Optimized segmentation of brain images using shuffled frog leaping algorithm -expectation–maximization framework

Sirisha P.G.K.* and Haritha D.
CSE Department, JNTU, Kakinada, INDIA
*sirishapgk@gmail.com

Abstract
Image segmentation has a significant part to play in the understanding as well as interpretation of medical images. Segmentation of medical images, particularly Magnetic Resonance Imaging (MRI) has obtained great deal of interest from research scholars who have suggested various methods, particularly in recent years. A new parametric as well as global image histogram thresholding technique is suggested. It has its basis in the prediction of statistical variables of 'object' as well as 'background' class by Expectation–Maximization (EM) protocol, under the presumption that the 2 classes follow Generalized Gaussian (GG) distribution. In this work an MRI image segmentation on the basis of modified EM protocol with GG distribution is proposed. In the modified Algorithm, Shuffled Frog Leaping Algorithm (SFLA) is used.

SFLA is a memetic meta-heuristic that has its basis in the evolution of memes, performed by interactive individuals as well as global transmission of data amongst the frog population.

Keywords: Magnetic Resonance Imaging (MRI), Image histogram thresholding, Expectation–Maximization (EM), Generalized Gaussian (GG) distribution, Shuffled Frog Leaping Algorithm (SFLA).

Introduction

Magnetic Resonance Imaging (MRI) is vastly utilized because of its good spatial resolution, tissue contrasts as well as non-invasive features. But, 3D images acquisition offers an improved way for diagnosing various disorders as well as for planning precise surgery. [1] Recently, modern methods offer a huge quantity of images to be analyzed. Study as well as evaluation of the images is typically built via visual rating carried out by experts on the basis of their own experiences as well as skills apart from other subjective, time-consuming steps which are fault-prone. Furthermore, the qualitative analysis is restricted by the human visual systems that is not able to differentiate more than several tens of gray levels; as opposed to recent MRI systems that offer till 16-bit quality depth images with 65,535 gray levels. Various segmentation protocols have been suggested as well as grouped into manual methods, semi-automated methods as well as automated methods. The first needs human experts for selecting voxels that belong to a particular object, and is the most famous method which has been utilized for several years.

Segmenting non-trivial images is a hard task within image processing. Segmentation accuracy defines the success or failure of computer analyses processes. Segmentation protocols are region-oriented rather than being pixel-oriented. Outcome of segmentation is the dividing of images into connected regions. Image segmentation is the basic step within image analyses, comprehension as well as interpretation and recognition tasks. Segmentation is the most significant stage in automatic recognition systems that have several applications in the domains of medical imaging, satellite imaging, movement detections, security as well as surveillance.[2, 3]

Image Segmentation splits an image into sets of regions. Regions represent meaningful regions in images or sets of border pixels grouped into a structure like line segment, edge and so on. Segmentation has 2 goals: 1) decomposing images into regions for more analysis and 2) performing change of representations of images for more rapid analyses. Various kinds of segmentation methods are utilized for segmentations. On the basis of the applications, one or a set of segmentation methods may be employed for solving the issue in an effective manner. Segmentation protocols have their basis in the characteristics of grey level values of pixels. Various kinds of segmentation methods include: 1) edge-based segmentation, 2) threshold-based segmentation, 3) region-based segmentation, 4) clustering as well as 5) matching.

Expectation–Maximization (EM) protocol refers to an unsupervised protocol which finds the estimated parameters of maximum likelihood or posterior probability when the data are incomplete. This method provides a systematic approach to find the determining parameters which describes the bias field. With an iteration of the EM protocol, the mixture model variablesas well as the bias field parameters are updated and perform the classification. Because the EM protocol might converge to local maxima or a minimum of a saddle point of likelihood function, correct initialization of unknown variables q is necessary.

Thresholding is a basic task within image processing as well as patterns recognition because it denotes a fundamental stage for image comprehension in various applications. It comprises the conversion of grey level images into binary images through selection of adequate decision thresholds. Grey level values less than or the same as the chosen

threshold are typically sorted as background, whereas the values greater than the threshold are sorted as object. Image thresholding is vastly utilized in several application fields, like bio-medical image analyses and so on.

