodology, KNN technique has given the best results; Naive Bayes and logistic regression have also performed well in diagnosis of breast cancer. SVM is a strong technique for predictive analysis, the creation of SVM classes Like LIBSVM has taken place. Accuracy score 0.991228, accuracy of SVM drastically improves.

Richaa Negi and Rejo Mathew [10] Results have shown that SVM has achieved an optimal execution with respect to accurateness and distinctness. Whereas RF comes up with maximum likelihood of appropriately categorizing a tumor. The level of proliferation, as reflected by mitotic count, does not serve as a prognostic factor for TNBC(triple negative breast cancer). Per 2 mm2 averaged manual mitotic counts ranged from 1 to 187 (mean 37.6, SD 23.4), hereas automatic counts ranged from 1 to 269 (mean 57.6; SD 42.2).

SubhamSadhukhan, Nityasree Upadhyay and Prerana Chakraborty[11] This early breast cancer cell detection can be predicted with the help of modern machine learning techniques. The efficiency (97.489%) of this entire model is quite high. It is observed that after thresholding the nuclei are segmented. After that finding the contours and tracing them on the greyscale image, the nuclei are distinctly separated. After getting the nuclei, we finally apply the obtained feature values to the proposed model. This confirms that the affected tumor is "BENIGN," that is "non–cancerous".

Shubham Sharma, Archit Aggarwal, Tanupriya Choudhury [12] It has been observed that each of the algorithms had an accuracy of more than 94%, to determine benign tumor or malignant tumor. KNN is the most effective in detection of the breast cancer as it had the best accuracy. They have utilized the 10 fold technique that is the

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

data set segregated in ten different chunks. Nine out of the folds used in the system are used for training and the last set is used for the purposes of testing and analysis. We have utilized 398 observations for training set and 171 observations for testing out of 569 observations.

YounessKhourdifi ,Mohamed Bahaj [13] The application of data mining technologies in the medical field is very important because they certainly help in the decision-making process. we used five learning Algorithms: SVM, Random Forest, Naive Bayes, and K-NN, SVM has proven its performance on several levels in front of others, especially by the lowest error rate and shortest turnaround time. SVM takes about 0.08 s to build its model unlike K-NN which takes only 0 s. This can be explained by the fact that K-NN is a lazy learner. The accuracy obtained by SVM (97.9%) is better than that obtained by RF, Naïve Bayes and k-NN which successively have an accuracy of 96%, 92.6%, and 96.1%.

Yao Lu, Jia-Yu Li, Yu-Ting Su, An-An Liu [14] The general detection process as follows: 1. image prepossessing to remove artefacts' and variance, 2. ROI area segmentation to extract candidates for further detection, 3. feature extraction and 4. Classification Evaluation of the detection methods are commonly based on the calculation of precision and recall (also known as sensitivity, on the detected results. Some statistical measures such as F1 score, Receiver Operating Characteristics (ROC) and its Area under Curve (AUC) are also commonly used for the evaluation of detection results. Deep learning based methods have achieved good results for detection on both image modalities; especially the ScanNet has achieved comparable accuracy with human performance on histological images.

Zexian Zeng, Sasa Espino, Ankita Roy, Xiaoyu Li [15] We applied NLP to breast cancer local recurrences identification by integrating a flexible number of medical concepts, their power sets, and the additional feature of the number of pathology reports. We then applied support vector machines (SVMs) to identify local recurrences in breast cancer patients. The predictions were compared with the annotated labels, and we obtained a precision of 0.50, a recall of 0.81, an f-measure of 0.62, and an AUC score of 0.87.In the cross-validation, the AUC for our model is 0.93 with a standard deviation of 0.01.The AUC score is 0.87 for our proposed model, which outperformed the other methods and is consistent with the cross-validation result.

Zhiqiongwang,Mo li, Huaxia wang [16] Main motive is to use deep feature extracted from cnn for 2nd stage of mass detection and diagnosis. Uses US-EML proposed algorithm, in mass detection we uses ELM CLASSIFIER. Detection results analysis. Mass Detection Based on Fusion Feature.

B.M.Gayathri ,C.P.Sumathi [17] This paper outlines the review on breast cancer using different ML methods and techniques, which are utilized to improve the exactness of malignant growth in one. RVM is applied for identifying optical cancer,and for malignant growth etc.Overall, if research on RVM continues,then all things considered, the utilization of RVM will turn out to be significantly more helpful in diagnosing breast disease.

