REGION MERGING STRATEGY FOR BRAIN MRI SEGMENTATION USING DEMPSTER-SHAFER THEORY

J. GHASEMI, M. R. KARAMI MOLLAEI, R. GHADERI AND S. A. HOJJATOLESLAMI

Abstract. Detection of brain tissues using magnetic resonance imaging (MRI) is an active and challenging research area in computational neuroscience. Brain MRI artifacts lead to an uncertainty in pixel values. Therefore, brain MRI segmentation is a complicated concern which is tackled by a novel data fusion approach. The proposed algorithm has two main steps. In the first step the brain MRI is divided to some main and ancillary cluster which is done using Fuzzy c-mean (FCM). In the second step, the considering ancillary clusters are merged with main clusters employing Dempster-Shafer Theory. The proposed method was validated on simulated brain images from the commonly used BrainWeb dataset. The results of the proposed method are evaluated by using Dice and Tanimoto coefficients which demonstrate well performance and robustness of this algorithm.

1. Introduction

There are many cases that accurate decision making by physicians is only possible using medical images, as they provide critical information about internal organs. Magnetic Resonance Imaging (MRI) is one of the most important non-invasive medical imaging techniques to present high resolution images which provides high intensity contrast between different soft tissues. It plays a key role in brain tissue visualization. Automatic segmentation of brain tissue plays an important role in the study of various abnormalities, brain development and evaluation of the progress of treatment. The major tissues in a normal brain are: Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF). Brain MRI segmentation (BMS) is one of the important processing operations in which a region of the image with a specific characteristics is labeled. It is an important processing step in many medical researches and clinical applications where decision making is critical. As brain MRI is a set of images with large volume information, manual BMS is a time consuming task. Therefore, an automatic segmentation system with acceptable speed, high accuracy and generalization capability is needed. Three main artifacts in MRI are: noise, partial volume effect (PVE) and intensity non-uniformity (INU). The main sources of noise are categorized to biological and scanner noises introduced in [9, 34, 42]. Tissue non-uniformity and limitation in hardware design are the main causes of biological and scanner noises respectively. PVE is recognized as

Received: November 2011; Revised: May 2012 ; Accepted: July 2012

Key words and phrases: MRI, Fuzzy c-mean, Brain MRI Segmentation, Dempster-Shafer Theory.

a mixture of intensities due to more than one tissue present at the pixel. Increasing the image resolution results in lower PVE and more detailed images, however in most cases causes a higher level of noise. INU, also called as bias field, is a low frequency smoothed artifact which is generated by inhomogeneity of magnetic field during scanning process [38,40]. These problems lead to an uncertainty about the pixel values which lead to a considerable overlap between different tissues [14]. Many methods have been presented for BMS [4,24,33]. Such methods can be classified to three main categories including boundary-based methods, region-based methods and hybrid methods [30,48]. In boundary-based strategies, the gradient features near an object boundary would be employed as a measure of discontinuity to guide for segmentation [8]. The main categories of this method are edge detection, deformable templates and active contours [1,18,27]. The BMS in region-based methods is normally carried out with the identification of a homogeneity feature representing one of the corresponding brain tissues [16]. The hybrid methods use a combination of similarity and discontinuity to segment the images. In all the three BMS categories, a decision is made based on one or more measurements against the criteria used to assign a class label to a pixel. Due to uncertainty of brain MRI pixels using rigid criteria can potentially limit the capability of the technique. For this reason the researchers have developed fuzzy approaches which allow different regions of the image to be considered as fuzzy sets. Given such strategy, each pixel may be assigned to potential multi-class tissues. Fuzzy segmentation provides more information than the other crisp methods [48]. Many types of fuzzy classifer systems have been applied in different classification problems [21, 25, 26, 29]. Fuzzy c-mean (FCM) is one of the prominent tools among different fuzzy approaches in BMS [10, 12, 13, 19, 23, 35, 37, 39, 50]. Different researches have been performed for FCM optimization to achieve better results. An adaptive FCM approach has been suggested in [31] to estimate INU by multiplying a distance function and clustering centers. This method is generalized for 3-D images in [32]. A regulative term has been proposed for FCM in which spatial information are interpreted to overcome noise effects [2].