Typically, thresholding methods on the basis of parametric models are vulnerable to some severe shortcomings (that might be critical in a complex image) like:

1) They have their basis in the Gaussian presumption to model the class distribution in the image that typically does not hold in a real image
2) They typically result in biased estimates of statistical variables of the object as well as background class
3) Their efficacy is highly decreased when prior probabilities of object as well as background class is unbalanced or when 2 classes overlap considerably

For addressing the above mentioned shortcomings, a new parametric as well as global image histogram thresholding method is proposed. It has its basis in the Expectation–Maximization (EM) protocol for estimating the variables of the object as well as background class, that is presumed to follow Generalized Gaussian (GG) distribution.

Optimization is a troublesome issue. Protocols which resolve these types of issues are various. Amongst them, the meta heuristic set of protocols which comprises the stochastic optimization protocols is cited here. SFLA is a novel meta heuristic which mimics the principle of the evolution of a set of frogs which search for discrete locations comprising maximal quantity of food. [4] SFLA merges benefits of particle swarm optimization that owes its inspiration to the swarming activity of species such as fish as well as from genetic algorithm, a research method developed by Holland [5] with features like excellent capability in global searches as well as simple execution. Genetic algorithms mimic the principle of biological evolution. SFLA has demonstrated efficacy in several optimization issues which are hard to resolve through other techniques like water distribution as well as ground water model calibration issues. [6]

Problem Formulation

Assume there is an image X of size M ×N pixels, wherein every pixel x of co-ordinates $(m, n) \in [1, M] \times [1, N]$ can take L possible grey-level values specified within $[0, L-1]$. Assume h(x) is the normalized histogram of the image X (it is assumed that h(x) is the sole information available regarding the image). The histogram h(x) may be observed as an estimate of the true probability density function p(x) of the image. Aimed at thresholding the image X for assigning every image pixel to one of 2 classes, which are object as well as background. The job may be simplified to an issue of choosing of threshold value T which optimizes a predefined condition. When T is calculated, threshold image

$Y = \{ y(m,n), 1 \le m \le M, 1 \le n \le N \}$ may be created through designating the values below:

$$y(m,n) = \begin{cases} 0, & \text{if } x(m,n) \le T \\ 255, & \text{if } x(m,n) > T \end{cases}$$

The above issue may be resolved as a binary classification issue, wherein the probability density function p(x) of the image X is a mix of 2 parametric density functions related to object as well as background class, that is:

$$p(X) = \sum_{i=1}^{2} P_i p_i(x \mid \theta_i)$$

wherein P_1 as well as P_2 are prior probabilities, while $p_1(x \mid \theta_1)$ as well as $p_1(x \mid \theta_2)$ represent the class-conditional densities related to the object as well as background class, correspondingly. 1 as well as 2 are the vectors of variables on which the 2 parametric class-conditional densities rely on. The Generation of the binary classification map Y needs:

- To adopt a statistical model for 2 class-conditional densities;
- An estimate of the statistical parameters associated with the 2 classes (that is, Pi as well as i, with i=1,2); and
- To apply a decision condition

Compared to works in literature which utilize various methods as well as other image kinds is hard because that would necessitate the usage of one dataset by various groups with evaluations carried out by similar metrics. Because of this, solely the outcomes of certain works are provided rather than a detailed comparative study. Though, utilizing manual expert segmentation as the ground truth, various performance metrics like Dice Overlap, Jaccard Index, False Positive as well as Negative Volume Fractions (FPVF, FNVF) were utilized in literature, (Dice) Overlap is utilized as a typical metric for comparing to earlier techniques here.

Literature Survey

Yang et al[7] considered the variable estimation for Linear Parameter Varying (LPV) models with arbitrarily missing output data. Multi-model LPV models are adopted and identification issue is formulated with the strategy of Generalized Expectation–Maximization (GEM) protocol. For dealing with missing output data, local models are initially taken to have Finite Impulse Response (FIR) model structures. For alleviating possible over parameterization issue, a prior on FIR model coefficients is enforced and GEM protocol is altered for deriving Maximum A Posterior (MAP) estimate of multimode LPV FIR model variables. Because FIR models are not adequate for generic control applications, a multimode LPV Output Error (OE) model is found through application of GEM protocol to the same

identification dataset with variables initialized on the basis of the predicted FIR model. A simulation as well as 2 experiments are provided for demonstrating the efficacy of the suggested technique.