Naresh Khuriwal ; Nidhi Mishra [18] they have utilized 12 features for finding breast malignant growth that we have got after pre-handling. But before train model they have applied some pre-processing techniques like Watershed Segmentation, Color based division and Adaptive Mean Filters for scaled dataset then applied model and accomplished precision. In this paper they contrast deep learning methods and other machine learning techniques and seen our proposed framework is demonstrated best from others machine learning techniques.

Moi Hoon Yap, Gerard Pons , Joan Martí [19] This paper proposes the utilization of deep learning approaches for breast ultrasound mass detection and find out three unique strategies: a Patch-based LeNet, a U-Net, and an transfer learning approach with a pretrained FCN-AlexNet. Their exhibition is compared against four best in class mass detection method (i.e., Radial Gradient Index, Multifractal Filtering, Rule-based Region Ranking, and Deformable Part Models).

Md. Milon Islam ; Hasib Iqbal ; Md. Rezwanul Haque [20] Prediction for the cancer is done by using two algorithms the Support Vector Machine and K-Nearest Neighbors which are the supervised techniques for malignant detection via preparing its training data set. The proposed framework utilizes 10-fold cross validation to get a optimized result. This strategies accomplished the accuracy of 98.57% and 97.14% by Support Vector Machine and K-Nearest Neighbors exclusively alongside the specificity of 95.65% and 92.31% in testing stage.

III. METHODOLOGY

In this paper we are designing a method for detecting breast cancer by using a biopsy-based dataset of 569 samples which is Wisconsin dataset. Machine learning classifier a KNN are used for classification.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

The proposed technique consists of three steps: 1. Wisconsin dataset i.e. an open source breast cancer dataset is uploaded. 2. We implement KNN and classify the performance of KNN classifier. 3. Extract the features and calculate the accuracy and all other performance matrics. 4

A. Dataset

In this framework the performance is calculated by using Wisconsin dataset which is an open-source dataset available it contains 569 samples with 30 features. These features are computed based on FNA biopsy [16]. Their class scattering is 357 benign and 212 malignant. Features used are Radius_mean (mean of distances from center to points on the perimeter), texture_mean (standard deviation of gray-scale values), Perimeter_mean (size of tumor), area_mean, smoothness_mean, compactness_mean, concavity_mean, concave points_mean,symmetry, fractal dimension. The mean, standard error and "worst" or largest (mean of the three largest values) of these features were calculated for respectively image, resultant in 30 features.

B. KNN (K-Nearest Neighbor)

KNN is a non-parametric learning technique that can be implemented without using any presupposition on the database. KNN predict on the basis of the entire training dataset. This algorithm is implemented by calculating the mathematical equation Euclidean distance to have the output in the form of the nearest neighbor based on the equation.

$$D(x1, x2) = \sqrt{(x2-x1)^2+(y2-y1)^2} \quad (1)$$

The efficiency of the KNN classifier depends on the K that is selected for the implantation and the distance formula is also depends on K that is applied for calculating distance. Theoretically in the big datasets, the value of K is large that is chosen for effective and accurate classification results. As the study says if the value of K is larger than the boundaries are smaller and better the generalization. K can also be selected through a process called cross-validation [17].

IV. EXPERIMENTAL SETUP AND RESULTS

We calculate the performance of our projected work to diagnose breast cancer using the Wisconsin dataset. Evaluated parameters are Accuracy,We firstly implement a KNN technique for this dataset results are:

A. Experimental Results for KNN

The implementation is done using KNN classifier for detection and hence the achieved accuracy is 96.49% and the optimal no. of neighbors is 17. If we increase the fold, then overfitting will also increase. The value of K in this is 17 to get the best results by avoiding overfitting [20]. The humblest method to custom cross-validation is to call the cross_val_score assistant function on the estimator and the samples. In the straightforward methodology, called k-fold CV, the preparation set is divided into k smaller sets. The succeeding technique is trailed for respectively of the k "folds":

1) A prototypical is trained to expend k−1 of the folds as training data.

2) The resultant model is authenticated on the residual portion of the data (i.e., it is recycled as a test set to calculate a performance quantity such as accuracy).

