International Journal of Pure and Applied Mathematics
Volume 117 No. 15 2017, 1-13
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

Architectural Distortion Detection in Mammograms based on Optimized Feature Selection using Fish Swarm Optimization

[1]S. Kowsalya and [2]D. Shanmuga Priyaa

[1]Department of Computer Science,
Karpagam University,
Karpagam Academy of Higher Education,
Coimbatore, Tamil Nadu, India.
kowsalyaca@kahedu.edu.in

[2] Department of Computer Science,
Karpagam University,
Karpagam Academy of Higher Education,
Coimbatore, Tamil Nadu, India.
shanmuga.d@kahedu.edu.in

Abstract

One of the foremost earlier sign of breast cancer is architectural distortion. This research work proposes the fish swam optimization based detection of architectural distortion in mammograms acquired prior to the diagnosis of breast cancer in the interval between scheduled screening sessions. The potential sites are obtained using node maps through the implementation of Gabor filter and portrait modeling namely linear phase for detecting its presence. The ROI is extracted after pre-processing and is characterized with entropy measures such as angle, coherence and orientation strength. The ROI is also represented using Fourier spectrum that includes Shannon's entropy and Renyi entropy. The outcome of the experiment reveals that, using the entropy measures with fish swarm optimization for feature selection performs better than the particle swarm optimization and ant colony optimization. The experimental results are proved by performing classification using Artificial Neural Network.

Key Words: Mammogram, architectural distortion, roc, fish swarm optimization, feature selection, Gabor filter.

1. Introduction

Architectural distortion is defined as distortion of the architecture of breast parenchyma without being accompanied by the increased density or mass [1]. It is the third most common mammographic sign of nonpalpable breast cancer and is an important finding in interpreting the manifestation of breast cancer on mammograms [2]. However, due to its subtlety and variability in presentation, this sign of abnormality is often missed during the screening. The detection of architectural distortion is performed by a radiologist through the identification of subtle.

In Architectural distortion problem a mammogram image contains deformation in its structure but there will be no specific mass. It shows speculations projecting from a specific point and the distortion occurs at the borders of the parenchymal tissue. Architectural Distortion also acts an important cause of malignancy in one-half to two-thirds of cancer cases under consideration [2]. But this symptom do not account for increased density as of masses and calcification. During the initial stages Architectural distortion looks a lot like a normal breast tissue [3]. Since these tissues are very delicate, they are missed to be identified and results in false-negative reports [2]. 12% - 45% of breast mammograms with Architectural Distortion are misinterpreted as benign [4].

2. Related Work

The proposed work on breast lesions and their features [5] describes briefly the computer oriented methodologies used to detect and diagnose each lesion. One of the most missed symptoms of breast cancer is Architectural Distortion (AD). The ArDist[6] method consists of two stages: i) usage of Gabor Filters to detect ROI with potential ADs and ii) detection of Architectural Distortions using 2D Fourier transform in polar coordinates. Comparison is done with the efficient results of Computer based systems and model methods. Detection of architectural distortion in paper [7] uses two methodologies i.e. i) countourlet transforms and ii) phase portrait methods. Filtering of image under multiple directions is done by contourlet transform and orientation is performed using phase portrait analysis. White spaces are concentrated more to eliminate false positives from the sliding window. The paper [8] aims to work on diagnosis of breast cancer with the symptom of architectural distortion interval cancer cases. Stepwise logistic regression and stepwise regression are the two methods used to select appropriate features. An roc curve is acquired that results in a fine result and the features are chosen by stepwise logistic regression.

Burnside et al. [9] and Sumkin et al. [10] described that when prior mammograms are used for analysis, specificity may be increased better but it does not improve sensitivity. Varela et al. [11] showed that when prior mammograms are used as reference, classification accuracy value between benign and malignant masses has been increased in a significant manner

A commercial CAD system was studied for detecting its performance by Evans et al. [12]. He found that 91% of sensitivity was achieved with screening mammograms and 77% with prior mammograms. Ikeda et al. [13] estimated the accuracy percentage of the CAD system using 172 cancer cases of prior mammograms and 42% of sensitivity was acquired.