Due to the need of accurate analyses of diseases, MRI segmentation persists as a tough issue, particularly in the presence of arbitrary noise. Ladgham et al [8] suggested a new meta heuristic protocol for MRI segmentation called Modified Shuffled Frog Leaping Algorithm (MSFLA), on the basis of the method of SFLA. In this novel paradigm, there is no requirement to filter original image. The novel fitness function suggested in the protocol assists in the rapid evaluation of particles for arranging them in descending order. The suggested method was contrasted with other meta heuristics like 3D-Otsu thresholding with SFLA as well as GA as well as with protocol of segmentation utilizing Rician Classifier (RiCE). Outcomes of experiments prove that the suggested MSFLA is capable of achieving better segmentation quality as well as implementation time than the other techniques.

Subashini et al[9] developed a non-invasive technique for tumor grade identification through usage of MRIs. The procedure includes pre-processing, image segmentations, tumor isolations, features extraction, features selection as well as classification. Analyses on performance of segmentation methods, features extraction techniques, automated features selection (SFLA) as well as built classifiers is carried out based on accuracy, efficacy as well as elapsed time. The analysis is motivated toward correct determination of tumor grade from MRIs rather than dependence on MRS or biopsies. Fuzzy c-means segmentation performed better than other segmentation methods, shape as well as size-based texture attributes promoted the marking of tumor grades. NB classifier was the best with regard to efficacy, error as well as elapsed time in contrast to SVM as well as LVQ. The research was performed with 200 images comprising 164 images in training set as well as 36 images in test set. Outcomes proved that the system was resilient as well as accurate (91%), took lesser time for identifying grades, and that it was an excellent alternate for biopsies or MRS in the diagnostic procedure of identifying brain tumor grades.

Ladgham et al[10] presented a new optimum protocol for MRI brain tumor recognition. For doing this, the novel meta heuristic MSFLA is used. Else, an adequate option of fitness function guarantees faster time of research with more chance of convergence to optimum value. Computation of used fitness functions are related to images. Images are to be scanned for calculating the function. For this, the functions assist in the rapid discovery of suitable region modeling the tumor. Computerized simulation outcomes prove the efficacy of the protocol.

Song et al[11] proposed a novel MRI segmentation technique on the basis of scale-space theory as well as EM protocol.

Fist, grey level density of brain MRIs is predicted; second, the corresponding fingerprint that includes initialization information for EM utilizing scale-space theory is got; finally, segmentation outcomes are obtained through initialized EM. In the initialization stage, constraints of clustering component weights reduce influence of noise or singular point. Brain MRIs segmentation outcomes denote that the technique determines more suitable initialization information and is capable of achieving more correct segmented tissues than the others.

Methodology
In this section, we briefly discuss the various techniques used in the proposed framework.

Expectation–Maximization (EM) Algorithm: EM is an iterative method of maximum likelihood estimation where each iteration has an Estimation (E) step and a Maximization (M) step. The EM algorithm starts with the optimization of the likelihood function /(e; X); step M maximizes the likelihood function which is redefined by step E in each iteration. The likelihood function /(e; X) can be simplified only if the missing variables are known and where X refers to the incomplete data. The probability model P for the missing variables to the actual data is given by (2):

$$P(y, z|x, B) \qquad (2)$$

The logarithm of the density P defines the "complete-data likelihood". The EM algorithm finds the expected value of the complete-data likelihood for the observed data and the current model. The E step is formulated by (3):

$$Q\left(\theta, \theta^{(p)}\right) = E\left[l_c\left(\theta; y\right) | x\right], \qquad (3)$$

Where $\theta^{(p)}$ refers to the value of the parameters of pth iteration and the expectation with respect to $\theta^{(p)}$ yielding a deterministic function Q. The new parameter estimates $\theta^{(p+1)}$ is computed in M step by (4):

$$\theta^{(p+1)} = \arg\max_\theta Q\left(\theta, \theta^{(p)}\right). \qquad (4)$$

The EM iterates to improve estimate of the complete likelihood by E step and increases the value of Q, the expectation of the complete likelihood. [16]

The EM protocol is sensitive to the issue of selecting the initial values of the variables to be predicted. If the initial values are not appropriately chosen, EM protocol might result in inadequate prediction of the class distribution. For addressing the issue, various techniques are presented. Few of those have their basis in the usage of several arbitrary initial conditions for generating several solutions and then choosing the one that yields best probability. [12] Others have their basis in initialization by clustering protocols [13] or under

a tree structure strategy. [14] In the current work, it is proposed to utilize SFLA as an alternate method for the issue of initialization. SFLA represents a popular family of techniques for global optimization which demonstrated itself as being extremely resilient due to:

(i)   At every cycle of the optimization procedure, they hold back huge quantities of potential solutions.

