Breast-Cancer Detection using Thermal Images with Marine-Predators-Algorithm Selected Features

Venkatesan Rajinikanth
Department of Electronics and Instrumentation Engineering
St. Joseph's College of Engineering
Chennai 600119, India
email: v.rajinikanth@ieee.org

Seifedine Kadry
Faculty of Applied Computing and Technology, Noroff University College, Kristiansand, Norway
email: skadry@gmail.com

David Taniar
Faculty of Information Technology
Monash University
Clayton, Victoria 3800, Australia
email: David.Taniar@monash.edu

Robertas Damaševičius
Faculty of Applied Mathematics
Silesian University of Technology
44-100 Gliwice, Poland
email: robertas.damasevicius@polsl.pl

Hafiz Tayyab Rauf
Centre for Smart Systems, AI and Cybersecurity, Staffordshire University,
Stoke-on-Trent, United Kingdom
email: hafiztayyabrauf093@gmail.com

2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII) | 978-1-6654-4126-1/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICBSII51839.2021.9445166

***Abstract*— Breast-cancer (BC) is one of the major diseases in women group and the early diagnosis and treatment is necessary to cure the disease. Early detection of BC is very essential to implement appropriate treatment and the proposed research aims to develop an automated BC detection system using Breast-Thermal-Images (BTI). The executed approach is as follows; (i) Recording the image for various breast orientation, (ii) Extracting the healthy/DCIS image patches, (iii) Treating the patches with image processing scheme, (iv) Feature extraction, (v) Feature optimization with Marine-Predators-Algorithm (MPA), and (vi) Two-class classification and validation. In this work, the essential image features, such as GLCM and LBP with varied weights are considered to classify the clinically collected BTI into healthy/DCIS class using a chosen two-class classifier. The result of this study confirms that the Decision-Tree (DT) classifier helps to achieve enhanced accuracy (>92%) compared to other methods adopted in this research.**

Keywords—Breast cancer, Thermal image, Marine-Predators-Algorithm, GLCM, LBP, Classification.

I. Introduction

In the current era, even though a number of modern detection and precautionary methods are available for the use, the diseases in humans are gradually rising because of their personal and environmental reasons [1,2]. Among the existing diseases, the cancer is one of the major illnesses leads to very painful death. The former research on cancer authenticates that; pre-mature detection will help to cure the disease with prescribed clinical treatment methods, such as chemotherapy, radiotherapy and mild/major surgery [3,4].

In women community, the cancer in breast region (Breast-Cancer) is rising due to a variety of reasons and the untreated Breast-Cancer (BC) will lead to various health problems []. The premature phase identification of BC will be completed by means of suggested medical practice. The final phase of BC can be predictable with the following indications; swelling in breast, variation in the silhouette/dimension of breast, transform of breast skin shade, uncontrollable pain, etc. Based on the location of the BC, it can be categorized into (i) Ductal-Carcinoma-in-Situ (DCIS) and (ii) Lobular-Carcinoma-in-Situ [5].

When the symptom of the BC is feat; then the patient should approach the doctor for through examination and confirmation of the disease with a variety of clinical procedures. The clinical level diagnosis of the BC involves in; (i) Visual examination by the disease exert, (ii) Examination of the suspicious section using the bio-imaging technique, and (iii) Extracting the sufficient breast tissues using the Core-Needle-Biopsy (CNB) in order to confirm the cancer and its stage using recommended clinical tests. The execution of the CNB is one of the hurting invasive practice in which a needle is inserted to collect the tissue for further evaluation. In most of the cases, the CNB will not be recommended and the BC will be diagnosed using the modern imaging techniques [6-8].

The earlier works on bio-imaging based breast abnormality detection considered various imaging modalities, such as X-ray (mammogram), ultrasound, thermal-imaging, Magnetic-Resonance-Angiogram (MRA), and MRI [4,7,8]. Recently, the thermal-imaging approaches are widely adopted in various clinics to detect the disease in organs using the infer-red (IR) radiation coming out of the skin section and due to its non-invasive nature and low cost, these approaches are widely adopted in BC diagnosis [9-12].

The thermal imaging of the barest section is recorded using a specialised thermal imaging device (camera) which converts the captured IR radiation in to RGB scaled image of chosen dimension. In this image, every colour demotes a thermal profile and by simply evaluating the thermal profile, the abnormality in the image can be detected. Thermal-Imaging (TI) supported BC diagnosis can be found in [13] and these works confirms that; this technique helps to detect the disease with better diagnostic accuracy.

