

WRAPPER BASED FEATURE SELECTION FOR CT IMAGE

D. Chitra¹ and G. M. Nasira²

¹Department of Computer Science, Government Arts College, Salem, India
E-mail: cithu74@gmail.com

²Department of Computer Science, Chikkanna Government Arts College, India
E-mail: nasiragm99@yahoo.com

Abstract

Diagnostic imaging is invaluable. Magnetic Resonance Imaging (MRI), digital mammography, Computed Tomography (CT), and others ensure effective noninvasive mapping of a subject's anatomy, and increased normal and diseased anatomy knowledge for medical research in addition to being a critical component in diagnosis and treatment. In this work various feature selection algorithms are investigated and a Swarm Intelligence Algorithm based on Bacterial Foraging is proposed. Features are extracted using wavelet and Gray-Level Co-occurrence Matrix (GLCM). The obtained features are fused using Median Absolute Deviation (MAD) after normalization and the feature selection techniques investigated. Results obtained show the improved performance of Bacterial Foraging based feature selection for different classifiers.

Keywords:

Computed Tomography (CT), Wavelet, Gray-Level Co-occurrence Matrix (GLCM), Median Absolute Deviation (MAD), Correlation Based Feature Selection (CFS), Bacterial Foraging Optimization (BFO)

1. INTRODUCTION

Medical images classification is fundamental, in different applications, in a medical image retrieval system [1]. Due to medical image data's high variability, it is important to use correct classification models. CT modality is applied to clinical diagnosis and helps radiologists detect and locate pathological changes accurately. CT images are distinguished for various tissues according to gray levels which present information that helps medical diagnosis. Medical images are from varied modalities like MRI, CT, Ultrasound and Positron emission tomography (PET). CT is more reliable for early detection of tumors and hemorrhages as this provides anatomical information to plan radio therapy [2].

Medical images provide diagnostic evidence and information about anatomical pathology. Medical images growth in database is enormous, in the past few years, when medical digital image equipment's like CT, MRI, and PET-CT are used in clinical work [3]. Medical information systems goals are defined to deliver needed information to right persons at right time and right place, to improve care process quality and efficiency. Digital image processing includes applications like remote sensing through satellites and spacecraft, image transmission /storage for business, medical processing sonar, radar and acoustic image processing, robotics, and automated industrial parts inspection and image recognition motivated by need for surveillance/security, telecommunication and digital libraries, human-computer interaction, and smart environments [4].

Basic CT measurement (developed in 1970s) involved rotation and displacing a collimated X-ray source (pencil beam)

around a patient. To speed it up, modern CT scanners have an X-ray source producing a fan-beam that penetrates a patient, impinging on detectors. This assembly rotates around a central core producing intensity measurements sequence over 360°. CT scanners measure Radon transform accurately, while MR imagers sample Fourier transform. While transforms are related through projection slice theorem, their properties which are different result in image quality [5] differences. CT scanners resample data to data points needed for image reconstruction. Most MR pulse sequences sample data at Cartesian lattice points required for image reconstruction through Fast Fourier Transform (FFT).

Texture identifies regions of interest in images. Grey Level Co-occurrence Matrix (GLCM) is a texture feature extraction method. GLCM extracts second order statistical texture features and is used in many applications. GLCM is a matrix where rows and columns are equal to gray levels, G , in an image [6]. Matrix element $P(i, j|\Delta x, \Delta y)$ is relative frequency where 2 pixels, separated by a pixel distance $(\Delta x, \Delta y)$, happen in a neighborhood, one with intensity i and the other with intensity j . One may say that matrix element $P(i, j|d)$ contain second order statistical probability values for changes between gray levels i and j at specific displacement distance d and at specific angle (Δ) . GLCM calculates the frequency with which a pixel with gray-level (gray scale intensity/Tone) value i occur horizontally, vertically or diagonally to value j adjacent pixels.

Feature selection selects an original features subset and its optimality is measured through an evaluation criterion. As domain dimensionality expands, features N increases. Finding optimal feature subset is intractable and problems related to feature selection are NP-hard [7]. A typical feature selection process has four basic steps; subset generation, evaluation, stopping criterion and result validation. Feature selection algorithms are either filter or wrapper models. The former relies on general characteristics of training data to select features without involving learning algorithms. It does not inherit any learning algorithm bias and are cheap computationally as they do not involve induction algorithm. Nevertheless, they risk selecting features subsets which fail to match chosen induction algorithm [8]. The wrapper model needs a predetermined learning algorithm in feature selection where its performance evaluates and determines selected features. For new subset features, a wrapper model learns a hypothesis (or classifier). It ensures superior performance as it locates features suited for a predetermined learning algorithm. On the other hand, it is expensive computationally.

