



IMAGE PROCESSING, SUPPORT VECTOR MECHANISM, AND GENETIC ALGORITHM A HYBRID TECHNIQUE TO DETECT CANCEROUS GROWTH IN DIGITAL MAMMOGRAMS

Priyadarshini Chatterjee¹, Dr. V Anandam¹,

Ch Mamatha¹, N. Chandra Sekhar Reddy²,

¹Department of Computer Science and Engineering,
Vardhaman College of Engineering,
Shamshabad, Hyderabad, Telangana, India.

¹bulan.babba@rediffmail.com,
velaanand@yahoo.com, mittumamatha03@gmail.com,

²Department of Computer Science and Engineering,
MLR Institute of Technology,
Hyderabad, Telangana, India.

²naguchinni@gmail.com

June 10, 2018

Abstract

It is very important to detect any malignant growth in human body correctly. Many times it is seen that incorrect detection due to misleading results, has led to trauma amongst patients. First step in correct detection of malignant growth is image segmentation. Genetic Algorithms fitness formula helps to produce correct offspring for correct image segmentation. This paper aims

to lessen the quantity of misclassified images depicting malignant cancerous growth. Image processing techniques such as enhancement of image is used for proper detection of lesion. As a part of pattern classification, support vector mechanism is applied as a tool to identify malignant tumors between different samples.

Keywords: machine learning, Image processing, software support vector mechanism, genetic algorithm.

1 INTRODUCTION

Segmentation of image is a technique whereby we bifurcate the image into various categories and regions. This process helps to classify different regions of the image into different objects of our interest.

Segmentation is done on grey scale level of the image pixels. Pixels having similar grey scale values are clustered as one zone. Then zones are classified as regions. Image segmentation can be applied directly or it can be applied after applying mathematical morphological tools to improve results.

Next part of image segmentation includes object identification, feature elicitation and object analysis and classification, if image segmentation is done correctly. Parameter selection is a vital area of image segmentation. If image segmentation is seen as an optimal problem then Genetic Algorithm can be applied as a search algorithm to search the correct parameter for image segmentation.

Classification is an important area of image area detachment using genetic algorithm. This classification can be differentiated into two different categories i.e. supervised and unsupervised.

Support vector mechanism, a machine learning techniques is used for finding the ROI (region of interest) or the abnormal region and also for feature recognition of this region. Appropriate size of the ROI is important for feature extraction. The pattern classification problem can be solved can be solved using support vector mechanism approach.

2 RELATED WORK

Medical image processing is a vital area to be analyzed with the increase in intricacies of healthcare. Ideally, medical image processing actually process vague, omitted, cryptic, conflicting, unnecessary, counteractive, askew data and also those data that have a strong anatomical character.

Apart from these, medical image processing studies the anatomical character of the data set and can be further processed using the techniques of Expert Systems ,Knowledge Based Systems, Neural Networks, Decision Support Systems, Fuzzy Logic and Systems, Genetic Algorithms, Knowledge Discovery, Data Mining, Symbolic Calculus for knowledge representation ,etc [1].

Digital mammograms are the most cogent method for the unmasking breast cancer It is always important to increase the symptomatic preciseness of biomedical imaging. So we can go for mach ine learning artistry for obtaining maximum result. The ambiguity between malignant and benign tumors can be decreased by using machine learning and artificial intelligence techniques. Traditional image processing algorithms can also be used along with thresh holding techniques and morphological operations for classifying malignant mammograms. The suctorial muscle in digital mammograms can divert the process of detection; so it should be removed from the mammograms.

For the purpose of feature differentiation between malignant and benign data, the optimal set of pattern is modeled using a machine learning tool called Support Vector Machine (SVM). The SVM made knowledgeable on two-third of the total data set and the left over one-third of specimen is used for validation and testing. The binary differentiation correctness of the matured system is determined by Receiver Operating Characteristic (ROC) analysis. Other parameters such as specificity, Area under the Curve, along with the sensitivity are also measured [3].

