A Hybrid Deep Learning Approach for Detecting and Classifying Breast Cancer Using Mammogram Images

K. Lakshminarayanan, Y. Harold Robinson, S. Vimal, and Dongwann Kang

1 Introduction

Breast cancer causes the most common cancer type in women in the world, and this disease is dangerous whenever there is a delay in the diagnosis process. The malignant tumor expands rapidly and spreads the tumor in abnormal shapes [1]. Figure 1 demonstrates the malignant breast cancer.

Mammography is the enhanced technique for detecting the breast cancer very early that will reduce the energy for X-ray to diagnosis the breast cancer [2]. It is the way of digital X-ray image of the breast that is utilized for disease identification. The Computer-Aided Diagnosis techniques were constructed to assist the radiologists for enhancing the accuracy for breast cancer detection [3]. The feature extraction technique is used for evaluation and classification. The deep learning–based hybrid technique is constructed for breast cancer diagnosis system.

K. Lakshminarayanan
Electronics and Communication Engineering, SCAD College of Engineering and Technology, Tirunelveli, India

Y. H. Robinson
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India

S. Vimal
Department of Computer Science and Engineering, Ramco Institute of Technology, Rajapalayam, Tamil Nadu, India

D. Kang (✉)
Department of Computer Science and Engineering, Seoul National University of Science and Technology, Seoul, South Korea
e-mail: dongwann@seoultech.ac.kr

H. R. Arabnia et al. (eds.), *Advances in Artificial Intelligence and Applied Cognitive Computing*, Transactions on Computational Science and Computational Intelligence, https://doi.org/10.1007/978-3-030-70296-0_64

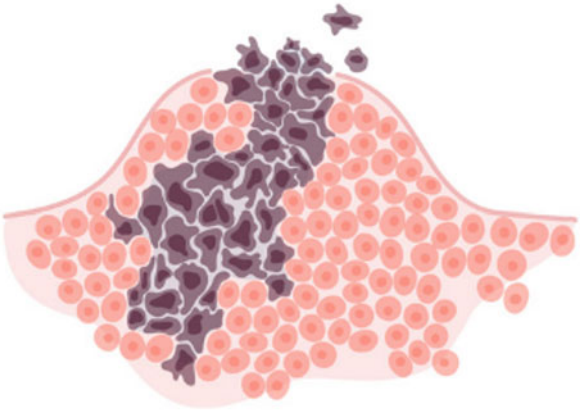

Fig. 1 Malignant breast cancer

2 Related Work

The classification of breast cancer from the input image has been completed using the k-means clustering technique that the feature extraction to apply the several parameters [4]. The Extreme Learning Machine Classifier [5] utilizes the classification by obtaining the specificity and sensitivity of the enhanced performance. The lesion annotations are utilized to implement the initial training time by performing the classification process [6]. The linear kernel function with the threshold value has provided the region-related segmentation process [7]. The machine learning algorithms have been utilized for providing the classification process from the image dataset to minimize the dimensionality feature for performance evaluation [8]. The uncertainty in the process of diagnosis has been eliminated with the median filtering and the histogram procedure [9]. The Feed Forward Neural Networks have been trained using the classifier to achieve the better sensitivity and specificity of the input image [10]. The MLP-based procedure [11] has been utilized to achieve the better breast cancer diagnosis. The principle component analysis [12] has been processed to identify the breast cancer from the dataset. The k-nearest algorithm [13] has implemented the classification procedure to diagnose the breast cancer from the ultrasound images. The efficient machine learning technique [14] and the enhanced supply vector machine [15] have been used for diagnosing the breast cancer more accurately.

3 Proposed Work

The novel hybrid technique proposed in this chapter joins the deep learning technique with the random forest classifier. The input image is preprocessed to eliminate the unwanted noise from the image and ROI is extracted with AlexNet.

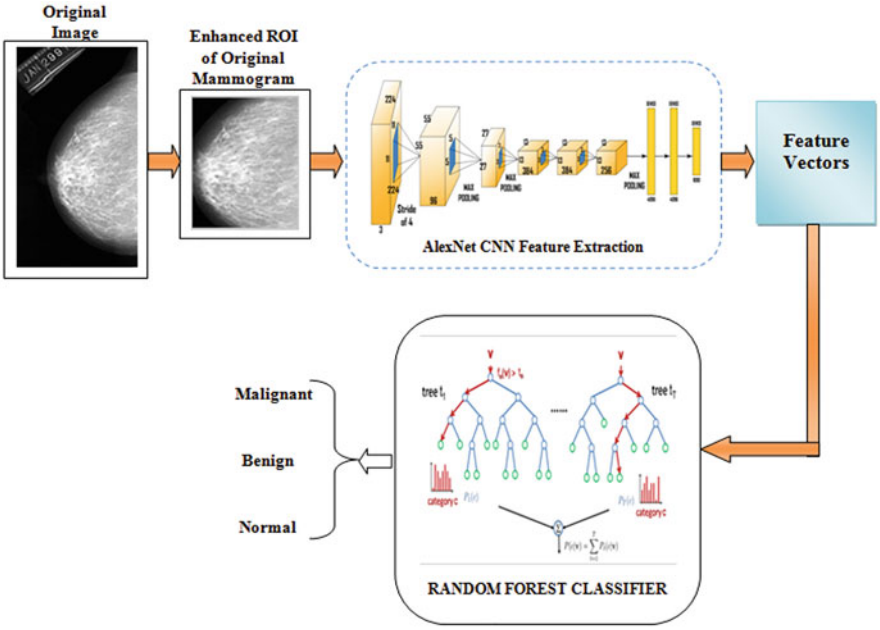

Fig. 2 Proposed system architecture

The feature vectors are generated for implementing the classification process, the random forest classifier is classifying the image and identifies the breast cancer type, and the whole process is demonstrated in Fig. 2.

