Research Journal of Computer Systems and Engineering
ISSN: 2230-8571, 2230-8563
Volume 03 Issue 01 - 2022 (January to June)
Page 90:94

Breast Cancer Detection by Feature Extraction and Classification Using Deep Learning Architectures

Dr. Sridhar Thota
Computer Science & Communication Engineering;VLSI and Embedded Systems
Alliance University
ORCHID ID-0000-0002-2047-4882
sridhar.t@alliance.edu.in

Dr.M.Sangeetha
Computer Science & Communication Engineering
Dr.NGP ARTS AND SCIENCE COLLEGE
ORCHID ID-0000-0001-6063-8709
sangeethaseenivasan@gmail.com

Roop Raj
Management
Education Department, Government of Haryana
ORCHID ID-0000-0001-9606-8031
Rooprajgahlot@gmail.com

Article History	***Abstract***
Received: 22 January 2022 Revised: 14 April 2022 Accepted: 19 May 2022	In the world, breast cancer is regarded as one of the leading causes of death for females between ages of 20 and 59. Machine learning is the method that is utilised in research the most frequently. There have been a lot of earlier machine learning-based studies.This research propose novel technique in breast cancer detection based on feature extraction and classification by deep learning techniques. here the input data is taken as breast cancer dataset and processed for noise removal and smoothening. In order to improve the accuracy of categorising microcalcifications as benign, malignant, or normal, textural features from the processed mammography picture have been retrieved using kernel independent component analysis.Utilizing optimization techniques, the tumour portion in the breast region is excised. Here, U-net convolutional learning (U-NetCL), which eliminates human labour, is suggested for diagnosing breast cancer. The U-NetCL framework is designed for effectively extracting features.This specifically created method recognises cancerous areas in mammography (MG) pictures and quickly categorises those areas as normal or abnormal. Keywords: Breast cancer, Microcalcifications, kernel independent component analysis, U-net convolutional learning, mammogram.

1. INTRODUCTION:

The development of the cancer tumour occurs as a result of aberrant cell growth that invades the body's surrounding tissues. The absence of a tumour in the breast is regarded as normal [1]. Tumors can be classified as benign or malignant. The non-cancerous cells that make up benign tumours can only proliferate locally and cannot invade other tissues.Cancerous cells can multiply uncontrollably, spread to various parts of body, and infiltrate nearby tissue to form malignant tumours [2]. One of the most prevalent malignancies and a serious global public health concern for women is breast cancer. With today's vast amounts of data and computing power, Deep Learning (DL) has demonstrated impressive performance in object detection and recognition [3] and natural language processing. Applying DL to the processing and interpretation of medical images has sparked attention as a result [4].

2. RELATED WORKS:

This section describes a few similar studies that used various optimization strategies to diagnose breast cancer. A summary of the most current cutting-edge DL based Computer Aid Design (CAD) systems created for mammography as well as breast histopathology pictures was provided by the authors in [5]. Similar to this, the authors of [6] have looked into the application of Local Quinary Patterns. Another article [7] discusses the drawbacks of magnetic resonance imaging (MRI) and CT, which have low sensitivity for sub-centimeter lesions due to their constrained spatial resolution.A few studies have revealed additional drawbacks of repeated mammograms performed over a ten-year period [8]. Women who have had a mammography every year for ten years have a 49.1% false positive diagnostic rate, according to a study by [9]. Breast cancer risk was shown to decrease when women were encouraged to undergo a NLS (Sentinel Lymph Node) biopsy [10]; furthermore, some authors [11] have cautioned against using breast cancer thermography data as the sole source of information for making decisions.

3. PROPOSED METHODOLOGY:

This section discussnovel technique in breast cancer detection based on feature extraction as well as classification by deep learning methods. In order to improve the accuracy of categorising microcalcifications as benign, malignant, or normal, textural features from the processed mammography picture have been retrieved using kernel independent component analysis. Here, U-net convolutional learning (U-NetCL), which eliminates human labour, is suggested for diagnosing breast cancer. Figure 1 below shows the methodology's process as it was developed.

