

# An Automated Approach for Segmentation of Brain MR Images using Gaussian Mixture Model based Hidden Markov Random Field with Expectation Maximization

Saurabh Shah<sup>1</sup> and N C Chauhan<sup>2</sup>

<sup>1</sup>*Department of Computer Science and Engineering, Babaria Institute of Technology, Vadodara and R.K. University, Rajkot, Gujarat, India.*

<sup>2</sup>*Department of Information Technology, A D Patel Institute of Technology, Anand, Gujarat, India. Email: saurabh\_er@rediffmail.com; narendracchauhan@gmail.com*

## ABSTRACT

Manual segmentation of brain tissues from MR images for diagnosis purpose is time consuming and requires much effort even by experts. This has motivated generation of automated segmentation techniques. Moreover, due to presence of noise in an image and its low contrast, it is difficult to correctly delineate tumour from brain MR images. In this paper, a novel hybrid method using Gaussian Mixture Model based Hidden Markov Random Field (HMRF) with Expectation Maximization (EM) has been proposed which segments tissues from MR brain images efficiently and helps to separate out tumour area easily. The proposed method minimizes energy function during each iteration of EM and gives comparable results with ground truth. The results obtained are also compared with the results of fuzzy c-means clustering algorithm for image segmentation.

**Keywords:** GMM, HMRF, EM, Segmentation, brain tumour

## 1 Introduction

Magnetic Resonance Imaging (MRI) is widely used to visualize soft tissues of brain and help medical experts to diagnose possible abnormality in brain structure. Brain image segmentation is one of the critical and challenging tasks and much effort has been invested by researchers and developers to automate this task [1]. Segmentation of a medical image helps to reveal important abnormality (if exists) which further helps in image analysis and classification. Image segmentation divides an image into regions such that pixels within a region are homogeneous with similar properties based on some predefined condition. All regions are mutually exclusive and aggregation of all the regions gives original image. Segmentation methods are challenged for MRI images as the tumours or any pathological area to be segmented have non-rigid anatomical structure, unknown noise, poor image contrast, complex shape which varies in size and position among MRI images [2,7,10].

Literature has shown studies which mainly focuses on segmentation techniques to separate out the intracranial volume into three overlapping parts, namely White Matter (WM), Grey Matter (GM) and Cerebro Spinal Fluid (CSF) [1,3,4]. Literature methods also cover segmentation of multiple sclerosis lesions [6], identification of desired pathological area or tumour [5]. Major challenges in segmentation are

DOI: 10.14738/jbemi.24.1411

Publication Date: 17<sup>th</sup> August 2015

URL: <http://dx.doi.org/10.14738/jbemi.24.1411>

presence of noise, partial volume effect caused because of the mixture of several tissue signals in a same voxel and bias field which is having spatial intensity inhomogeneity [4]. According to Corso, et al., automatic detection of brain tumour and its quantification is very difficult and challenging task [7].

A variety of brain tumour segmentation techniques have been proposed by researchers. These techniques produce reasonable results for specific imaging applications, however, they are not suitable for all kinds of imaging applications. Wu et al. proposed automated segmentation methods of human brain images which mainly include machine learning based segmentation, probabilistic approach and atlas based segmentation [9]. Segmentation techniques are mainly categorized into four classes by authors namely, threshold-based techniques, region-based techniques, pixel classification techniques and model-based techniques [2]. Out of all mentioned techniques, with threshold-based techniques it is impossible to segment Glioblastoma Multiforme (GBM) brain tumour and generally it fails [7, 10, 16]. Many segmentation methods like simple watershed methods, clustering, thresholding and adaptive thresholding do not perform well with noisy and low resolution images having intensity inhomogeneity [16]. Fuzzy based clustering approach is also popular approach for automatic tumour segmentation, however, research results conveys that alone fuzzy clustering method also face difficulties in separating out tumour and normal tissues and their intensity distribution generally overlaps. Hence, researchers recommend to incorporate knowledge based fuzzy clustering techniques for better segmentation results [7]. Ahmed, et al., proposed fuzzy c-means clustering and hybrid approaches and have demonstrated success in segmenting MR images [8].

Generally, intensity distribution of any region to be segmented in an image is assumed to follow Gaussian distribution. MR signal can be modelled as Gaussian distribution [4,9]. But, it is insufficient to model intensity distribution of real-world objects with single one Gaussian distribution because of their complexity. So, Gaussian mixture model (GMM) based approach is preferred and it is much more powerful for modelling the complex intensity distributions of an image compared to single Gaussian distribution [3,11].

Classification methods are generally categorized as parametric and non parametric where in case of parametric approach, the model parameters are usually estimated using a maximum likelihood (ML) or maximum a posteriori (MAP) approach and its optimisation process is carried out using the expectation–maximisation (EM) algorithm [4].

The organization of the paper is as follows. The proposed novel hybrid method with theoretical details of the techniques used has been presented in section 2. Section 3 shows the the experiments performed and results. Finally, section 4 concludes the paper with necessary discussions.

