Evaluation and Classification of the Brain Tumor MRI using Machine Learning Technique

R. Pugalenthi*, M.P. Rajakumar*, J. Ramya*, V. Rajinikanth**

* *Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai 600 119, Tamilnadu, India. (e-mail: rpugalsir@gmail.com)*

***Department of Electronics and Instrumentation Engineering, St. Joseph's College of Engineering, Chennai 600 119, Tamilnadu, India (e-mail: rajinikanthv@stjosephs.ac.in).*

Abstract: The proposed work implements a Machine-Learning-Technique (MLT) to evaluate and classify the tumor regions into low/high grade based on the analysis carriedout with the brain MRI slices. The MLT implements a sequence of procedures, such as pre-processing, post-processing and classification procedures. The pre-processing enhances the tumor section based on Social Group Optimization (SGO) algorithm assisted Fuzzy-Tsallis thresholding. The robustness of the proposed thresholding is also confirmed by considering the noise corrupted MRI slices. The post-processing implements the Level-Set Segmentation (LSS) to mine the tumor region. The performance of the LSS is validated with segmentation procedures, like Active-Contour (ACS) and Chan-Vese (CVS) technique. The fundamental data of the tumor section is then extracted using the Gray Level Co-occurrence Matrix (GLCM) and most dominating features are then chosen with a statistical test. Finally, a two-class classifier is implemented using the Support Vector Machine with Radial Basis Function (SVM-RBF) kernel and its performance is then validated with other classifiers, like the Random-Forest and k-Nearest Neighbor. The outcome of the proposed work confirms that, implemented tool with the SVM-RBF helps to achieve an accuracy of >94% on the benchmark BRATS2015 database.

Keywords: Brain tumor, Social group optimization, Fuzzy-Tsallis thresholding, Level-Set segmentation, SVM-RBF classification.

1. INTRODUCTION

Brain Cancer (BC) is habitually begins due to various reasons and which can be identified by inspecting the brain sections during the medical screening process. During the screening process, Magnetic Resonance Imaging (MRI) techniques are normally considered to record the various sections of the brain and later it can be reconstructed as a three-dimensional (3D) image. The recorded brain sections are then analyzed using the dedicated imaging procedures, in 3D form or the 2D form. Evaluation of brain MRI in 3D form is more complex and it may augment the computation complication and the evaluation time. Further, extracting the essential data from 3D image is quite complex compared to the 2D case. Hence, in recent years, most of the brain MRI evaluation techniques are implemented by considering the 2D slices (Rajinikanth et al., 2017; 2017a; Roopini et al., 2018; Rajinikanth et al., 2018).

The brain MRI is a commonly adopted technique in the medical field to record the brain structures in different views, like the axial, coronal and sagittal. Further, the MRI also supports varied modalities, like T1, T1C, T2, Flair, DW, and fMRI (Pereira et al., 2016; Menze et al., 2015; Kistler et al., 2013; Tian et al., 2018). Every modality has its own merit and demerits and in the computerized examination purpose; the MRI with T2 and Flair modalities are widely adopted by the researchers due to its better contrast and visibility (Rajinikanth et al., 2018a).

In recent years, a collection of semi- and fully-automated bio-image examination approaches are discussed and implemented by the researchers (Gudigar et al., 2019; Raja et al., 2018; Subudh et al., 2018; Talo et al., 2019; Gudigar et al., 2019; Mohsen et al., 2018). These techniques can be adopted in diagnosing clinics and doctors to acquire the vital information during the brain image evaluation process. The early detection and examination of the brain irregularity also helps to decrease the morbidity and mortality levels. Semi-automated computerized tools are extensively implemented to asses a variety of medical images due to its accuracy (Pereira et al., 2016; Menze et al., 2015). The automated methods are also implemented to enhance the outcome during the medical image evaluation (Gudigar et al., 2019; Mohsen et al., 2018).

Due to its clinical importance, a considerable number of Machine-Learning-Technique (MLT) and Deep-Learning (DL) procedures are employed to assess the brain MRI to identify a class of cognitive disorders. Implementation of the DL technique is computationally complex and it requires a large amount of the dataset to train, test and validate the proposed technique. Hence, the MLT is normally considered to examine the less complex database with greater examination accuracy.

In the proposed work, a MLT is implemented to examine the 2D brain pictures of T2 and Flair modality. This MLT implements a series of evaluation procedures to attain a better result during the low and high grade class examination. The thresholding techniques are widely adopted in image processing applications to enhance the visibility based on a chosen threshold value. The methods, such as Otsu, Kapur, Tsallis, Fuzzy-Tsallis, and Renyi based thresholding methods are already implemented by the researchers to improve the visibility of classical and medical pictures (Rajinikanth and Couceiro, 2015; Satapathy et al., 2018; Sezgin and Sankar, 2004). This work initially implements a pre-processing technique based on the Fuzzy-Tsallis entropy (Rajinikanth and Satapathy, 2018). The post-processing practice usually implements a mining methodology to mine the required segment from the test picture. A significant amount of automated and semi-automated evaluation techniques, such as watershed method, level-set, active-contour, Chan-Vese, region-growing, and morphology based techniques are implemented to extract the segment from the threshold and conventional images (Rajinikanth and Satapathy, 2018).. The classifiers are normally adopted by the researchers to classify the medical images and pictures into a two or multiclass groups according to the vital data extracted from the test cases. For medical image evaluation task, the essential details, like the texture and the shape features are normally adopted to extract the vital data with the help of chosen approaches (Chandra and Gupta, 2011; Acharya et al., 2012). The unprocessed feature size is normally large and implementing the classifier based on the raw features are more complex and time consuming. Hence, a feature selection practice is to be employed to recognize the prime features and the recognized features are then considered to train and test the classifier systems (Acharya et al., 2012). The classifiers, such as Support-Vector-Machine (SVM) with various kernels, Random-Forest and k-Nearest Neighbor techniques are widely employed and the performances are evaluated with various image datasets (Bousseta et al., 2016; Lin and Jeon, 2006; Altman, 1992). The performance of the implemented approaches is evaluated according to the classifier accuracy.

The work of the paper implements a Social-Group-Optimization (SGO) assisted Fuzzy-Tsallis thresholding (SGO+FTT) to improve the considered MRI slices. The tumor fragment from the improved picture is then mined using a semi-automated segmentation called the Level-Set (LSS) method. The features of the tumor section are then mined with the GLCM technique and t-test based feature selection is then implemented to identify the prime feature-set from the extracted features. These features are then considered to train and test the SVM-RBF classifier system employed to classify the considered dataset. In this work, separate evaluation techniques are employed to validate the performance of the post-processing and classifier sections. The overall outcome of the proposed MLT confirms that, proposed technique helps to achieve a classification accuracy of >94% with the adopted BRATS2015 database.