Accuracy is shown in below figure

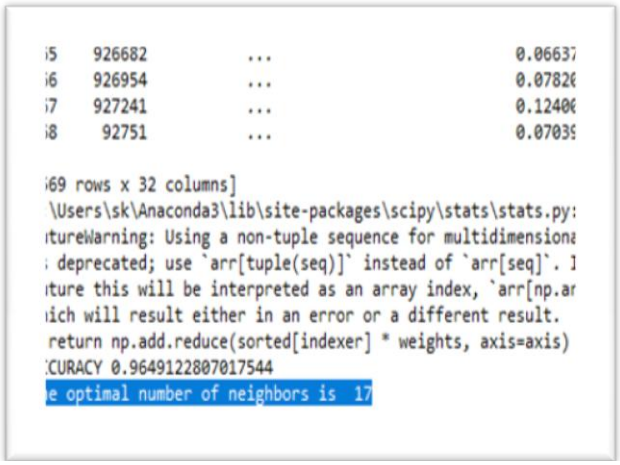


Fig. 1. Accuracy achieved After implementing KNN

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

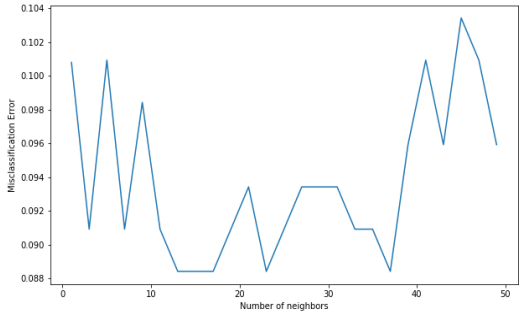


Fig. 2. Graph of number of Neighbors vs Misclassification Error

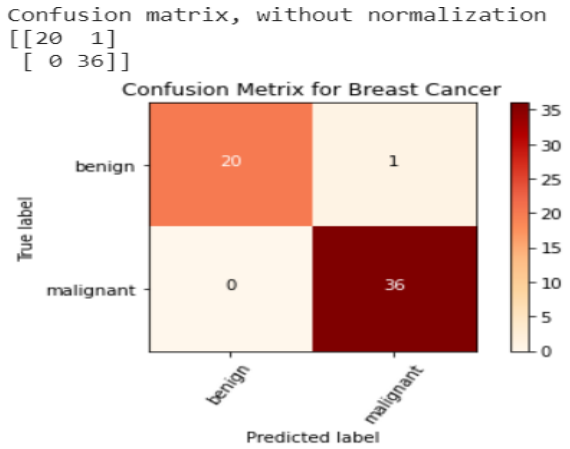


Fig. 7 Confusion Matrix of Breast cancer

The confusion matrix is used for analyzing the misclassification that is done while implementing. Every row shows the event in the prediction class and column shows the event of actual. Correctly classified classes are represented on diagonals. This matrix shows us the classes that are being misclassified and also what is being misclassified [24].

V. DISCUSSION

In this research we propose a scheme to analyze breast cancer using medical Biopsy computerized data, in this we are using K-Nearest Neighbor. Wisconsin database is used to train the prototype and to assess the performance, evaluation is done on various parameters like Accuracy,This experiment is conducted on the database WBCD (Wisconsin breast cancerdatabase) [10, 3] obtained by the university of Wisconsin. Thisdatabase contains information about the breast cancer which aretaken by the Fine Needle Aspirate (FNA) of human breast tissue.These data correspond to699clinical cases where458(65,50%)are a benign cases and241(34,50%) are malignant cases.The WBCD database contains missing data for16observations,which limited this experimentation to683clinicalcases.
After implementing we get the result as KNN accuracy is 96.49% where the optimised value of k is 17.

VI. CONCLUSION

The main objective of this paper is to help scientists that are developing CAD system for detection and prediction of breast cancer and mainly for their early treatment at stage I or stage II rather than on stage III and stage IV.In this paper, we characterized a productive and powerful structure for the determination of breast disease utilizing WBCD Dataset. Features are being separated with the assistance of K-Nearest neighbor at that point assist the model with preparing the framework with features to improve the arrangement strategy. With the assessment of proposed method, we can accomplish high accuracy from the technique and any remaining performance parameters.
This proposed method is fundamentally intended to be utilized as clinical use by oncologist specialists to effectively analyze breast malignant growth at beginning phase, which will assist with expanding the death pace of disease sufferers. Our technique can accomplish 96.49% accuracy with KNN.