The AlexNet CNN has been pretrained by the dataset to complete the classification process. The classification is more complicated for the medical images that the transfer learning procedure is used to provide the better solution for the deep learning–based issues. The dense part is eliminated from the image using the AlexNet and the random forest classifier to require the training process. The fully connected network is being trained according to the newly included layers. The final fully connected layer is eliminated from the CNN and to be replaced by the random forest classifier. The complexity has been minimized by the reduced amount of time and produces the enhanced accuracy. The AlexNet CNN architecture for feature extraction is demonstrated in Fig. 3. Here, the random forest classifier is included after the completion of the Max pooling layers in the convolutional neural network.

The ROI extraction from the input image after performing the preprocessing is illustrated in Fig. 4. The adaptive histogram equalization is generated to be capable of increasing the contrast value from the image for implementing the feature extraction procedure.

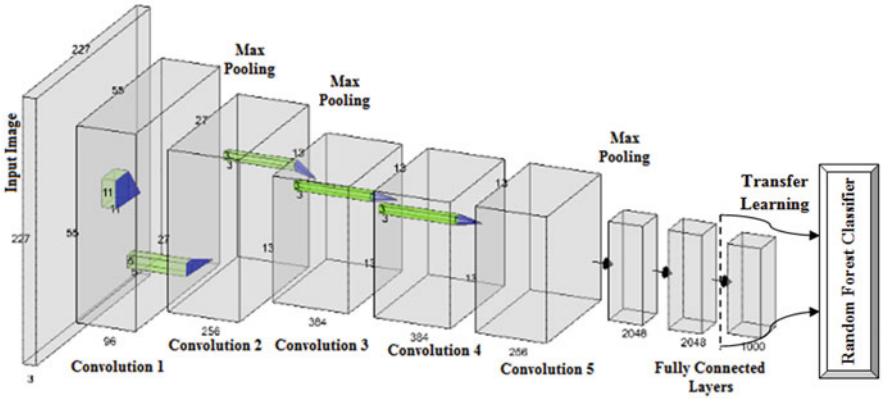

Fig. 3 AlexNet CNN architecture

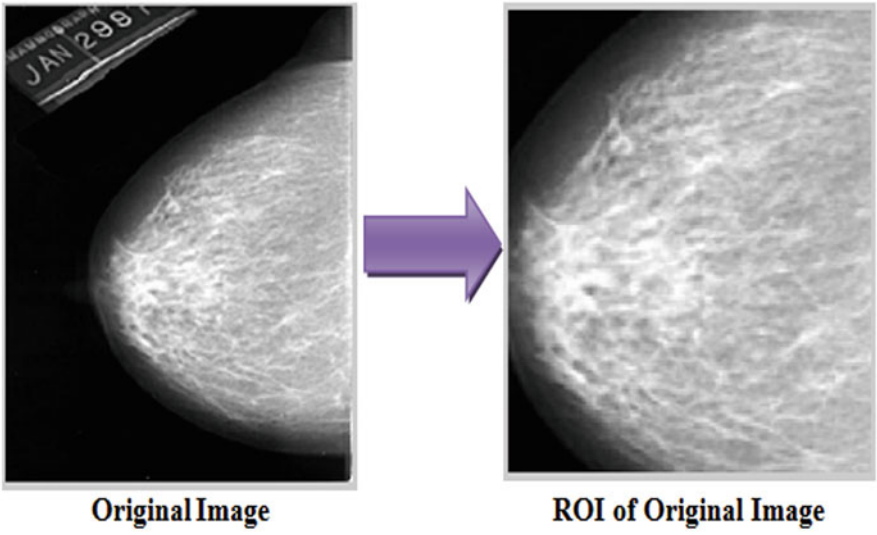

Fig. 4 ROI extraction

The random forest classifier is the supervisor learning–based technique to classify the images with the CNN-based classification process. The ensemble technique has been combined with the Bootstrap Aggregation mechanism to provide the sample replacement. The bagging technique is utilized to minimize the variance without reducing the bias value.