2. RELATED EARLIER WORKS

Examination of the brain abnormality based on the imaging technique is a key research domain and a considerable number of the research works are proposed with the benchmark as well as clinical image dataset.

Traditional, Machine-Learning (ML) and Deep-Learning (DL) based brain tumor examination with 2D MRI slices is normally adopted by the researchers, because of its simplicity (Rajinikanth et al., 2017; 2017a; Roopini et al., 2018; Rajinikanth et al., 2018). Most of the traditional methods need an experienced doctor to recognize the tumor severity and orientation to implement a suitable treatment process to cure the patient. The traditional techniques are time consuming and less accurate. Further, the availability of modern data analyzing procedures will motivate the researchers to develop the ML and DL approaches to inspect the abnormalities in brain sections.

Implementation of the ML with traditional and heuristic algorithm supported technique is quite simple and it requires a small computation cost. The works of (Rajinikanth et al., 2017; 2017a; 2018) implemented a heuristic-algorithm assisted thresholding and segmentation procedures to mine the tumor section from the benchmark 2D brain MRIs. The results of the above works confirms that, the heuristic algorithm with a chosen thresholding and segmentation technique will helps to achieve better image similarity values. (Roopini et al., 2018) executed fuzzy thresholding and Level-Set Segmentation (LSS) to mine the tumor section with greater accuracy. The recent work of (Satapathy and Rajinikanth, 2018) implements a hybrid image examination tool to mine the tumor section from the brain MRI with and without the skull section. The skull stripping is an essential task in MRI with the skull section. The tumor extraction from the MRI with the skull is a challenging task, since the skull and the tumor threshold level is approximately similar in most of the cases. The threshold of the edema will be lesser than the tumor/skull. Hence, recently a number of methods are proposed and implemented by the researchers to eliminate the skull in MRI. The skull section elimination with a thresholding filter can be found in (Raja et al., 2018). After removing the skull section, the ML techniques can be easily executed with the brain MRIs and possibilities of getting better accuracy is more (Gudigar et al., 2019; Subudhi et al., 2018). Implementation of the machine learning requires very less computational cost compared to the DL approaches (Pereira et al., 2016).

The DL practices are the recent advancements in the medical image domain, in which a pre-trained or a customized Convolutional-Neural-Network (CNN) methodology is implemented to support a two/multi-class separation. The work of (Pereira et al., 2016) implements a CNN model to categorize the brain MRI with a greater accuracy. The work of (Talo et al., 2019) implements a transfer-learning procedure to classify the brain MR images. The work by (Mohsen et al., 2018) also implements a DL technique to classify the tumor sections. The other works, such as (Paul et al., 2017; Ari and Hanbay, 2018; Amin et al., 2018) implemented a DL technique to classify the tumor sections with the MRIs and better classification accuracies are obtained. The work by (Amin et al., 2018) considers the BRATS2013 dataset and they implemented a big-data based analysis to achieve better classification accuracy.

From the earlier works, it can be noted that, implementing the DL approach in medical domain is computationally complex and it requires a dedicated execution and processing device. But the ML technique is quite simple and it requires only a fewer tasks to achieve better result. Further, the ML accuracy can be improved with continuous learning. Hence, in the proposed work, a MLT is planned to classify the considered brain tumor dataset into the low/high grade tumor with greater accuracy. The overall computation time and the accuracy are quite considerable compared to the existing DL techniques.

3. METHODOLOGY

The superiority in the semi- or automated-disease assessment system depends on the procedures2 implemented to extraction and evaluation of the image related data. The system implemented in this work is depicted in Fig 1.

Firstly, the existing 3D MRI of T2 and Flair modalities and the related Ground-Truth (GT) are converted in to 2D slices using the (ITK-Snap, 2006; 2018). The 2D slices are then grouped into low and high class and are separately evaluated with the proposed tool. The Social Group Optimization (SGO) assisted Fuzzy-Tsallis thresholding (SGO+FTT) is employed to enhance the tumor section by dilating the other brain sections. The improved tumor segment is then mined with the semi-automated segmentation procedure known as Level-Set (LSS) method. The performance of the initial processing is then validated with a comparative assessment between the mined tumor and the GT and the similarity parameters are then computed. Later, the texture and the area related data are then extracted with the GLCM practice. A feature selection practice to select the dominant feature-set is executed and the selected features are then considered to train and test the classifiers. The SVM-RBF classifier is initially employed to categorize the BRATS2015 database into low and high grade and its performance is then validated with the RF and KNN classifiers existing in the literature. The outcome of the proposed MLT confirms that, implemented approach offers better result on the adopted database.

3.1 Pre-processing

This is the initial stage in proposed MLT, which employs SGO+FTT approach to enhance the test picture. It is a well known soft-computing methodology, which already executed and validated on Ischemic-Stroke-Database (ISD) by (Rajinikanth and Satapathy, 2018).

The SGO is a swarm technique proposed by (Satapathy and Naik, 2016) and it requires a very few algorithm parameters to be optimized.

Mathematical representation of the SGO is presented as;

$$G_{bestj} = \text{maximize } \{ F(X_a) \text{ for } a = 1, 2, \ldots q \} \quad (1)$$

where X_i=introductory information of persons in a group, $a = 1, 2, \ldots q$ denote entire persons of group, and F=fitness function.

The enhancing phase is considered to update locations every individuals as in Eq. (2):

$$X_{new\,a,b} = \Im * X_{old\,a,b} + Rand * (G_{best\,b} - X_{old\,a,b}) \quad (2)$$

where X_{new}=efficient location, X_{old}=early location, G_{best}=united peak location, Rand=arbitrary digit [0,1] and $\Im$ =self-introspection constraint fixed as 0.2 (Naik et al., 2016).