In future, we will enhance the working of our proposed strategy. To improve the exactness as well the accuracy of the diagnosis process we will going to focus on additional element extraction matters from disease pictures or H&E stained pictures. To inspect the particular regions for highlight to deal with cancer.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

ABBREVIATIONS

ANN	Artificial Neural Network
SVM	Support Vector Machine
KNN	K-nearest Neighbor
DT	Decision Tree
LR	Logistic Regression
RF	Random Forest
NLP	Natural Language Processing
RBF	Radial basis function
EEG signals	Electroencephalogram signal
LDA	Linear Discriminate Analysis
ROC	ReceiverOperatingCharac-teristics
AUC	Area under Curve
BPNN	Back Propagation Neural Network
WDBC	Wisconsin Diagnostic Breast Cancer
CAD	Computer-aided Design
FNA	Fine-needle aspiration
RVM	Relevance Vector Machine

REFERENCES

1. Siegel, L.Rebecca, D. Miller Kimberly, and A. Jemal. "Cancer statistics, 2020." Wiley Online Library , Vol. 70., 2020.
2. D Kunio. "Computer-aided diagnosis in medical imaging: historical review, current status and future potential." Elsevier , Vol. 31., 2007.
3. D Kunio. "Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology." IOP Publishing Vol. 51., 2006.
4. V. Ginneken, B., Armato III, S. G., Hoop, B., Amelsvoort-van de Vorst, S., Duindam, T., Niemeijer, M., et al. " Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the ANODE09 study " Vol. 14, pp. 707-722. 2010.
5. Yassin, I. R. Nisreen, O Shaimaa, Enas M. F. El Houby, and A. Hemat. "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review." Elsevier, Vol. 156. 25-45 . 2018.
6. C.K. Katharina, et al. "Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions?" Radiological Society of North America, Vol. 211. 101-110. 1999.
7. Corsetti, Vittorio, et al. "Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: interval breast cancers at 1 year follow-up." Elsevier, Vol. 47. 1021-1026. 2011.
8. R. Negi , and M. Rejo. "Machine Learning Algorithms for Diagnosis of Breast Cancer." Springer 928-932. 2018.
9. S. Sadhu khan , U.Nityasree, and C.Prerana. "Breast Cancer Diagnosis Using Image Processing and Machine Learning." Springer, 113-127. 2020.
10. A.Alarabeyyat, A.Mohannad, et al. "Breast cancer detection using k-nearest neighbor machine learning algorithm." IEEE, 35-39. 2016.
11. Kavita Mittal, Gaurav Aggarwal, Prerna Mahajan, "Performance study of K-nearest neighbor classifier and K-means clustering for predicting the diagnostic accuracy" in Singapore, vol. 4, 2018.
12. Moh'dRasoulA Al-Hadidi, Mohammed Y Al-Gawagzeh, Bayan A Alsaaida, "Solving mammography problems of breast cancer detection using artificial neural networks and image processing techniques". India, 2012.
13. RikuTurkki,DmitriiByckhov,Mikael Lundin, "Breast cancer outcome prediction with tumour tissue imades and machine learning " in Finland, 2019.
14. Dina A. Ragab, MahaSharkas, "A Comparison between Support Vector Machine and Artificial Neural Network for Breast Cancer Detection," in uk2013.
15. David A. Omondiagbe, Shanmugam Veeramani, "Machine Learning Classification Techniques for Breast Cancer Diagnosis," *IOP conferrence.*, Australia 2019.
16. Karthikeyan Kandasamy, Jacqueline Kai, Chin Chuah, Ran Su, Peng Huang, "Prediction of drug indused nephrotoxicity and injury mechanisms with human induced pluripotent stem cell-derived cells and mechanism learning methods" in 2015.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 4440 - 4448
Received 20 January 2021; Accepted 08 February 2021.