Wrapper methods are superior alternatives in supervised learning as they use inductive algorithm to evaluate alternatives they consider the algorithm's biases. Implementing a wrapper is straightforward in supervised learning, as there are external

validation measures available. One executes a classifier and gets estimation prediction accuracy in a known class label [9]. Wrapper model uses classification to measure a features set's importance and so the selected feature depends on classifier used. Wrapper methods lead to better performance than filter methods as feature selection is optimized for classification algorithm used. But, wrapper methods are expensive for large dimensional database regarding computational complexity and time as every feature set is evaluated by a classifier algorithm [10].

This study uses classification process for CT images based on feature extraction using Gaussian wavelet and GLCM. Feature fusion is by MAD after normalization. Feature selection is by wrapper technique using CFS and Bacterial foraging algorithm. Section 2 reviews related work, section 3 describes methodology, section 4 discusses experimental results and section 5 concludes the work.

2. RELATED WORK

A CAD system for classification and automated detection of pulmonary nodules in 3D CT images was proposed by Namin et al., [11]. The method achieved 88% sensitivity for nodule detection with around 10.3 False-Positive (FP)/subject; achieved nodules classification was concordant with radiologists' opinion. An automatic labeling solution for manual labeling issues using radiology report associated with medical images was proposed by Gong et al., [12]. The experiment was conducted with traumatic brain injury CT images; but the proposed framework of automatically labeling and classifying medical cases can be used for medical images in other modalities or anatomical parts.

Two sets of shape based features for every segmented hematoma region was proposed by Gong et al., [13]. Overall classification accuracies for hematoma region from CT slices was 80.7%, 81.3%, and 81.1% with primitive features, geodesic distance features, or both features. For volumetric hematoma classification, overall accuracies were 80.9%, 81.5%, and 81.5% respectively and results were promising for radiologists and neurosurgeons in this field.

Importance of parameters set in classification of lung CT images was evaluated by Vasconcelos et al., [14], so that ROIs size, quantization level, and textural features were used in classification. SVM was the classifier. Performance of the classifier was evaluated with 10-fold cross validation and results based on overall accuracy, sensibility and specificity were compared. Textural features had good discriminatory power in lung emphysema classification in CT images.

An automated method to detect and classify an abnormality into acute infarct, chronic infarct and hemorrhage at slice level of non-contrast CT images was presented by Chawla et al., [15]. The new method gave 90% accuracy and 100% recall in detecting patient abnormality achieving average precision of 91% and 90% recall at slice level. An automatic detection and classification method to improve and accelerate physicians' decision-making was attempted by Shahangian and Pourghassem [16] where multilayer neural network was more successful than KNN classifier due to higher accuracy (93.3%). Finally, more than 90% accuracy was achieved in brain hemorrhages detection and classification.

Evaluation of potential role of adaptive hybrid segmentation algorithm, Contourlet transform and Extreme Learning Machine in liver tumors differential diagnosis in CT images was proposed by Selvathi et al., [17]. Segmentation results and experts' results were compared and analyzed. The classifier differentiated tumors with relatively high accuracy ensuring a second opinion for radiologists. A classification system to differentiate malignant pulmonary nodules from the benign in CT images based on fractal features from Fractional Brownian Motion (FBM) model was presented by Huang et al., [18]. The result revealed that classification accuracy and area under Receiver Operating Characteristic (ROC) curve were 83.11% and 0.8437, respectively, when using the new fractal-based feature set and a SVM classifier thereby demonstrating that the new classification system had highly satisfactory diagnostic performance by analyzing lung nodules fractal features in CT images from one post-contrast CT scan.

Automated systems to detect and classify hematomas type using Artificial Neural Network (ANN) algorithm for CT images of patients was dealt with by Sharma and Venugopalan [19]. The methodology had four phases, first preprocessing on brain CT images, second, histogram based centroids initialization for K-means clustering algorithm to segment image into clusters based on pixels intensity values. Third phase includes features extraction from segmented images. The fourth phase has ANN created and trained according to features from image to classify hematoma types according to features.

A new detection and classification method to process SPECT-CT images of breast/prostate lymph nodes was presented by Papp et al., [20]. Results showed that the new method was effective in supporting physicians working with related images in nuclear medicine. A pseudo-stochastic variation of Metropolis dynamics using Markov Random Fields was presented by Liang et al., [21]. Results were compared with those from Metropolis algorithm, Gibbs sampler and Iterated Conditional Mode (ICM) which indicated that using MMD reflected interior spatial distribution of concrete materials on deformation, affording a method of concrete meso-structure CT image study.

Properties of sound logic which is based on flexible logic were discussed by Jialin and Huacan [22]. Compared to traditional extracting algorithms that aimed at entire or a portion, experiments through examination and clinical cases analysis revealed the new algorithm obtained higher diagnostic accuracy.

A new segmentation algorithm for 3D brain PET-CT images, that classified every voxel by fusing its memberships which is estimated from four points of view using smoothness prior, PET information, CT information, and probabilistic brain atlas was proposed by Xia et al., [23]. Results in 11 clinical brain PET-CT studies proved the new algorithm providing accurate and reliable segmentation.