In this work, the clinical grade BTI images collected from the DITI-India [14] is considered for the evaluation and the earlier works on this database can be found in [4,8,13]. The recording of the image is carried out in different orientations and after the recording, the suspicions sections having abnormal/normal thermal patterns are cropped and resized to 256x256x3 pixels. The cropped images are then evaluated using; (i) Saliency detection, morphological segmentation and GLCM feature extraction, and (ii) Local-Binary-Pattern (LBP) generation with varied weights (W=1 to 4) and LBP feature extraction. All these extracted features are serially integrated to form a one-dimensional (1D) feature vector. Then the dominant features from this feature vector is selected using Marine-Predators-Algorithm (MPA) and these values are then considered to train and validate the binary classifier. During the classification task, a 10-fold cross validation is implemented and the best value attained is then considered as the classifier output. The task of this section is to categorize the TI into healthy/DCIS group. In this work, the performance evaluation of different classifiers

978-1-6654-4126-1/21/$31.00 ©2021 IEEE

are executed and based on the classification accuracy, the merit of the chosen binary classifier is confirmed.

Remaining part of this research are prearranged as below; Section 2 demonstrates the methodology, Section 3 presents the experimental outcome, Section 4 and 5 demonstrates the discussion and conclusion of this work.

II. METHODOLOGY

The detection of the BC using TI is one of the proven approach, in which a chosen image processing schemes are employed to extract and evaluate the abnormal section in the breast. The implementation of various segmentation techniques to extract the abnormal section from the breast TI can be found in [13]. Further, the image executed using the DITI-India dataset can be found in [14]. All the above said techniques are implemented the segmentation techniques.

In the proposed work, a methodology is suggested to extract and evaluate the abnormal breast segment from TI using the MLS. The various phases available in this system can be found in Figure 1. The patient with the breast cancer is completely tested by the doctor and suggests the image supported examination. The BT for the patient are recorded for the patients in a controlled environment using a special camera, which converts the IR radiation into RGB scale images with varied thermal patterns (colours ranging from blue to white). After collecting the TI; cropping and resizing is performed to get the recommended dimension image patches. The image patches are then evaluated using two different feature extraction pipelines, such as (i) saliency enhancement, morphological segmentation and GLCM feature mining, and (ii) LBP enhancement and feature extraction. After collecting these features, serial feature integration is then implemented. Feature optimization is then implemented using the MPA in order to choose the dominant features and the optimized features are then considered to train and validate the binary classifier using a 10-fold cross validation.

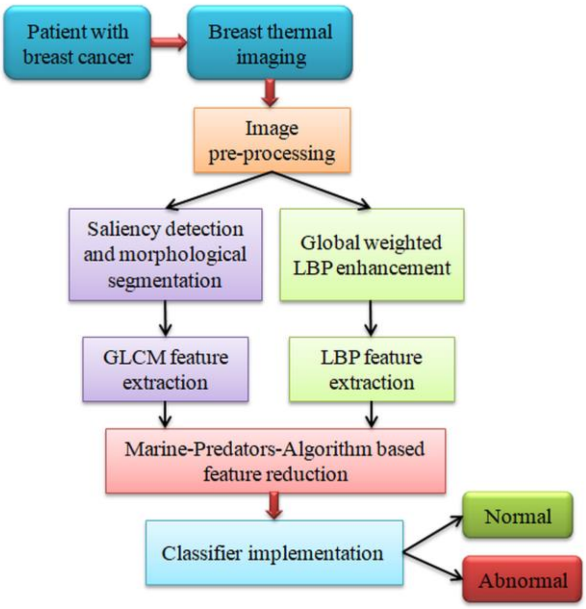


Fig.1. Structure of the proposed thermal image examination scheme

A. Therma Image Database

Recently, TI assisted disease detection method is widely adopted due to its non-invasive nature. In this work, the real-time images of TI discussed in [4,8,13] is adopted for the examination. This dataset consist both healthy/DCIS class images collected from the volunteers using various angle positions, such as $\theta = 0^{o}$, $\theta = \pm 45^{o}$ and $\theta = \pm 90^{o}$ as depicted in Figure 2. In this work, only the thermal patches from the normal and abnormal image section is then extracted and resized to 256x256x3 for the assessment. The number of image slices considered in this work for both the class is depicted in Table I. The converted images are then considered for the feature extraction and classifier validation task.