17. Manjishaet al.,"Automated Pre-Seizure Detection for Epileptic Patients Using Machine Learning Methods" *International journal of image, Graphics and Signal Processing(IJIGSP), Jul. 2017.*
18. Abien Fred Agarap, "Application of machine learning on breast cancer" in Phu Quoc Island, vol. 4, 2018.
19. Anusha Bharat,pooja N, R Anishka Reddy.," Machine Learning algorithms for risk prediction" *International Journal of Innovative Technology and Exploring Engineering (IJITEE) Vol.8, Jul. 2018.*
20. Richaa Negi and Rejo Mathew, "Review of machine learning algorithms for diagnosis of breast cancer" *International Conference on Computer Networks, Big Data and IoT* (*ICCBI) ,2018.*
21. SubhamSadhukhan, Nityasree Upadhyay and PreranaChakraborty , "Breast Cancer Detection Using Machine Learning Algorithms", *Springer,* 2018.
22. YounessKhourdifi ,MohamedBahaj, "The hybrid machine learning model based random forest optimized by pso and predicting heart disease" , in Morocco , *International Conference on Systems Computing and Wireless Communication System (ICCWCS),* 2019.
23. Yao Lu, Jia-Yu Li, Yu-Ting Su, An-An Liu "A Review of Breast Cancer Detection in Medical Images", *IEEE 2018*.
24. Shubham Sharma, Archit Aggarwal, Tanupriya Choudhury, "Breast Cancer Detection Using Machine Learning Algorithms", 2018.
25. Zexian Zeng, Sasa Espino, Ankita Roy, Xiaoyu Li "Using natural language processing and machine learning to identify breast cancer local recurrence " in China, *ICIBM*, 2018, .
26. Zhiqiong wang, Mo li, Huaxia wang, Hanyu jiang, "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features " in China , *IEEE,* 2019.
27. B.M.Gayathri ,C.P.Sumathi, "Breast cancer diagnosis using machine learning algorithms –a survey" , Chennai,,India , 2013
28. Naresh Khuriwal ; Nidhi Mishra, "Breast Cancer Detection From Histopathological Images Using Deep Learning" , 2018
29. M. H. Yap et al., "Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks", *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 4, pp. 1218-1226, July 2018.
30. H. Song, A. Men, Z. Jiang, "Breast tumor detection using empirical mode decomposition features", *IEEE Access*.
31. M. M. Islam, H. Iqbal, M. R. Haque, M. K. Hasan, "Prediction of breast cancer using support vector machine and K-Nearest neighbors", *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, pp. 226-229, 2017.
32. Arbab Masood Ahmad, Gul Muhammad, Khan, S. Ali Mahmud, Julian F. Miller, "Breast Cancer Detection Using Cartesian Genetic Programming evolved Artificial Neural Networks", *Philadelphia Pennsylvania USA GECCO'12*, pp. 7-11, July 2012
33. Fadzil Ahmad, Nor Ashidi Mat Isa, Zakaria Hussain, Siti Noraini Sulaiman, "A genetic algorithm-based multi-objective optimization of an artificial neural network classifier for breast cancer diagnosis *Springer*, October 2013.
34. D. Bazazeh, R. Shubair, "Comparative study of machine learning algorithms for breast cancer detection and diagnosis", *2016 5th International Conference on Electronic Devices Systems and Applications (ICEDSA)*, 2016.
35. EminaAličković, Abdulhamit Subasi, "Breast cancer diagnosis using GA feature selection and Rotation Forest", *Neural Computing and Applications Springer*, April 2017.
36. A. Alzubaidi, G. Cosma, D. Brown, A. G. Pockley, "Breast Cancer Diagnosis Using a Hybrid Genetic Algorithm for Feature Selection Based on Mutual Information", *2016 International Conference on Interactive Technologies and Games (ITAG)*, 2016.
37. A. I. Pritom, M. A. R. Munshi, S. A. Sabab, S. Shihab, "Predicting breast cancer recurrence using effective classification and feature selection technique", *2016 19th International Conference on Computer and Information Technology (ICCIT)*, 2016.
38. A. Qasem et al., "Breast cancer mass localization based on machine learning", *2014 IEEE 10th International Colloquium on Signal Processing and its Applications*, 2014.
39. D. T. Saleh, A. Attia, O. Shaker, "Studying combined breast cancer biomarkers using machine learning techniques", *2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2016.
40. D. Bazazeh, R. Shubair, "Comparative study of machine learning algorithms for breast cancer detection and diagnosis", *2016 5th International Conference on Electronic Devices Systems and Applications (ICEDSA*), 2016.