International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

MRI based Breast Cancer Segmentation and Classification using Machine Learning Techniques

Nagaraja Rao P.P., D. Venkatasekhar, P. Venkatramana, Chowdavarapu Usha Rani

Abstract*: Presently, the death rate of breast cancer among women is in dangerous proposition in both developing and developed countries. This threat is addressed by the effective detection of breast cancer in earlier stages. Henceforth, the early detection of breast cancer enhances the probability of cure and survival rate. So, it is vital to develop an automated system for detecting the breast cancer in earlier stages. Magnetic Resonance Imaging (MRI) is the regularly utilized diagnosis tool for detecting and classifying the normalities and abnormalities of breast. This paper analysis the previous research carried-out in breast cancer detection and also explores the issues faced by the researchers in existing works. In addition, this paper assists the researchers for attaining better solution to the current problems faced in breast cancer detection.

Index Terms*: Breast cancer detection, machine learning technique, magnetic resonance imaging, mammogram segmentation and classification.

I. INTRODUCTION

In recent decades, breast cancer is considered as the primary cause of death among women and more than 9% of the women suffers from breast cancer during their lifetime [1-2]. Early detection of breast cancer is one of the key factor for reducing the death rate (>40%), since the cause of breast cancer still remains unknown [3]. Though, the detection of breast cancer at early stage needs a reliable and precise diagnosis for distinguishing malignant and benign cancers [4]. Henceforth, neuroimaging is the most hopeful area of research, where multiple breast imaging methods are used for determining the normal and abnormal portions. The MRI, computed tomography, X-rays, positron-emission tomography, ultrasound scans are widely utilized in neuroimaging procedures. Among these techniques, MRI received more attention among the researchers, because it is very effective in diagnosing minor abnormalities and the resultant image has high spatial resolution. Recently, several techniques are developed for breast cancer detection such as artificial neural networks [5 - 6], Convolutional Neural Networks (CNN) [7- 8], Support Vector Machine (SVM) [9-10] etc. These methods are classified based on the supervised learning with pixel-wise segmentation labels and also the feature descriptors are data-driven or hand-designed manner. In this paper, a detailed research on MRI based breast cancer detections accomplished for analysing the performance and problems faced by the researchers in prior works. This paper helps the researchers to do further research work in breast cancer detection. The paper is arranged as follows, Section II represents the process of breast cancer detection. Section III presents a few recent papers on MRI breast cancer detection topic. Summary is done in the section IV.

Revised Manuscript Received on December 30, 2019.
* Correspondence Author
Nagaraja Rao P.P.*, Department of Electronics & Communication Engineering, Annamalai University, Annamalainagar, India. Email: Nagaraj9s@gmail.com
Dr. D. Venkatasekhar, Department of Information Technology, Annamalai University, Annamalainagar, India. Email: ramavenkatasekar@yahoo.co.in
Dr. P. Venkatramana, Department of Electronics & Communication, Sree Vidyanikethan Engineering College, Tirupati, India. Email: pvramana2604@gmail.com
Chowdavarapu Usha Rani, Department of Computer Science & Engineering, Annamachaya Institute of Technology & Sciences, Tirupati, India. Email: Ch.usha86@gmail.com

II. PROCESS OF BREAST CANCER DETECTION

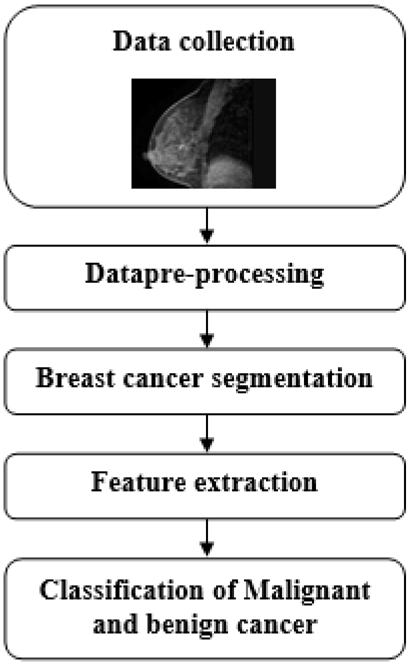

Fig. 1. The workflow of breast cancer detection

In recent periods, the MRI based breast cancer detection is an emerging clinical practice to the early diagnosis of breast cancer. According to American cancer society, in 2019 around 90000+ women died from the breast cancer.