Cases of Mammograms with interval-cancer cases were investigated by Garvican and Field [14] using a commercial CAD system. Most sensitive cases were not identified by this system but it was able to detect normal areas of masses. Moberg et al. [15] in their study concluded that CAD system has no consequence on the sensitivity or the specificity of interval case cancer mammograms. Most mammograms are single or double read for the presence of breast tumour by radiologists. A comparison study on this single, double and CAD based reading of interval cancer cases is done by Ciatto et al. [16]. They affirmed that specific and significant results were produced by CAD method and it is as sensitive as double reading. The efficiency of various classifiers was analysed [25], [26] and it is proved that Artificial Neural Network performs well.

3. Proposed Architecture of Architectural Distortion

Architectural distortion is a difficult symptom to be detected in prior mammograms. In this study, an entropy based fractal dimension using Renyi Dimension spectrum is used to detect this architectural distortion.

Proposed Framework for Architectural Distortion

The following stages are accomplished to detect Architectural Distortion:

- Breast segmentation – Use region growing method to slice the mammogram image into breast region and background;
- Preprocessing of mammogram – reduce noise in the image by filtering and sharpen the spicules of Architectural Distortion;
- Enhancement of spicules – analyses and extracts line structures with different orientation that is texture orientation.
    - Various angles of text orientation are filtered with a collection of Gabor filters.
- ROI selection – extracted texture orientation is examined
    - phase portrait modeling – examine the oriented texture that creates three map values, in which only node map is analyzed further to perceive potential sites of ADs.
- Estimate AD features that are of importance which reduces false positives:
    - Multifractal dimension using Renyi Dimension Spectrum
- Feature Selection:
    - Fish Swarm Optimization is used for selecting potential attributes in an optimized manner
- Features classification:
    - Artificial Neural Networks (ANN) is used to separate boundaries of normal breast tissues and Ads.

The following Fig. 1 provides a detailed overview of detection of Architectural Distortion.