## 2 Proposed Methodology

In this paper, a method is proposed that approximates intensity distribution of brain MR image using Gaussian Mixture Model. As our approach is automated and unsupervised, there is no prior knowledge of background/foreground available; hence for estimation of parameters we use hidden markov random field (HMRF) where segmentation is carried out in an unsupervised manner. Our approach is modification of method suggested in [13] to optimize initial conditions to be used by HMRF-EM framework. When initial condition is very different than normal, mainly initial segmentation then EM may not give proper results. Hence, to optimize initial conditions k-means clustering is used to generate initial clusters and for

approximation of intensity distribution in each segment Gaussian Mixture Model has been used instead of single Gaussian component. Expectation maximization method has been used for learning parameter sets and label configurations to solve the problem of HMRF.

The Gaussian mixture model is widely used statistic method in parametric approaches to model complex intensity distribution instead of single Gaussian distribution and widely used to model MR signal. Intensity of MR brain image is given as input for Gaussian mixture model. Initially, each pixel is assigned to known classes. Generally iterative Expectation-Maximization (EM) algorithm or Maximum A Posteriori (MAP) estimation from a well-trained prior model are used for estimating GMM parameters from training data [12]. It has been found that combining the Hidden Markov Random Field (HMRF) model with the Expectation-Maximization (EM) algorithm produces more accurate and robust segmentation results [3].

## 2.1 Gaussian Mixture Model (GMM)

A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features extracted. A Gaussian mixture model is a weighted sum of N component Gaussian densities as given by the equation,

$$p(x/\lambda) = \sum_{i=1}^N w_i g(x | \mu_i, \sigma_i) , \quad (1)$$

where  $x$  is a D-dimensional continuous-valued data vector (i.e. measurement or features),  $w_i$ ,  $i = 1, \dots, N$ , are the mixture weights, and  $g(x | \mu_i, \sigma_i)$ ,  $i = 1, \dots, N$ , are the component Gaussian densities. Each component density is a D-variate Gaussian function of the form,

$$g(x | \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi} \sigma_i} \exp\left(-\frac{(x - \mu_i)^2}{2 \sigma_i^2}\right) \quad (2)$$

with mean vector  $\mu_i$  and covariance matrix  $\sigma_i$ . The mixture weights satisfy the constraint that

$$\sum_{i=1}^N w_i = 1. \quad (3)$$

The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation,

$$\lambda = \{w_i, \mu_i, \sigma_i\} \quad i = 1, \dots, N \quad (4)$$

In our experiment, GMM has been used to model intensity distribution of tissues and tumour in MR brain images where its parameters are represented using mean, covariance and Gaussian components which are represented by  $w_i$  weights. Each Gaussian component represents individual object. Estimated parameters of GMM are then given to HMRF-EM framework to predict class label.

## 2.2 Hidden Markov Random Fields

Hidden Markov Random Fields (HMRF) is known to maintain neighbourhood information and spatial relations between pixels of neighbourhood. So, it has been widely used in segmentation process. In case of medical images, neighbourhood pixels have similar properties and thus belong to same target class. So for segmentation of various regions of medical images, HMRF is generally used which maintains spatial relation and it is also more robust against noise. For estimation of parameters of HMRF, EM is usually used [3].

Applications of Markov Random Fields (MRFs) have been found in many computer vision problems, like surface reconstruction, image segmentation and depth inference [11]. For segmentation of brain MR images, the HMRF-EM framework was first proposed in [13]. It is assumed that the image is 2D gray-level and the intensity distribution of each region to be segmented follows a Gaussian distribution [4,9,11]. Given an image  $Y = (y_1, \dots, y_N)$  where  $N$  is the number of pixels and each  $y_i$  is the gray-level intensity of a pixel, we want to find out labels  $X = (x_1, \dots, x_N)$  where  $x_i \in L$  and  $L$  is the set of all possible labels.. According to the MAP criterion, we are interested in labelling  $X^*$  which satisfies:

$$X^* = \underset{X}{\operatorname{argmax}} \{P(Y|X, \theta)P(X)\} \quad (5)$$

The prior probability  $P(X)$  is a Gibbs distribution, and the joint likelihood probability is given by

$$P(Y|X, \theta) = \prod_i P(y_i|X, \theta) \quad (6)$$

$$P(Y|X, \theta) = \prod_i P(y_i|x_i, \theta_{x_i}) \quad (7)$$

where  $P(y_i|x_i, \theta_{x_i})$  is a Gaussian distribution with parameters  $\theta_{x_i} = (\mu_{x_i}, \sigma_{x_i})$ . The parameter set  $\theta = \{\theta_l \mid l \in L\}$  are learned from the training data in case of MRF as prior knowledge is available while in case of HMRF, these parameter set are learned in an unsupervised manner. Mainly, it is learned using EM algorithm.

According to the Hammersley–Clifford theorem, an MRF can equivalently be characterized by a Gibbs distribution [13]. So, Prior probability can be written as,

$$P(X) = \frac{1}{Z} \exp(-U(X)) \quad (8)$$

where  $U(X)$  is the prior energy function and  $Z$  is a normalizing constant called the partition function.  $U(X)$  is defined in terms of clique potentials as,

$$U(X) = \sum_{c \in C} V_c(X) \quad (9)$$

which is a sum of clique potentials  $V_c(X)$  over all possible cliques and  $C$  is set of all possible cliques . A clique is defined as a subset of sites in  $S$  in which every pair of distinct sites are neighbours, except for single-site cliques. The value of  $V_c(X)$  depends on the local configuration of clique  $c$  [13] .

