

CONVOLUTION NEURAL NETWORK AND MACHINE LEARNING FOR THE PREDICTION OF BREAST CANCER: A COMPARATIVE STUDY

Suriya priyadharsini M¹ Sathiaseelan J G R²

¹Research scholar, Department of Computer Science, Bishop Heber College, Trichy, Tamilnadu, julimca.sigc@gmail.com

²Associate Professor & Head, Department of Computer Science, Bishop Heber College, Trichy, Tamilnadu, jgrsathiaseelan@gmail.com.

-----***-----

Abstract

Breast cancer has been a common occurrence in recent years. Mammograms are available in hospitals to detect breast cancer. Waiting a long time for a breast cancer diagnosis will raise the chances of the cancer spreading. As a result, a computerised breast cancer diagnosis has been created to reduce the time it takes to diagnose breast cancer and to increase the accuracy of the diagnosis. This paper summarises the results of a study on breast cancer diagnosis using a variety of machine learning algorithms and conventional approaches to increase cancer prediction accuracy. This survey will also support others in determining which approaches are most effective in diagnosing breast cancer.

Keywords. Breast Cancer, Machine Learning, Deep Learning, Convolutional Neural Network, SVM.

1.Introduction

One of the most common diseases among women is breast cancer. While the precise cause of breast cancer is unknown, several studies have shown that screening for breast cancer is a successful way to increase early detection. Clinical breast examination, imaging diagnosis, and histopathological biopsy are all standard

methods of breast cancer diagnosis. Tumor classification may be used to detect breast cancer. Tumors are classified into two categories: malignant and benign tumours. To differentiate between these tumours, doctors need a reliable diagnosis technique. In general, however, it is very difficult to distinguish tumours even by experts. As a result, tumour diagnosis requires the automation of diagnostic systems.

Many researchers have attempted to use machine learning and deep learning algorithms to detect cancer survival rate in humans, and they have shown that these algorithms are more successful at detecting cancer diagnoses. Machine learning is now in such high demand that it is being offered as a service. experts in machine learning, on the other hand, allocate a significant amount of time to improving their models. Pre-processing, segmentation, feature selection, and classification processes require a collection of skills and experience to design an effective machine learning model. The methods and parameters used in the pre-processing and classification stages are automatically selected in the proposed

method. A machine learning specialist selects the best methodology for the problem at hand.

In order to diagnose breast cancer, researchers have moved to deep learning in recent years. Deep learning is a subfield of machine learning that uses algorithms to model the human brain's structure. Deep learning can effectively identify images by extracting the high-dimensional representation of the image by convolution, down-sampling, and other techniques.

Convolutional Neural Network (CNN) is a popular deep learning model in image classification due to its quick and easy implementation, it has capacity to extract small bit of information from large amount of data, learned automatically significant features hierarchies from raw data given directly to it, it can achieve better accuracy with large datasets and many more [1].

2 Literature review

This section contains background on research methods that has already been completed. Breast cancer is detected using two different methods. Machine learning is the first, and deep learning is the second. Many studies are carried out with the aid of machine learning. However, machine learning methods have several flaws that are addressed by deep learning. This section gives the information about most popular breast breast cancer detection methods namely Naive Bayes Classifier, Support

Vector Machine Classifier, Bi-Clustering and Ada boost Techniques, Convolution Neural Network and Recurrent Neural Network.

Naive Bayes Classifiers are statistical classifiers that can predict class membership probabilities, such as the possibility that a given sample will belong to a specific case. It performs better in spam classification, medical diagnosis, and weather forecasting.

SVM is a supervised machine learning algorithm that can be used to solve classification and regression problems. It is, however, primarily used for classification purposes. The goal of SVM is to divide datasets into groups in order to find the best marginal hyperplane (MMH). SVM is suitable for small datasets.

AdaBoost is a well-known ensemble technique that improves classification accuracy by combining a number of weak classifiers. For improved generalisation performance, the bi-cluster oriented classifiers can be combined with a strong ensemble classifier. Various weights are assigned during training, and decisions are made using "weighted majority voting."