Generally in combination methodology, classification tasks involve multisource data fusion and evidential reasoning in order to achieve higher classification accuracy and robustness. The information extracted from any source in favor of, or against, a given class assignment must be combined with that of other sources to infer the likelihood of this assignment [6]. This idea is also valid for image segmentation as a special case of classification task. In [3] the authors proposed a method for brain-tissue segmentation by fusing the information in structural MRI and diffusion tensor (DT) images in a statistical framework. Dempster-Shafer Theory (DST) is considered as one of the main important tools in information fusion area. This theory is a general extension of the Bayesian theory which offers a number of advantages compared to Bayesian theory [49]. DST can robustly deal with incomplete information which has wide applications in image processing and pattern recognition area [43–45, 47]. In [15] a multispectral BMS based on data fusion has been proposed in which a genetic fuzzy system and DST were used for fuzzy modeling and fusing the results, respectively. In this method the results of individual

classifiers which were applied to three well-known modalities, T1, T2 and PD, are combined by using DST. It should be noted that these modalities are not accessible in the same coordinate system in real data acquisition. As a result the performance of the technique is highly dependent on the quality of the data or the result of an algorithm used for registering the images. The other weakness of this method is that it does not use any spatial information which reduces its application.
Here we propose a novel information fusion approach based on Fuzzy c-mean and Dempster-Shafer Theory for Brain BMS. The proposed method applies only to one of the mentioned modalities (T1-weighted brain MR images). In the proposed method, firstly, all pixels are classified with uncertainty consideration in the pixel values. This process is done by employing the ancillary clusters beside the main clusters. The ancillary clusters are merged with the main clusters in the final step of the proposed method. In fact, FCM and DST are utilized for primary clustering and merging homogeny areas, respectively. Also in the feature extraction step, the appropriate combines of the spatial information are used as feature sets.
The outline of the paper is as follows: In Section 2, a classic fuzzy c-mean, Dempster-Shafer theory and Singular Value Decomposition are introduced. Proposed Method, called FCMMDS is explained in Section 3. The experimental results are presented in Section 4 which follows with concluding remark in Section 5.

2. **Fundamental Definition**

2.1. **Fuzzy c-Mean (FCM).** In FCM, as a development of hard K-means algorithm, every input data will be assigned to all existing clusters [46]. Pattern membership in a class is based on the similarity of the pattern to the class with respect to all classes. Objective function of the standard FCM which segments an image to c clusters can be defined as:

$$J_q(u,v) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{q} d^2(x_j, v_i) \; subject\ to \; \sum_{i=1}^{c} u_{ij} = 1 \;\; \& \;\; u_{ij} \in [0,1] \tag{1}$$

In which $\mathbf{X} = (x_1, x_2, \ldots, x_{\mathrm{i}}, \ldots, x_{\mathrm{n}})$ is a $p \times n$ data matrix, p is the dimension of the feature vector of each x_j and n is the number of feature vectors. In BMS, p=1 (intensity value), n is the image pixels number, q is the fuzziness index (in this study, q is chosen 2). u_{ij} is the membership of j^{th} pattern in i^{th} cluster, v_i is the center of i^{th} fuzzy cluster. d represents the similarity of the feature vector of x_j to cluster center of v_i in the features space which can be calculated with Euclidean norm as:

$$d^2(x_j, v_j) = \|x_j - v_j\|^2 \tag{2}$$

To achieve a minimum objective function, a higher membership value (MV) should be assigned to patterns close to the cluster centers and also lower values needs to be assigned to the patterns far from cluster centers. By applying the derivative of J_q with respect to u and v and equal to zero, the conditions to minimize the objective function is achieved as:

$$u_{ij} = \left(\sum_{k=1}^{c} \left(\frac{d\,(x_j, v_i)}{d\,(x_j, v_k)} \right)^{2/(q-1)} \right)^{-1} \tag{3}$$

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^{q}.x_j}{\sum_{j=1}^{n} u_{ij}^{q}} \tag{4}$$

According to (3) and (4), the patterns are assigned to all existing clusters based on the associated MVs and then, new cluster centers are calculated. The FCM algorithm reaches a solution by an iterative process until a termination criterion is met, ie.: $v(t) - v(t-1) < \varepsilon$. Finally, by assigning the pattern to a cluster with the highest MV, a segmentation of the data can be done.

2.2. **Dempster-Shafer Theory.** Dempster-Shafer Theory (DST), also called evidence theory, is a mathematical theory of evidence which allows a combination of evidence weight from different sources to new evidence weight [36, 49]. DST is basically a generalized Bayesian statistical theory. For instance, probabilities in the Bayesian approach can be assigned only to the singleton subsets (i.e. elements) of Ω (which is defined below) but in DST those are assigned to all subsets of Ω.