So, the early diagnosis of breast cancer is a vital mechanism to detect and evaluate several cancerous mammograms automatically. Generally, the breast cancer detection includes five major phases such as data collection, pre-processing, segmentation, feature extraction, and classification. The process of breast cancer detection is graphically illustrated in Fig. 1.

A. Data collection

Presently, there are several databases available for breast cancer detection such as QIN Breast DCE-MRI, RIDER Breast MRI, Breast-MRI-NACT-Pilot, etc. These are the datasets majorly used in breast cancer detection because these datasets include ground truth image for each breast image that helps to compare the performance of the planned system with the prevailing systems. While utilizing real-time images, at first need to create ground truth images (manually by physician experts) that consumes more time and it is created for only a limited set of images. To address these difficulties, online available datasets are preferred highly. The sample MRI breast image is graphically depicted in Fig. 2.

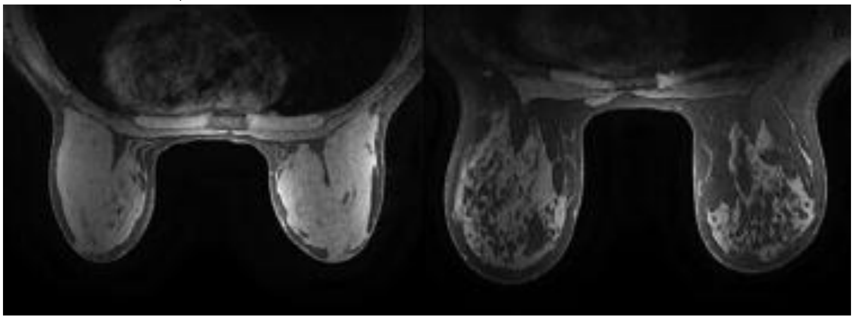

Fig. 2. Sample MRI breast images

B. Data pre-processing

After collecting the images, an important step in breast cancer detection is image pre-processing. Pre-processing is used for enhancing the input breast images or reducing the noise that acquired in the images. There are numerous techniques used for image de-noising such as, median filter, contrast-limited adaptive histogram equalization, wiener filter, etc.

Most of those mammogram images were captured or collected using imaging instruments like MRI, the captured images mainly comprised of two different noises namely impulse noise & the machinery noise (Gaussian noise and adaptive white Gaussian noise). Henceforth, the filtering methods are very effective in removing the above noises, and it also helps very significantly to enhance the quality of the image.

C. Breast cancer segmentation

After data pre-processing, segmentation of the image is being carried-out to segment the normal and cancer regions from the denoised image. Image segmentation is defined as the mechanism of categorizing the medical images into dissimilar regions that are homogeneous concerning some image features. Image segmentation aims to extract and recognize the regions that constitute a medical image. In present times, several segmentation methods are developed, which are explained below.

1) Thresholding based segmentation

Thresholding based segmentation is one of the simplest image segmentation technique, which works based on the clip level or threshold value to convert the grayscale image into a binary image. The main idea behind this segmentation technique is to choose a single threshold value or multiple threshold values. In recent times, several methods are used in thresholding based segmentation that includes maximum entropy method, Otsu's method (maximum variance), etc.

2) Region-based segmentation

The region-based segmentation groups the medical image regions on the basis of image characteristics, which includes original image value, unique texture or patterns of the region, spectral profile of multi-dimensional image data. For instance, region growing.

3) Model-based segmentation

The core idea of model-based segmentation methods is to find the interested structures, which have a propensity towards a particular shape. While segmenting a medical image, the following constraints need to be imposed such as, statistical inference between the image and the model, probabilistic variation of the registered samples, and the registration of the training examples to a common pose. Active appearance and active shape models are other important methodologies in model-based segmentation.

4) Feature-based segmentation

Feature-based segmentation finds the image displacements and features (for instance, image corners, edges etc.) that tracks the object for each medical images. Usually, the feature-based segmentation comprises of two stages. At first, the features are extracted for two or more successive images and then the extracted features are related between the images. In simplest case, two medical images are utilized and two set of features are related to a single-set of motion vectors.

5) Clustering-based segmentation

Clustering-based segmentation is an exceedingly one amongst the unsupervised learning issues that deals with finding a structure in an unlabelled data collection.