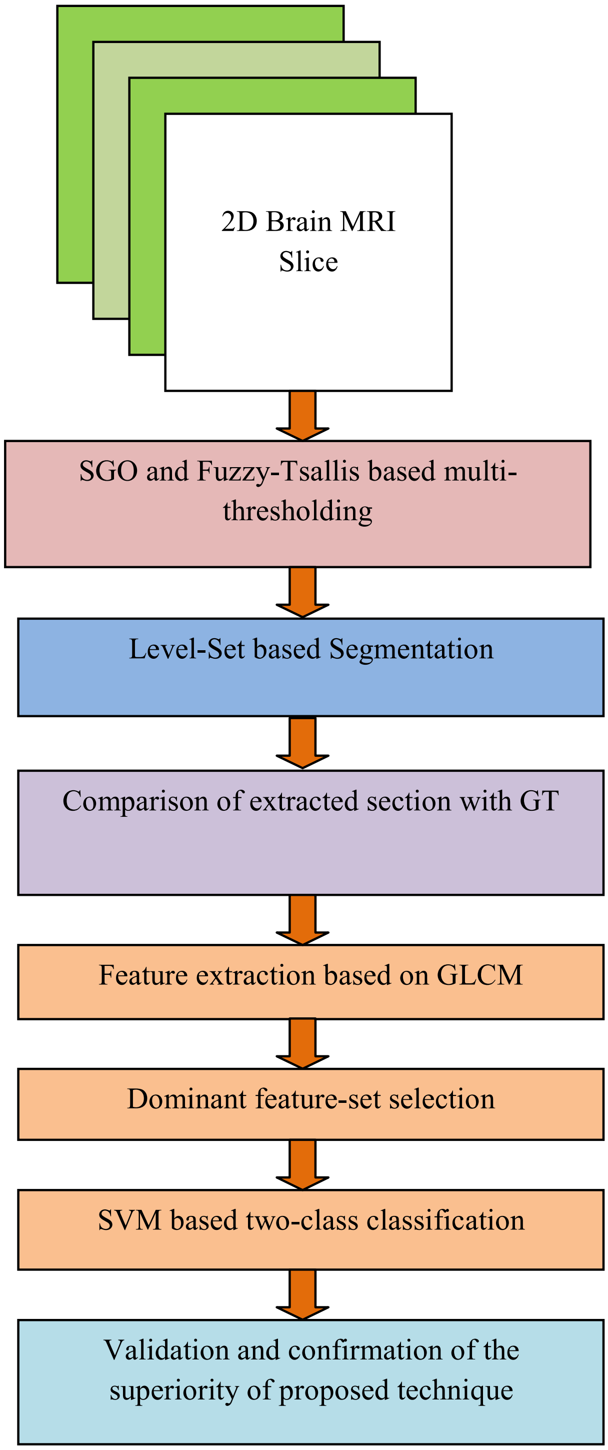


Fig. 1. Structure of planned system.

At improving stage, individuals will decide G_{best} based on information renewal as in Eq. (3).

$$X_{new\,a,b} = X_{old\,a,b} + Rand_1 * (X_{a,b} - X_{r,b}) + Rand_2 * (G_{best\,b} - X_{a,b}) \quad (3)$$

where Rand_1 and Rand_2 are random values and $X_{r,b}$ =subjectively chosen position of an individual in group. Detailed description and working of SGO can be found in (Dey et al., 2018).

The SGO is considered to choose the optimal threshold which maximizes the Fuzzy-Tsallis Entropy (FTE). Consider a test-image P with a dimension of MxN pixels and T thresholds.

Then, the test-image is represented by: $P = \{(c,d) : c = 0,1,...,M-1; d = 0,1,...,N-1\}$ with threshold allocation as $k = 0,1,...,L-1$.

Let, $P_g(c,d)$ is the gray-levels at pixels (c, d), then;

$$P_k = \{(c,d) : P_g(c,d) = k, (c,d) \in P\} \quad (4)$$

If a bi-level threshold (T=2) is assumed, then image can be separated as background and object according to pixel density according to its probability density function. Other related on the FTE based thresholding can be found in (Satapathyand Naik, 2016; Tang et al., 2008; Sadek and Al-Hamadi, 2015; Sarkar et al., 2014; Tsallis, 1988).

3.2 Post-processing

Post-processing is normally adopted to mine the essential information from the enhanced test image. This work implements a semi-automated Level-Set segmentation (LSS) technique to mine the tumor section and then a comparative appraisal among the mined tumor segment and the Ground-Truth (GT) is then performed to confirm the performance of the LSS approach (Li et al., 2010; Vaishnavi et al., 2014). The LSS is then validated against the related semi-automated methods, like Active-Contour (ACS) and Chan-Vese (CVS) segmentation procedures (Bresson et al., 2007; Rajinikanth et al., 2017b; Chan and Vese, 2001). The LSS is initiated with a bounding-box and the box is allowed to converge towards the tumor section. This convergence will identifies all possible similar pixel groups of the tumor segment and the LSS stops its convergence when the energy function of LSS reaches a minimal value. Similar technique is implemented in ACS and CVS and the details of LSS, ACS and CVS can be available in (Rajinikanth et al., 2017; 2017a; Roopini et al., 2018; Rajinikanth et al., 2018).

3.3 Classifier implementation

Two or multi-class classifier plays a major role in the MLT and a number of earlier works confirms the need of the classifiers. The earlier works also confirms that, the SVM algorithm with the Radial-Basis-Function (SVM-RBF) offers better result compared to SVM with other kernels. Further, the categorization accuracy also depends on the feature extracting actions employed to mine the important features from the MRI and the feature selection procedure adopted to recognize the dominant features. This work employs the GLCM practice to extract the texture and the area related features from the tumor segment. Later, a two-tailed t-test feature selection technique is employed to select the leading features based on the p-value attained during the statistical test. The feature selection further helps to reduce the complexity during the training and testing the SVM-RBF classifier and helps to achieve better classification accuracy.

3.3.1 GLCM feature extraction

GLCM features are one of the widely considered texture and shape features for the classification task. The GLCM is also known as the Haralick features and was initially invented in 1973 (Haralick et al., 1973). This practice helps to extract nearly 25 number of texture (Wang and Ren, 2014; Gebejes et al., 2016) and 7 number of shape features such as area, centroid, major-axis, minor-axis, eccentricity, diameter and perimeter as discussed in (Lakshmi et al., (2016). The total number of features adopted in this work is 32 and training of the chosen classifier with this technique is considerably a larger data and it may increase the training time. Hence, a feature selection technique is to be adopted in this work to reduce the complexity.

3.3.2 Principle feature-set selection

The dominant feature-set from the 32 number of features are then chosen using the two-tailed t-test with a three-fold cross validation (Manickavasagam et al., 2014). During this test, the features are sorted in ascending order according to the p-value attained with the t-test and the features with lesser p-value are to be adopted to form the feature-set. This t-test helped to identify the features, like area, entropy, energy, variance, correlation, sum entropy, homogeneity, and sum average are chosen as the dominant features and the selection procedure is clearly discussed in (Acharya et al., 2012).

