

A New Hybrid Method for Effective Computer Aided Diagnosis to Predict Breast Tumor in Elderly Patients

Balaji Samiraj, Arunprasath Thiyagarajan, Yudong Zhang, Pallikonda Rajasekaran Murugan,
Vishnuvarthanan Govindaraj, Vigneshwaran Senthilvel

Abstract: Breast cancer related morbidity has surged in recent years due to lack of awareness and failure to have an early prediction. With technological advancements available, doctors have put-forth their fullest effort to downsize the death-rate. Despite, certain areas of patient diagnosis, specifically with the case of breast cancers need a symbiotic improvement, and should greatly help the patients to lead a better and healthy life. For such extents, we authors like to propose a new hybrid approach that encompasses the functioning of Particle Swarm Optimization (PSO) and Improved Fuzzy C-Means (IFCM) through this paper. The technique recommended has tendered an elevation in the values of PSNR, TC and DOI, which most of the contemporary algorithms/techniques have failed to achieve, thus assisting the Doctors to have a profound idea regarding the prevalence of breast cancer and tumor.

Keywords : Improved Fuzzy C-Means (IFCM), Particle Swarm Optimization (PSO).

I. INTRODUCTION

Breast Cancer has caused high mortality rate of women, and it needs to be downsized to a greater extent. Several diagnostic techniques have been extensively used to detect breast cancers, and this research paper focuses on augmenting the breast cancer analysis by segmenting the breast tumor portion that have been imaged through X-ray mammograms and MRI Mammograms, having the algorithm capable of segmenting multimodal images. In X-ray Mammograms, low dosage levels of X-rays are used to vision the tumor region affecting the breast of the patient, where there is a patient exposure to ionizing radiations. Large and malignant breast tumors occurring in the lymph nodes of the

patient can be perfectly observed with X-ray Mammograms. On the contrary, MRI Mammogram does not use ionizing radiations and they literally help in identifying the micro-level calcifications spread within the breast of the patient. The tumor portions in the breast region of the patient are to be tediously detected during clinical practices. Radiologist has to exhibit labour-intense activity to assess the presence and analyzing its nature during clinical diagnosis. Based on the suggestions and feedback from the radiologist, an oncologist further proceeds with a motive to provide/ or achieve better prognosis. The segregation of tumor portion which may be of malignant, benign or widespread calcifications is the main concern of this research paper, and the algorithm recommended in this work is constructed in such a way that it helps in the ease identification of breast tumors in both modalities, even with the variations in the topographical locations of such tumors.

The algorithm developed through this paper has the caliber to segment the breast tumors present in individual left and right breast slices of X-ray mammograms, further, the algorithm has the potentiality to segment the breast tumors present in clinical MRI Mammograms comprising left and right breast portions in a single slice (3D view). The algorithm helps the radiologist with proper tumor region segregation, which will support the oncologist through a promising/confident clinical report, as both the procedures followed by the doctors are hand-in-hand process which will increase the life span of the patient. The algorithm developed in this paper has been tuned in such a way that it satisfies the above said expectancies.

II. DESCRIPTION OF DATASET

The Mammographic Image Analysis Society (MIAS) images taken from the National Breast Screening Programme in United Kingdom were recorded used a Joyce-Loebl scanning microdensitometer which delivers pixel edges digitized up to 50 microns. The device renders an optical density range of 0-3.2 and assists in each pixel representation with an 8-bit word. The dataset contains 323 digitized films which includes 64 benign, 52 malignant, 207 normal images and is available on 2.3GB 8mm (ExaByte) tape. It is inclusive of radiologist's "truth"-markings on the locations of any deformity that may be present. All the images are resized to the pixel size of 200 micron, so that they have a size uniformity of 1024x1024.

Revised Manuscript Received on December 05, 2019.