Usually, clustering methods are classified into two types such as supervised clustering and un-supervised clustering. The supervised clustering methods requires human intervention for selecting the clustering criteria and the unsupervised clustering methods selects the clustering criteria automatically. The best example of supervised clustering is relevance feedback method. The density-based clustering technique is considered as one of the best examples of unsupervised clustering. In current era, a lot of clustering-based methods are available such as, log-based clustering, retrieval dictionary-based clustering, hierarchical clustering etc. In that, fuzzy c means and k-means are efficient and well-known algorithms in clustering large data.

D. Feature extraction

After segmentation, the segmented mammogram images were carried out for feature extraction. The feature extraction maps the image pixels into feature subsets. Hence, the features are determined to quantify the properties of segmented mammogram images. General feature types are detailed below.

1) *Global features*

Global features consist of contour representations, texture features, local features, and shape descriptors that indicates the texture in an image patch. Shape matrices, histogram oriented gradients are a few examples of global descriptors. Besides, local binary pattern, scale-invariant feature transform, speeded up robust features are a few examples of local descriptors.

2) *Local features*

Features that are determined from the outcomes of sub-division of an image band are named as local image features. For example, texture, colour and shape of a mammogram image. The local features were extracted from the segmented images, whereas the global image features mainly depends on the local image features. At last, classification is accomplished to categorise the normal and abnormal images of the breast.

E. Classification

Classification is defined as the process of sorting objects into separate classes in images, which plays an essential role in medical imaging, particularly in breast cancer detection. Generally, the classification methodologies are sub-grouped into three types such as supervised learning, unsupervised learning, and semi-supervised learning classifiers. The detailed explanation about the classification methodologies is detailed below.

1) *Supervised learning*

It identifies only the known patterns in the collected data along with pre-existing the labels. The supervised learning methods explores the training data and develops an inferred function that is utilized to map new examples. SVM, decision tree, are some examples of supervised learning.

2) *Unsupervised learning*

It effectively finds the unknown patterns in the collected data without pre-existing the labels. In addition, unsupervised learning is named as self-organization, which allows modelling probability densities of given data. Unsupervised learning is one of the four main groups of machine learning, along with reinforcement, supervised, and semi-supervised learning. Self-organizing map, CNN, are some examples of unsupervised learning.

3) *Semi-supervised learning*

The semi-supervised learning method falls between supervised learning methods (where the training of data is completely marked or labelled) and the unsupervised learning method (without the training data marked or labelled). In this scenario, a few collected data are labelled but most of the data are unlabelled. Here, a combination of both unsupervised and supervised methodologies are applied for breast cancer detection. Heuristic models, Generative methods, are some common examples of semi-supervised learning.

III. LITERATURE WORKS

Numerous methodologies were presented by different researchers in breast cancer detection. In this portion, a description of some important contributions to the existing works of literature is presented.

Author	Method	Advantage	Dataset	Limitation	Performance measure
L. Zhang et al. [11]	Deep and transfer learning models (Unet and Segnet) for segmentation	In this literature paper, Unet and Segnet algorithms are developed for normal and abnormal breast cancer segmentation in MRI. The independent and internal tests shows that the developed algorithms achieved promising segmentation on DCE-MRI data.	Breast DCE-MRI	Two major problems in Unet model are, (i) computationally expensive and (ii) sensitive to noise.	Dice coefficient