2.2.1. **Basic Concepts.** Assume that there is a finite set of propositions or mutually exclusive and exhaustive hypotheses or answers $\Omega = \{C_1, C_2, \ldots, C_n\}$ which is called "frame of discernment". In DST a basic belief assignment (BBA) μ, also called mass function, is a mapping function $\mu : 2^{\Omega} \to [0, 1]$ which satisfies the two conditions as follows:

$$\sum\nolimits_{A \subseteq \Omega} \mu(A) = 1,\ \mu(\varnothing) = 0 \tag{5}$$

The notation 2^{Ω} relates to the power set of Ω. If A is considered as a subset of Ω, the value $\mu_i(A)$ for the available evidence i, is interpreted as a degree of belief that is assigned to the exact A and not to any subset of A. Focal elements are sets receiving a non-null mass value. It is attractive to overcome the restriction of conventional probability theory by representation of both imprecision and uncertainty through the definition of two functions: plausibility (Pls) and belief (Bel), both are derived from a mass function μ [36]. The measures of belief and plausibility for A will be determined as follows:

$$Bel(A) = \sum_{B \subset A} \mu(B),\ \forall \mathrm{A} \subseteq \Omega\ \ \&\ \ \mathrm{A} \neq \varnothing\ ,\ Bel(\varnothing) = 0 \tag{6}$$

$$Pls(A) = \sum_{B \cap A \neq \varnothing} \mu(B)\ \forall A \subseteq \Omega\ \ \&\ \ \mathrm{A} \neq \varnothing\ ,\ Pls(\varnothing) = 0 \tag{7}$$

Bel(A) represents the confidence that a proposition f lies in A or any subset of A. Clearly Pls(A) represents the extent to which we fail to disbelieve A [5]. DST presents an explicit measure of ignorance about an event A and its complementary $\bar{A}$ as a length of the interval [Bel(A), Pls(A)] called belief interval. It can also be interpreted as an imprecision of the "true probability" of A [7].

2.2.2. **Combination of BBAs.** One of the main advantages of DST is the flexible combination operator it offers to combine information by Dempster's rule of combination. Suppose that μ_1 and μ_2 are two independent, un-contradictory belief structures on a frame of discernment Ω, with focal elements $\Psi_i, i = 1, 2, ..., n_1$, and $\Psi_j, i = 1, 2, ..., n_2$. The combination of μ_1 and μ_2 (also called the joint mass (μ_{12})) are defined as:

$$\mu_1 \oplus \mu_2(\zeta_k) = \mu_{12}(\zeta_k) = \frac{\sum_{\Psi_i \bigcap \Psi_j = \zeta_k} \mu_1(\Psi_i) \times \mu_2(\Psi_j)}{1 - \sum_{\Psi_i \bigcap \Psi_j = \varnothing} \mu_1(\Psi_i) \times \mu_2(\Psi_j)} \tag{8}$$

Where μ_{12} is another BBA whose focal element is $\zeta_k = \Psi_i \bigcap \Psi_j$. Dempster' rule is commutative and associative [49].

2.2.3. **Decision Making.** After combining process, a decision has to be made by using the final combined belief structure. In this study pignistic transformation which is defined in transferable belief model (TBM) is used [41]. By this strategy, for all $A \subseteq \Omega$, the $\mu(A)$ is distributed amongst its elements uniformly, the pignistic probability distribution is defined as:

$$BetP(\omega) = \sum_{\{A \subseteq \Omega, \omega \in A\}} \frac{1}{|A|} \times \frac{\mu(A)}{1 - \mu(\varnothing)},\ \omega \in \Omega \tag{9}$$

Where$|A|$ is the cardinality of subset A. In normal BBAs (with $\mu(\varnothing) = 0$), $\mu(A)/(1 - \mu(\varnothing)) = \mu(A)$.