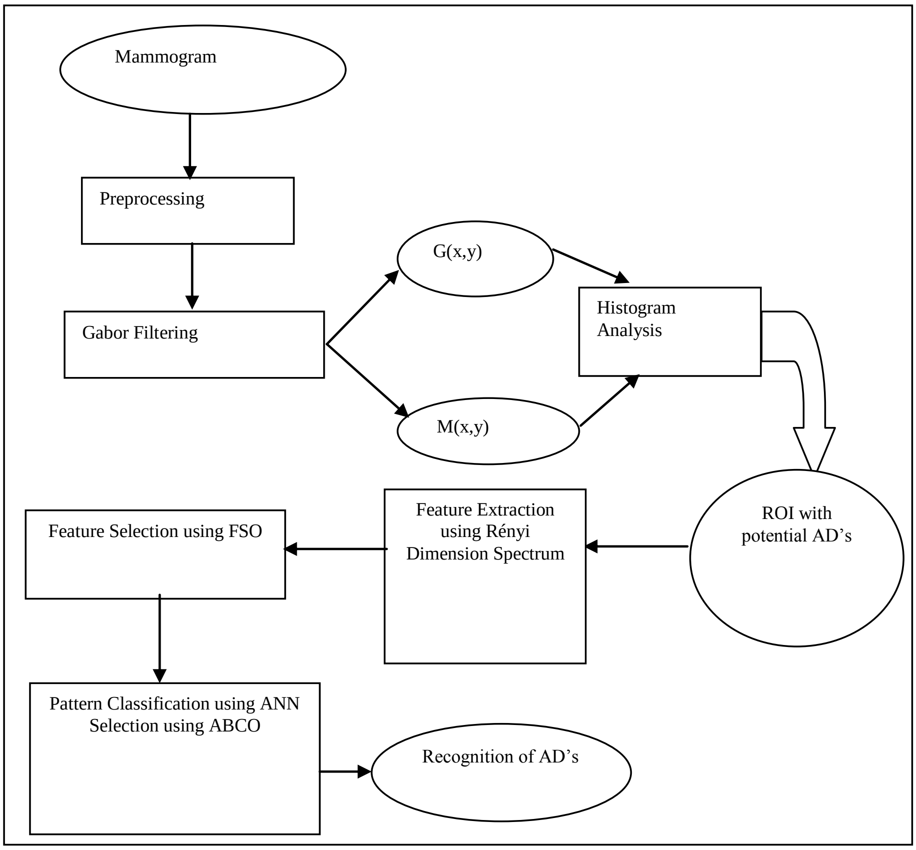

Figure 1: Detailed Overview of the Architectural Distortion

Entropy based Fractal Dimensions

When fractals are distributed in a heterogeneous fashion, then Entropy-based fractal dimensions can be used. But in the case of morphological dimensions, only the shape of a projection of the fractal is made to be known. Since morphological dimensions are purely metric concepts, this is explicable.

Feature Extraction using Renyi Entropy based Approach

A node-like prototype with speculation becomes apparent to radiate from a common point in case of Architectural distortion. But in the projected image of mammograms curvilinear structures of ligaments, ducts, vessels, and parenchymal region boundaries might provide the increase like patterns. Node analysis results in revealing a large number of FPs to detect significant areas of architectural distortion. Renyi entropy is a widespread measure for subsystems which are independent of statistical information. It characterizes the most important concepts of Shanonn's entropy method. Renyi entropy method curves only for value of q between 0 and 1. It is a special case of Shanonn's entropy where $q \rightarrow 1$. Renyi entropy [18, 19] is given by

$$H_R(q) = \frac{1}{(1-q)} \log_2 (\sum_i p_i^q) \qquad (1)$$

Through the elementary information provided, Renyi's measure gives the exponential mean value while Shannon's entropy is the simplified method of finding the averaged measure. This exponential mean of Renyi's entropy represents the equivalent elementary information gains of log2(1/pi) [19]. Renyi entropy has been widely used in multifractal theory [20], texture classification [21], pattern recognition, and image segmentation [22,23].

Analysis of Angular Spread

To detect the existence of architectural distortion, three measures that is orientation field, coherence and orientation strength are considered from Gabor magnitude response. These measures may reveal noteworthy information about the presence of architectural distortion. Among the three measures, only the orientation field is used in the preliminary stage of phase portrait analysis. Architectural distortion can also be detected using the Gabor magnitude reaction in some cases.

Angular Spread of Power in the Frequency Domain

A 2D Fourier spectrum, S(f, θ) is obtained initially. Then to obtain the angular power, this 2D value is converted into one dimensional function S(θ). This is done by combining as a function into the angle θ for the angle range of 0 - 179◦. The frequency point has a radial distance of f = [6, 96] pixels. For the next level of computation, regions with low and high frequency limits were purged. This is because these regions are treated as noise and are not related in detecting architectural distortion.

Orientation Strength

To calculate the orientation strength for each pixel, a function G(θ) is represented to provide the potential strength of orientation. This computation is done for each angle θ of the image in consideration. In the present work, this measure of strength is computed as a mean of Gabor magnitude responses, Gk (m, n), for all directions of the filters used, θ_k , k = 0, 1, 2, . . . , 179, as follows

$$\left[\alpha(m,n) = \sqrt{\frac{\left[\sum_{k=0}^{179} G_k(m,n)\cos(2\theta_k)\right]^2 + \left[\sum_{k=0}^{179} G_k(m,n)\sin(2\theta_k)\right]^2}{\left[\sum_{k=0}^{179} G_k(m,n)\right]^2}} \right] \qquad (2)$$

Coherence

To compute coherence, the values of magnitude and orientation is represented as orientation information for each pixel. For each neighbourhood pixel, this computation involves the average dominant orientation and extent of alignment of orientation information. This is done with regard to dominant orientation always.

$$\gamma(p,q) = G(p,q)\frac{\sum_{m=1}^{P}\sum_{n=1}^{P}|G(m,n)\cos[\theta(m,n)-\Psi(p,q)]|}{\sum_{m=1}^{P}\sum_{n=1}^{P}G(m,n)} \qquad (3)$$