A.Q. Al-Faris et al. [13]	Modified automatic seeded region growing	A new region based segmentation algorithm is developed in this research work for effective segmentation of normal and cancer regions. The statistical analysis and evaluation results showed improved performance of developed algorithm.	RIDER Breast MRI	Region based segmentation includes three major drawbacks (i) Deciding stopping criteria for image segmentation is very difficult, (ii) Computational time and memory is quiet expensive in nature, and (iii) selection of noisy seed mostly leads to weak segmentation.	Sensitivity, Specificity, false positive fraction, true negative fraction and sum of true volume fraction
A. Gubern-Mérida et al. [14]	Atlas and expectation maximization based metho	In this research paper, a fully automatic algorithm is developed for detecting the breast cancer in DCE-MRI. The developed algorithm effectively supports the clinicians during the analysis of DCE-MRI data by habitually prompting the suspicious areas.	Breast DCE-MRI	The atlas based segmentation methods lack the flexibility in locally tuning the segmentation boundary that mostly leads to flawed segmentation.	Sensitivity
Q. Yu et al. [15]	SVM and random forest	The developed research work effectively identifies appropriate classifiers and also analysis the applicability of computer aided diagnosis to relate MR-CAD with MRI for classifying the malignant region from the benign region.	Breast DCE-MRI	The classifiers like random forest is only suitable for structural medical datasets in light of vertical splitting. It does not supports unstructured medical datasets.	Accuracy, sensitivity and sensibility
R. Ha et al. [12]	CNN	In this research work, a new neural network method is developed for breast cancer diagnosis. This method performed comparatively well on DCE-MRI dataset related to the existing state of the art methods.	Breast DCE-MRI	The learning performance of CNN is affected by the number of images in the training set, while the number of images are low that highly affects the detection of breast cancer.	True positive rate and false positive rate
R. Rasti et al. [16]					Positive predicted value, sensitivity, negative predicted value, accuracy, and specificity
A.H. Yurttakal et al. [17]	Deep-CNN (DCNN)	In this work, a new user independent DCNN method is developed for characterizing the lesions as benign or malignant cancers utilizing MRI images. The developed methods is utilized in clinical practice for preventing the needless biopsies.			Accuracy and training loss

IV. CONCLUSION

Currently, breast cancer is the emerging causes of death among women, which is developed from the uncontrolled growth of cells. Though, several methodologies are developed by the researchers for the early diagnosis of breast cancer. This paper explores about the process of breast cancer detection and also examines the existing research works in light of benefit, limitation, dataset and performance measures. From the analysis, it is identified that still more work need to be carried-out in the breast cancer detection topic as shown in the section III in order to attain better outcome. This paper motivates the researchers to do more meaningful works in the future.

REFERENCES

1. R. Vijayarajeswari, P. Parthasarathy, S.Vivekanandan, and A.A. Basha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," Measurement, vol.146, pp. 800-805, 2019.
2. H. Li, S. Zhuang, D.A. Li, J.Zhao, and Y. Ma, "Benign and malignant classification of mammogram images based on deep learning", Biomedical Signal Processing and Control, vol. 51, pp. 347-354, 2019.
3. V.K. Singh, H.A. Rashwan, S. Romani, F.Akram, N. Pandey, M.M.K. Sarker, A. Saleh, M. Arenas, M.Arquez, D.Puig, and J. Torrents-Barrena, "Breast Tumor Segmentation and Shape Classification in Mammograms using Generative Adversarial and Convolutional Neural Network", Expert Systems with Applications, vol. 139, pp. 112855, 2019.
4. Y. Zhang, J.H. Chen, K.T.Chang, V.Y. Park, M.J. Kim, S. Chan, P. Chang, D. Chow, A. Luk, T.Kwong, and M.Y. Su, "Automatic Breast and Fibroglandular Tissue Segmentation in Breast MRI Using Deep Learning by a Fully-Convolutional Residual Neural Network U-Net," Academic radiology, vol. 26, no. 11, pp. 1526-1535, 2019.
5. S. Beura, B. Majhi, and R. Dash, "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer," Neurocomputing, vol. 154, pp. 1-14, 2015.
6. C. Muramatsu, T. Hara, T. Endo, and H. Fujita, "Breast mass classification on mammograms using radial local ternary patterns," Computers in Biology and Medicine, vol. 72, pp. 43-53, 2016.
7. S. Khan, N. Islam, Z.Jan, I.U. Din, and J.J.C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," Pattern Recognition Letters, vol. 125, pp. 1-6, 2019.
8. X. Zhang, Y. Zhang, E.Y. Han, N.Jacobs, Q.Han, X.Wang, and J. Liu, "Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks," IEEE transactions on nanobioscience, vol. 17, no. 3, pp. 237-242, 2018.
9. F.S.S.de Oliveira, A.O. de Carvalho Filho, A.C.Silva, A.C. de Paiva, and M.Gattass, "Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM," Computers in biology and medicine, vol. 57, pp. 42-53, 2015.
10. T.N. Cruz, T.M. Cruz, and W.P. Santos, "Detection and classification of lesions in mammographies using neural networks and morphological wavelets," IEEE Latin America Transactions, vol. 16, no. 3, pp. 926-932, 2018.
11. L. Zhang, R. Chai, A.A. Mohamed, B. Zheng, Z. Luo, and S. Wu, "Automated deep-learning method for whole-breast segmentation in diffusion-weighted breast MRI," In Medical Imaging 2019: Computer-Aided Diagnosis, International Society for Optics and Photonics, vol.10950, pp.109502R, 2019.