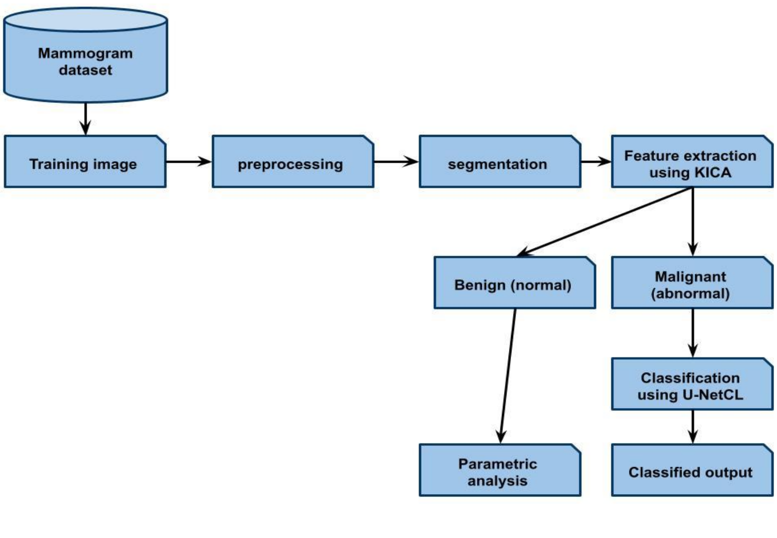

Figure1: Architecture of the Proposed technique

3.1 *Kernel independent component analysis-based feature extraction:*

We discover that V A R mm is orthogonal when the sources si are assumed to have a zero mean and unit variance. Consequently, the ICA unmixing model with whitened noise becomes in eq. (1)

$$Y = X^{\top} W$$

$$O(m) := \{X \in \mathbb{R}^{m \times m} \mid X^{\top} X = I\} \qquad (1)$$

With linked marginal measures Pru and Prv, let Pru,v be a joint measure on (U U,) (here, and are Borel algebras on U). The definition of the covariance operator Cuv: G F is given by eq. (2), (3)

$$\langle f, C_{uv}(g)\rangle_{\mathcal{F}} = \mathbb{E}[f(u)g(v)] - \mathbb{E}[f(u)]\mathbb{E}[g(v)]$$

$$\|C_{uv}\|_{\mathrm{HS}}^2 = \mathbb{E}_{u,u',v,v'}\left[\psi(u,u')\hat{\psi}(v,v')\right] + \mathbb{E}_{u,u'}[\psi(u,u')]\mathbb{E}_{v,v'}\left[\hat{\psi}(v,v')\right] - 2\mathbb{E}_{u,v}\left[\mathbb{E}_{u'}[\psi(u,u')]\mathbb{E}_{v'}[\hat{\psi}(v,v')]\right] \qquad (2)$$

$$H: O(m) \to \mathbb{R},$$

$$H(X) := \sum_{1\leq i<j\leq m}^{m} \mathbb{E}_{k,l}\left[\phi\left(x_i^{\top}\bar{w}_{kl}\right)\phi\left(x_j^{\top}\bar{w}_{kl}\right)\right] + \mathbb{E}_{k,l}\left[\phi\left(x_i^{\top}\bar{w}_{kl}\right)\right]\mathbb{E}_{k,l}\left[\phi\left(x_j^{\top}\bar{w}_{kl}\right)\right] - 2\mathbb{E}_k\left[\mathbb{E}_l\left[\phi\left(x_i^{\top}\bar{w}_{kl}\right)\right]\mathbb{E}_l\left[\phi\left(x_j^{\top}\bar{w}_{kl}\right)\right]\right] \qquad (3)$$

The difference between the kth and lth samples of the whitened observations is indicated in equation (7c), where X:= [x1,...,xm] O(m), wkl = wkwl R m, and the empirical expectation over all k are all represented. Suppose X = [x1,...,xm] O. (m) by eq. (4).

$$\mathrm{D}\, H(X) \equiv = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(H \circ \gamma_X)(\varepsilon)\bigg|_{\varepsilon=0} = \sum_{i,j=1;i\neq j}^{m} \mathbb{E}_{k,l}\left[\phi'\left(x_i^{\top}\bar{w}_{kl}\right)\xi_i^{\top}\bar{w}_{kl}\phi\left(x_j^{\top}\bar{w}_{kl}\right)\right] + \mathbb{E}_{k,l}\left[\phi'\left(x_i^{\top}\bar{w}_{kl}\right)\xi_i^{\top}\bar{w}_{kl}\right]\mathbb{E}_{k,l}\left[\phi\left(x_j^{\top}\bar{w}_{kl}\right)\right] - 2\mathbb{E}_k\left[\mathbb{E}_l\left[\phi'\left(x_i^{\top}\bar{w}_{kl}\right)\xi_i^{\top}\bar{w}_{kl}\right]\mathbb{E}_l\left[\phi\left(x_j^{\top}\bar{w}_{kl}\right)\right]\right] \qquad (4)$$