From eq.(7),

$$P(Y|X, \theta) = \prod_i P(y_i|x_i, \theta_{x_i}) \quad (10)$$

Let us represent in terms of Gaussian distribution function as,

$$P(Y|X, \theta) = \prod_i g(y_i; \theta_{x_i}) \quad (11)$$

$$= \frac{1}{Z'} \exp(-U(Y|X)) \quad (12)$$

In our experiment, parameters of HMRF model are mean and standard deviation of each GMM. MAP and EM algorithms have been used to learn HMRF parameters and class labels alternatively as both are dependent on each other. Parameters are learned by maximizing probability of class labels and by minimizing total posterior energy as shown in eq.(5) and eq.(12).

### 2.3 Expectation Maximization (EM) algorithm

The basic idea of the EM algorithm is to begin with an initial model, to estimate a new model  $\theta'$ , such that  $p(X | \theta') \geq p(X | \theta)$ . The new model then becomes the initial model for the next iteration and the process is repeated until some convergence condition is satisfied. In our experiment when no significant change in total energy observed or maximum EM iterations are reached, algorithm converges. EM algorithm has been used in our experiment to estimate the parameter set  $\theta = \{\theta_l \mid l \in L\}$  for HMRF model. The detailed EM algorithm is mentioned as below:

1. **Start:** Assume that initial parameter set is  $\theta^{(0)}$ .

2. **E-step:** At the  $t^{\text{th}}$  iteration, parameter set is  $\theta^{(t)}$  and conditional expectation is calculated as,

$$Q(\theta | \theta^{(t)}) = E [\ln P(x, y | \theta) | y, \theta^{(t)}] \quad (13)$$

$$Q(\theta | \theta^{(t)}) = \sum_{x \in X} P(x | y, \theta^{(t)}) \ln P(x, y | \theta) \quad (14)$$

where,  $X$  indicates set of labels.

3. **M-step:** Now, maximize  $Q(\theta | \theta^{(t)})$  to obtain the next estimate:

$$\theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}} Q(\theta | \theta^{(t)}) \quad (15)$$

then let  $\theta^{(t+1)} \rightarrow \theta^{(t)}$  and repeat from the E-step.

In the EM algorithm, we do MAP estimation and solve for  $x^*$  that minimizes the total posterior energy

$$X^* = \underset{x}{\operatorname{argmin}} \{U(Y|X, \theta) + U(X)\} \quad (16)$$

$$X^* = \underset{x}{\operatorname{argmax}} \{P(Y|X, \theta)P(X)\} \quad (17)$$

## 2.4 Proposed MR Image Segmentation Algorithm

The proposed MR image segmentation algorithm is given in the following Fig. 1

---

**Algorithm:** Segmentation of MR brain images using GMM based HMRF

---

Input: Brain MRI images of different modalities

(T1, T1c, T2, T2FLAIR)

Output: Segmented images

Steps:

1. Perform contrast enhancement by Histogram normalization.
2. Perform initial intensity based clustering using k-means depending upon number of clusters specified.
3. Estimate intensity distribution of MR brain image using Gaussian Mixer Model with parameter set  $\theta_{x_i} = (\mu_{x_i}, \sigma_{x_i})$  as explained in section 2.1 and fit GMM to the data generated in step 2.
4. Estimate the class labels  $X^*$  by HMRF-MAP estimation as explained in section 2.2, eq.(5) which satisfies the following condition.

$$X^* = \underset{X}{\operatorname{argmax}} \{P(Y|X, \theta)P(X)\}$$

5. Update parameter set  $\theta = \{\theta_l \mid l \in L\}$  using EM algorithm as explained in section 2.3 until it converges where L is the set of all possible labels. Do MAP estimation and solve for  $x^*$  that minimizes the total posterior energy as per eq.(16) and (5),

$$X^* = \underset{X}{\operatorname{argmin}} \{U(Y|X, \theta) + U(X)\}$$

$$X^* = \underset{X}{\operatorname{argmax}} \{P(Y|X, \theta)P(X)\}$$

6. Repeat step 4 till maximum EM iterations reached or there is no significant change in the value of energy function,  $U(Y|X, \theta) + U(X)$ .
- 

**Figure 1 Proposed MR image segmentation algorithm**

As suggested in [13], initial conditions in algorithm play very important role in final segmentation, mainly when initial conditions are far from ground truth. So, to ensure reasonably good initial parameter estimation, k-means clustering algorithm has been applied in our work. For each resultant cluster, intensity distribution of brain tissues are estimated using Gaussian Mixture Models and parameterized with the mean, covariance and mixture weights from all component densities, represented as  $\lambda = \{w_i, \mu_i, \sigma_i\}$ . The GMM model parameters have also been estimated using an EM-algorithm. In the E-step, we determine which data belongs to which Gaussian component and in the M-step, GMM parameters are recomputed. Model fitting is the process of estimating the unknown parameters. So, the problem is to define parameters and functional form of the HMRF model where the parameters are mean and standard deviation of each GMM. The major objective of the proposed algorithm is to incorporate spatial information using HMRF in addition to intensity values and to get robust and accurate segmentation results. As parameters and class labels are unknown but dependent on each other, they are learned using MAP and EM algorithm by maximizing probability of class labels and by minimizing total posterior energy. EM algorithm starts with the current estimation, conditional expectation is calculated as shown in eq. (14) and then it maximizes this condition to get new estimate as explained in subsection 3.3. Process continues until it converges and, no significant change in total energy observed or maximum EM iterations are reached.