Body segment classification i.e., classifying figures in CT modality into different body segments like head, abdomen, pelvis, or thorax was done by Xue et al., [24]. The new method's evaluation on a dataset of 2465 figures from the National Library of Medicine's (NLM) PubMed Central[®] repository's open access biomedical articles subset achieved classification accuracy of more than 90% proving its effectiveness and potential to be vital in biomedical document retrieval systems

like the NLM developed OpenI, a multimodal biomedical literature search system.

A new Fisher criterion and Genetic optimization, based feature selection method called FIG, to tackle CT Imaging Signs of Lung Diseases (CISLs) recognition problem was proposed by Liu et al., [25]. To evaluate the new feature selection method and CISL recognition approach, 5-fold cross validation experiments were conducted on 511 ROIs from real lung CT images. The new FIG method ensured improved recognition performance than the full original features set and also single type features. The new method was then compared to feature selection based on classification Accuracy Rate and Genetic optimization (ARG). FIG's advantages on computation effectiveness and efficiency over ARG are seen through experiments.

A feature-based learning framework for correct prostate localization in CT images was proposed by Liao and Shen [26]. The method was evaluated on 24 patients CT prostate dataset to localize the prostate, where patients had more than 10 longitudinal images scanned at varied treatment times. It was compared to many state-of-the-art prostate localization algorithms in CT images, and results proved that the new method had higher localization accuracy of all methods compared. A Computer-Aided Detection (CADe) system for Hepatocellular Carcinoma (HCC) using Sequential Forward Floating Selection (SFFS) with Linear Discriminant Analysis (LDA) was presented by Xu and Suzuki [27]. An SFFS method was developed and coupled to LDA which maximized Area Under receiver-operating-characteristic curve (AUC) value.

3. METHODOLOGY

This work investigates the classification problem for medical images. The greatest challenge in medical image classification is the large feature space created during feature extraction. In this work two different feature vectors namely GLCM for texture features and Wavelet for image energy coefficients are generated. The obtained features are fused using MAD. To reduce the feature space existing statistical techniques like Information Gain and Correlation Based Feature Selection are investigated. A novel Bacterial foraging algorithm is proposed to improve the feature selection. The block diagram of the proposed system is shown in Fig.1.

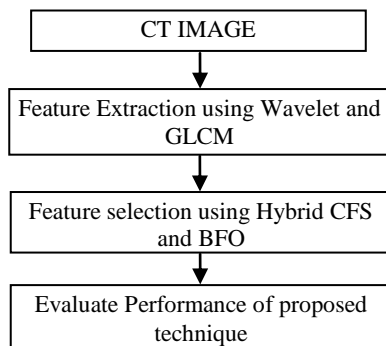


Fig.1. Flowchart for the proposed method

3.1 WAVELET

Wavelets satisfy some mathematical requirements and represent data or other functions. Wavelet analysis is a common

tool to analyze localized variations of power in time series. By decomposing time series into time–frequency space, one can determine dominant modes of variability and how they vary in time. Wavelet transform was regarded as an interesting diversion producing colorful pictures and purely qualitative results. This misconception is due to wavelet analysis which involves a transform from a one-dimensional time series (or frequency spectrum) to diffuse two-dimensional time-frequency image [28]. Diffuseness was exacerbated through arbitrary normalizations and lack of statistical significance tests.

Wavelets are appropriate to analyze non-stationary time series, while Fourier analysis is not. They are applicable to time series as a fusion (compromise) between filtering and Fourier analysis. Wavelets compress information in two dimensional images from satellites/ground based remote sensing techniques like radars. Wavelets are useful when highest frequencies are removed, as local information is retained and image resembles a low resolution version of full pictures [29]. With Fourier analysis/global functional fits, images can lose resemblance to picture, after removal of few harmonics. This is due to wavelets being a hierarchy of local fits retaining time localization information, whereas Fourier or polynomial fits are usually global fits.

The wavelet transform is given by [30]:

$$X_W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \left(\frac{t-b}{a} \right) x(t) dt \quad (1)$$

where, a and b are wavelet function parameters and $x(t)$ is signal to be transformed. Prototype wavelet function is given by

$$h_{a,b}(t) = \frac{1}{\sqrt{a}} h\left(\frac{t-b}{a}\right) \quad (2)$$

3.2 GRAY-LEVEL CO-OCCURRENCE MATRIX

Gray-Level Co-occurrence Matrix (GLCM) is a statistical method to examine textures that consider pixels spatial relationship. GLCM functions characterize an image's texture by calculating how often a pair of pixels with specific values and in specified spatial relationship happen in an image thereby creating a GLCM, and extracting statistical measures from the matrix [31]. MATLAB's graycomatrix function creates a GLCM through calculating how many times a pixel with intensity (gray-level) value i occurs in specific spatial relationship to a pixel with value j . Each element (i, j) in resultant GLCM is sum of times that pixel with value i occurs in specified spatial relationship to a pixel with value j in input image.