3.3.3 SVM-RBF implementation

SVM based classification is widely adopted in the literature to implement a two and multi-class classifier. The earlier works also confirms that, properly chosen and properly tuned SVM can help to achieve a better result compared to other classifiers. The SVM with varied kernels are considered by the researchers to implement the classifier system. In the proposed work, the SVM with RBF kernel is adopted for the classification task and the tuning and implementation of SVM-RBF is well famous among the researchers and it can be found in the recent articles (Li et al., 2014; Wang et al., 2019; Natteshan and Jothi, 2015). The tuning of the SVM-RBF relies on a tuning parameter ‘σ’, which is varied from 0.2 to 1.5 in steps of 0.1 and for the considered dataset, the tuning parameter with a value of 1.3 offered better classification accuracy for the BRATS2015 database. The performance of the SVM-RBF is then confirmed against the classifiers, like RF and KNN existing in (Acharya, 2012; Lin and Jeon, 2006; Altman, 1992).

3.4 Performance Assessment

Evaluation of the performance is essential to confirm the advantage of the proposed MLT. In this work, two separate performance evaluation techniques are proposed to confirm the merit of the proposed technique. The first assessment is considered to confirm the performance of the LSS technique. This is achieved by computing the image similarity parameters, such as Jaccard-Index (JI), Dice (DI), Accuracy (AC), Precision (PR), Sensitivity (SE) and Specificity (SP) based on a relative assessment among tumor-segment and

GT. The second assessment is also computes the AC, PR, SE and SP to confirm the performance of the SVM-RBF. The overall performance of the proposed MLT is confirmed based on the AC attained with the SVM-RBF classifier.

The following equation depicts the adopted performance measures (Lu et al., 2004);

$$JI(P_{GT}, P_{TS}) = P_{GT} \cap P_{TS} / P_{GT} \cup P_{TS} \tag{5}$$

$$DI(P_{GT}, P_{TS}) = 2(P_{GT} \cap P_{TS}) / |P_{GT}| \cup |P_{TS}| \tag{6}$$

$$AC = (T_{+ve} + T_{-ve}) / (T_{+ve} + T_{-ve} + F_{+ve} + F_{-ve}) \tag{7}$$

$$PR = T_{+ve} / (T_{+ve} + F_{+ve}) \tag{8}$$

$$SE = T_{+ve} / (T_{+ve} + F_{-ve}) \tag{9}$$

$$SP = T_{-ve} / (T_{-ve} + F_{+ve}) \tag{10}$$

where, P_{GT} is GT, P_{TS} is tumor segment, T_{+ve}, T_{-ve}, F_{+ve} and F_{-ve} denotes true positive, true negative, false positive and false negative; respectively.

4. RESULT AND DISCUSSIONS

This section presents the experimental outcome of the present work and its discussions. The entire work of the developed MLT is implemented as in Fig 2. Initially, the required 2D slices form the existing 3D MRI database is extracted using the ITK-Snap (Yushkevich et al., 2006) and the 2D slices are then adopted for the investigation. The brain MRI with modalities, like T2 and Flair are considered in this investigation due to its better visibility. Further, the robustness of the proposed technique is also confirmed by introducing a noise signals, such as salt & pepper, Gaussian and Speckle. These signals are considered to degrade the visibility of the 2D brain slices considered in this study. An evaluation of performance degradation with the noise is also discussed and the performance of the proposed MLT is tested on the normal and the noise stained MRIs. This MLT consist a sequence of techniques, such as thresholding, segmentation, feature extraction & selection, classification and validation as shown in Fig 2.

Fig 3 shows the sample 2D MRI slices of Flair and T2 modality considered for the demonstration. Fig 3(a) presents the pseudo name and Fig 3(b)-(d) depicts the slices of Flair, T2 and GT respectively. These are gray-scale pictures with a dimension of 216x160 pixels. The evaluation of gray-scale MRI is less complex than other medical pictures (Rajinikanth et al, 2018). The advantage of the BRATS2015 dataset is that, it provides the MRIs without the skull section. So the MRI slices are directly considered to test the performance of proposed MLT. Later, noise signals are added to the considered slices to check the robustness of implemented procedure. Fig 4 depicts the noise stained images and its impact on the visibility. Fig 4(a) presents the image class, Fig 4(b) and 4(c) presents the image and its histogram. This figure confirms that, the impact of the Gaussian and Speckle noise is high and it degrades the histogram of the picture significantly.

The role of the histogram is essential during the thresholding process and it is always necessary to have a smooth histogram in order to evaluate the image data with the traditional or intelligent technique (Rajinikanth et al., 2017; 2017a). Further, the quality of heuristic algorithm assisted image data analysis and enhancement relies mainly on the chosen image and its histogram. The details on the histogram based examination can be found in (Satapathy and Rajinikanth, 2018; Rajinikanth et al., 2018). The proposed MLT implements an Intelligent Data Analysis (IDA) technique, which helps to enhance the tumor section of the test picture by choosing optimal threshold with SGO+FTT. During this search, SGO is allowed to arbitrarily vary the threshold of the test-picture until the Fuzzy-Tsallis entropy reaches a maximal level. This work employs a three-level thresholding technique discussed in (Rajinikanth and Satapathy, 2018).