### 3 Experiments

In this work, mainly two sets of experiments have been performed, one on real MR brain images, and other on simulated MR brain images. A total 30 datasets containing high grade tumours collected from BRATS 2012 dataset [14] have been used as the dataset. Segmentation results of proposed method are compared with ground truth available with dataset. Results of the proposed method are also compared with known Fuzzy c-Means clustering algorithm [2,4,8,10]. Results of both the methods are compared based on smoothness of segmentation. Also, plot of total energy signifies quality of segmentation during each iteration of EM-MAP. Reduction in total energy after each iteration indicates improvement in segmentation results. Fig. 2 and Fig. 3 presents real brain images, ground truth for tumour region, segmentation results and plot of total energy while Fig. 4, Fig. 5 and Fig. 6 presents simulated brain images along with complete ground truth of segmentation, segmentation results and plot of total energy respectively. As a performance measurement of segmentation, Dice overlap metric has been used in many literature [1,6,7,15], but it is not considered to be most reliable parameter as it depends on size and shape of an object and major errors occur at object boundaries. So, results of smaller objects are compromised compared to larger objects [15]. Hence, for quantitative assessment of segmentation, only smoothness of segmentation has been considered in our experiments.

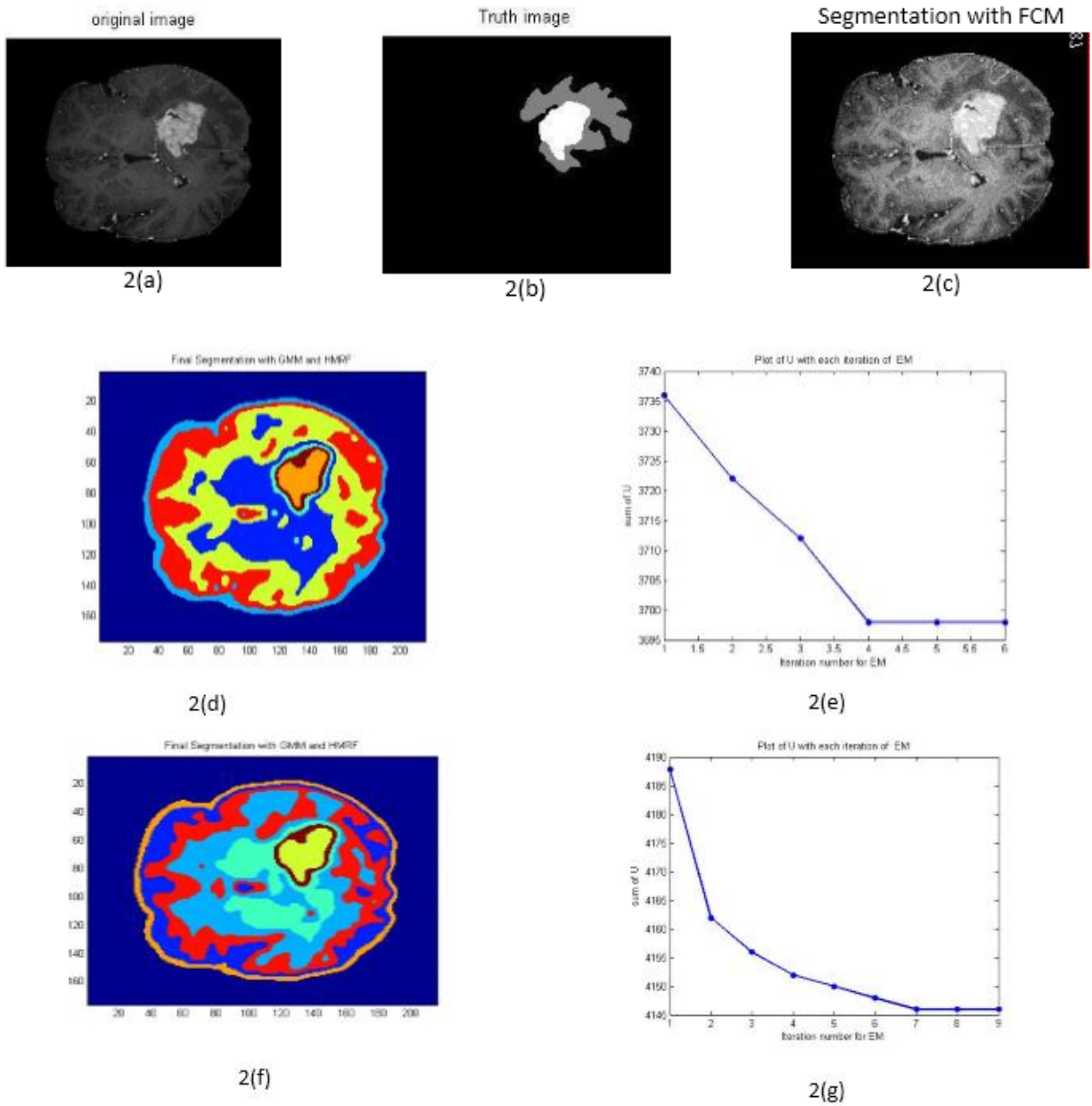
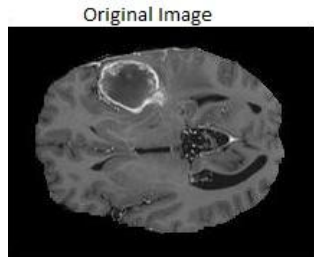