CNN is referred to as Convolutional layers in the architecture. These layers are primarily used for detecting image features. CNN learns the component of the images with each development of the part on the images. Every neuron in this layer was used to extract the image's close structure features.

RNNs are a type of Neural Network (NN) that has a high sequential dimension and has been widely used in time sequence prediction. RNNs can process data points where the activation at each step is based on the previous step.

Machine Learning application on medical image classification

R. Vijayarajeswari et al [2] explores the methods for classification and extraction of features. Hough transform is used here to identify the characteristics of the image mammograms and It is evaluated by SVM. The classification accuracy is more by the use of SVM classifier. Accuracy- 94%. Chang Shuran et al [3] presents the diagnostic and prediction model of PSO-SVM (GP-SVM) on breast cancer based on gray relational tests (GRA) for a dataset composed of traditional sign and blood test results. Gray relational analysis will perform correlation analysis on each function and target factor to obtain more efficient input, minimize classification time, and use inertia weights and learning factors After dynamic modification, the PSO algorithm optimizes the parameters of the SVM and increases the precision of the forecast. Accuracy (95.65%). Sheng Cai et al [4] suggests a microcalcification algorithm for early breast cancer mammography. Next, the contrast properties of mammograms are improved by transformation of contourlets and morphology (CTM). Second, split the ROI with the better K-means algorithm. Third, measure the Grayscale Function, the Form

Feature and the Directed Gradient Histogram (HOG) for the ROI area. The accuracy of the automated classification effects of the adaptive support vector machine (ASVM) algorithm is considered to be 94 %. The authors used the most popular BC detection methods, including the Naive Bayes Classifier, the Support Vector Machine (SVM) Classifier, the Bi-clustering and Ada boost Techniques, the R-CNN (Convolutional Neural Networks) Classifier, and the Bidirectional Recurrent Neural Networks (HA-BiRNN) [5] Deep learning does not require manual adjustment of features; rather, it learns from data sets using a general-purpose learning approach. [6] Deep learning based on Convolution Neural Network (CNN) has achieved great success in the field of biomedical image analysis in recent years, such as the detection of mitosis cells from microscopic images. V Sansya Vijayam et al. [7] proposed the model based on deep learning . The author focused on Lloyd's algorithm for clustering and CNN for classification. The 96% accuracy was achieved by the proposed methods. It used the histopathology images for the diagnosis purpose. This paper also explained the image processing and deep learning.

Table 1

Machine Learning Techniques for Breast Image Classification

Author and reference	Dataset	Method	Findings
Shubham Sharma et al. [8]	Wisconsin Breast Cancer dataset from UCI Repository	Random Forest, KNN and Naïve Bayes.	KNN was a good classifier in terms of accuracy
R. Preetha et al. [9]	wisconsin breast cancer dataset	DataMining techniques	Detect the hidden cancer associated for classification.
Muhammed et al. [10]	MIAS	Classification techniques like SVM, Ada Boost, Decision Tree, Logistic regression ,Random forest, Gradient Boost and KNN.	Study observation reflects that SVM Classifier is best with 90% accuracy.
Aswathy et al. [11]	UCSB Dataset	SVM	Automated Classification of Breast cancer image is based on SVM model on GLCM features and achieved an accuracy 91.1%

Table 2. Deep Learning based Techniques for Breast Image Classification.

Author and reference	Dataset	Method	Findings
Nasser Edinne et al. [12]	Mini-MIAS database	Four different types of Classifiers namely,SVM, ANN RF and NB	SVM achieved an average accuracy 99.5%.
Ajay kumar et al. [13]	BCDW11 and WBCD32 dataset from UCI Repository.	Classification techniques like SVM, KNN, Naïve Bayes and Decision Tree.	By using BCDW11, it gave 97.13% accuracy and using WBCD32, SVM gave 97.89% accuracy