Genetic Algorithm is established on natural selection process. This algorithm takes a parameter as input and then calculates the output. The final output is based on considering different probabilistic output. The Genetic Algorithm is structured to replicate evolution naturally necessary for development.

These algorithms search the global spaces efficiently to get a maximal solution. Genetic Algorithm can effectively bring about contrast intensification. It is an optimal tool to generate image with natural heterogeneity. These algorithms go for a well-ordered desultory search. These can provide us with a universal and very simple method to find solution of multifarious aggrandizement problems.

Genetic Algorithm is a inferred-free and an equivocal development method. This algorithm requires little preexistent data about the issues to be solved than the traditional methods. Considering individuals fettle value, this algorithm uses the function like crossover reproduction, and mutation to obtain the next offspring that contains chromosomes with better fitness [14].

Recently, lots of work is done to up lift digital mammograms that can help radiologists to find and identify benign and malignant anomalies. This category of service provided by digital mammograms is termed as computer-aided detection.

In the yester year, computer-controlled systems noted and marked apprehensive areas on mammograms for the radiologist. These machines reduced searching time, conception and calculation errors even if radiologists failed to identify.

The computer-controlled identification is used after the radiologists have done with their analysis of the mammograms. If the radiologist has identified any abnormality in the mammogram initially but his finding does not match with the result of computer controlled system, the radiologist is supposed to give a positive result to the patient.

3 PROPOSED SYSTEM

First step of our approach will segment the image using genetic algorithm and unsupervised learning technique. Mathematical Morphology can also be applied along with Genetic Algorithm to obtain a better result. The image is first partitioned into non overlapping sub images. Then we adopt genetic algorithm along with morphological operations to each of the sub images to get the final segmented image.

Foremost step in mammogram preprocessing is noise removal.

Background Suppression follows Noise removal. The surrounding muscles depicted in the mammogram needs to be also suppressed.

Support vector machines are computerized classification technique to detect malignant and benign lesions in digital mammograms. SVM is a much better way for classification and provides an accurate result when compared with other methods over noisy data. The proposed system uses SVM to detect mass lesion in digital mammography. Support vector machine is used in the proposed system for pattern codification of the textures of malignant and benign mammograms.

We will build a dual classifier for character classification.

4 RESULT ANALYSIS

The segmentation process of our system is compared with the results of classical approach applied for segmentation on mammograms. The comparison shows that our system shows much better results.

Using the trained data set and applying SVM hyper parameter D , we see how it affects the correctness of SVM engine. As there are 20 bounded Software Vectors in the trained data set, so we take, $\tilde{a}_i = D$ in

$$\tilde{a}_i = D \Rightarrow z_i \quad f(y_i) \leq 1 \quad \& \quad o_i \geq 0 \quad (1)$$

The Grid Search method uses the parameter D that depicts a choicest correlation amongst training data set error and the intricacy of the model, with parameter $D=60$. This signifies how much noise the schooled data set has. A small value of the parameter D is used, and obtains the results of SVM classification with reduced noise and much smother. The Radial Basis Function kernel parameter, used in the Support Vector Machine classification controls the asset parameter of the RBF kernel. The parameter is in a way related to, defined as:

$$\acute{a} = \frac{1}{2} \quad \tilde{a}^2 \quad (2)$$

We define $\tilde{a}^2$ as the variance of the resulting Gaussian hyper sphere. The optimal value of the SVM hyper parameter $\acute{a}$ is

computed using by taking $\acute{a} = 0.001953125$. So, the value of δ can be computed using the following expression:

$$\delta = \sqrt{\frac{1}{2}\acute{a}} \quad (3)$$

The value of δ is computed to be $\delta = 16$ using $\acute{a} = 0.001953125$. This value of δ for the erudite binary classifier is acceptable, for all value of δ below 0.01 is taken to be small and all value of δ above 100 is granted large.