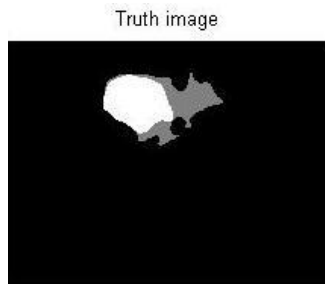
Figure 2 (a) Original image - Real brain dataset 1 (b) Ground truth (c) Segmented image with FCM (d) Segmentation with proposed method with GMM =3, Number of Clusters =8 (e) Plot of U with EM iterations = 6 (f) Segmentation with proposed method GMM =6, Number of Clusters =8 (g) Plot of U with EM iterations =9



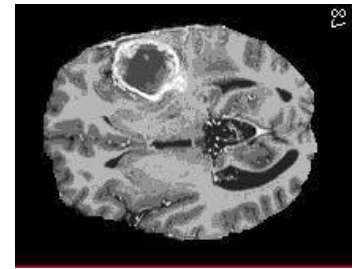
Segmentation with FCM



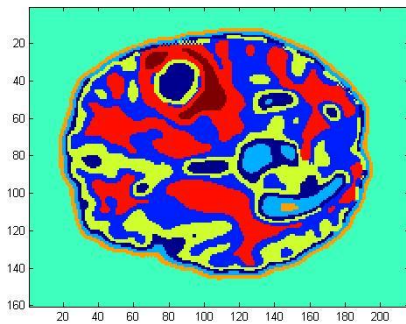
3(a)



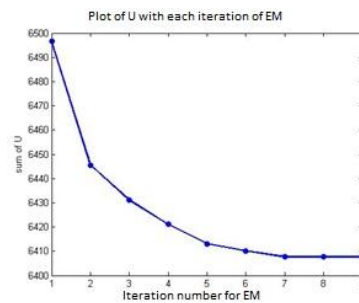
3(b)



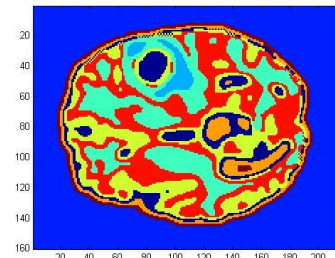
3(c)



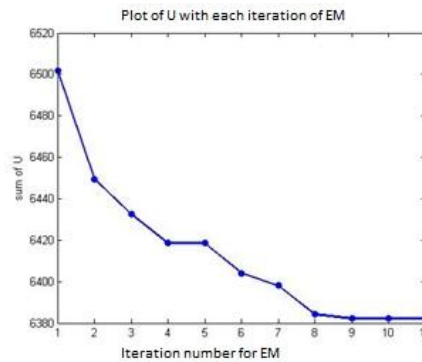
3(d)



3(e)