SriHari Nallamala et al. [14]	Wisconsin Breast Cancer dataset.	Machine learning techniques	It achieved the 98.50% precision.
R.Chtihrakka nnan, P.Kavitha et al. [15]	Mammogram images.	Machine learning techniques	It achieved 96% accuracy by using DNN
Nikita Rane et al [16]	Wisconsin Breast Cancer Dataset.	Machine learning techniques	According to this, enhancement in machine learning gave better results.
Chang Shuran et al [3]	UCL	GP-SVM	After dynamic modification, the PSO algorithm optimizes the parameters of the SVM and increases the precision of the forecast. Accuracy (95.65%)
Majid Nawaz et al [25].	BreakHis Dataset is used	Deep Learning Convolution neural network	It got 95.4% accuracy when compared with state-of-art models and Dense CNN model used for this.
NareshKhuri wal et al. [17]	Mammogram MIAS database.	Deep learning	It achieved 98% accuracy by using CNN.
Weal E.Fathy et al [18]	Digital Database for Screening Mammography dataset	Deep learning	It achieved 96% area under ROC and 99.8% sensitivity and 82.1% specificity.
PanuwatMek ha et al. [19]	Breast Cancer Wisconsin dataset.	Deep Learning	The author compared the machine learning techniques and deep learning. It achieved the 96.99% accuracy with deep learning.

Author and reference	Dataset	Method	Findings
Hungleminh et al [20].	Histopathology images dataset	Deep feature fusion method	It achieved 95% accuracy in comparison of 4 cancer classes and 97.5% for differentiating two combined groups of classes
Luqman Ahmed et al [21]	MIAS and CBIS DDSM	Deep lab v3 (Xception65) and Mask RCNN (ResNet101)	we observe that the Mask RCNN method achieves classification accuracy 98%.
Nithya et al [22]	mini-MIAS (Mammogram Image Analysis Society)	convolution neural network (CNN) is used for automated classification of mammogram density as fatty, dense and glandular	They obtained an average accuracy of 98.5%. So, the proposed CAD system aids the clinicians in the classification of mammogram density .
V.Sansya Vijayam et al. [7]	histopathology images.	Deep learning techniques	focused on Lloyd's algorithm for clustering and CNN for classification. The 96% accuracy was achieved by the proposed methods. It used the histopathology images for the diagnosis purpose.
Vo et al [23]	Histopathology image	Deep Learning	This DL model For binary classification of breast histopathological images, the proposed model also had accuracy of 99.5 %

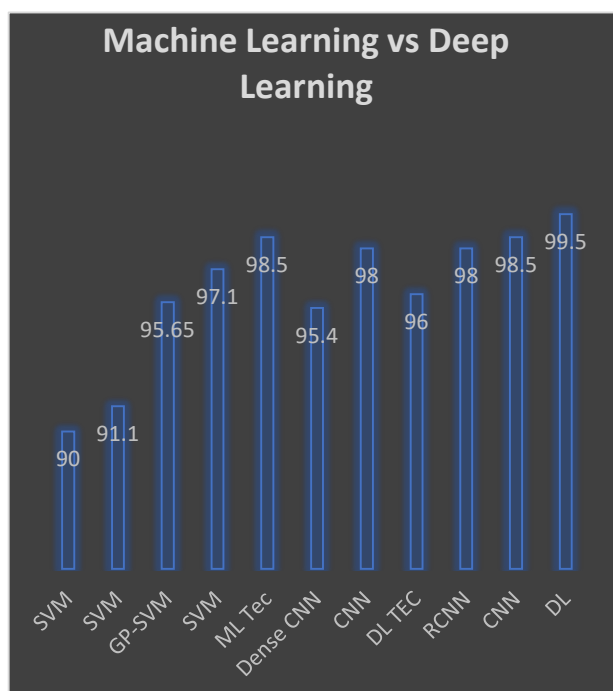


Figure 1 : A classification Accuracy of machine Learning vs Deep Learning Algorithm

Conclusion

Breast cancer detection is a difficult problem because it is the most common and dangerous disease. Breast cancer is becoming more common each year, and there is less chance of recovery. Machine learning and deep learning techniques are used to detect breast cancer. According to previous research, machine learning techniques produce better results in their respective fields. These studies used a variety of machine learning techniques, with some dataset enhancements for improved performance. However, it has been determined that machine learning produces better results on linear data. Previous research has also concluded that when the data is in the form of images, the machine

fails. An innovative technique is used to solve the problem of machine learning techniques. Deep learning is a relatively new technique that is widely used in data science. CNN, a deep learning-based technique, is used to classify breast cancer image data. CNN primarily operates on image datasets. researcher has also concluded that CNN produces superior results when compared to other methods.