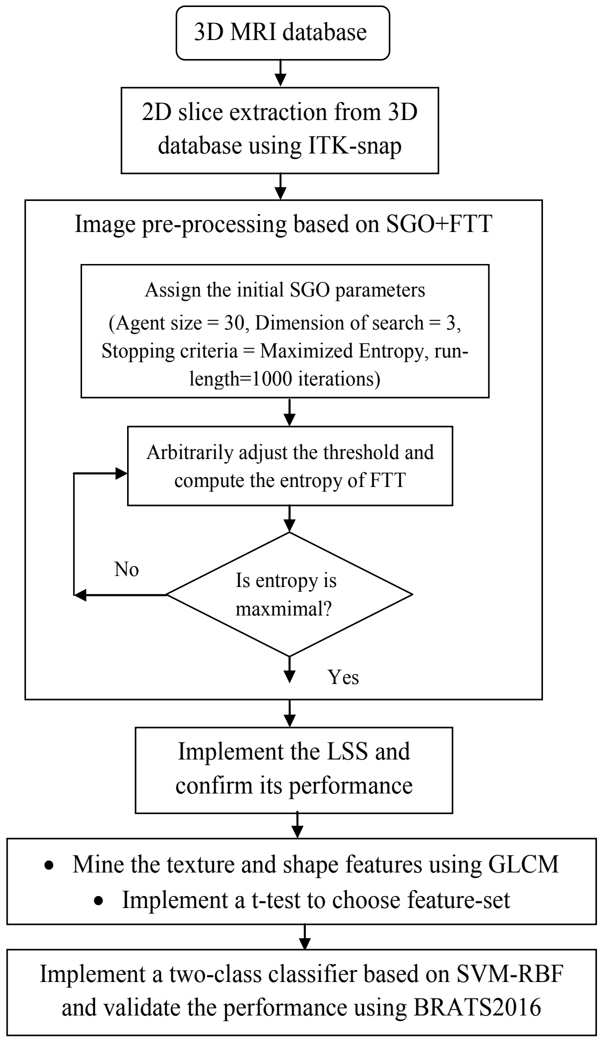
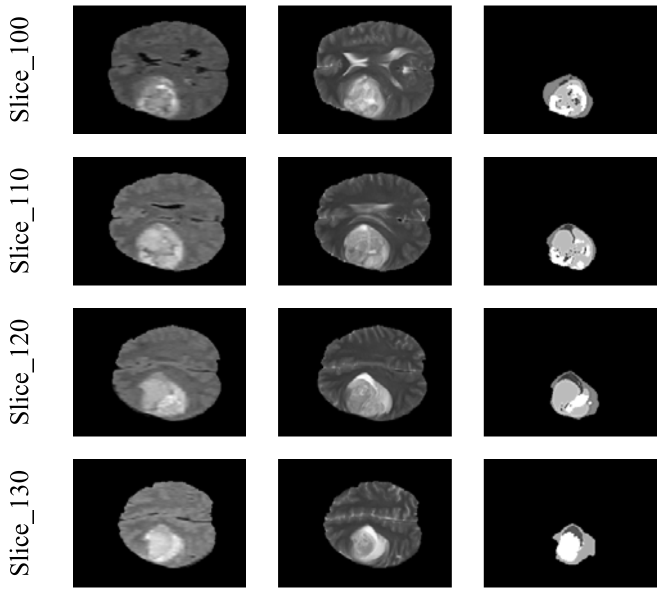
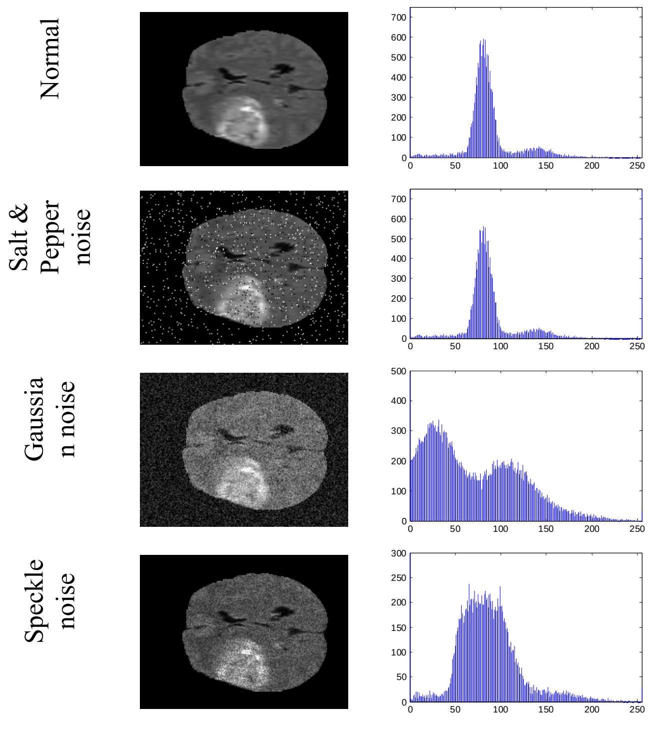


Fig. 2. The structure of the implemented MLT.



(a) (b) (c) (d)

Fig. 3. Sample test images of Brats2015 considered for the demonstration.



(a) (b) (c)

Fig. 4. Pre-processing of the sample test picture.
(a) Image case (b) Gray scale image, (c) Gray histogram

Fig 5 shows the implementation of LSS on the threshold picture. Fig 5(a) presents the outcome of the SGO+FTT and Fig 5(b) to (d) depicts the results on LSS such as initialised bounding-box, converged box with respect to iteration and extracted tumor. Similar work is implemented on the noise stained test pictures and the outcomes are recorded.

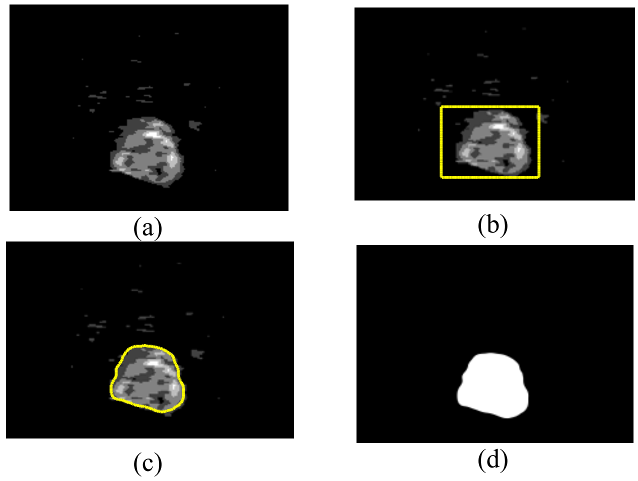
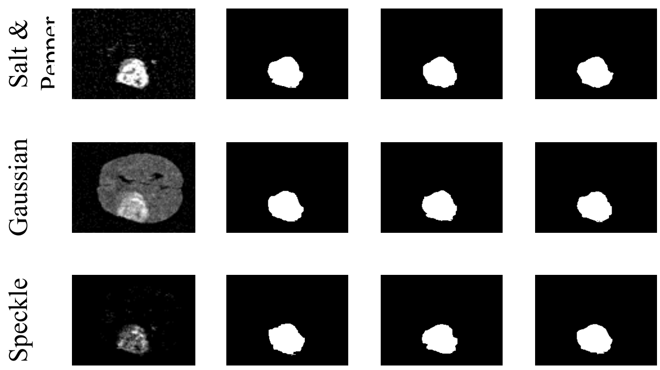


Fig. 5. Results obtained with the LSS.
(a) Pre-processed image, (b) Initial LSS, (c) Converged LSS, (d) Extracted tumor section

Like the LSS, other considered segmentation techniques, such as ACS and CVS based tumor mining is implemented and the results are depicted in Fig 6. Fig 6(a) shows the thresholded noise stained pictures. Fig 6(b) to (d) presents the outcome of LSS, ACS and CVS correspondingly. The performance of LSS, ACS and CVS are then assessed with an analysis between the mined tumor and GT and the essential values are then computed as presented in Table 1 and 2.