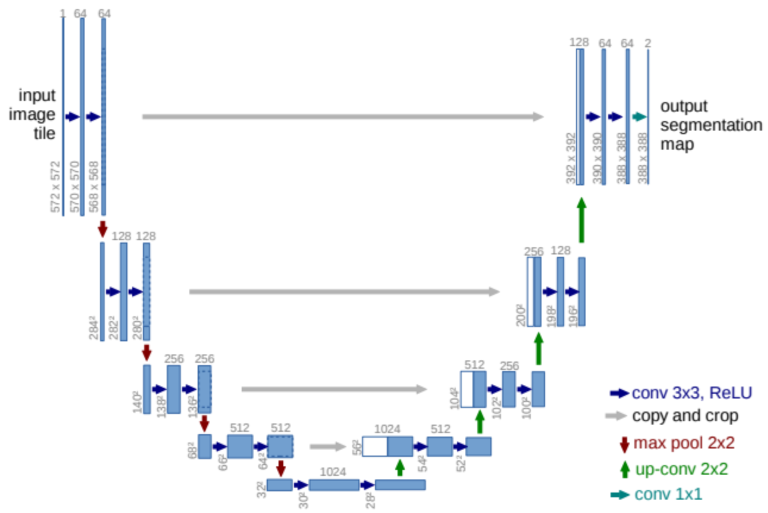

Figure 2 The U-Net-like convolutional neural network (CNN) architecture.

Convolutional (Conv2D) and deconvolutional (Deconv2D) operations present the central blocks in the CNN architecture (Figure 2) framework. An array of input data is multiplied by a two-dimensional array of weights during convolution, which entails multiplying a set of weights with input provided by RGB intensities. The kernel, a window with a specific size, slides over the input image during convolution.Finally, two-dimensional output arrays are used to hold the values that were obtained for each sliding window location. Because of this, the output two-dimensional arrays are narrower and taller than the source image. The opposite technique, deconvolution, operates very identically but enlarges the input arrays.

3.2 Performance analysis:

For research purposes in breast cancer detection and classification systems, Digital Database for Screening Mammography (DDSM) is a freely accessible online resource [24]. South Florida University compiled this data set [25]. With an average size of 3000 4800 pixels, a resolution of 42 microns, and 16 bits, it is gathered to represent genuine breast data. 2,620 digitised film mammography studies totaling 43 volumes make up the DDSM database.The four breast pictures in each case consist of two Mediolateral Oblique (MLO) views and two Cranio-Caudal (CC) views of each breast. Expert radiologists are able to identify and annotate benign and malignant masses in every mammogram.

Table-1 Comparative analysis between proposed and existing technique

Parameters	NLS	MRI_CT	BCD_DLA
Accuracy	91	95	98
Precision	78	82	85
Recall	82	85	88
F1_Score	76	78	82

From above table-1, the analysis has been carried out based on number of epochs. Accuracy calculation is done by the general prediction capability of projected DL method. For calculating F-score, number of images processed are EEG signal for both existing and proposed technique. The F-score reveals each feature ability to discriminate independently from other features. For the first feature, a score is generated, and for the second feature, a different score is obtained. However, it says nothing about how the two elements work together. Here, calculating the F-score using exploitation has determined the prediction performance. It is created by looking at the harmonic component of recall and precision. If the calculated score is 1, it is considered excellent, whereas a score of 0 indicates poor performance. The actual negative rate is not taken into consideration by F-measures.