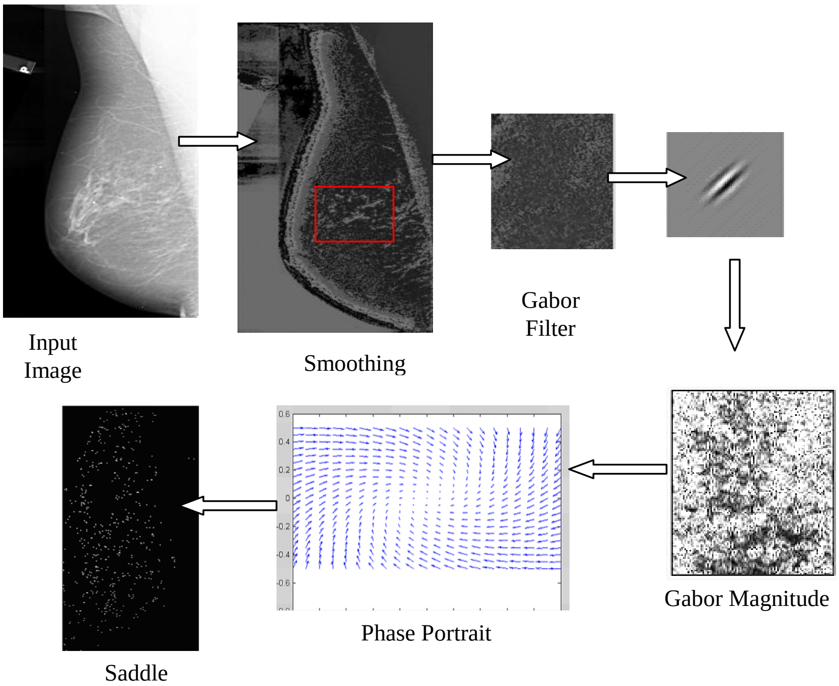

Figure 2: Output of Proposed Architectural Distortion

The Fig.2 provides the output produced by the proposed Architectural Distortion techniques for detecting breast cancer.

Feature Selection using Fish Swarm Optimization

Li et al. [24] in 2002 proposed a new technique called The Fish Swarm Algorithm (FSA) stimulated by natural behaviour of fish. This technique is a new population based progression technique. Using the FSA algorithm, the nearby minimum values are eliminated to achieve overall optimization. $X_i = (x_1, x_2, \ldots, x_k, \ldots, x_D)$ represents the position of the fish with dimensional D. FSi imparts the food satisfaction acquired by the fish. $d_{ij} = \|X_i - X_j\|$ gives the Euclidean distance between the spots of two fishes. The FSA emulates three common characteristics:

- Explore food source.
- Swarming in case of threat.
- Enhance successful result rate.

The objective of exploration of food source is to minimize the value of FS(Food Satisfaction). Exploring process involves in the search of food source by the fish in a random manner. Swarming holds the objective to gratify the requirements of intake of food, satisfying the existing and new members of the swarm. Xi indicated the position of a fish within its neighbor's visibility. X_c locates the midpoint of neighbor's locations which can be used to represent the entire swarm neighbors.

A fish shifts from its current position X_i to next position X_{i+1} towards the midpoint X_c, if this midpoint of swarm concentrates more on food source rather than the current position X_i (i.e., FSc < FSi), and if (X_c) is not over crowded ($ns/n < \delta$). The fish in the neighborhood position also follows the same behavior.

Under the viewpoint of a fish, certain fish may be recognized to get more food than others.

This may naturally convince others to follow the best one (X_{min}) to improve the food satisfaction level [$FS_{min} < FS_i$] and less crowding [$nf/n < \delta$]). The number of fishes under the viewpoint of X_{min} is denoted as nf.. The three main constraints involved in the Fish Swarm Algorithm are i) visual distance ii) highest step length (step) and a iii) crowd factor. Among these three constraints, visual distance and step length stimulates the efficiency of FSA.