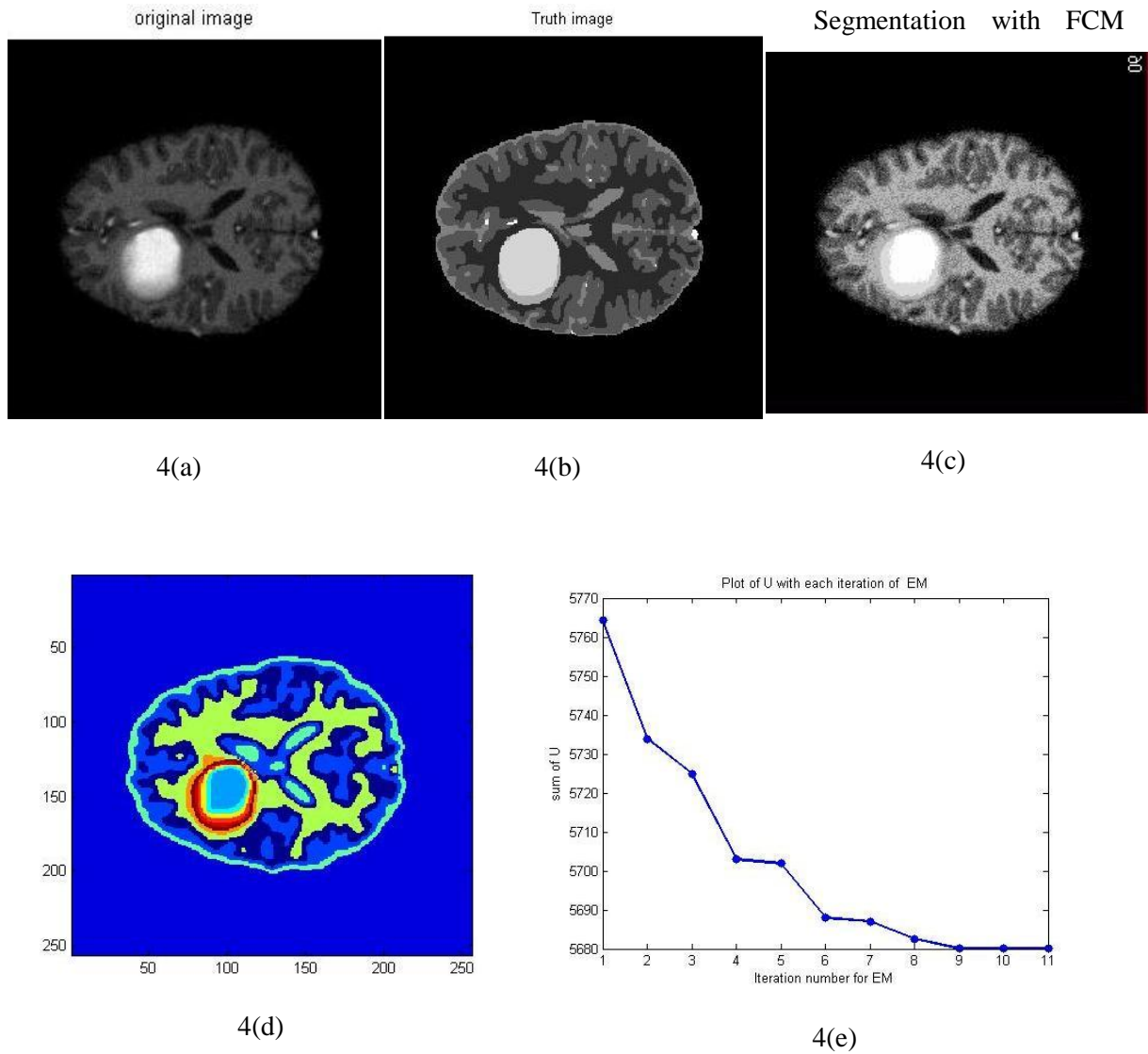


3 (f)

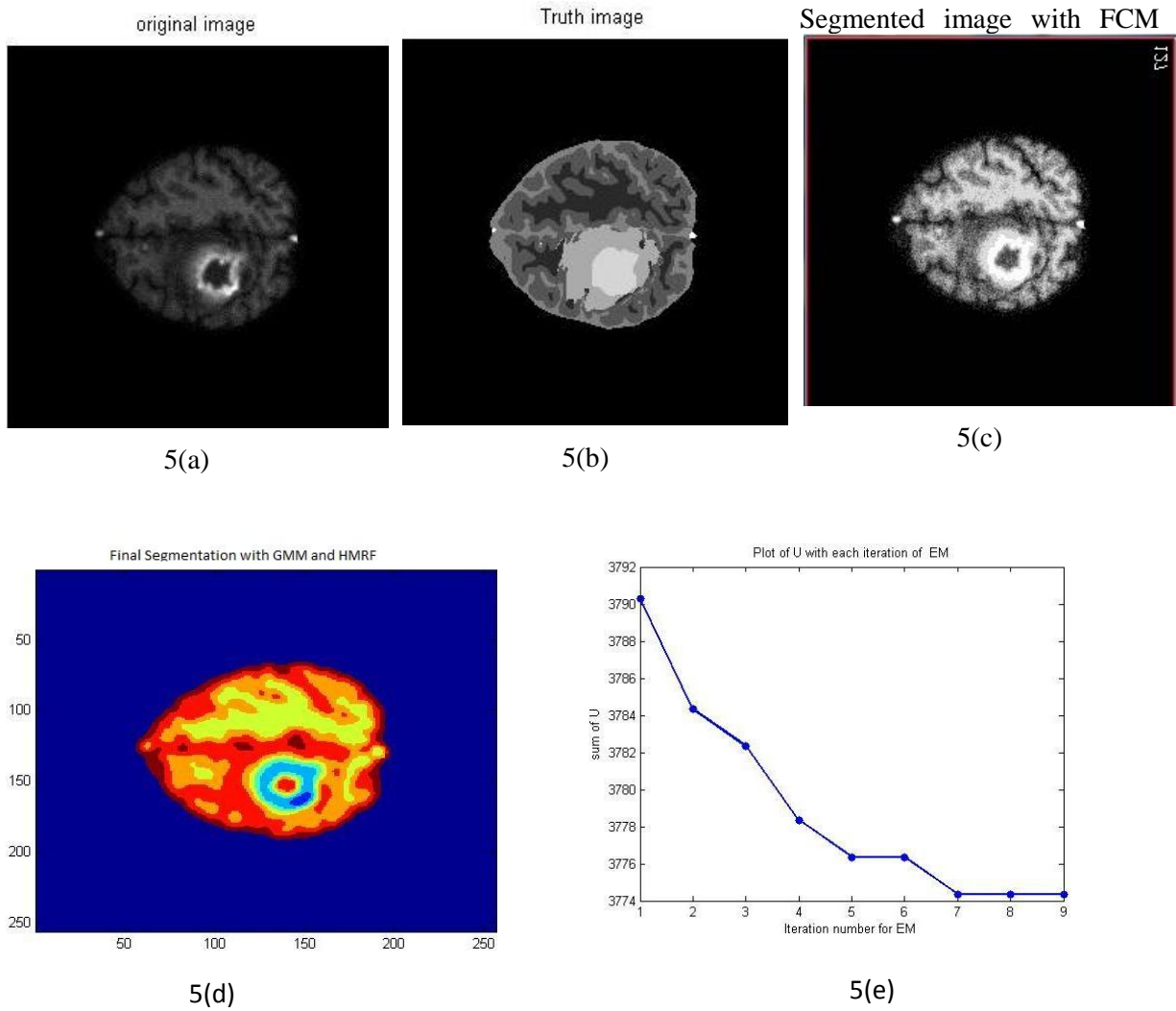


3(g)