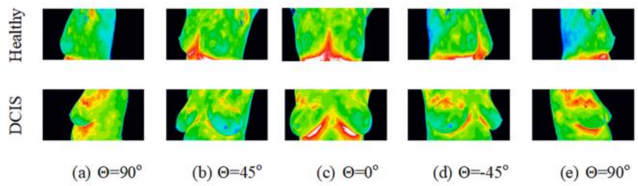


Fig.2. Sample test images collected from volunteers

TABLE I. THERMAL IMAGE PATCHES CONSIDERED

Database	Dimension	Images		
		Total	*Training*	*Validation*
Normal (clinical)	256x256x3	300	200	100
Abnormal (clinical)	256x256x3	300	200	100

B. Feature Extraction

Automated disease detection is important to lessen the diagnostic burden during the mass screening task. In most of the cases, features based disease categorization is employed due to its proven potential and every MLS and deep-learning methods works based on the features. In most of the MLS; the handcrafted-features such as GLCM and LBP are widely adopted to categorize the test pictures.

1. GLCM features mining

GLCM features helps to get the shape and texture information of the images under study. The texture information can be easily attained using the raw/processed images (gray level) and the shape features are collected from the binary image. In this work, the GLCM features are extracted from the pre-processed image; which is achieved using the saliency enhancement and morphological segmentation.

The essential information regarding the saliency supported image enhancement can be found in [15] and the morphology based segmentation scheme is presented in [16].

The GLCM features extracted in the proposed for healthy/DCIS picture is depicted in Eqn. (1),

$$F_{GLCM\ (1x25)} = GLCM_{(1,1)}, \ldots, GLCM_{(1,25)} \quad (1)$$

The GLCM features extracted using this work can be found in the earlier works [17-19].

2. LBP features mining

The LBP is also a handcrafted-feature widely adopted in MLS and in this work, the global weighted LBP invented by

Gudigar et al. [20] is adopted to enhance the image with varied weights ranging from W=1 to 4 and every technique helped to get a one-dimensional features of value 1x59 and the total features extracted with the proposed LBP will be 4*59 = 1x236. Eqn. (2) depicts the LBP features of this research;

$$F_{LBP}\,(1x236) = F_{LBP1}\,(1x59) + F_{LBP2}\,(1x59) + F_{LBP3}\,(1x59) + F_{LBP4}\,(1x59) \tag{2}$$

The essential work on LBP and its features can be found in [].

C. Feature Selection

The dominant feature selection is one of the essential tasks in MLS to reduce the problem of over fitting. The traditional feature reduction procedures existing in the literature can be accessed from [18]. In this work, heuristic algorithm based feature selection is employed.

The raw features existing for the assessment is depicted in Eqn. (3) and the feature reduction process will help to get a reduced feature vector value;

$$F_{Total\,(1,261)} = F_{GLCM\,(1,25)} + F_{LBP\,(1,236)} \tag{3}$$

The MPA is a nature-inspired meta-heuristic method invented by Faramarzi et al [21] and this approach combines the Lévy/Brownian-search to find the optimal solutions for a chosen task. As depicted in Figure 3, the MPA consist three search phases in which it uses a Lévy-search when the available prey is less and utilizes the Brownian-search when prey concentration is more. It continuously adopts both the search tactics till it discovers optimal solution. The sample 2D and 3D search strategy in MPA is depicted in Figure 4 and this strategy is continuously adopted in all the phases of MPA till optimal solution is achieved.

Initially, all the agents are arbitrarily dispersed in the investigate space based on Eqn. (4);

$$X_0 = X_L + \Re(X_U - X_L) \tag{4}$$

where $\Re$ is random value [0,1], X_U and X_L are upper and lower bounds.

The MPA uses the survival-of-fittest (SOF) theory to choose the top predator, which survives over a long duration. In the MPA, every search iteration ($Iter_{max}$) is divided into three phases, such as PhaseI = upto $\frac{1}{3} Iter_{max}$, PhaseII= $\frac{1}{3} Iter_{max} < Iter < \frac{2}{3} Iter_{max}$ and PhaseIII = greater than $\frac{2}{3} Iter_{max}$

PhaseI: It is the initial section called the high-velocity-phase, in which the predator is assumed faster than the prey.

PhaseII: It is the mid section called the unit-velocity-phase, in which the predator and prey are moving at same velocity.