* Correspondence Author

Balaji Samiraj*, Research Scholar, Department of ECE, Kalasalingam Academy of Research and Education, Srivilliputtur, India. balajisamiraj@gmail.com

Arunprasath Thiyagarajan, Associate Professor, Department of BME, Kalasalingam Academy of Research and Education, srivilliputtur, India. arun.aklu@gmail.com

Yudong Zhang, Professor, F26 Informatics Building, Department of Informatics, University of Leicester, University Road, Leicester, LE1 7RH, UK. yudongzhang@ieee.org

Pallikonda Rajasekaran Murugan, Professor, Department of ECE, Kalasalingam Academy of Research and Education Srivilliputtur, India. m.p.raja@klu.ac.in

Vishnuvarthanan Govindaraj, Associate Professor, Department of BME, Kalasalingam Academy of Research and Education, Srivilliputtur, India. gvvarthanan@gmail.com

Vigneshwaran Senthilvel, Research Scholar, Department of ECE, Kalasalingam Academy of Research and Education, Srivilliputtur, India. a.s.vigneshwaran@klu.ac.in

III. METHODOLOGY

The block diagram of the proposed method is given in the fig.1. The input images chosen here is breast MRI which have been subjected to preprocessing steps, where to this purpose, Contrast limited Adaptive Histogram Equalization (CLAHE) is used in this work.

CLAHE has the advantage over the Adaptive Histogram Equalization (AHE), and CLAHE has the ability to avoid the noise amplification accumulating in homogeneous regions. The IFCM deals the poor image contrast setback, which is one of the major problems in Breast mammogram segmentation. The pre-processed image is inputted to PSO algorithm for the optimization process and the optimized results are clustered by using IFCM algorithm.

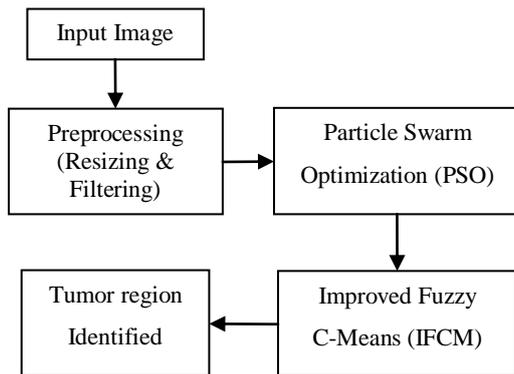


Fig. 1. Proposed Method.

IV. PARTICLE SWARM OPTIMIZATION (PSO)

Eberhart and Kennedy introduced the Particle Swarm Optimization which ensembles social and cognitive behavior of a bird. The individuals, so called as the particles flow across the multifaceted search space, whereunto a possible solution is arrived for a multifaceted problem as the particle propagates through the problem space. Two factors affect the motion of the each particles; one, the best solution of the particle (*pbest*), second, the particle's global best solution (*gbest*) identified by all the particles which adjusts velocity of the particle's through the search space by developing an attractive force. It results; the particle interacts with all the neighbors and stores the information of optimal location in its memory. On and after iterations, the *pbest* and *gbest* are updated correspondingly, and as a result, a more optimal solution is found from the population or particle. This operation is contiguous for several iterations, and it is progressive till the attainment of the expected results or on the completion of an exhaustive search operation.

The formulae of PSO defines each particle in the D-dimensional space as $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, where the subscript *i* denotes the number of particle, and the dimension intimates in the second subscript. The last stored best position is memorized and denoted as position (*pbest*), and it is denoted as $p_i = (p_{i1}, p_{i2}, \dots, p_{iD})$ and the velocity of particle in each dimension is specified as $\vec{v}_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ whereof the velocity gets updated

after every iteration. The motion/movement of the particle is influenced by its own best position (*pbest*) and the global best position (*gbest*) as well.