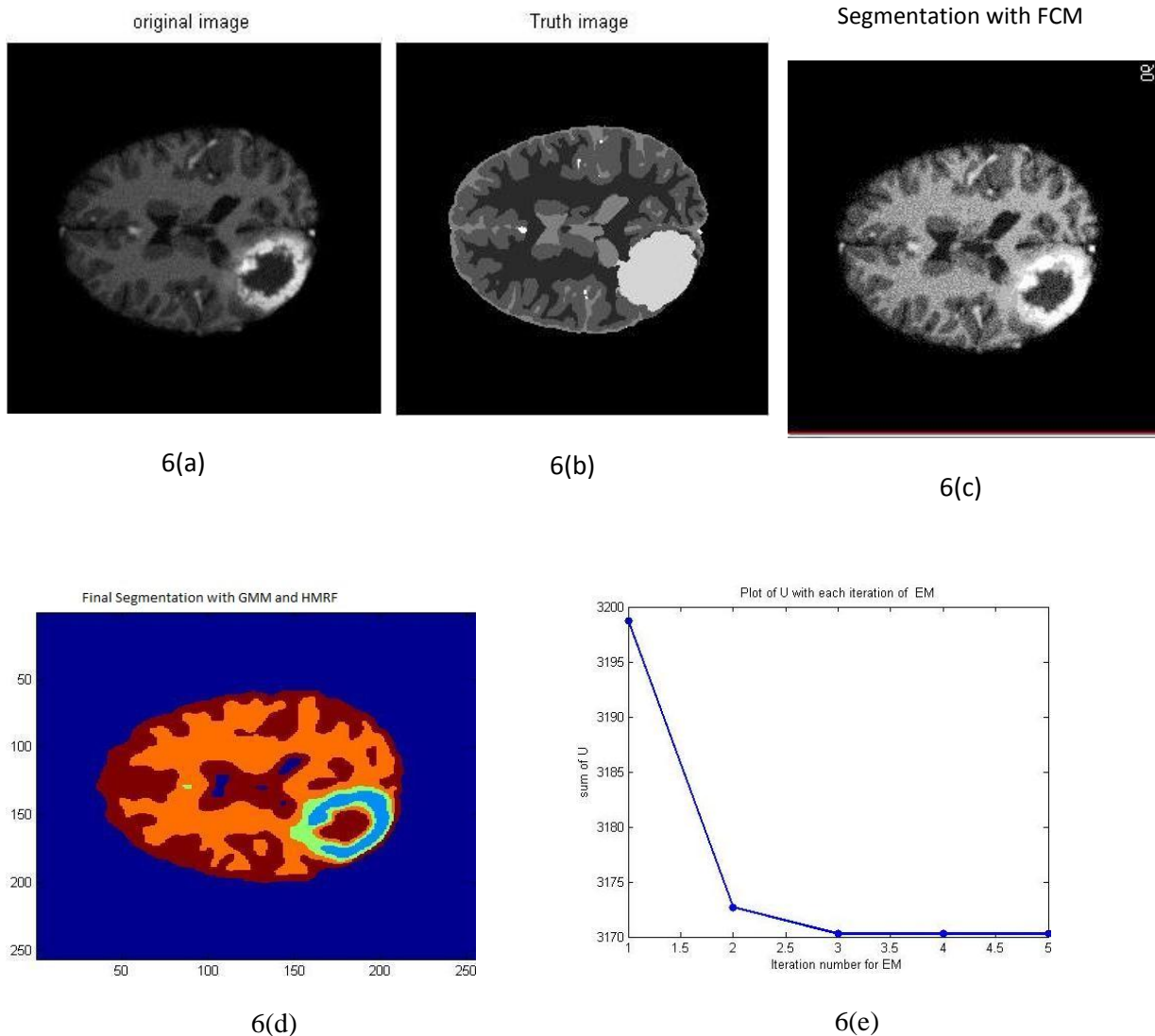
Figure 3(a) Original image - Real brain dataset 2 (b) Ground truth (c) Segmented image with FCM (d) Segmentation with proposed method with GMM =3, Number of Clusters =8 (e) Plot of U with EM iterations = 9 (f) Segmentation with proposed method GMM =6, Number of Clusters =8 (g) Plot of U with EM iterations = 11



**Figure 4 (a) Original image -Simulated brain dataset 1 (b) Complete ground truth (c) Segmented image with FCM (d) Segmentation with proposed method with GMM =3, Number of Clusters =8 (e) Plot of U EM iterations =11**



**Figure 5 (a) Original image - Simulated brain dataset 2 (b) Complete ground truth (c) Segmented image with FCM (d) Segmentation with proposed method with GMM =3, Number of Clusters =8 (e) Plot of U EM iterations =9**



**Figure 6 (a) Original image - Simulated brain dataset 3 (b) Complete ground truth (c) Segmented image with FCM (d) Segmentation with proposed method with GMM =3, Number of Clusters =8 (e) Plot of U with EM iterations =5**