PhaseIII: It is the final section called the low-velocity-phase, in which prey moves faster than the predator.

When the MPA search is initiated, it will explores the entire search space till the objective-value of the chosen problem is maximized. The complete information of the MPA can be found in [22,23].

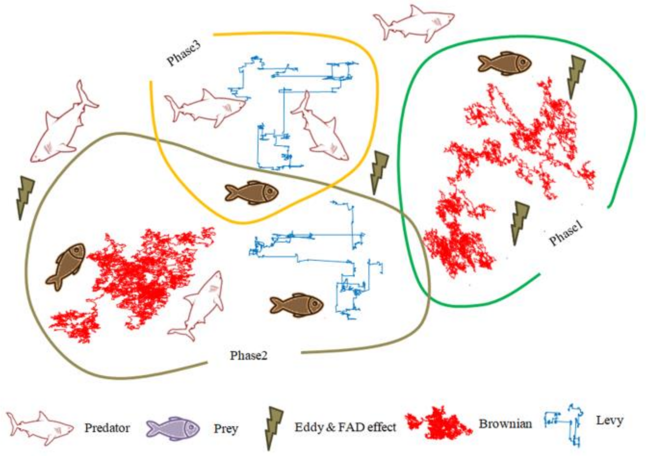
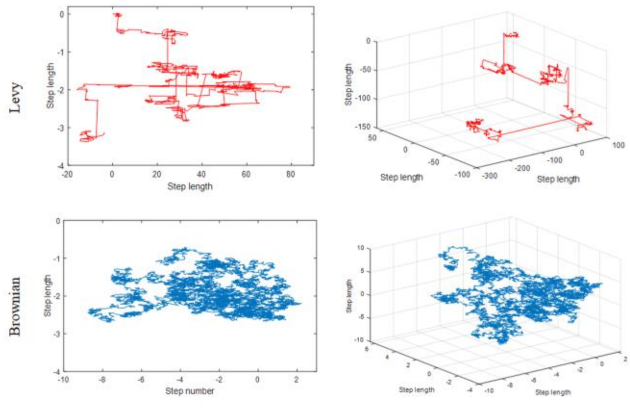


Fig.3. Various phases in the MPA optimization search



(a) 2D search (b) 3D search

Fig.4. Sample 2D and 3D search pattern

In this work, the following values are considered for the MPA; number of agents = 25, search dimension = features to be selected, objective-value = maximal Hamming-distance, maximum iterations ($Iter_{max}$) = 3000 and stopping criteria = $Iter_{max}$.

D. Implementation

The implementation of the feature optimization is depicted in Figure 5 and in this task, every feature (healthy/DCIS case) is individually compared against each other and the feature offers the maximal Hamming-distance is selected as the best feature. Similar procedure is employed for all the 261 features of GLCM and LBP and the features which have the clear difference between the healthy/DCIS cases are selected as the optimized features. In this work, the MPA helps to identify (1x74) optimized values and these features are then considered to evaluate the classifier performance.

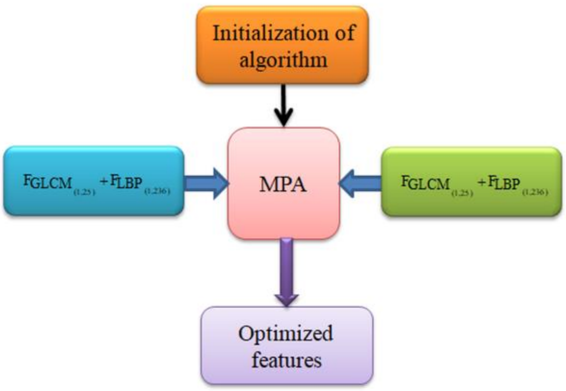


Fig.5. Feature optimization with MPA

E. Classifier Implementation and Validation

A binary classification with a 10-fold cross validation is then implemented in this work to validate the performance of the BC detection scheme. The binary classification is implemented using SVM variants, such as linear (SVM-L), quadratic (SVM-Q), cubic (SVM-C), Fine-Gaussian (SVM-FG), Medium-Gaussian (SVM-MG), Coarse-Gaussian (SVM-CG) are considered and the results are computed [24].

The performance of the proposed BC detection approach is verified by computing the performance-measures, like True-Positive (TP), True-Negative (TN), False-Positive (FP) and False-Negative (FN), From these PV, additional measures, such as accuracy (AC), precision (PR) sensitivity (SE), specificity (SP) and F1-Score (FS) [17-19].