Pseudo Code for Fish Swarm Optimization

```
Start t = 0
xi(t) (i = 1, . . .,m) ← Initialize While (stopping criteria are not met)
do For (each xi (t))
If ("visual scope" is empty)
yi(t)← Random(xi(t))
else If ("visual scope" is crowded)
yi(t)← Search(xi(t))
else yi(t)← best of Swarm(xi(t)) and
Chase(xi(t))
End for
xi(t + 1)(i = 1, . . .,m) ← Select(xi(t), yi(t) (i = 1,….,m))
If ("stagnation" occurs) x rand(t + 1) ← Leap(xrand (t + 1))
xbest(t + 1) ← Local(xbest(t + 1)) t = t + 1
End while
End
```

4. Experimental Result

Mini-MIAS database is used to for experimental analysis of the proposed work and Matlab is used for implementation. A total of 44 cases of mammograms are chosen from the Mini-MIAS database, out of which 8 cases of architectural distortion, 14 cases of bilateral asymmetry, and 22 normal cases are taken into consideration. 19 mammograms out of which 8 cases with architectural distortion are taken into the consideration.

Performance Analysis of Proposed of Architectural Distortion

Table 1: List of Features with Accuracy Values

Symbol	Feature	Accuracy
Node	Node Value	0.65
HR_1	Power in the frequency domain	0.64
HR_2	Gabor magnitude response	0.69
HR_3	Gabor angle response	0.71
HR_4	Coherence	0.67
HR_5	Orientation strength	0.68

For interval cancer cases and normal cases the accuracy values and ROIs for all the features listed in the Table. 1 is computed. HR_1- HR_5 extracted using Renyi Entrophy based on angular spread. The potential features are selected using Fish Swarm Optimization, PCO and ACO. The FSO selects four sets out of four features: Node value, HR_2, HR_3, HR_5. The PCO selects five feature set of features Node, HR1 HR_2, HR_3, $HR_{5.}$ The ACO selects three attributes Node, $HR_{1,}$ $HR_{5.}$ The Fig.3 shows the performance chart shows the performance chart for these features.

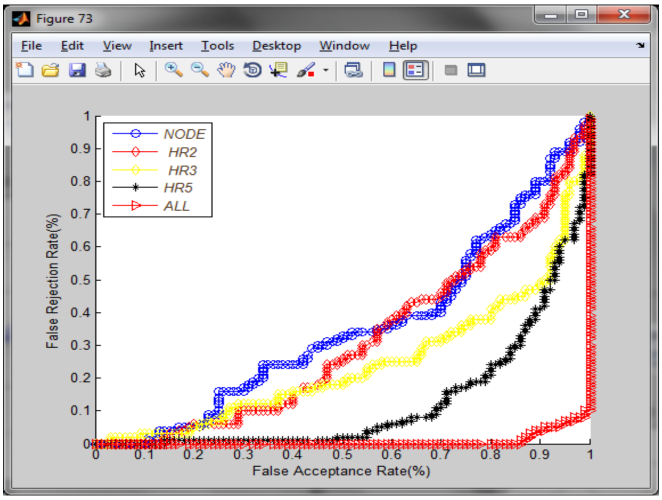

Figure 3: Performance Chart for False Acceptance and False Rejection Rate

The Artificial Neural Network is used as the classifier to find the performance Precision, Recall and F-measure of both the selected feature sets of FSO, PCO and ACO.

Artificial Neural Networks is used for classification after implementing three different feature selection techniques Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO) and Artificial Bee Colony Optimization based on Precision, Recall and F-measure.

Table 2: Precision, Recall and F-Measure Using PSO, ACO and FSO

Method for feature selection	Selected Feature Set	Precision	Recall	F-measure
PSO	Node, HR1 HR_2, HR_3, HR_5	0.867	0.860	0.853
ACO	Node, $HR_{1,}$ HR_5	0.912	0.928	0.909
FSO	Node, HR_2 ,HR_3, HR_5	0.949	0.967	0.940

The Table 2. Shows that the Fish Swarm Optimization (FSO) selects four different features Node values with three variables of Renyi entropy related to the angular spread, performs better than the existing two evolutionary approaches Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) with the highest precision value of 0.949 and its recall value is 0.967. The F-measure for PSO is 0.853 which is less than the ACO and FSO. The highest F-measure value is obtained by proposed FSO 0.940 and it is proved using ANN classification to detect the presence of architectural distortion in the potential sites which consist of four different features. The performance chart for these feature selection methods is shown in Fig. 4.