(a) (b) (c) (d)

Fig. 6. Experimental outcome of various segmentation procedures.
(a) Sample picture, (b) LSS, (c) ACS, (d) CVS

Table 1. Performance measure attained with the sample images of Figure 6.

Image	Method	T_{+ve}	T_{-ve}	$T_{+verate}$	$F_{-verate}$	$T_{-verate}$	$F_{+verate}$
Salt & Pepper	LSS	2233	32040	0.9197	0.0803	0.9972	0.0028
	ACS	2228	31963	0.8912	0.1088	0.9970	0.0030
	CVS	2186	32036	0.9166	0.0834	0.9957	0.0043
Gaussian	LSS	2272	31986	0.9012	0.0988	0.9984	0.0016
	ACS	2224	32017	0.9107	0.0893	0.9969	0.0031
	CVS	2229	32011	0.9087	0.0913	0.9971	0.0029
Speckle	LSS	2281	32021	0.9142	0.0858	0.9987	0.0013
	ACS	2233	32018	0.9114	0.0886	0.9972	0.0028
	CVS	2214	32028	0.9145	0.0855	0.9966	0.0034

Table 2. Image similarity values attained for Figure 6 images.

Image	Method	JI	DI	AC	PR	SE	SP
Salt & Pepper	LSS	0.8868	0.9400	0.9918	0.9613	0.9197	0.9972
	ACS	0.8586	0.9239	0.9894	0.9591	0.8912	0.9970
	CVS	0.8668	0.9286	0.9903	0.9410	0.9166	0.9957
Gaussian	LSS	0.8834	0.9381	0.9913	0.9780	0.9012	0.9984
	ACS	0.8752	0.9335	0.9908	0.9574	0.9107	0.9969
	CVS	0.8751	0.9334	0.9908	0.9595	0.9087	0.9971
Speckle	LSS	0.8991	0.9469	0.9926	0.9819	0.9142	0.9987
	ACS	0.8791	0.9357	0.9911	0.9613	0.9114	0.9972
	CVS	0.8751	0.9334	0.9909	0.9531	0.9145	0.9966

These tables confirm that, for the chosen image, the proposed technique offers better values of performance measures in the noise stained MRIs and for normal MRI case, these values are superior. Similar procedure is implemented for all the images (75x2=150 low-grade and 75x2=150 high-grade tumours) and the related outputs are recorded for LSS, ACS and CVS cases. Fig 7 confirms that, the % performance (mean value) attained with the LSS is better compared to ACS and CVS.

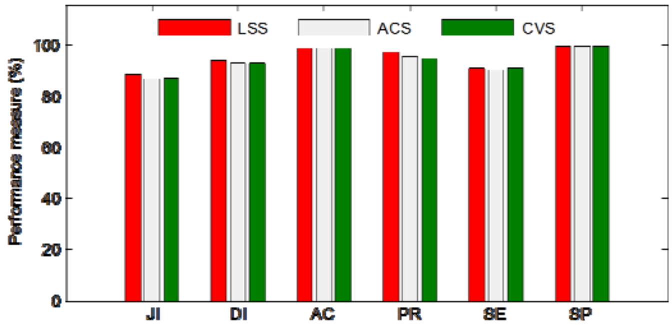


Fig. 7. Average performance measure attained for the considered test images.

Table 3. Dominant feature-set identification with t-test.

Chosen features	Low-grade		High-grade		p-value	t-value
	Mean	SD	Mean	SD		
Area	1396	0.4645	2794	0.5752	6.17E-28	15.9536
Entropy	0.2858	0.0266	0.2935	0.0241	2.84E-25	15.5427
Energy	0.7826	0.0453	0.8622	0.0742	2.37E-23	15.1106
Sum Variance	3.6813	0.1075	3.7049	0.1153	5.62E-21	14.5422
Correlation	0.9405	0.0994	0.9531	0.1185	4.21E-21	13.8427
Sum entropy	0.2881	0.0783	0.2932	0.0958	2.36E-21	13.6385
Homogeneity	0.9217	0.1013	0.9961	0.1074	9.73E-19	13.4482
Sum average	2.1074	0.0253	2.3398	0.0643	4.72E-18	12.9432

After extracting the tumor segment from the considered image dataset (300 MRI slices of T2 and Flair modality), the essential texture and the shape features are then extracted using the GLCM approach (25+7=32 features) and from this feature, 8 dominant features are then chosen with a t-test and the chosen features along with its statistical values are presented in Table 3. These values are then arranged according to its p-value and t-value.

The performance of classifier is based on the data considered during the training and testing task. In this work, 30% of the data (90 images) are adopted to train the classifier and 70% (210 images) are considered to test the classifier. Table 4 presents the classification result obtained with the SVM-RBF for various 'σ' and selected features. This table confirms that, when number of chosen features are 6 and σ=1.3, the classifier performance is superior. For lesser and higher values of the chosen features, the accuracy attained is less. Hence, in the proposed MLT, the SVM-RBF is trained and tested with 6 dominant features along with σ=1.3 is considered. The performance of SVM-RBF is then validated against RF and KNN and the results are depicted in Table 5.This table confirms that, the PR (95.83%) attained with KNN is better than SVM-RBF. But, other performance indicators like, AC, SE and SP obtained with SVM-RBF is better than RF and KNN.

Table 4(a). Performance appraisal of SVM-RBF with different feature dimensions.

σ	No. of Features	T_{+ve}	T_{-ve}	F_{+ve}	F_{-ve}
0.4	2	188	54	38	20
0.7	3	218	19	30	33
0.9	4	220	44	19	17
1.2	5	226	47	10	17
1.3	***6***	***231***	***52***	***11***	***6***
1.4	7	227	48	13	12
1.5	8	228	49	11	12

Table 4(b). Performance appraisal of SVM-RBF with different feature dimensions.

σ	No. of Features	AC (%)	PR (%)	SE (%)	SP (%)
0.4	2	80.66	83.18	90.38	58.69
0.7	3	79.00	87.90	86.85	38.77
0.9	4	88.00	92.05	92.83	69.84
1.2	5	91.00	95.76	93.00	82.46
1.3	***6***	***94.33***	***95.45***	***97.47***	***82.54***
1.4	7	91.66	94.58	94.98	78.69
1.5	8	92.33	95.40	95.00	81.67

Table 5(a). Performance evaluation of adopted classifiers.

Method	No. of Features	T_{+ve}	T_{-ve}	F_{+ve}	F_{-ve}
SVM-RBF	6	231	52	11	6
RF	4	219	47	19	15
KNN	7	230	48	10	12

Table 5 (b). Performance evaluation of adopted classifiers.