The velocity is updated as:

$$V^{(k+1)} = \omega V^{(k)} + c_1 rand_1 (pbest^{(k)} - X^{(k)}) + c_2 rand_2 (gbest^{(k)} - X^{(k)}) \quad (1)$$

The particle position is updated by:

$$X^{(k+1)} = V^{(k)} + X^{(k)} \quad (2)$$

Here, the constants *c*₁ and *c*₂ give the relative influence of *pbest* and *gbest*, each are set to have the same value. The factor '*ω*' is used to control the memory of PSO, where *rand*₁ and *rand*₂ are the values between 0 and 1, which are randomly generated. With the repeated application of equation (1) and (2), execute the PSO until the iterations exceed the specified number. Also, the operation ends once the velocity updates near to zero over the sequence of iterations.

A. Particle Swarm Optimization Steps

Step 1: Start the particle swarm as well as the initial position, velocity of particles and population size, etc.

Step 2: All particle needs to be calculated for its fitness, and the best position of the particle is connoted (*pbest*) using the fitness value. Among the best stored position will be the best fitness particle (*gbest*).

Step 3: Based on step (1) and step (2), the position and velocity of each particle have been updated.

Step 4: Comparing the updated position of each particle's fitness with the previous fitness (*pbest*) of the same to have a specific optimized fitness value, which is depicted as *pbest* for the new updated/optimized particle search position.

Step 5: Compare the every particle fitness with the previous of global best fitness. If it is better, then designate the position as the current best fitness (*gbest*).

Step 6: with the help of search algorithm to determine whether the results meet the conditions. If pre conditions are not met, again do the process from step (3) else stop the process, and find the optimized result.

V. FUZZY C-MEANS ALGORITHM (FCM)

In 1981 Bezdek extended the Fuzzy C-Means algorithm which was initially introduced by Dunn in 1973. For the analysis of data and construction of models, fuzzy clustering is used, and it is a robust unsupervised method too. Fuzzy clustering is most common than the hard clustering in most of the cases. Objects in an image aren't coerced into a single class, i.e., either 0 or 1; rather they are given the freedom to which particular class they can belong to, which is achieved through multiple membership function assignment. The FCM invokes fuzzy partitioning so that a data point gets classified in any group/class possessing different



membership grades between 0 and 1.

Consider $X = \{X_i\}$ is the image, $i = \{1,2,3,\dots\}$ the pixels of X is denoted as X_i and the total number of pixels is denoted by ' n '. The objective function of FCM gets reduced to a minimum value in most of the cases of Fuzzy clustering.

$$J_{FCM} = \sum_{i=1}^c \sum_{k=1}^n \mu_{ki}^m d^2(x_k, v_i) \quad (3)$$

From the objective function given in equation (3), membership function matrix is μ_{ki} , the Euclidian distance is denoted as d , and it is the distance between x_k and the cluster center v_i . Degree of fuzziness is denoted as ' m ', which is always greater than 1. The core of the membership function is μ_{ki} , and accordingly the FCM membership degrees are given by:

$$\mu_{ki} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ij}}{d_{jk}} \right)^{\frac{2}{m-1}}} \quad (4)$$

Centers of the cluster are:

$$V_i = \frac{\sum_{k=1}^n \mu_{ki}^m X_k}{\sum_{k=1}^n \mu_{ki}^m} \quad (5)$$

In clustering problems, the choice of metric plays a vital role. The Euclidean distance evaluation method is used because it is simple to construct; but it may consider that the data are uncorrelated and the clusters have the shape of super spherical geometry. This cannot occur in every image or most of the cases, and the choice of Mahalanobis distance is preferable for clustering complicated structures.

A. Steps involved in FCM algorithm:

Step 1: initialize the clusters number c ($4 \leq c < n$) and assign the number of clusters ' c ' which relates to the tissue classes. For fuzzy index ' m ', it is $m > 1$, which is instantiated using the membership in image matrix, and maximum iterations are set be as ' n '.

Step 2: calculate the fuzzy membership

Step 3: find the centroid (fuzzy centers)

Step 4: re-do the step 2 and 3 until the maximum number of iterations are finished or minimum objective function (10^{-6}) is reached.