GLCM proved a popular statistical method to extract images textural features. According to co-occurrence matrix, Haralick defines 14 textural features measured from probability matrix to extract texture statistics characteristics of remote sensing images [32]. In reality, for each d , resulting values for 4 directions are averaged.

To show computation, for image I , let m represent pixels gray level (x, y) and n represent gray level of pixels $(x \pm d\phi_0, y \pm d\phi_1)$ with gray tones L level of where $0 \leq x \leq M-1$, $0 \leq y \leq N-1$ and $0 \leq m, n \leq L-1$. From these representations, the gray level co-occurrence matrix $C_{m, n}$, for distance d and direction ϕ is defined as [33],

$$C_{m,n,\phi} \sum_x \sum_y P\{I(x,y)=m \& I(x \pm d\phi_0, y \mp d\phi_1)=n\} \quad (3)$$

where, $P\{\cdot\} = 1$ if argument is true and otherwise, $P\{\cdot\} = 0$.

3.3 FEATURE CONCATENATION

Feature concatenation involves concatenating features of extracted from different sources. In this work, wavelet and the GLCM features are concatenated. A fused feature set is obtained by concatenating the two features point sets. The descriptors from both the feature vectors are normalized using min-max normalization technique. Thus, both the wavelet and GLCM feature vectors consists of values ranging from 0-1. Then construct an augmented feature vector by concatenating its rows or columns. As the resulting concatenated vector encompasses all the key representations of both the features, it has more discriminative power.

3.4 FEATURE SELECTION

Feature selection refers selection of input variables which are relevant in predicting an instance. It is used to rank relevant input variables or to build a good classifier. Feature selection is a search problem, where a subset of features relates to different state in the search space [34]. It improves prediction accuracy of the classifier and reduces computation time by reducing the feature space. This is achieved by removal of irrelevant, redundant and noisy features [35].

Feature selection methods can be broadly divided into filter and wrapper approaches. In the filter approach, relevance of the features is assessed based on the intrinsic properties of the data for feature subset selection and does not consider the classifier. Whereas in wrapper method, the classifier is embedded in the process of feature subset selection as it determines the quality of the subset [36].

3.4.1 Information gain (IG):

Information gain (IG) measures the amount of information in bits about the class prediction, if the only information available is the presence of a feature and the corresponding class distribution. Concretely, it measures the expected reduction in entropy. Given S_X the set of training examples, x_i the vector of i^{th} variables in this set, $|S_{x_i=v}|/|S_X|$ the fraction of examples of the i^{th} variable having value v [37]:

$$IG(S_X, X_i) = H(S_X) - \sum_{v=values(x_i)} \frac{|S_{x_i=v}|}{|S_X|} H(S_{X_i=v}) \text{ with entropy : } (4)$$

3.4.2 Correlation based Feature Selection (CFS):

Univariate filters downside is information gain as it does not take into account the interactions between features, which is overcome by the multivariate filters for example Correlation based Feature Selection (CFS). CFS evaluates a subset of attributes worth considering each feature's individual predictive ability with degree of redundancy between them. Correlation coefficients estimate correlation between subset of attributes and class, and with inter-correlations between features. Relevance of a features group grows with correlation between features and

classes, decreasing with increased inter-correlation. CFS determines best feature subset and combines with search strategies like backward elimination, forward selection, bi-directional search, best-first search and genetic search [38].

CFS algorithm uses correlation based objective function to evaluate features' usefulness. The objective function $J_{cfs}(\lambda)$, known as Pearson's correlation coefficient, is based on heuristic that a good feature subset has high correlation with class label but remains uncorrelated among themselves [39].

$$J_{cfs}(\lambda) = \frac{\lambda \psi_{cr}}{\sqrt{\lambda + \lambda/(\lambda-1) \psi_{rr}}} \quad (5)$$

The above equation illustrates merit of λ features subset where ψ_{cr} is average feature to class correlation and ψ_{rr} is average feature to feature correlation within class. CFS based feature selection algorithm uses $J_{cfs}(\lambda)$ to search feature subsets using best first search.

3.4.3 Bacterial Foraging Optimization (BFO):

BFO algorithm is a new evolutionary computation algorithm based on of *Escherichia coli* (E. coli) bacteria foraging behavior, in the human intestine. BFO algorithm is a biologically inspired computing technique which mimics E. coli bacteria's foraging behaviour. Natural selection removes animals with poor foraging strategies favoring circulation of genes of animals with successful foraging strategies, as they are more likely to enjoy reproductive success. After generations, poor foraging strategies are either removed or shaped into good ones. This foraging activity is used in optimization [40].

E. coli bacteria in the intestines have a foraging strategy governed by four processes; chemotaxis, swarming, reproduction, and elimination and dispersal [41].