III. Experimental Result

This part of the research presents the experimental outcome attained using the MATLAB software. Initially, the essential sections from the TI are extracted and resized into 256x256x3 pixels and the trial picture of healthy/DCIS case is depicted in Figure 6. These images are then considered for the evaluation. Fig 6(a) depicts the thermal profiles of healthy class and Fig 6(b) shows the DCIS category.

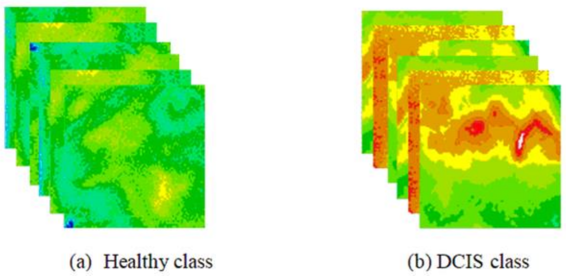


Fig.6. Sample trial images of healthy/DCIS class

Initially, the saliency enhancement and morphological extraction based technique is implemented to mine the abnormal section from the chosen breast TI section and then the GLCM features are extracted. Figure 7 shows the outcome attained with the implemented technique. Fig 7(a) depicts chosen TI segment, Fig 7(b) and (c) shows saliency and feature maps of the enhanced image, Fig 7(d) depicts the saliency feature plot and Fig 7(e) illustrates the extracted DCIS section. The GLCM features are then extracted and then integrated with the LBP features to form the essential handcrafted-features vector.

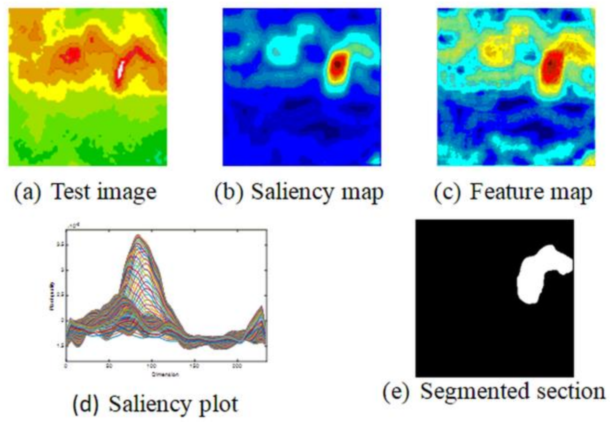


Fig.7. Saliency and morphological segmentation based evaluation of thermal pictures

The image slices are then considered for the LBP supported enhancement with varied weights (W=1 to 4) as in Figure 8 and every image helps to extract 59 LBP features and the total LBP features extracted from this image group is 4x59=236. The corresponding LBP histogram is depicted in Figure 9 for healthy and DCIS class.

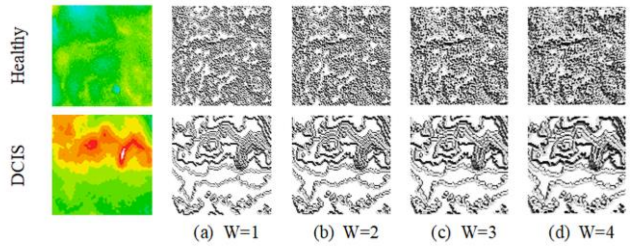


Fig.8. LBP enhanced thermal images for W=1 to 4

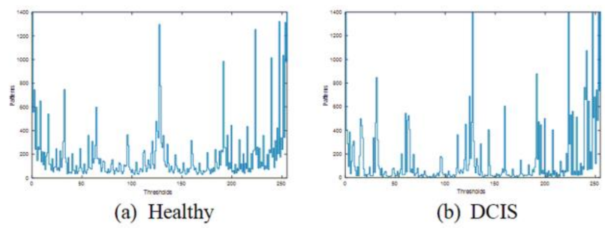


Fig.9. LBP histogram of healthy and DCIS class

IV. Discussion

The integration of the GLCM and LBP features helps to get a feature vector with a dimension 1x261 and these features are then reduce to a lower value (1x74) using the MPA. These features are then considered to train and validate the binary classifiers considered in this research.