VI. PROPOSED ALGORITHM

FCM clustering has better efficiency but requires more convergence time because it considers all the neighboring pixels for grouping/associating with the centroid. The same philosophy is extended for clustering an image. The proposed clustering algorithm, i.e., the IFCM is the combination of PSO and FCM, and the PSO gives the best optimal result to FCM, hence the IFCM has the phenomenon of PSO and FCM as well.

A. IFCM algorithm Steps:

Step 1: initialize the parameters including the maximum iterations n_{max} , weight factor exponent m ; number of clusters c ; termination/truncation threshold ϵ ($\epsilon = 10^{-6}$); population of particles p_{size} ; inertia weight w ; coefficients for acceleration c_1 and c_2 .

Step 2: Set the population size p_{size} .

Step 3: Set the count for iterations as $Gen_1 = 0$.

Step 4: Set other count for iterations as $Gen_2 = Gen_3 = 0$.

Step 5: Commence the working of PSO algorithm.

Step 5.1: Apply the global and local operators of PSO for p_{size} particles update.

Step 5.2: update Gen_2 as $Gen_2 = Gen_2 + 1$. Check that if $Gen_2 < 8$, and then go for Step 5.1 to proceed further.

Step 6: (FCM method) for each particle i in the image do:

Step 6.1: Take the position (both p_{best} and g_{best}) of particle i as the start of the cluster center for FCM algorithm.

Step 6.2: Re-evaluate each cluster center using the FCM algorithm.

Step 6.3: Assign $Gen_3 = Gen_3 + 1$, if $Gen_3 < 4$ is the condition, then move on to Step 6.2.

Step 7: Let $Gen_1 = Gen_1 + 1$. If the condition $Gen_1 < n_{max}$ occurs, then travel to Step 4.

Step 8: Group data point d_i to the cluster j which has the highest membership function μ_{ki} .

VII. RESULTS

Table- I: Performance analysis using comparison metrics

Metrics	PSO	IFCM
MSE	0.0002	0.0062
PSNR in dB	86.2543	70.41
Jaccard Index	0.216	0.223
Dice Overlap Index	0.3544	0.3641
Computational Time in sec	7.072	49.764
Memory requirement in byte	4.49E+08	4.48E+08

Input images

IFCM Results

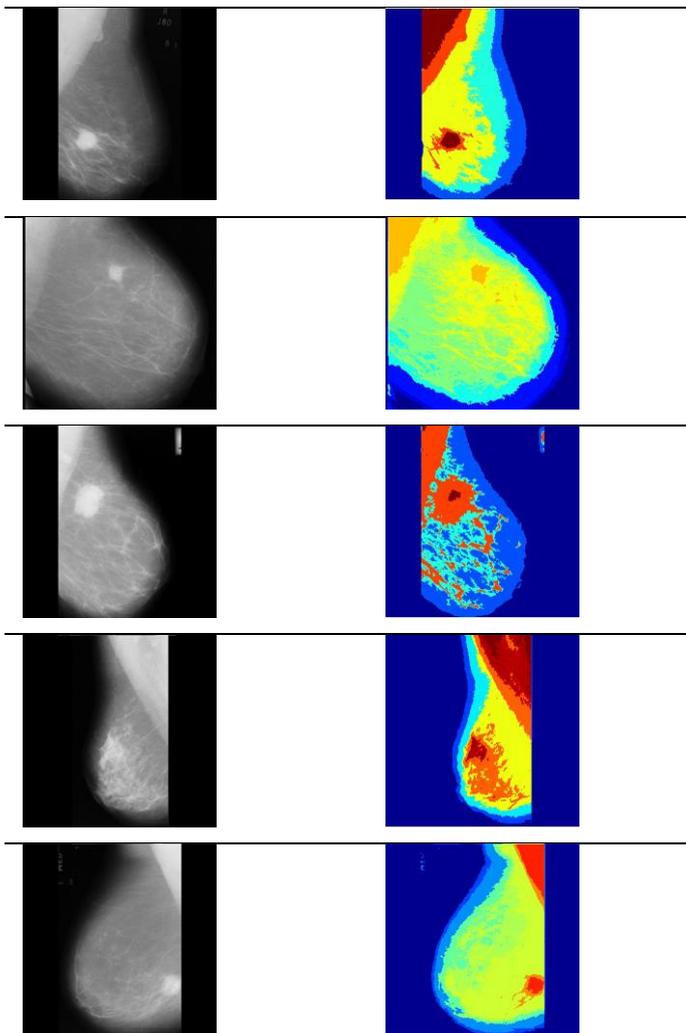


Fig. 2. Segmented results of MIAS dataset.