Chemotaxis: This is achieved through swimming and tumbling. Depending upon flagella rotation in each bacterium, it decides if it should move in a specific direction (swimming) or a different direction (tumbling), in the bacterium's entire lifetime.

Swarming: It is desired that bacterium that has searched optimum path of food should attract other bacteria to ensure that they reach desired place rapidly.

Reproduction: Least healthy bacteria die, and other healthy bacteria split into two, and are placed in same location thereby ensuring constant bacteria population.

Elimination and Dispersal: It is possible that in local environment, bacteria population life changes gradually by consuming nutrients or suddenly due to other influence.

Extracted features are reduced through BFO to remove redundancy and irrelevant features and the resulting feature subset (got through BFO) are most representative subset. The position of the bacteria is either 1 or 0 based on whether the feature is selected or not in the search space. During Chemotaxis, tumbling leads to new random position determining whether the feature is selected or not in the next iteration. Fitness is evaluated for each bacterium and the position is updated if the fitness is better. The bacteria with least health are removed and bacteria with best health are reproduced. At the completion of the iterations, the position of the bacteria represents the best feature subset achieved.

BFO parameters are given in the Table [41] below,

Table.1. BFO Parameters

Parameter Name	Description
J_{cc}	Cost function value
J_{Health}^i	Health of bacterium i
L	Counter for elimination-dispersal step
P_{ed}	Probability of occurrence of elimination-dispersal events
S	Population of the E. coli bacteria
$\omega_{attract}$	Width of attractant
$\omega_{repellant}$	Width of repellent

4. EXPERIMENTAL RESULTS

The experimental results are shown below: Table.2 shows the results for classification accuracy.

Table.2. Classification Accuracy

Classification accuracy	Furia	Ripper	OneR
Feature selection using CFS	91.33	84.53	85.47
Feature selection using IG	92.73	86.87	87.60
Feature selection using proposed CFS -BFO	93.73	87.87	89.20

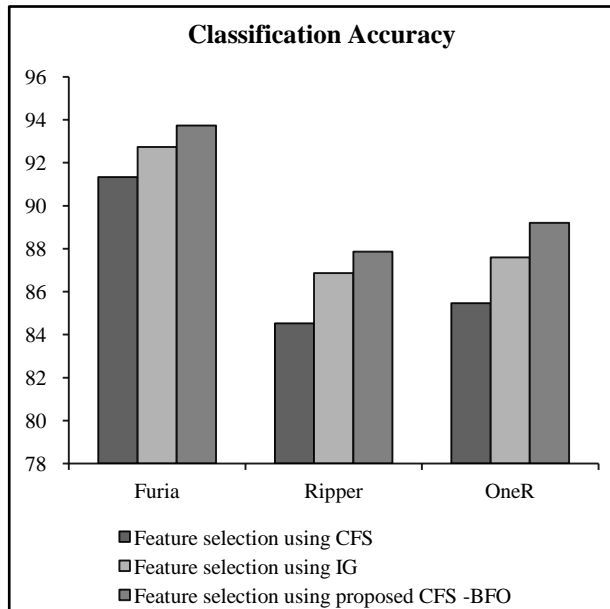


Fig.2. Classification Accuracy

The proposed CFS-BFO in OneR method increased classification accuracy by 4.18% when compared with CFS in OneR. The proposed CFS-BFO in RIPPER method increased classification accuracy by 3.80% when compared with CFS in RIPPER. The proposed CFS-BFO in FURIA method increased

classification accuracy by 2.56% when compared with CFS in FURIA.

Table.3. Precision

Precision	Furia	Ripper	OneR
Feature selection using CFS	0.9170	0.8455	0.8556
Feature selection using IG	0.9281	0.8728	0.8799
Feature selection using proposed CFS -BFO	0.9380	0.8824	0.8918

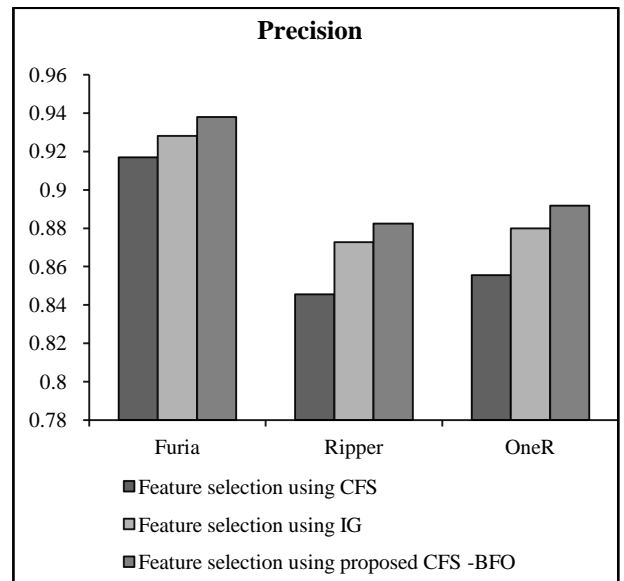


Fig.3. Precision

The proposed CFS-BFO in OneR method increased precision by 4.05% when compared with CFS in OneR. The proposed CFS-BFO in RIPPER method increased precision by 4.18% when compared with CFS in RIPPER. The proposed CFS-BFO in FURIA method increased precision by 2.23% when compared with CFS in FURIA.