#### 4 Discussion and conclusion

The major goal of the work is to perform segmentation of MR images in order to delineate tumour region. As suggested in [13], the initial condition plays important role in the segmentation results. Hence, in this paper, initial segmentation has been performed through k-means which is further optimized using GMM based HMRF-EM model. The proposed method is unsupervised, fully automated and parametric approach for segmenting tissues and tumour in brain MR images. The major objective of segmentation to separate tumour from other tissues is well obtained in all the experimental results which is also visible from Figs. 2(d), 2(f), 3(d), 3(f), 4(d), 5(d), 6(d) for both the set of experiments, one with real brain and other with simulated brain MR images. The segmented tumour is found to be very close to the tumour highlighted in the corresponding ground truth image. The segmentation results obtained using proposed method as shown in Figs. 2-6 are better than FCM clustering algorithm in terms of smoothness as seen visually and

also it is very near to ground truth segmentation. Also segmented regions are not intermixed with neighbouring segments as HMRF imposes strong spatial constraint based on neighbourhood property which results into smooth segmentation. The adjacent pixels of the same tissue have nearly similar properties and that is why HMRF is able to segment similar tissues based on this principle. In both cases, MR images of real brain, and simulated brain with varying level of noise, performance is better than FCM. Also results obtained through proposed method are better in presence of noise which proves the proposed method to be robust for noisy images. Experiments are also carried out for different Gaussian components to verify its effect on segmentation. The selection of number of Gaussian components play important role, and generally it is decided based on experiments; however, the more number of components also increase computational complexity. Hence, moderate number of components has to be selected. The plot of sum of  $U$ , total energy during each iteration of EM as shown in Figs. 2(e), 2(g), 3(e), 3(g), 4(e), 5(e), 6(e) represents reduction in energy function and finally it gets converged before it reaches to maximum number of EM iteration which was set to 15 for all the set of experiments. In some of the experiments, it converged early while in others it required more iterations.

#### REFERENCES

- [1] Yousefi Sahar, Reza Azmi, and Morteza Zahedi. "Brain tissue segmentation in MR images based on a hybrid of MRF and social algorithms.", Elsevier, *Medical image analysis* 16, no. 4 (2012): 840-848.
- [2] Gordillo Nelly, Eduard Montseny and Pilar Sobrevilla. "State of the art survey on MRI brain tumor segmentation." Elsevier, *Magnetic resonance imaging* 31.8 (2013): 1426-1438.
- [3] Balafar, M. A. "Gaussian mixture model based segmentation methods for brain MRI images." Springer, *Artificial Intelligence Review* 41, no. 3 (2014): 429-439.
- [4] Caldaïrou Benoît, Nicolas Passat, Piotr A. Habas, Colin Studholme and François Rousseau. "A non-local fuzzy segmentation method: application to brain MRI." , Elsevier, *Pattern Recognition* 44, no. 9 (2011): 1916-1927.
- [5] Khotanlou Hassan, Olivier Colliot, Jamal Atif and Isabelle Bloch. "3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models." *Fuzzy Sets and Systems* 160, no. 10 (2009): 1457-1473.
- [6] Lladó Xavier, Arnau Oliver, Mariano Cabezas, Jordi Freixenet, Joan C. Vilanova, Ana Quiles, Laia Valls, Lluís Ramió-Torrentà and Àlex Rovira. "Segmentation of multiple sclerosis lesions in brain MRI: a review of automated approaches.", Elsevier, *Information Sciences* 186, no. 1 (2012): 164-185.

- [7] Jason J. Corso, Eitan Sharon, Shishir Dube, Suzie El-Saden, Usha Sinha and Alan Yuille, "Efficient multilevel brain tumor segmentation with integrated bayesian model classification." *Medical Imaging, IEEE Transactions on* 27.5 (2008): 629-640.
- [8] Mohamed N. Ahmed, Sameh M. Yamany, Nevin Mohamed, Aly A. Farag and Thomas Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data." *Medical Imaging, IEEE Transactions on* 21.3 (2002): 193-199.
- [9] Wu T, Bae MH, Zhang M, Pan R and Badea A, "A prior feature SVM-MRF based method for mouse brain segmentation." *Elsevier, NeuroImage* 59.3 (2012): 2298-2306.
- [10] Balafar, M. A., Ramli A. R., Saripan M. I., & Mashohor S., "Review of brain MRI image segmentation methods." *Artificial Intelligence Review* 33.3 (2010): 261-274.
- [11] Wang, Quan. "GMM-Based Hidden Markov Random Field for Color Image and 3D Volume Segmentation." *arXiv preprint arXiv:1212.4527* (2012).
- [12] Reynolds, Douglas A., Thomas F. Quatieri and Robert B. Dunn. "Speaker verification using adapted Gaussian mixture models." *Digital signal processing* 10.1 (2000): 19-41.
- [13] Zhang Yongyue, Michael Brady and Stephen Smith. "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm." *Medical Imaging, IEEE Transactions on* 20.1 (2001): 45-57.
- [14] MRI Brain Datasets available on BRATS 2012 - <http://www2.imm.dtu.dk/projects/BRATS2012/>
- [15] Greenspan Hayit, Amit Ruf and Jacob Goldberger. "Constrained Gaussian mixture model framework for automatic segmentation of MR brain images." *Medical Imaging, IEEE Transactions on* 25.9 (2006): 1233-1245.
- [16] Mahapatra Dwarikanath and Ying Sun, "Integrating segmentation information for improved MRF-based elastic image registration." *Image Processing, IEEE Transactions on* 21.1 (2012): 170-183.