5 CONCLUSION

Digital mammograms are an effective computer controlled mechanism for the detection of breast cancer. The computer-controlled tools based on image processing and machine learning algorithms help systems to find doubtful region of the breast that are susceptible to mass lesions. We need a tool that will suppress the surrounding of the doubtful region and bring into focus the suspicious region. Our system provides a meaningful way by which efficient segmentation can be done, whereby the doubtful region is clearly identified by the radiologists.

By taking the SVM parameter D as 60, the noise in the image is reduced further. For feature extraction and detection along with pattern differentiation between malignant and benign data, the optimal set of features are modeled by a Support Vector Machine. Segmentation is done using Genetic Algorithm along with some morphological operators if at all needed to get a better result.

References

- [1] Hariton Costin and Cristian Rotariu, Medical image processing by means of some artificial intelligence methods, Institute of Computer Science of Romanian Academy, Iasi Branch, Received: May 8, 2013, Accepted for publication: June 10, 2013.
- [2] Adams, R. Bischof, L. 1994, Seeded Region Growing, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 6, pp. 641647.

- [3] Adler, D.D. and Helvie M.A. 1992, Mammographic biopsy recommendations, *Curr. Opin. Radiol.*, vol. 4, pp. 123-129.
- [4] Alexander, F.E. Anderson, T.J. Brown, H.K. Forrest, A.P. Hepburn, W. Kirkpatrick, A.E. Muir, B.B. Prescott, R.J. Smith, A. 1999. 14 years of follow-up from the Edinburgh randomised trial Of breast-cancer screening, *Lancet*, vol. 353, pp.1903-1908.
- [5] American Cancer Society, 2003a. Breast Cancer: Facts and Figures 2003-2004, Atlanta, GA: American Cancer Society.
- [6] American Cancer Society, 2003b. Cancer Prevention and Early Detection Facts and Figures 2003-2004, Atlanta, GA: American Cancer Society.
- [7] American Cancer Society, 2009. Breast Cancer: Facts and Figures 2009-2010, Atlanta, GA: American Cancer Society.
- [8] Anastasio, M.A. Yoshida, H. Nagel, R. Nishikawa, R.M. Doi, K. 1998, A genetic algorithm-based method for optimizing the performance of a computer-aided diagnosis scheme for detection of clustered microcalcifications in mammograms, *Medical Physics*, vol. 25, no. 9, pp. 1613-1620.
- [9] Anttinen, I. Pamilo, M. Soiva, M. and Roiha, M. 1993, Double reading of mammography screening films: one radiologist or two? *Clinical Radiology*, vol. 48, pp. 414-421.
- [10] Aylward, S.R. Hemminger, B.M. Pisano, E.D. Johnston, R.E. 1998, Mixture modeling for digital mammogram display and analysis. In Karssemeijer, N. Thijssen, M.A.O. Hendriks, J.H.C.L. van Erning, L.J.T.O. editors, *Digital Mammography*, pp. 305-312. Kluwer, Dordrecht.
- [11] Bajger, M. Ma, F. and Bottema, M.J. 2005, Minimum Spanning Trees and Active Contours for Identification of the Pectoral Muscle in Screening Mammograms, in *Proc. of Digital Image Computing: Techniques and Applications*, Dec. 2005, pp. 323-329.

- [12] Baker, J.A. Kornguth, P.J. Lo, J.Y. and Floyd Jr., C.E. 1996, Artificial neural network: improving the quality of breast biopsy recommendations, Radiology, vol. 198,pp. 131136.
- [13] Ballard, D.H. and Brown, C.M. 1982, Computer Vision. Englewood Cliffs, New Jersey:Prentice-Hall.
- [14] Miss. Komal R. Hole , Prof. Vijay S. Gulhane , Prof. Nitin D. Shellokar, Application of Genetic Algorithm for Image Enhancement and Segmentation, International Journal of Advanced Research in Computer Engineering Technology (IJARCET), Volume 2, Issue 4, April 2013.