Table.4. Recall

Recall	Furia	Ripper	OneR
Feature selection using CFS	0.9133	0.8453	0.8547
Feature selection using IG	0.9273	0.8687	0.8760
Feature selection using proposed CFS -BFO	0.9373	0.8787	0.8920

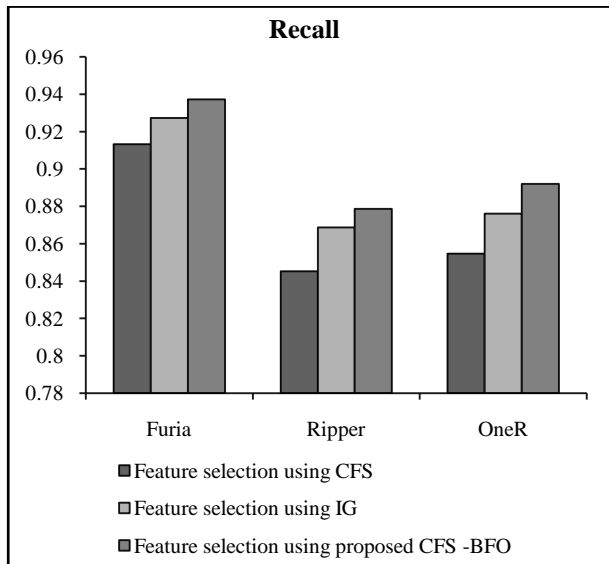


Fig.4. Recall

The proposed CFS-BFO in OneR method increased recall by 4.18% when compared with CFS in OneR. The proposed CFS-BFO in RIPPER method increased recall by 3.80% when compared with CFS in RIPPER. The proposed CFS-BFO in FURIA method increased recall by 2.56% when compared with CFS in FURIA.

Table.5. F-Measure

F-Measure	Furia	Ripper	OneR
Feature selection using CFS	0.9151	0.8454	0.8551
Feature selection using IG	0.9277	0.8707	0.8779
Feature selection using proposed CFS -BFO	0.9376	0.8805	0.8919

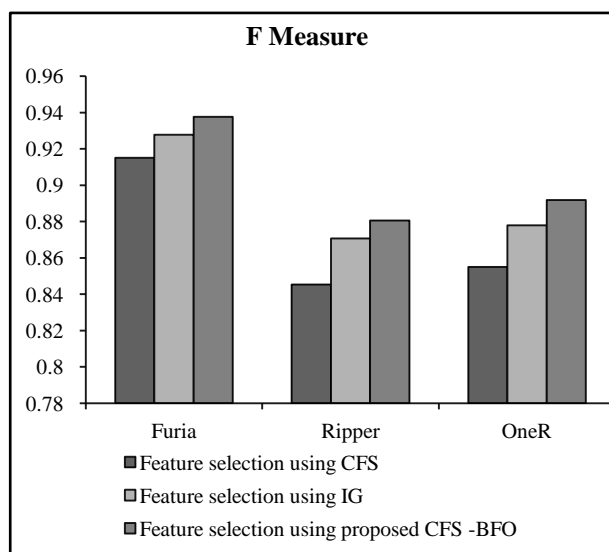


Fig.5. F Measure

The proposed CFS-BFO in OneR method increased F measure by 4.12% when compared with CFS in OneR. The proposed CFS-BFO in RIPPER method increased F measure by 3.98% when compared with CFS in RIPPER. The proposed CFS-BFO in FURIA method increased F measure by 2.39% when compared with CFS in FURIA.

5. CONCLUSION

Medical Image analysis/processing are important in medical diagnoses, especially in non-invasive treatment or clinical study. Imaging helps doctors visualize and analyse images to understand internal structure abnormalities and identify medical conditions. Bio-medical devices' medical image data using imaging techniques like Computed Tomography (CT), Mammogram, and Magnetic Resonance Imaging (MRI) reveal presence or absence of lesion with patient history and is an important diagnostic factor. In this work a novel feature selection mechanism using Bacterial Foraging algorithm is investigated. The proposed techniques improve the classification accuracy compared to Correlation Based Feature Selection and OneR. Further work can be carried out in the direction of improved objective function considering multiple objectives. Investigations can also be carried out using soft computing classifiers.