In this work, 200 TI slices are considered to train the SVM classifiers and 100 TI are considered to validate the merit of proposed scheme. The experimental results attained in this work are depicted in Table II. The results of this table confirm that the accuracy of SVM-C and SVM-CG is similar and superior to other methods. The overall performance presented in the Glyph-plot (Fig. 10) confirms that the overall performance of CSVM-C and SVM-CG are approximately similar. If the proposed work is implemented with any one of the above mentioned classifiers, the result will be superior.

TABLE II. PERFORMANCE VALUES ATTAINED WITH TWO-CLASS CLASSIFIERS

Method	TP	FN	TN	FP	AC (%)	PR (%)	SE (%)	SP (%)	FS (%)
SVM-L	92	8	90	10	91	90.19	92	90	91.09
SVM-Q	92	8	94	6	93	93.88	92	94	92.93
SVM-C	93	7	94	6	93.50	93.94	93	94	93.47
SVM-FG	91	9	92	8	91.50	91.92	91	92	91.46
SVM-MG	92	8	93	7	92.50	92.93	92	93	92.46
SVM-CG	94	6	93	7	93.50	93.07	94	93	93.53

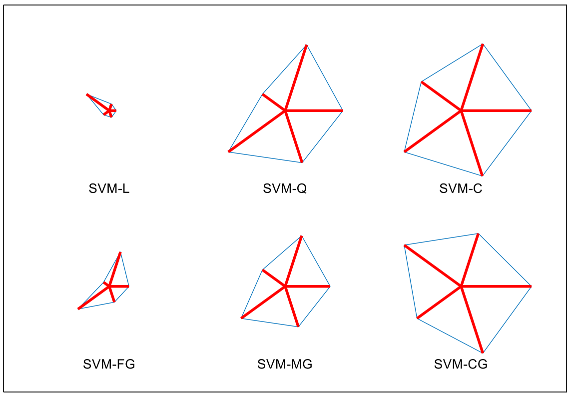


Fig.10. Glyph-plot of the computed performance values

In this research, the clinical grade TI is considered for the assessment and the attained result confirms that this work is clinically noteworthy. In future, the MPA can be replaced with the recent approach called the Red-Fox-Optimization [25] method existing in the literature.

V. CONCLUSION

Assessment of breast TI is necessary to identify the BC in women community and in this research; a methodology is suggested to examine the TI using chosen handcrafted-features. The proposed work implements saliency based enhancement and morphology segmentation to extract the GLCM features. Later, the LBP features are extracted from the pictures enhanced with W=1 to 4 and these features are then combined to form the hand-crafted feature vector. The dominant features are then selected using the MPA and these features are then considered to test the performance of the considered SVM classifiers. In this work, the result attained with SVM-C and SVM-CG are approximately similar. The proposed scheme works well on the clinical grade breast TI and provides better disease detection accuracy.

REFERENCES

[1] https://www.who.int/cancer/detection/breastcancer/en/

[2] N. Dey, V. Rajinikanth, and A.E. Hassanien, "An Examination System to Classify the Breast Thermal Images into Early/Acute DCIS Class," In *Proceedings of International Conference on Data Science and Applications*, pp. 209-220. Springer, Singapore, 2020.

[3] R. Elanthirayan, K. S. Kubra, V. Rajinikanth, N.S.M. Raja, and S.C. Satapathy, "Extraction of Cancer Section from 2D Breast MRI Slice Using Brain Strom Optimization," In Intelligent Data Engineering and Analytics, pp. 731-739. Springer, Singapore, 2021. https://doi.org/10.1007/978-981-15-5679-1_71.

[4] N.S.M. Raja, V. Rajinikanth, S.L. Fernandesand S.C. Satapathy, “Segmentation of breast thermal images using Kapur’s entropy and hidden Markov random field,” J. Med. Imaging Health Info., vol.7, no.8, pp. 1825–1829, 2017. https://doi.org/10.1166/jmihi.2017.2267.

[5] https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470

[6] H.S. Sheshadri, and A. Kandaswamy, "Breast tissue classification using statistical feature extraction of mammograms," Medical Imaging and Information Sciences, vol. 23, no. 3, pp. 105-107, 2006.

[7] R.I.R.Thanaraj, B. Anand, J. A. Rahul, and V. Rajinikanth, "Appraisal of Breast Ultrasound Image Using Shannon’s Thresholding and Level-Set Segmentation," In Progress in Computing, Analytics and Networking, pp. 621-630. Springer, Singapore, 2020.