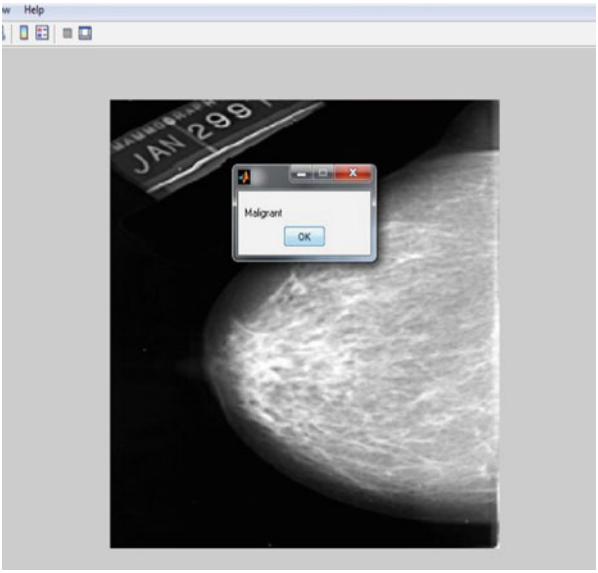

Fig. 5 Image classified as malignant

4 Performance Evaluation

The hybrid CNN–based random forest classifier is implemented in MATLAB to validate the dataset of large amount of images. The ROI as the input deep learning–based features are extracted with the AlexNet and the classification is completed using the random forest classifier to segregate the input images as the breast cancer types. The classification output is demonstrated in Fig. 5 as the input image has the malignant type of breast cancer. The confusion matrix is evaluated to specify the breast cancer classification that provides the details about the types of breast cancer in the actual stage and also the prediction stage.

5 Conclusion

In this chapter, the proposed hybrid deep learning technique has analyzed the root cause for the breast cancer of the women according to the mutations and also aging conditions. The proposed technique has classified the mammogram images and produced the efficiency from the pretrained CNN for feature extraction through the connected layer using the random forest classifier. This method has achieved the better accuracy compared with the related techniques. The experimental results show that the proposed technique providing the best solution for early detection of breast cancer.

Acknowledgments This work has supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2019R1F1A1064205).

References

1. S. Khan, N. Islam, Z. Jan, I.U. Din, J.J. Rodrigues, A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognition Letters **125**, 1–6 (2019)
2. A.M. Solanke, R. Manjunath, D.V. Jadhav, A novel approach for classification of mammograms using longest line detection algorithm and decision tree classifier. Int. J. Innov. Technol. Explor. Eng. **8**(8) (2019)
3. A.S. Eltrass, M.S. Salama, Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitised mammograms. IET Image Process. **14**(3), 495–505 (2020)
4. A. Gopi Kannan, T. Arul Raj, R. Balasubramanian, Comparative analysis of classification techniques in data mining on breast cancer dataset. Int. J. Inf. Comput. Sci. **9**(12) (2019)
5. C. Hemasundara Rao, P.V. Naganjaneyulu, K. Satyaprasad, Automated detection, segmentation and classification using deep learning methods for mammograms-a review. Int. J. Pure Appl. Math. **119**(16), 5209–5249 (2018)
6. H.D. Cheng, S. Juan, J. Wen, G. Yanhui, Z. Ling, Automated breast cancer detection and classification using ultrasound images: A survey. Pattern Recogn. **43**, 299–317 (2010). https://doi.org/10.1016/j.patcog.2009.05.012
7. D. Nagthane and A. Rajurkar, "A novel approach for mammogram classification," 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, pp. 2613–2617
8. T.G. Debelee, A. Gebreselasie, F. Schwenker, M. Amirian, D. Yohannes, Classification of mammograms using texture and CNN based extracted features. J. Biomimetics Biomater. Biomed. Eng. **42**, 79–97 (2019)
9. H. Dhahri, E. Al Maghayreh, A. Mahmood, W. Elkilani, M. Faisal Nagi, Automated breast cancer diagnosis based on machine learning algorithms. J. Healthc. Eng. **2019**, 4253641 (2019). https://doi.org/10.1155/2019/4253641
10. E. Karthikeyan, S. Venkatakrishnan, Breast cancer classification using SVM classifier. Int. J. Recent Technol. Eng. **8**(4) (2019)., ISSN: 2277–3878
11. F.J.M. Shamrat, M.A. Raihan, A.S. Rahman, I. Mahmud, R. Akter, An analysis on breast disease prediction using machine learning approaches. Int. J. Sci. Technol. Res. **9**(02) (2020)
12. S. Gardezi, A. Elazab, B. Lei, T. Wang, Breast cancer detection and diagnosis using mammographic data: Systematic review. J. Med. Internet Res. **21**(7), e14464 (2019). https://doi.org/10.2196/14464
13. P. Giri, K. Saravanakumar, Breast cancer detection using image processing techniques. Orient. J. Comp. Sci. Technol **10**(2)
14. R. Guzmán-Cabrera, J.R. Guzmán-Sepúlveda, M. Torres-Cisneros, D.A. May-Arrioja, J. Ruiz-Pinales, O.G. Ibarra-Manzano, G. Aviña-Cervantes, A.G. Parada, Digital image processing technique for breast cancer detection. Int. J. Thermophys. **34**, 1519 (2013)
15. H. Lu, E. Loh, S. Huang, The classification of mammogram using convolutional neural network with specific image preprocessing for breast cancer detection, in *2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, (Chengdu, China, 2019), pp. 9–12