(ii)   Searches are not based on gradient principle and rather directly on the function to be optimized.

Shuffled Frog Leaping Algorithm (SFLA): Typically, when employing SFLA to an optimization problem, every frog possesses a differing solution from the rest as per its adaptability as assessed by fitness function. The total population of frogs is split into a pre-defined quantity of sub-sets named memeplexes. Frogs of every memeplex have their own scheme for exploring the environment in various directios. After a pre-defined set of memetic evolutions, the sharing of information between memeplexes occurs in a process of shuffling. This process ought to guarantee that the evolution towards a certain interval is free from all prejudices. Memetic evolution as well as shuffling are carried out alternatively till fulfilling convergence condition or till a terminating condition is fulfilled. Stages of SFLA are as follows:

Step 1: Initial population: Initial population. $X_i = (i = 1, 2, ..., F)$ of F frogs, wherein individual frogs are analogous to GA chromosomes, is generated arbitrarily.

Step 2: Ranking as well as distribution: Every frog is ranked in descending order on the basis of their fitness value and split into m memeplexes, with every memeplex comprising p frogs (that is, $F = m \cdot p$); the frog which is placed 1st moves to 1st memeplex, the 2nd one to the 2nd memeplex, the pth one to the pth memeplex, and the (p+1)th goes back to the 1st memeplex and so on.

Step 3: Memeplex evolution: In every memeplex, frogs possessing best as well as worst fitness values are detected, respectively, by X_b and X_w. Frog with best fitness in the entire population is marked by global best X_g.

At the time of evolution of memplexes, worst frogs jump toward best frogs through usage of equations (5) as well as (6), that are like PSO ones:

$$S = rand(X_b - X_w)$$
$$IX_w = X_w + S; \ S \prec S_{max}$$
$$(5, 6)$$

Where in S represents jumping step of worst frog, IX_w represents improved worst solution, rand refers to a random number within [0,1], while S_{max} refers to the maximal jump

distance. Equation (1) as well as (2) are iterated for a pre-defined set of iterations for obtaining improved results than X_w. If they do not enhance the worst solution, X_b is substituted with X_g and adapted to Equation (7):

$$S = rand(X_g - X_w)$$
$$(7)$$

If Equations (5) as well as (6) do not enhance the worst solution, a novel position is created randomly.

Step 4: Shuffling: After a set quantity of memeplex evolution phases, every frog of every memeplex is gathered and ranked in descending order once more on the basis of fitness. Step 2 splits frogs into various memeplexes once more, and then step 3 is attained.

Step 5: Terminating criterion: If a pre-defined solution or a set iteration count is arrived at, the protocol ends. Figure 1 displays the pseudo code of SFLA.

```
Start;
Create arbitrary set of P solutions;
For every individual frog in the population:  compute fitness (i);
Rank entire population P in descending order of fitness;
Split P into m memeplexes;
For every memeplex;
Figure out best as well as worst positions;
Enhance the worst individual position
Iterate for a certain quantity of cycles;
End;
Merge the evolved memeplexes;
Rank the population P in descending order of fitness;
Check if terminating condition = true;
End;
```

Figure 1: Pseudo code for SFLA

Initialization process for EM protocol: The option to adopt SFLA [15] to initialize variables of the distribution of classes primarily relies on the earlier mentioned characteristics of the optimization technique. Because the initial values of the statistical variables of classes considerably impact both estimates got through EM protocol at convergence as well as thresholding outcomes, it is necessary to utilize a correct as well as resilient initialization process that is capable of accurately exploring solution space. In SFLA, efficacy in exploration of solution space is fused with the possibility to define evaluatory criteria on the basis of fitness functions which implement the notions taken up both by EM protocol (maximizing log-likelihood function of estimates) as well as by threshold selection protocol (minimizing estimated error probability).