REFERENCES

- [1] Daniel Keyzers, Christian Gollan and Hermann Ney, "Classification of Medical Images Using Non-Linear Distortion Models", *Bildverarbeitung für die Medizin Springer Berlin Heidelberg*, pp. 336-370, 2004.
- [2] R. J. Ramteke and Monali Y Khachane, "Automatic Medical Image Classification and Abnormality Detection Using K-Nearest Neighbour", *International Journal of Advanced Computer Research*, Vol. 2, No. 4, pp. 186-192, 2012.
- [3] B. Jyothi, Y. Madhavee Latha, P. G. Krishna Mohan, and V. S. K. Reddy, "Medical Image Retrieval Using Moments", *International Journal of Application or Innovation in Engineering & Management*, Vol. 2, No. 1, pp. 195-200, 2013.
- [4] Sanwta Ram Dogiwal, Y.S.Shishodia, Abhay Upadhyaya, Hanuman Ram and Satish Kumar Alaria, "Image Preprocessing Methods in Image Recognition", *International Journal of Computers and Distributed Systems*, Vol. 1, No.3, pp. 96-99, 2012.
- [5] S. K. Patch, "k-Space Data Preprocessing for Artifact Reduction in MR Imaging", *Radiological Society of North America 2005 Scientific Assembly and Annual Meeting*, pp 73-87, 2005.
- [6] Fritz Albreghsen, "Statistical Texture Measures Computed From Gray Level Cooccurrence Matrices", Image Processing Laboratory, Department of Informatics, University of Oslo, 2008.
- [7] Huan Liu and Lei Yu, "Toward integrating feature selection algorithms for classification and clustering", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, No. 4, pp. 491-502, 2005.

- [8] S. Senthamarai Kannan and N. Ramaraj, "A Novel Hybrid Feature Selection via Symmetrical Uncertainty Ranking Based Local Memetic Search Algorithm", *Journal on Knowledge-Based Systems*, Vol. 23, No. 6, pp. 580-585, 2010.
- [9] Luis Talavera, "An evaluation of filter and wrapper methods for feature selection in categorical clustering", *In Advances in Intelligent Data Analysis VI*, Vol. 3646, pp. 440-451, 2005.
- [10] Asha Gowda Karegowda, M. A. Jayaram and A. S. Manjunath, "Feature subset selection problem using wrapper approach in supervised learning", *International journal of Computer applications*, Vol. 1, No. 7, pp. 13-17, 2010.
- [11] S. T. Namin, H. A. Moghaddam, R. Jafari, M. Esmaeil-Zadeh and M. Gity, "Automated detection and classification of pulmonary nodules in 3D thoracic CT images", *IEEE International Conference on Systems Man and Cybernetics (SMC)*, pp. 3774-3779, 2010.
- [12] Tianxia Gong, Shimiao Li, Jie Wang, Chew Lim Tan, Boon Chuan Pang, C. C. T. Lim and Zhuo Zhang, "Automatic labeling and classification of brain CT images", *18th IEEE International Conference on Image Processing*, pp. 1581-1584, 2011.
- [13] Tianxia Gong, Nengli Lim, Li Cheng, Hwee Kuan Lee, Bolan Su, Chew Lim Tan, Boon Chuan Pang and Cheng Kiang Lee, "Finding Distinctive Shape Features for Automatic Hematoma Classification in Head CT Images from Traumatic Brain Injuries", *IEEE 25th International Conference on Tools with Artificial Intelligence*, pp. 242-249, 2013.
- [14] V. Vasconcelos, J. S. Silva, L. Marques and J. Barroso, "Statistical textural features for classification of lung emphysema in CT images: A comparative study", *5th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1-5, 2010.
- [15] M. Chawla, S. Sharma, J. Sivaswamy and L. T. Kishore, "A method for automatic detection and classification of stroke from brain CT images", *IEEE 2009 Annual International Conference on the Engineering in Medicine and Biology Society*, pp. 3581-3584, 2009.
- [16] B. Shenanigan and H. Pourghassem, "Automatic brain hemorrhage segmentation and classification in CT scan images", *8th Iranian Conference on Machine Vision and Image Processing*, pp. 467-471, 2013.
- [17] D. Selvathi, C. Malini and P. Shanmugavalli, "Automatic segmentation and classification of liver tumor in CT images using adaptive hybrid technique and Contourlet based ELM classifier", *International Conference on Recent Trends in Information Technology*, pp. 250-256, 2013.
- [18] Po - Whei Huang, Phen - Lan Lin, Cheng - Hsiung Lee and C. H. Kuo, "A classification system of lung nodules in CT images based on fractional Brownian motion model", *International Conference on System Science and Engineering*, pp. 37-40, 2013.
- [19] B. Sharma and K. Venugopalan, "Classification of hematomas in brain CT images using neural network", *International Conference on Issues and Challenges in Intelligent Computing Techniques*, pp.41-46, 2014.
- [20] L. Papp, N. Zsoter, C. Loh, B. Ole, B. Egeler, I. Garai, and U. Luetzen, "Automated lymph node detection and classification on breast and prostate cancer SPECT-CT images", *IEEE Annual International Conference on the Engineering in Medicine and Biology Society*, pp. 3431-3434, 2011.
- [21] Zhao Liang, Li Changhua, Chen Dengfeng and Dang Faning, "Based on Modified Metropolis Dynamics Concrete CT Image Classification", *1st International Conference on Information Science and Engineering*, pp. 1180-1182, 2009.
- [22] Chen Jialin and He Huacan, "Feature extraction and classification of brain CT image based on sound logic", *2nd IEEE International Conference on Broadband Network & Multimedia Technology*, pp. 868-872, 2009.
- [23] Yong Xia, S. Eberl and Dagan Feng, "Dual-modality 3D brain PET-CT image segmentation based on probabilistic brain atlas and classification fusion", *17th IEEE International Conference on Image Processing*, pp. 2557-2560, 2010.
- [24] Zhiyun Xue, Sameer Antani, L. Rodney Long, Dina Demner-Fushman, D and George R. Thoma, "Classification of CT Figures in Biomedical Articles Based on Body Segments", *IEEE International Conference on Healthcare Informatics*, pp. 264-268, 2013.
- [25] Xiabi Liu, Ling Ma, Li Song, Yanfeng Zhao, Xinming Zhao and Chunwu Zhou, "Recognizing Common CT Imaging Signs of Lung Diseases through a New Feature Selection Method based on Fisher Criterion and Genetic Optimization", *IEEE Journal of Biomedical and Health Informatics*, Vol. 19, No. 2, pp. 635- 647, 2014.
- [26] S. Liao and D. Shen, "A feature-based learning framework for accurate prostate localization in CT images", *IEEE Transactions on Image Processing*, Vol. 21, No. 8, pp. 3546-3559, 2012.
- [27] Jian Wu Xu and K. Suzuki, "Computer-aided detection of hepatocellular carcinoma in hepatic CT: False positive reduction with feature selection", *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1097-1100, 2011.
- [28] Christopher Torrence and Gilbert P. Compo, "A practical guide to wavelet analysis", *Bulletin of the American Meteorological Society*, Vol. 79, No. 1, pp. 61-78, 1998.
- [29] D. L. Hartmann, "Wavelets", Notes, Department of Atmospheric Sciences, University of Washington, pp. 204-221, 2014.
- [30] Dhaval Shah, "Wavelets and Filter Banks Theory and Design", Available at <https://www3.nd.edu/~jkliwer/EE67003/project2/WaveletFilterbank.pdf>
- [31] Nitish Zulpe, N and Vrushen Pawar, "GLCM textural features for brain tumor classification", *International Journal of Computer Science Issues*, Vol. 9, No. 3, 2012.
- [32] P. Mohanaiah, P. Sathyanarayana and L. GuruKumar, "Image Texture Feature Extraction Using GLCM Approach", *International Journal of Scientific and Research Publications*, Vol. 3, No. 5, 2013.
- [33] M. M. Mokji and S. A. R. Abu Bakar, "Gray level co-occurrence matrix computation based on haar wavelet",