References

- [1] Joshi, "Deep learning, in machine learning and artificial intelligence,," *in machine learning and artificial intelligence*, pp. 117-126, 2020.
- [2] vijarajeswari, ": Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," 2019.
- [3] c. shuran, "Breast cancer diagnosis and prediction model based on improved PSO-SVM based on gray relational analysis," *19th International symposium on distributed computing and application* , vol. bussiness engineering and science, 2020.
- [4] sheng, "Breast Microcalcification Detection Algorithm Based on Contourlet and ASVM," 2019.
- [5] anji, "breast cancer identification and diagnosis techniques," *machine*

- learning for Intelligent Decision making, 2020.
- [6] lecun, "deep learning," 2015.
- [7] V SansyaVijayam, "Deep learning based prediction of breast cancer in histopathology,," *International Journal of Engineering Research & Technology*, vol. 8, pp. 148-152, 2019.
- [8] S. Sharma, ""Breast cancer detection using Machine Learning Algorithm," *International Conference on Comutational Techniques,Electronics and Mechanical System(CTEMS)*, ., pp. 114-118, 2018.
- [10] muhammed, "Application of Machine Learning and Image Processing for detection of Breast cancer," vol. Innovation in health information , 2020.
- [11] Aswathy, "Performance Analysis of Segmentation Algorithms for the detection of Breast Cancer," *International conference on Computational Intelligence and Data science(ICCIDS)*, 2020.
- [12] N. Edinne, "A New CAD System for Breast Cancer Classification Using Discrimination Power Analysis of Wavelets Coefficients and Support Vector Machine," *Journal of Mechanics in Medicine and Biology*, 2020.
- [13] A. Kumar, "" Comparative study of classification techniques for breast cancer diagnosis,," *International Journal of Computer Science and Engineering*, , vol. 7, no. no 1, pp. 234-240, 2019.
- [14] S. H. Nallamala, "Breast cancer detection using machine learning way," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2s3, pp. . 1402-1405, , 2019.
- [15] .Chtihrakkannan, " Breast cancer detection using machine learning,," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 11, pp. 3123-3126, 2019..
- [16] N. Rane, "Breast cancer classification and prediction using machine learning,," *International Journal of Engineering Research & Technology*, , vol. 9, no. 02, pp. 576-580, 2020.
- [17] “. Naresh Khuriwal, ""Breast cancer detection from histopathological images using deep learning,," *3rd International Conference and Workshops on Recent Advances and Innovations in Engineering*, pp. 22-25, 2018.
- [18] weal, ""A deep learning approach for breast cancer mass detection""," " *International Journal of Advanced*

- Computer Science and Applications*, vol. 10, pp. 175-182,, 2019..
- [19] P. Mekha, ""Deep learning algorithms for predicting breast cancer based on tumor cells," *4th International Conference on Digital Arts, Media and Technology and 2nd ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering*, pp. 343-346, pp: 343-346, 2019..
- [20] M. M. V. T. V. L. HUNGLE LE MINH, "Deep feature fusion for breast cancer diagnosis on histopathology images," *11th International Conference on Knowledge and Systems Engineering (KSE), Da Nang, ,* pp. 1-6, 2019.
- [21] L. ahmed, "Images data practices for Semantic Segmentation of Breast Cancer using Deep Neural Network.," *Journal of Ambient Intelligence and Humanized Computing*, 2020.
- [22] nithya, "Mammogram density classification using deep convolutional neural network," 2021.
- [23] N. N. Vo DM, "Classification of breast cancer histology images using incremental boosting convolution networks.," *Inf Sci 482*, Vols. 123-138, 2019.
- [24] k. c. burcak, "A new deep convolutional neural network model for classifying breast cancer histopathological images and the hyperparameter optimisation of the proposed model," *The Journal of Supercomputing*, 2020.
- [25] majid, ""Multi-class breast cancer classification using deep learning convolution neural network.," *International Journal of Advanced Computer Science and Application*, , vol. 9, no. 6, pp. 316-322, 2018.