The comparison of various benchmark metrics are given in Table I. and the Fig.2. depicts the achievements of IFCM algorithm in segmenting the breast images.

VIII. CONCLUSION

FCM algorithm is one among the clustering method that favours the clear analysis through perfect theoretical foundation and wide applications. Despite all these features, the decreasing gradient optimization approach is used in the existing FCM to optimize the given data, which is prone to initialization and might enter into local optimum vulnerably. Later identify the PSO applications, the PSO and the IFCM methods are compared, which show that the IFCM method has generated the better segmentation results. Future studies will include the examination of mammogram images with more soft computing techniques. We would also like to extend the IFCM clustering algorithm for the dynamic prediction of optimum number of clusters for accruing better segmentations.

ACKNOWLEDGMENT

The authors thank the School of Electronics and Electrical Technology of Kalasalingam Academy of Research and Education, Tamilnadu, India for permitting to use the computational facilities available in open source research laboratory.

REFERENCES

1. Qinghai Bai, "Analysis of Particle Swarm Optimization Algorithm" Computer and Information Science February, 2010.
2. Li Wang, Yushu Liu and Xinxin Zhao, YuanqingXu "Particle Swarm Optimization for Fuzzy c-Means Clustering" Proceedings of the 6th World Congress on Intelligent Control and Automation, June 21 - 23, 2006, Dalian, China.
3. TolgaBerbera, AdilAlpkocakb, Pinar Balci, OguzDicle, "Breast mass contour segmentation algorithm in digital mammograms" computer methods and programs in biomedicine,110(2013)150–159.
4. AbdenourMekhmoukh, KarimMokrani "Improved Fuzzy C-Means based Particle SwarmOptimization (PSO) initialization and outlier rejection with level set methods for MR brainimage segmentation " computer methods and programs in biomedicine(2015).
5. JianZhang,LingShen "An Improved Fuzzy c-Means Clustering Algorithm Based on Shadowed Sets and PSO" Hindawi Publishing corporation Computational Intelligence and Neuroscience Volume 2014.
6. Guang Hu, Zhenbin Du "Adaptive Kernel-Based Fuzzy C-Means Clustering with Spatial Constraints for Image Segmentation" International Journal of Pattern Recognition and Artificial Intelligence Vol. 33, No. 1 (2019).
7. Brad M. Keller, Diane L. Nathan, Yan Wang, Yuanjie Zheng, James C. Gee, Emily F. Conant, and Despina Kontos "Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation" Medical Physics 39, 4903 (2012);
8. JiayinKang, Wenjuan Zhang "Combination of Fuzzy C-Means and Particle Swarm Optimization for Text Document Clustering"Advances in Electrical Engineering and Automation, AISC 139, pp. 247–252.
9. QiangNiu,Xinjian Huang "An Improved Fuzzy C-means Clustering Algorithmbased on PSO" journal of software, vol. 6, no. 5, may 2011.

AUTHORS PROFILE



Mr. S. Balaji is an Assistant professor in Department of Electronics and Communication Engineering at Kalasalingam Academy of Research and Education. he received his M.Tech., in VLSI Design from Kalasalingam University, Krishnankoil in 2013, his B.E. in Electronics and Communication Engineering from Anna University in 2011 (R.V.S College of Engineering and Technology) and his Diploma in Electronics and Communication Engineering in 2008 (P.A.C.R Polytechnic College).