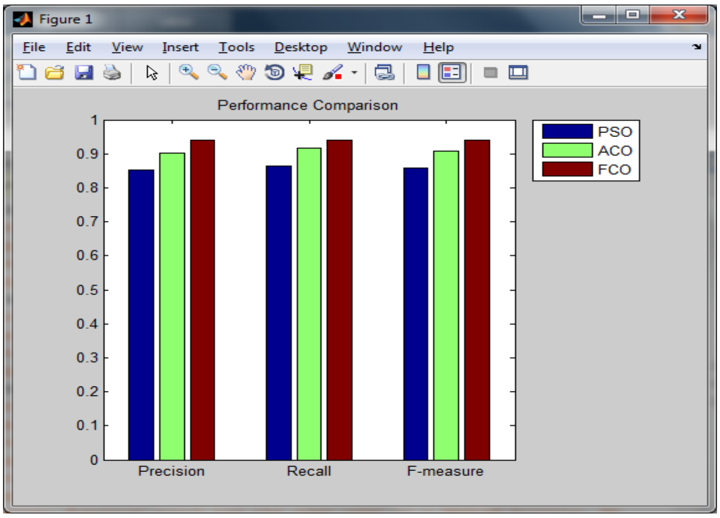

Figure 4: Performance Comparison for Precision, Recall and F-Measure using PSO, ACO and FSO

5. Conclusion

For recognition of architectural distortion in the potential sites, the features are extracted using the Renyi entropy based on angular spread. The FSO is used for selecting the potential features and retrieve the four features, one feature from node map and other three belongs to the entropy features specifically Gabor magnitude response, Gabor angle response and Orientation strength have the higher accuracy value than the others. After applying ANN classification on the mammogram images based on the selected set of features using PSO, ACO and FSO, the result shows that the performance of the FSO is better than the remaining ones by reducing the False positives in the automatic ROI detection of the suspected sites

References

[1] Homer M.J., Mammographic interpretation: a practical approach, McGraw-Hill, 766 (1997).

[2] Matsubara T., Ichikawa T., Hara T., Fujita H., Kasai S., Endo T., Iwase T., Novel method for detecting mammographic architectural distortion based on concentration of mammary gland, In International Congress Series 1268 (2004), 867-871.

[3] Knutzen A.M., Gisvold J.J., Likelihood of malignant disease for various categories of mammographic ally detected, non palpable breast lesions, In Proceedings of Mayo Clinic. Elsevier 68 (5) (1993), 454-460.

[4] Yankaskas B.C., Schell M.J., Bird R.E., Desrochers D.A., Reassessment of breast cancers missed during routine screening mammography: a community-based study, American Journal of Roentgenology 177 (3) (2001), 535-541.

[5] Bandyopadhyay S.K., Diagnosis of Breast Abnormalities in Mammographic Image, IJCST 2 (1) (2011).

[6] Jasionowska M., Przelaskowski A., Rutczynska A., Wroblewska A., A two-step method for detection of architectural distortions in mammograms, Information Technologies in Biomedicine, AISC 69 (1) (2010), 73-84.

[7] Anand S., Rathna R.A.V., Detection of architectural distortion in mammogram images using contourlet transform. IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN) (2013), 177-180.

[8] Rangayyan R.M., Banik S., Desautels J.L., Detection of architectural distortion in prior mammograms via analysis of oriented patterns, Journal of visualized experiments: JoVE 30 (2013).

[9] Burnside E.S., Sickles E.A., Sohlich R.E., Dee K.E., Differential value of comparison with previous examinations in diagnostic versus screening mammography, American Journal of Roentgenology 179 (5) (2002), 1173-1177.

[10] Sumkin J.H., Holbert B.L., Herrmann J.S., Hakim C.A., Ganott M.A., Poller W. R., Gur, D., Optimal reference mammography: a comparison of mammograms obtained 1 and 2 years before the present examination, American Journal of Roentgenology 180 (2) (2003), 343-346.

[11] Varela C., Karssemeijer N., Hendriks J.H.C.L., Holland R., Use of prior mammograms in the classification of benign and malignant masses, European Journal Radiology 56 (1) (2005), 248–255.