- IEEE Conference on Computer Graphics, Imaging and Visualisation*, pp. 273-279, 2007.
- [34] Shay Cohen, Gideon Dror and Eytan Ruppim, "Feature selection via coalitional game theory", *Neural Computation*, Vol. 19, No. 7, pp. 1939-1961, 2007.
- [35] M. Vasantha, V. Subbiah and R. Dhamodharan, "Medical image feature, extraction, selection and classification", *International Journal of engineering science and technology*, Vol. 2, No. 6, 2010.
- [36] Sunitha Beniwal and Jitender Arora, "Classification and feature selection techniques in data mining", *International Journal of Engineering Research and Technology*, Vol. 1, No. 6, pp. 1-6, 2012.
- [37] Danny Roobaert, Grigoris Karakoulas and Nitesh V. Chawla, "Information gain, correlation and support vector machines", *Feature Extraction*, Vol. 207, pp. 463-470, 2006.
- [38] Asha Gowda Karegowda, A. S. Manjunath and M. A. Jayaram, "Comparative study of attribute selection using gain ratio and correlation based feature selection", *International Journal of Information Technology and Knowledge Management*, Vol. 2, No. 2, pp. 271-277, 2010.
- [39] Ajay Kumar and David Zhang, "Biometric recognition using feature selection and combination", *Audio- and Video-Based Biometric Person Authentication*, Vol. 3546, pp. 813-822, 2005.
- [40] S. Narendhar and T. Amudha, "A Hybrid Bacterial Foraging Algorithm For Solving Job Shop Scheduling Problems", *International Journal of Programming Languages and Applications*, Vol. 2, No. 4, pp. 1-11, 2012.
- [41] T. Datta, I. S. Misra, B. B. Mangaraj and S. Imtiaj, "Improved adaptive bacteria foraging algorithm in optimization of antenna array for faster convergence", *Progress In Electromagnetics Research C*, Vol. 1, pp. 143-157, 2008.