Dr. T. Arunprasath is an Associate professor in Department Biomedical Engineering at Kalasalingam Academy of Research and Education. he received his Ph.D. in Electronics and Communication Engineering from Kalasalingam University, Krishnankoil in 2015, his M.E. in Applied Electornics from Anna University in 2009 (Mohamed Sathak Engineering College) and his B.E. in Electrical and Electronics Engineering from Anna University (Syed Ammal Engineering College) in 2006. His research interests include biomedical instrumentation, image processing, image segmentation cloud computing, image segmentation. He has published 16 technical journals and 20 technical papers in refereed conferences in these areas. He is a life member of ISTE.



Prof. Yu-Dong Zhang received his PhD degree from Southeast University in 2010. He worked as a postdoc from 2010 to 2012 in Columbia University, USA, and as an assistant research scientist from 2012 to 2013 at Research Foundation of Mental Hygiene (RFMH), USA. He served as a full professor from 2013 to 2017 in Nanjing Normal University, where he was the director and founder of Advanced Medical Image Processing Group in NJNU. Now he serves as Professor in Department of Informatics, University of Leicester, UK. He was included in "Most Cited Chinese researchers (Computer Science)" by Elsevier from 2014 to 2018. He was the 2019 recipient of "Highly Cited Researcher" by Web of Science. He won "Emerald Citation of Excellence 2017" and "MDPI Top 10 Most Cited Papers 2015". He was included in "Top Scientist" in Guide2Research. He published over 160 papers, including 16 "ESI Highly Cited Papers", and 2 "ESI Hot Papers". His citation reached 10096 in Google Scholar, and 5362 in Web of Science.



He is the fellow of IET (FIET), and the senior members of IEEE and ACM. He is the editor of Scientific Reports, IEEE Transactions on Circuits and Systems for Video Technology, etc. He served as the (leading) guest editor of Information Fusion, Neural Networks, IEEE Transactions on Intelligent Transportation Systems, etc. He has conducted many successful industrial projects and academic grants from NSFC, NIH, Royal Society, and British Council.



Dr.M.PallikondaRajasekaran, Born in Srivilliputhur, Virudhunagar District of Tamil Nadu in 1980, he had his schooling in the same town and graduated in Electronics and Instrumentation Engineering in 2001 from Shanmugha College of Engineering, Thanjavur and completed his M.Tech. degree in 2002 with second Rank in SASTRA University. He pursued his doctoral programme in Anna University, Chennai. Starting as a Lecturer in 2003, he became Asst. Professor in 2008, Associate Professor in 2009 and Professor in 2012 in Kalasalingam Academy of Research and Education. He had a deep involvement in Bio-signal Processing research. His work on the Image Segmentation for identification of brain tumour and image reconstruction and compression using medical images for diagnosis. Over 150 B.Tech students, 75 M.Tech students, and 8 Doctorates stand testimony for his productivity in Image Processing, Wireless Sensor Networks, and Biomedical Instrumentation research. He has so far published more than 50 papers in national and international journals and conferences. He is a Fellow of Indian Society For Technical Education (ISTE), Institute of Electrical and Electronics Engineers (IEEE), Asia-Pacific Chemical, Biological & Environmental Engineering Society (APCBEEES), Institution of Engineers (India)(IE), International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT).



Dr.G.Vishnuvarthanan, born in 1986, has research stints in the avenues of medical image processing and artificial intelligence. He was awarded PhD in the year 2015 and bachelor's degree in Instrumentation and Control Engineering by 2007, and Master's Degree in VLSI by 2009. He has more than ten years of teaching and research experience and has his affiliation as Associate Professor with the Department of Biomedical Engineering of School of Bio and Chemical Sciences in the Kalasalingam Academy of Research and Education, Tamilnadu, India.



Mr.S.Vigneshwarn, is a full time research scholar in the Department of Electronics and Communication Engineering at Kalasalingam Academy of Research and Education. He has completed his M.E in Applied Electronics at Anna University in 2015, and B.E. in Electronics and Communication Engineering at Sudharsan Engineering College in 2012, both located at Tamilnadu, India. His primary research focus is upon medical image processing.