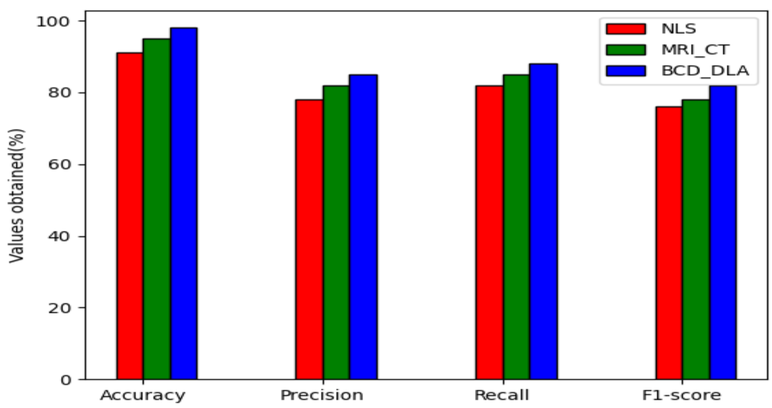

Figure 3- Comparison

The graph for F-1 score, recall, accuracy, and precision is illustrated below in figures 5, 6, 7, and 8. Comparisons of parameters between existing and suggested techniques are shown in the graphs below.

From above figure 3 shows comparative analysis between proposed and existing technique. the proposed technique attained accuracy of 96%, precision of 81%, recall of 71%, F-1 score of 65%..

4. CONCLUSION:

The goal of this research is to increase the CAD technique's accuracy in the identification of breast cancer. A framework was contributed with this goal in mind, along with the framework's flow and simulation-related parameters. The usefulness of the algorithm for classifying normal and abnormal breast photos of various individuals is examined using publicly available dataset.Weighted particle swarm optimization (WPSO) with CNN is used in this study to extract the features and assess the error between the estimated and true density using a kernel density estimation based classifier for identifying breast cancer.Performance of WPSO-CNN is amazing compared to existing techniques, according to the results. Since existing techniques used to diagnose breast cancer are offline, future work may involve developing an online breast cancer detection method.

REFERENCES:

[1] Ganggayah, MoganaDarshini, "Predicting factors for survival of breast cancer patients using machine learning techniques", *BMC medical informatics and decision making*, vol.19, no.1, pp.1-17, 2019.

[2] Hamian M, Darvishan A, Hosseinzadeh M, Lariche MJ, Ghadimi N, Nouri A,"A framework to expedite joint energyreserve payment cost minimization using a custom-designed method based on mixed integer genetic algorithm",*Engineering Applications in Artificial Intelligence*,pp.203-212, 2018.

[3] Leng H, Li X, Zhu J, Tang H, Zhang Z, Ghadimi N,"A new wind power prediction method based on ridgelet transforms, hybrid feature selection and closed-loop forecasting", *Advances in Engineering Information*,pp.20-30,, 2018.

[4] Qi X, Zhang L, Chen Y, "Automated diagnosis of breast ultrasonography images using deep neural networks",*Medical Image Analysis,pp*.185-198, 2019.

[5] Nilashia M, Ibrahima O, Ahmadic H, Shahmoradi L,"A knowledge-based system for breast cancer classification using fuzzy logic method",*Telematics and Informatics*, vol.34, no.4, pp.133-144, 2017.

[6] Jeyasingh S, Veluchamy M,"Modified bat algorithm for feature selection with the wisconsin diagnosis breast cancer (WDBC) dataset",*Asian Pacific Journal of Cancer Prevention*, vol.18, no.5, pp.1257-1264, 2017

[7] Sahu B, Nandan S, Mohanty Kumar Rout S,"A hybrid approach for breast cancer classification and diagnosis",*EAI Endorsed Transactions on Scalable Information Systems*, pp.1-8, 2019.

[8] Gao F, Wu T, Li J, Zheng B, Patel B,"SD-CNN: a shallow-deep CNN for improved breast cancer diagnosis",*Computing Medical Imaging Graph*,pp.53-62, 2018.

[9] Ting FF, Tan YJ, Sim KS,"Convolutional neural network improvement for breast cancer classification",*Expert Systems and Applications*,pp.103-115, 2019.

[10] Khan, S., Islam, N., Jan, Z., Din, I. U., & Rodrigues, J. J. C. (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recognition Letters*, *125*, 1-6.

[11] Kassani, S. H., Kassani, P. H., Wesolowski, M. J., Schneider, K. A., & Deters, R. (2019, October). Breast cancer diagnosis with transfer learning and global pooling. In *2019 International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 519-524). IEEE.