[12] Evans W.P., Burhenne L.J.W., Laurie L., O'Shaughnessy K.F., Castellino R.A., Invasive Globular carcinoma of the breast: Mammographic characteristics and computer-aided detection, Radiology 225 (1) (2002), 182–189.

[13] Ikeda D.M., Birdwell R.L., O'Shaughnessy K.F., Sickles E.A., Brenner R.J., Computer-aided detection output on 172 subtle findings on normal mammograms previously obtained in women with breast cancer detected at follow-up screening mammography, Radiology 230 (1) (2004), 811–819.

[14] Garvican L., Field S., A pilot evaluation of the R2 image checker system and users' response in the detection of interval breast cancers on previous screening films, Clin. Radiol., 56 (1) (2001), 833–837.

[15] Moberg K., Bjurstam N., Wilczek B., Rostgard L., Egge E., Muren C., Computer assisted detection of interval breast cancers, Eur. J. Radiol., 39 (1) (2001), 104–110.

[16] Ciatto S., Turco M.R.D., Burke P., Visioli C., Paci E., Zappa M., Comparison of standard and double reading and computer-aided detection (CAD) of interval cancers at prior negative screening mammograms: Blind review, Br. J. Cancer 89 (2003), 1645–1649.

[17] Burrell H., Evans A., Wilson A., Pinder S., False-negative breast screening assessment: What lessons we can learn?, Clin. Radiol., 56 (2001), 385–388.

[18] Renyi A., On measures of entropy and information, Proceedings of the 4th Berkeley symposium on mathematics, statistics and probability, University of California Press, Berkeley, CA (1961), 547–561.

[19] Masi M., A step beyond Tsallis and Renyi entropies, Phys Lett A 338 (3–5) (2005), 217–224.

[20] Kinsner W., A unified approach to fractal dimensions, Proceedings of the fourth IEEE international conference on cognitive informatics (ICCI), Irvine, CA, (2005), 58–72.

[21] Gabarda S., Cristobal G., Discrimination of isotrigon textures using the Renyi entropy of Allan variances, J Opt Soc Am A., 25 (9) (2008), 2309–2319.

[22] Li Y., Fan X., Li G., Image segmentation based on Tsallisentropy and Renyi-entropy and their comparison, IEEE international conference on industrial informatics, Singapore (2006), 943–948.

[23] Sahoo P., Wilkins C., Yeager J., Threshold selection using Renyi's entropy, Pattern Recognition 30 (1) (1997), 71–84.

[24] Li X. L., Shao Z.J., Qian J.X., An Optimizing Method based on Autonomous Animate: Fish Swarm Algorithm, System Engineering Theory and Practice 22 (11) (2002), 32-38.

[25] Thara L., Gunasundari R., Significance of data mining techniques in disease diagnosis and biomedical research – a survey, IIOABJ 7 (1) (2016), 284-292.

[26] Gunasundari R., Thara L., Helicobacter Pylori Infection and Associated Stomach Diseases: Comparative Data Mining Approaches for Diagnosis and Prevention Measures, IEEE International Conference on Advances in Computer Applications (ICACA), 2017.

[27] RAJESH, M. "A SYSTEMATIC REVIEW OF CLOUD SECURITY CHALLENGES IN HIGHER EDUCATION." The Online Journal of Distance Education and e‐Learning 5.4 (2017): 1.

[28] Rajesh, M., and J. M. Gnanasekar. "Protected Routing in Wireless Sensor Networks: A study on Aimed at Circulation." Computer Engineering and Intelligent Systems 6.8: 24-26.

[29] Rajesh, M., and J. M. Gnanasekar. "Congestion control in heterogeneous WANET using FRCC." Journal of Chemical and Pharmaceutical Sciences ISSN 974 (2015): 2115.

[30] Rajesh, M., and J. M. Gnanasekar. "Hop-by-hop Channel-Alert Routing to Congestion Control in Wireless Sensor Networks." Control Theory and Informatics 5.4 (2015): 1-11.

[31] Rajesh, M., and J. M. Gnanasekar. "Multiple-Client Information Administration via Forceful Database Prototype Design (FDPD)." IJRESTS 1.1 (2015): 1-6.