Method	No. of Features	AC (%)	PR (%)	SE (%)	SP (%)
SVM-RBF	6	94.33	95.45	97.47	82.54
RF	4	88.67	92.02	93.59	71.21
KNN	7	92.67	95.83	95.04	82.76

From the above results, it can be confirmed that, the proposed MLT works well on the considered BRATS2015 MRI dataset (T2 and Flair modality) and classifies the considered 2D slices into low-grade and high-grade tumor with a classification accuracy >94%. In future, the proposed technique can be tested on the clinical grade MRI slices and brain MRI with the skull section. Further, this technique can also be tested on other modalities, such as T1, T1C, DW, and fMRI. The performance of the SVM classifier can also be validated against, SVM with other kernels, Decision Tree, and Neural-Network classifiers existing in the literature.

5. CONCLUSION

The aim of this work is to develop a Machine-Learning-Technique (MLT) to classify the abnormal 2D brain MRI slices into low/high grade tumors. This MLT employs a chain of actions, such as 2D slice extraction from 3D database, pre-processing with SGO+FTT, segmentation with LSS, GLCM feature extraction, feature selection with two-tailed t-test, training and testing of SVM-RBF, and validation. The chosen 2D slice is initially treated using a three-level thresholding process to enhance the visibility of the tumor section. The enhanced tumor is then mined Level-Set technique. The LSS is a semi-automated approach extracts all the possible pixel of the tumor with a greater accuracy. The texture and shape features (32 features) are then extracted using the GLCM approach and a t-test is implemented to choose 8 prevailing features based on the attained p-value. Finally an SVM-RBF classifier is implemented to classify the considered image dataset in to low/high grade tumor. The robustness of proposed MLT also confirmed by considering the brain MRI slices corrupted with the noise. Further, separate validation actions are implemented to verify the superiority of segmentation procedure and the classifier system. The final outcome attained with the proposed MLT authenticates a classification accuracy of >94%.

REFERENCES

Rajinikanth, V., Satapathy, S.C., Fernandes, S.L. and S. Nachiappan, S. (2017). Entropy based segmentation of tumor from brain MR images–a study with teaching learning based optimization, *Pattern Recognition Letters*, vol. 94, pp. 87-95.

Rajinikanth, V., Raja, N.S.M. and Kamalanand, K. (2017a). Firefly algorithm assisted segmentation of tumor from brain MRI using Tsallis function and Markov random field. *Journal of Control Engineering and Applied Informatics*, vol.19, no.3, pp. 97–106.

Roopini, I.T., Vasanthi, M., Rajinikanth, V., Rekha, M. and Sangeetha, M. (2018). Segmentation of tumor from brain MRI using fuzzy entropy and distance regularised level set, *Lecture Notes in Electrical Engineering*, vol.490, pp.297-304.

Rajinikanth, V., Fernandes, S.L., Bhushan, B. and Sunder, N.R. (2018). Segmentation and analysis of brain tumor using Tsallis entropy and regularised level set, *Lecture Notes in Electrical Engineering*, vol.434, pp.313-321.

Satapathy, S.C. and Rajinikanth, V. (2018). Jaya algorithm guided procedure to segment tumor from brain MRI, *Journal of Optimization,* vol. 2018, pp.12. https://doi.org/10.1155/2018/3738049.

Pereira et al. (2016). Brain tumor segmentation using convolutional neural networks in MRI images, *IEEE Transactions on Medical Imaging*, vol. 35, no.5, pp.1240-1251.

Menze et al., (2015) The multimodal brain tumor image segmentation benchmark (BRATS), *IEEE Transactions on Medical Imaging*, vol.34, no.10, pp.1993-2024.

Kistler et. al, (2013). The virtual skeleton database: an open access repository for biomedical research and collaboration. JMIR.

Tian, Z., Dey, N., Ashour, A.S., McCauley. P. and Shi, F. (2018). Morphological segmenting and neighborhood pixel-based locality preserving projection on brain fMRI dataset for semantic feature extraction: an affective computing study, *Neural Computing and Applications,* vol. 30, no.12, pp. 3733-3748.

Rajinikanth, V., Dey, N., Satapathy, S.C. and Ashour, A.S. (2018a). An approach to examine magnetic resonance angiography based on Tsallis entropy and deformable snake model, *Future Generation Computer Systems*, vol.85, pp.160-172.

Gudigar, A., Raghavendra, U., Ciaccio, E.J., Arunkumar, N., Abdulhay, E. and Acharya, U.R. (2019). Automated categorization of multi-class brain abnormalities using decomposition techniques with MRI images: a comparative study, *IEEE Access*, vol.7, pp.28498-28509.

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

Subudhi, A., Acharya, U.R., Dash, M., Jena, S. and Sabut, S. (2018). Automated approach for detection of ischemic stroke using Delaunay Triangulation in brain MRI images, *Computers in Biology and Medicine*, vol.103, pp.116-129.

Talo, M., Baloglu, U.B., Yıldırım, Ö. and Acharya, U.R. (2019). Application of deep transfer learning for automated brain abnormality classification using MR images, *Cognitive Systems Research*, vol.54, pp.176-188.

Gudigar, A., Raghavendra, U., San, T.R., Ciaccio, E.J. and Acharya, U.R. (2019). Application of multiresolution analysis for automated detection of brain abnormality using MR images: A comparative study, *Future Generation Computer Systems,* vol.90, pp.359-367.

Mohsen, H., Dahshan, E-S.A.E., Horbaty, E-S. M.E. and Salem, A-B. M. (2018). Classification using deep learning neural networks for brain tumors, *Future Computing and Informatics Journal*, vol. 3, no.1, pp. 68-71.

Rajinikanth, V. and Couceiro, M.S. (2015). RGB histogram based color image segmentation using firefly algorithm, *Procedia Computer Science*, vol.46, pp. 1449-1457.

Satapathy, S.C., Raja, N.S.M., Rajinikanth, V., Ashour, A.S. and Dey, N. (2018). Multi-level image thresholding using Otsu and chaotic bat algorithm, *Neural Computing and Applications*, vol. 29, no.12, pp. 1285–1307.

Sezgin, M. and Sankar, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, vol. 13, no.1, pp.146–165.

Rajinikanth, V. and Satapathy, S.C. (2018). Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and Fuzzy-Tsallis entropy, *Arabian Journal for Science and Engineering*, vol.43, no.8, pp.4365-4378.

Chandra, B. and Gupta, M. (2011). An efficient statistical feature selection approach for classification of gene expression data, *Journal of Biomedical Informatics*, vol.44, no.4, pp. 529–535.