Most basic form of SFLA includes the stages given below: 1) initial population of individuals is created arbitrarily, 2)

excellence of chromosomes is tested as per pre-defined fitness function that permits retaining best memeplex after discarding the worst (better the fitness, greater the likelihood of being chosen). This selection procedure is significant for the subsequent stage that is focused on reproduction of the population. This process is repeated till user-specified convergence condition is arrived at. Figure 2 shows the flowchart of the suggested SFLA -based initialization process.
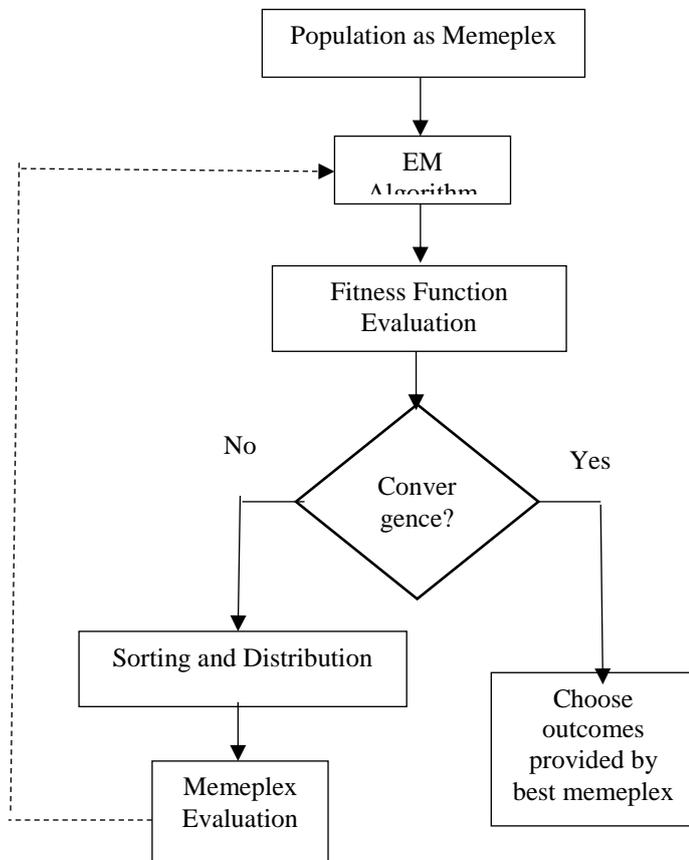
$$D(A,B) = 2 \times \frac{J(A \cap B)}{1 + J(A,B)} \tag{10}$$

Table 1 presents the overview of outcomes. Figure 3 to 8 displays the Total Correct Fraction, Sensitivity of expert and SFLA-EM, Specificity of expert and SFLA-EM, Dice Coefficient and volume error respectively.
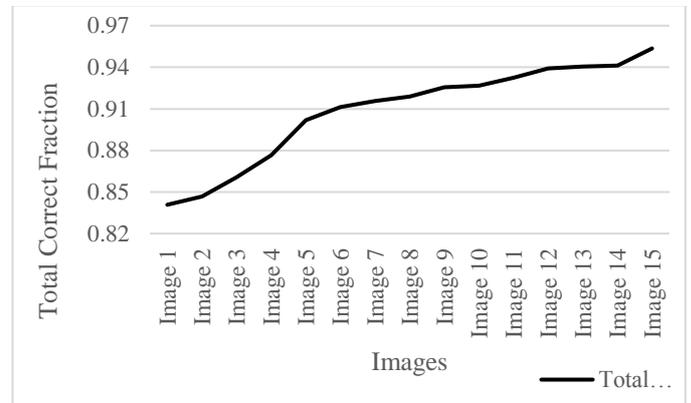