[8] V. Rajinikanth, N.S.M. Raja, S.C. Satapathy, N. Dey, and G.G. Devadhas, “Thermogram assisted detection and analysis of ductal carcinoma in situ (DCIS),” In: International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), IEEE, pp.1641–1646, 2018. https://doi.org/10.1109/icicict1.2017.8342817.

[9] A.Q. Al-Faris, U. K. Ngah, N. A. M. Isa, and I. L. Shuaib, "Breast MRI tumour segmentation using modified automatic seeded region growing based on particle swarm optimization image clustering," In Soft Computing in Industrial Applications, pp. 49-60. Springer, Cham, 2014.

[10] N. Shrivastava, and J. Bharti, "Breast tumor detection and classification based on density," Multimedia Tools and Applications, vol.79, no. 35, pp.26467-26487, 2020.

[11] S.P.S. Raj et al., "Examination of digital mammogram using otsu's function and watershed segmentation", In. Fourth International Conference on Biosignals Images and Instrumentation (ICBSII), pp. 206-212, 2018. DOI: 10.1109/ICBSII.2018.8524794.

[12] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth, "Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation," Journal of Ambient Intelligence and Humanized Computing, pp.1-12, 2018. https://doi.org/10.1007/s12652-018-0854-8.

[13] S.L. Fernandes, V. Rajinikanth, and S. Kadry, “A hybrid framework to evaluate breast abnormality using infrared thermal images,” IEEE Consum. Electron. Mag., vol. 8, no.5, pp.31–36, 2019. https://doi.org/10.1109/MCE.2019.2923926.

[14] http://www.thermography.co.in/

[15] X. Hou, and L.Zhang, "Saliency detection: A spectral residual approach," In *2007 IEEE Conference on computer vision and pattern recognition*, pp. 1-8. IEEE, 2007.

[16] F. Meyer, and S. Beucher, "Morphological segmentation," *Journal of visual communication and image representation, vol.* 1, no. 1, pp.21-46, 1990.

[17] N. Dey et al., "Social-Group-Optimization based tumor evaluation tool for clinical brain MRI of Flair/diffusion-weighted modality," *Biocybernetics and Biomedical Engineering, vol.* 39, no. 3, pp. 843-856, 2019. https://doi.org/10.1016/j.bbe.2019.07.005.

[18] A. Bakiya, K. Kamalanand, V. Rajinikanth, R.S. Nayak, and S. Kadry, "Deep neural network assisted diagnosis of time-frequency transformed electromyograms," *Multimedia Tools and Applications*, vol. 79, no. 15, pp.11051-11067, 2020.

[19] N. Dey, Yu-Dong Zhang, V. Rajinikanth, R. Pugalenthi, and N.S.M. Raja, "Customized VGG19 architecture for pneumonia detection in chest X-rays," Pattern Recognition Letters, vol.143, pp.67-74, 2021. https://doi.org/10.1016/j.patrec.2020.12.010.

[20] A. Gudigar et al., "Global weighted LBP based entropy features for the assessment of pulmonary hypertension," Pattern Recognition Letters, vol.125, pp.35-41, 2019. https://doi.org/10.1016/j.patrec.2019.03.027.

[21] A. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, "Marine Predators Algorithm: A nature-inspired metaheuristic," Expert Systems with Applications, vol. 152, pp.113377, 2020.

[22] M.A.A. Al-Qaness, A. A. Ewees, H. Fan, L. Abualigah, and M. A. Elaziz, "Marine predators algorithm for forecasting confirmed cases of COVID-19 in Italy, USA, Iran and Korea," International journal of environmental research and public health, vol. 17, no. 10, pp.3520, 2020.

[23] M. Abdel-Basset, R. Mohamed, M. Elhoseny, A.K. Bashir, A. Jolfaei, and N. Kumar, "Energy-aware marine predators algorithm for task scheduling in IoT-based fog computing applications," IEEE Transactions on Industrial Informatics, 2020. DOI: 10.1109/TII.2020.3001067.

[24] M.A. Khan et al., "Computer-aided gastrointestinal diseases analysis from wireless capsule endoscopy: A framework of best features selection," *IEEE Access, vol.* 8, pp.132850-132859, 2020. Doi: 10.1109/ACCESS.2020.3010448.

[25] D. Połap, and M. Woźniak, "Red fox optimization algorithm," Expert Systems with Applications, vol. 166 pp.114107, 2021. https://doi.org/10.1016/j.eswa.2020.114107.