12. R. Ha, P. Chang, S. Mutasa, J. Karcich, S. Goodman, E. Blum, K. Kalinsky, M.Z. Liu, and S. Jambawalikar, "Convolutional neural network using a breast MRI tumor dataset can predict Oncotype Dx recurrence score," Journal of Magnetic Resonance Imaging, vol. 49, no. 2, pp. 518-524, 2019.
13. A.Q. Al-Faris, U.K. Ngah, N.A.M. Isa, and I.L. Shuaib, "Computer-aided segmentation system for breast MRI tumour using modified automatic seeded region growing (BMRI-MASRG)," Journal of digital imaging, vol. 27, no. 1, pp. 133-144, 2014.
14. A. Gubern-Mérida, R. Martí, J. Melendez, J.L. Hauth, R.M. Mann, N. Karssemeijer, and B. Platel, "Automated localization of breast cancer in DCE-MRI," Medical image analysis, vol. 20, no. 1, pp. 265-274, 2015.
15. Q. Yu, K. Huang, Y. Zhu, X. Chen, and W. Meng, "Preliminary results of computer-aided diagnosis for magnetic resonance imaging of solid breast lesions," Breast cancer research and treatment, vol. 1777, no. 2, pp.1-8, 2019.
16. R. Rasti, M. Teshnehlab, and S.L. Phung, "Breast cancer diagnosis in DCE-MRI using mixture ensemble of convolutional neural networks," Pattern Recognition, vol. 72, pp. 381-390, 2017.
17. A.H. Yurttakal, H. Erbay, T. İkizceli, and S. Karaçavuş, "Detection of breast cancer via deep convolution neural networks using MRI images," Multimedia Tools and Applications, pp.1-19, 2019.

AUTHORS PROFILE

Mr. Nagaraja Rao P.P., Research Scholar at Annamalai University, In Department of Electronics & Communication Engineering pursuing research (Ph.D) in medical image processing. He received B.Tech degree in Electronics & Communication Engineering from Priyadarshini College of Engineering, JNT University, Hyderabad in 2006 and Completed M.Tech in VLSI Design, from Annamachaya Institute of Technology & Sciences, JNT University, Hyderabad in 2008. His areas of interest are VLSI, Image Processing, Computer Vision, Medical image processing & Embedded Systems. He is working as an Assistant Professor at Sri Venkatesa Perumal College of Engineering & Technology in Electronics & Communication Engineering department. He has 11 years of teaching experience and guided around 9 PG projects in his area of interest. He is a life Member of ISTE & IAEng.

Prof. D. Venkatasekhar has completed his B.Tech degree in Electronics and Communication Engg, from JNTU COLLEGE OF ENGINEERING, Kakinada, in 1987, M.S (Engg) by Research in Information Technology from Annamalai University, in 2002, and Ph.D in Computer science and Engineering in 2015 from Annamalai University. He has Teaching Experience of 30+ years. His field of interest includes among others Image processing, Computer Networks, and Data Communication. Currently he is guiding 7 Ph.D Research Scholars..

Dr. P. Venkat Ramana received his B.Tech. from Madras University, Chennai, M.Tech. and Ph.D. from Sri Venkateswara University, Tirupati. He is currently Professor & Head, Department of Electronics and Communication Engineering at Sree Vidyanikethan Engineering College, Tirupati. He has 25 years of experience in teaching and research. He is a life member of ISTE, fellow of IETE. He has published more than 35 papers in various National and International papers/conferences and guided 24 PG students. His current interests include Signal Processing, Cognitive Radio, and wireless communication systems.

Mrs. Chowdavarapu Usha Rani Received B.Tech degree in Computer Science and Engineering from Annamachaya Institute of Technology & Sciences, JNT University, Hyderabad in 2006 and Completed M.Tech, Computer Science and Engineering, from Madanapalle Institute of Technology & Science, JNT University, Anathapuramu, in 2010. Her interested areas are Image Processing, Computer Vision, Data Mining, Software Engineering, and Software Architecture. she is worked as an Assistant Professor at Annamacharya Institute of Technology and Sciences, Tirupati.