Figure 4: Total Correct Fraction

From figure 4 it can be seen that on an average 91% of the tumor pixels are extracted correctly by the proposed algorithm.
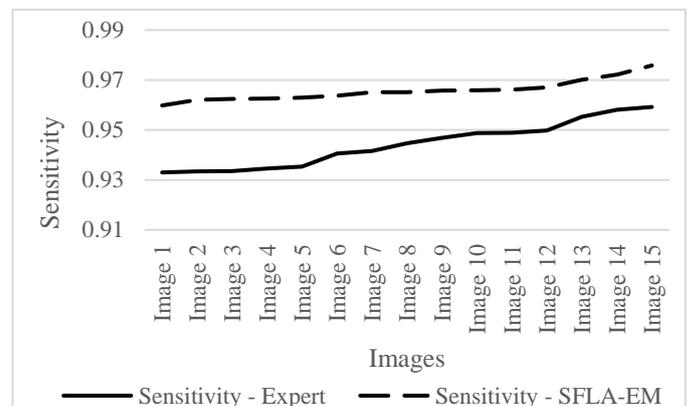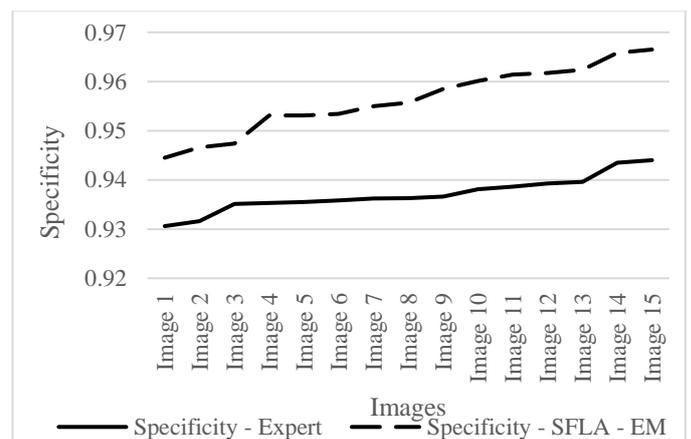

Figure 5: Sensitivity

Figure 2: Flowchart of the suggested SFLA-based initialization process

Results and Discussion

Typically utilized metrics for reporting segmentation overlap results in brain tumour segmentation literature include Dice Overlap (D), Jaccard Index (J), False Positive as well as Negative Volume Fraction (FPVF, FNVF) pairs. Dice Overlap (D) between 2 label sets A as well as B is given in equation (8):

$$D(A,B) = 2 \times \frac{s(A \cap B)}{s(A) + s(B)} \tag{8}$$

Jaccard Index (J) of 2 sets A as well as B is given as in equation (9):

$$J(A,B) = \frac{s(A \cap B)}{s(A \cup B)} \tag{9}$$

And can be transformed to Dice Overlap score through equation (10):

Figure 6: Specificity

Table 1
Summary of Results

	Total Correct Fraction	Sensitivity - Expert	Sensitivity - SFLA-EM	Specificity - Expert	Specificity - SFLA - EM	Dice Coefficient	Volume error
Image 1	0.8408	0.933	0.9598	0.9306	0.9445	87.14	19.66
Image 2	0.8468	0.9334	0.962	0.9316	0.9466	87.64	18.64
Image 3	0.8605	0.9336	0.9624	0.9351	0.9474	87.9	18.19
Image 4	0.8765	0.9345	0.9625	0.9353	0.9531	88.53	17.96
Image 5	0.9019	0.9353	0.9629	0.9355	0.9531	88.96	16.68
Image 6	0.9113	0.9406	0.9637	0.9358	0.9534	89.21	16.26
Image 7	0.9156	0.9416	0.9651	0.9362	0.955	90.06	16.13
Image 8	0.9188	0.9447	0.9651	0.9363	0.9557	91.68	14.42
Image 9	0.9256	0.9468	0.9658	0.9366	0.9585	92.13	13.98
Image 10	0.9267	0.9487	0.9659	0.9381	0.9601	92.16	12.59
Image 11	0.9324	0.9489	0.9661	0.9386	0.9614	92.63	12.51
Image 12	0.9391	0.9498	0.967	0.9393	0.9617	92.82	12.21
Image 13	0.9405	0.9553	0.9701	0.9396	0.9624	92.92	11.72
Image 14	0.9412	0.9581	0.9721	0.9435	0.9658	93.22	11.59
Image 15	0.9536	0.9592	0.9758	0.944	0.9665	93.57	11.36

It is noted in table 1 as well as figure 5 that the sensitivity of SFLA-EM performs better by 2.83% than Expert for image 1. Similarly, for image 15, the sensitivity of SFLA-EM performs better by 1.72% than Expert.