Acharya, U.R., Sree, S.V., Ang, P.C.A., Yanti, R. and Suri, J.S. (2012). Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals, *International Journal of Neural Systems,* vol.22, no.2, pp. 1250002.Doi: 10.1142/ S0129065712500025.

Bousseta et al. (2016). EEG efficient classification of imagined hand movement using RBF kernel SVM, In. 11th International Conference on Intelligent Systems: Theories and Applications (SITA), *IEEE*. https://doi.org/10.1109/SITA.2016.7772278.

Lin, Y. and Jeon, Y. (2006). Random forests and adaptive nearest neighbours, *Journal of the American Statistical Association*, vol.101, no.474, pp.578–590.

Altman, N.S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression, *The American Statistician*, vol.46, no.3, pp. 175–185. https://doi.org/10.1080/00031305.1992.10475879.

Paul, J.S., Plassard, A.J., Landman, B.A. and Fabbri, D. (2017). Deep learning for brain tumor classification, Proc. SPIE 10137, Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging, pp.1013710. https://doi.org/10.1117/12.2254195

Ari, A. and Hanbay, D. (2018). Deep learning based brain tumor classification and detection system, *Turkısh Journal Of Electrıcal Engıneerıng Computer Scıences*, vol. 26, pp. 2275-2286.

Amin, J. Sharif, M. Yasmin, M. and Fernandes, S.L. (2018). Big data analysis for brain tumor detection: Deep convolutional neural networks, *Future Generation Computer Systems*, vol. 87, pp.290-297.

Rajini, N.H. and Bhavani, R. (2013). Computer aided detection of ischemic stroke using segmentation and texture features, *Measurement*, vol.46, pp.1865–1874.

Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G. Ho, S., Gee, J.C. and Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability, *NeuroImage*, vol.31, no.3, pp. 1116–1128.

ITK-SNAP (http://www.itksnap.org/pmwiki/pmwiki.php)

Satapathy, S. and Naik, A. (2016). Social group optimization (SGO): a new population evolutionary optimization technique, *Complex & Intelligent Systems*, ol.2, no.3, pp. 173–203.

Naik, A., Satapathy, S.C., Ashour, A.S. and Dey, N. (2016). Social group optimization for global optimization of multimodal functions and data clustering problems, *Neural Computing and Applications*, vol.30, no.1, pp.271-287.

Dey, N., Rajinikanth, V., Ashour, A.S. and Tavares, J.M.R.S.(2018). Social group optimization supported segmentation and evaluation of skin melanoma images, *Symmetry*, vol.10, no.2, pp.51.

Tang, Y., Di, Q., Guan, X. and Liu, F. (2008). Threshold selection based on Fuzzy Tsallis entropy and particle swarm optimization, *NeuroQuantology*, vol. 6, no.4, pp. 412–419.

Sadek, S. and Al-Hamadi, A. (2015). Entropic image segmentation: a fuzzy approach based on Tsallis entropy, *International Journal of Computer Vision & Signal Processing*, vol. 5, no.1, pp.1–7.

Sarkar, S., Paul, S., Burman, R., Das, S. and Chaudhuri, S.S. (2014). A fuzzy entropy based multi-level image thresholding using differential evolution, *Lecture Notes in Computer Science*, vol. 8947, pp.386–395.

Tsallis, C. (1988). Possible generalization of Boltzmann–Gibbs statistics, *Journal of Statistical Physics*, vol.52, no.1, pp.479–487.

Li, C., Xu, C. Gui, C. and Fox, M.D. (2010). Distance regularized level set evolution and its application to image segmentation, *IEEE Transactions on Image Processing*, vol.19, no.12, pp. 3243–3254.

Vaishnavi, G., Jeevananthan, K., Begum, S.R. and Kamalanand, K. (2014). Geometrical analysis of schistosome egg images using distance regularized level set method for automated species identification, *Journal of Bioinformatics and Intelligent Control,* vol.3, pp.147–152.

Bresson, X. Esedoḡlu, S. Vandergheynst, P. Thiran, J-P. and Osher, S. (2007). Fast global minimization of the active contour/snake model, *Journal of Mathematical Imaging and Vision*, vol. 28, no.2, pp.151–167.

Rajinikanth, V., Raja, N.S.M., Satapathy, S.C. and Fernandes, S.L. (2017b). Otsu's multi-thresholding and active contour snake model to segment dermoscopy images, *Journal of Medical Imaging Health Informatics*, vol.7, no.8, pp. 1837–1840.

Chan, T.F. and Vese, L.A. (2001). Active contours without edges, *IEEE Transactions on Image Processing*, vol. 10, no.2, pp. 266–277.

Haralick, R., Shanmugam, K. and Dinstein, I. (1973). Textural feature for image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, vol.3, no.6, pp. 610–621.

Gebejes et al. (2016). Texture characterization based on GLCM, In. 24th Color and Imaging Conference, pp.271-277. https://doi.org/10.2352/ISSN.2169-2629.2017.32.271.

Wang, J-S. and Ren, X-D. (2014). GLCM based extraction of flame image texture features and KPCA-GLVQ recognition method for rotary kiln combustion working conditions, *International Journal of Automation and Computing*, vol.11, no.1, pp.72-77.

Lakshmi, V.S., Tebby, S.G., Shriranjani, D. and Rajinikanth, V. (2016). Chaotic cuckoo search and Kapur/Tsallis approach in segmentation of T. cruzi from blood smear images, *International Journal of Computer Science and Information Security*, vol.14, no.CIC 2016, pp. 51–56.

Manickavasagam, K., Sutha, S. and Kamalanand, K. (2014). Development of systems for classification of different plasmodium species in thin blood smear microscopic images, *Journal of Advanced Microscopy Research*, vol. 9, no.2, pp.86–92.

Li et al. (2014). Classification of EEG signals using a multiple kernel learning support vector machine, *Sensors*, vol.14, pp.12784-12802. https://doi.org/10.3390/s140712784.

Wang et al. (2019). Morphological segmentation analysis and texture-based support vector machines classification on mice liver fibrosis microscopic images, *Current Bioinformatics*, vol.14, no.4, pp. 282-294.

Natteshan, N.V.S. and Jothi, J.A.A. (2015). Automatic classification of brain MRI images using SVM and neural network classifiers, *Advances in Intelligent Systems and Computing*, vol. 320, pp.19-30.

Lu, H., Kot, A.C. and Shi, Y.Q. (2004). Distance-reciprocal distortion measure for binary document images, *IEEE Signal Processing Letters,* vol. 11, no.2, pp. 228–231.