It is observed for table 1 and figure 6 that the specificity of SFLA-EM performs better by 1.48% than Expert for image 1. Similarly, for image 15, the specificity of SFLA-EM performs better by 2.36% than Expert.
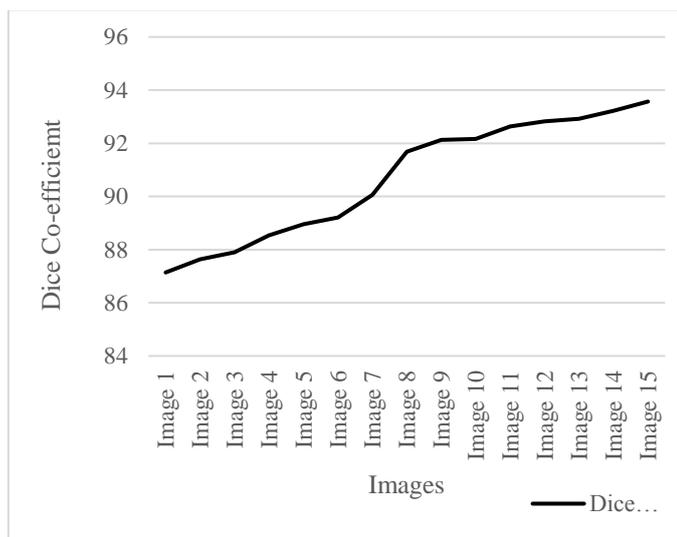
$$Volume\ error = \frac{T_s}{T_a}$$

where

T_s *are the selected pixels and*

T_a *are the actual tumor pixels*

The total correct fraction is given by:

$$Total\ correct\ fraction = \frac{T_t}{T_{tot}}$$

Where T_t are the selected tumor pixels

T_{tot} *are the total tumor pixels*

Figure 7: Dice Co-efficient

From figure 7 it can be observed that congruence is very good and in par with other work in literature. Dice coefficients for the individual tumor sub-regions are relatively higher for all the 15 images.

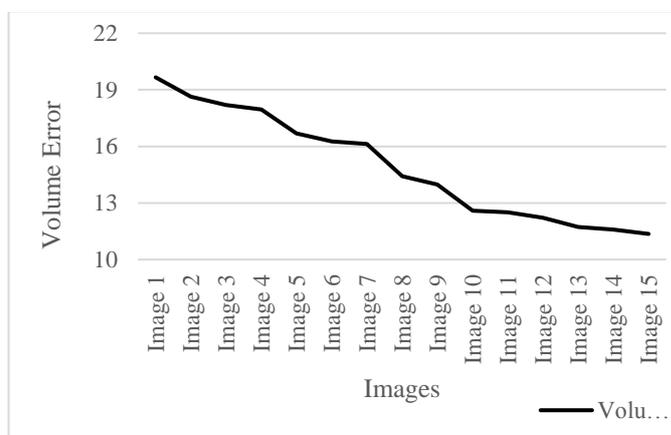Figure 8 shows the volume error obtained for each image. The volume error is computed by

Figure 8: Volume Error

Some of the images output is shown in figure 9.

The first and last image shows ideal segmentation whereas the middle image includes noise. The mean volume error was less than 14.9% and computed using
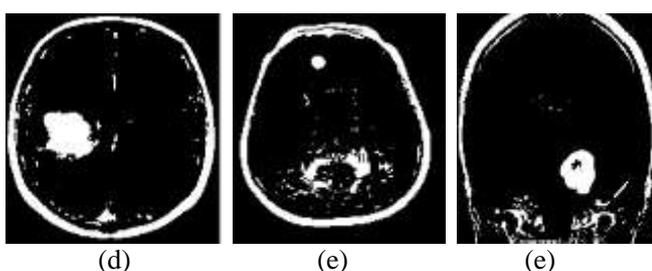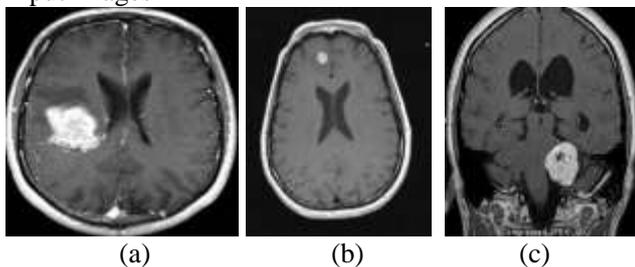
$$Mean\ Volume\ Error = \frac{\sum v_i}{N}\ i=1\ to\ N$$

where

v_i *is the volume error for each image*

N *is the total number of images*

Input images

(a)   (b)   (c)

(d)   (e)   (e)

Output images

Figure 9: Input and output Images

Conclusion

Expectation Maximization (EM) protocol is an unsupervised clustering protocol, however initialization information, particularly the quantity of clusters is critical to its performance. Thresholding method is sensitive to noise as well as intensity homogeneities. For addressing shortcomings of EM, a SFLA based optimization is proposed. SFLA performs searches through evolution of a set of potential individuals modelled with "Memeplex". From one generation to the subsequent one, the set of potential individuals is enhanced through methods which owe their inspiration to genetics, that is, by the usage of both deterministic as well as non-deterministic operators. The proposed SFLA, improves the performance of segmentation.

References

1. Ortiz A., Gorriz J.M., Ramirez J. and Salas-Gonzalez D., Improving MR brain image segmentation using self-organising maps and entropy-gradient clustering, *Information Sciences*, **262**, 117-136 **(2014)**

2. Chaos D., Chacón J., Lopez-Orozco J.A. and Dormido S., Virtual and remote robotic laboratory using EJS, MATLAB and LabVIEW, *Sensors*, **13(2)**, 2595-2612 **(2013)**

3. Sujji G.E., Lakshmi Y.V.S. and Jiji G.W., MRI brain image segmentation based on thresholding, *International Journal of Advanced Computer Research*, **3(1)**, 97-101 **(2013)**

4. Eusuff M.M. and Lansey K.E., Optimization of water distribution network design using the shuffled frog leaping algorithm, *J. Water Resour. Plan. Manag.*, **129(3)**, 210–225 **(2003)**

5. Goldberg D.E., Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA **(1989)**

6. Huynh T.H., A modified shuffled frog leaping algorithm for optimal tuning of multivariable PID controllers, In IEEE International Conference on Industrial Technology, Chengdu, 21–24 April 2008, 1–6 **(2008)**

7. Yang X., Huang B., Zhao Y., Lu Y., Xiong W. and Gao H., Generalized expectation–maximization approach to LPV process identification with randomly missing output data, *Chemometrics and Intelligent Laboratory Systems*, **148**, 1-8 **(2015)**

8. Ladgham A., Sakly A. and Mtibaa A., MRI Brain tumor Recognition Using Modified Shuffled Frog Leaping Algorithm, In Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2014 15th International Conference, IEEE, 504-507 **(2014)**

9. Subashini M.M., Sahoo S.K., Sunil V. and Easwaran S., A non-invasive methodology for the grade identification of astrocytoma using image processing and artificial intelligence techniques, *Expert Systems with Applications*, **43**, 186-196 **(2016)**

10. Ladgham A., Hamdaoui F., Sakly A. and Mtibaa A., Fast MR brain image segmentation based on modified Shuffled Frog Leaping Algorithm, *Signal, Image and Video Processing*, **9(5)**, 1113-1120 **(2015)**

11. Song Y., Bao X., Liu Z., Yuan D. and Song M., An Improved Brain MRI Segmentation Method Based on Scale-Space Theory and Expectation Maximization Algorithm, In Pacific Rim Conference on Multimedia, Springer International Publishing, 516-525 **(2015)**

12. Fwu, J.K. and Djuric P.M., EM algorithm for image segmentation initialized by a tree structure scheme, *IEEE Trans. Image Process*, **6**, 349–352 **(1997)**

13. Hastie T. and Tibshirani R., Discriminant analysis by Gaussian mixtures, *J. R. Stat. Soc. B*, **58**, 155–176 **(1996)**

14. Roberts S., Husmeier D., Rezek I. and Penny W., Bayesian approaches to Gaussian mixture modeling, *IEEE Trans. Pattern Anal. Mach. Intell.*, **20**, 1133–1142 **(1998)**

15. Bazi Y., Bruzzone L. and Melgani F., Image thresholding based on the EM algorithm and the generalized Gaussian distribution, *Pattern Recognition*, **40(2)**, 619-634 **(2007)**

16. Jordan M.I. and Jacobs R.A., Hierarchical mixtures of experts and the EM algorithm, *Neural Computation*, **6(2)**, 181-214 **(1994)**.