2.2.4. **Discussion.** In this section, the uncertainty decrement in DST-based three elements is discussed. We aim to define a set of sufficient conditions which guarantee a reduction in uncertainty. It also does not necessarily mean that the uncertainty will not decrease without meeting these conditions. Assume that for a given problem the frame of discernment is $\Omega = \{\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3\}$ in which evidences A and B have the following belief structures:

Focal element	Evidence A	Evidence B
$\mu\{C_1\}$	a	h
$\mu\{C_2\}$	b	i
$\mu\{C_3\}$	c	j
$\mu\{C_1, C_2\}$	d	k
$\mu\{C_1, C_3\}$	e	l
$\mu\{C_2, C_3\}$	f	m
$\mu\{C_1, C_2, C_3\}$	g	n

By combining two beliefs using DST, the non-normalized joint masses are given by:

$$\mu_{AB}\{\emptyset\} = h.(b + c + f) + i.(a + c + e) + j.(a + b + d) + k.c + l.b + m.a$$

$$\mu_{AB}\{C_1\} = a.(h + k + l + n) + d.(h + l) + e.(h + k) + g.h$$

$$\mu_{AB}\{C2\} = b.(i + k + m + n) + d.(i + m) + f.(i + k) + g.i$$

$$\mu_{AB}\{C_3\} = c.(j + l + m + n) + e.(j + m) + f.(j + l) + g.j$$

$$\mu_{AB}\{C_{\mathbf{1}},\ C_{\mathbf{2}}\} = d.k + d.n + g.k\ ,\quad \mu_{AB}\{C_{\mathbf{1}},\ C_{\mathbf{3}}\} = e.l + e.n + g.l$$

$$\mu_{AB}\{C_{\mathbf{2}},\ C_{\mathbf{3}}\} = f.m + f.n + g.m\ ,\quad \mu_{AB}\{C_{\mathbf{1}},\ C_{\mathbf{2}},\ C_{\mathbf{3}}\} = g.n$$

It is clear that the uncertainty term related to the belief of evidences A and B are d+e+f+g and k+l+m+n respectively. Also, the normalized uncertainty term for the combination of two beliefs can be written as:

$$\frac{\mu_{AB}\{C_1,C_2\}+\mu_{AB}\{C_1,C_3\}+\mu_{AB}\{C_2,C_3\}+\mu_{AB}\{C_1,C_2,C_3\}}{1-\mu_{AB}(\varnothing)}$$
$$=\frac{d.k+d.n+g.k+e.l+e.n+g.l+f.m+f.n+g.m+g.n}{1-\mu_{AB}(\varnothing)}$$
$$=d.\frac{k+n}{1-\mu_{AB}(\varnothing)}+e.\frac{l+n}{1-\mu_{AB}(\varnothing)}+g.\frac{k+l+m+n}{1-\mu_{AB}(\varnothing)}+f.\frac{m+n}{1-\mu_{AB}(\varnothing)} \tag{10}$$

To find the conditions in which the uncertainty of the combined beliefs is lower than the uncertainty of an individual belief (e.g: evidence A), consider:

$$d+e+f+g>d.\frac{k+n}{1-\mu_{AB}(\varnothing)}+e.\frac{l+n}{1-\mu_{AB}(\varnothing)}$$
$$+g.\frac{k+l+m+n}{1-\mu_{AB}(\varnothing)}+f.\frac{m+n}{1-\mu_{AB}(\varnothing)} \tag{11}$$

On the left hand side of (11), the coefficients of each BBA (d, e, f and g) are equal to one but on the right hand side, all of BBAs have coefficients not equal to one. Note that $\frac{k+l+m+n}{1-\mu_{AB}(\varnothing)}$ is the largest coefficient. If $\frac{k+l+m+n}{1-\mu_{AB}(\varnothing)}\leq 1$ is satisfied, then all of the other coefficients on the right hand side are lower than one and consequently (11) is valid.

$$\frac{k+l+m+n}{1-\mu_{AB}(\varnothing)}\leq 1\Rightarrow k+l+m+n\leq 1-\mu_{AB}(\varnothing) \tag{12}$$

Since the sum of all BBAs is equal to one, we have:

$$k+l+m+n=1-(h+i+j) \tag{13}$$

Substitution of (13) into (12), the following condition is achieved:

$$h+i+j\geq\mu_{AB}(\varnothing) \tag{14}$$

The same scenario is also valid for other evidence (in this case B) which leads to the following condition:

$$a+b+c\geq\mu_{AB}(\varnothing) \tag{15}$$

Accordingly, if the sum of certainty terms of the belief of evidence (B or A) is equal or greater than the null mass of the combination of two beliefs (Equations (14) and (15)), then the uncertainty of the combined belief is always lower than the uncertainty of the other belief of the evidence (A or B respectively). Therefore, the uncertainty of combined beliefs decreases.

This set of conditions can be used in two general ways:

1- To verify the guarantee of uncertainty decrement in a specific problem.

2- To find the area of the parameters in which the uncertainty reduction can be assumed.

Furthermore, these conditions can be used as a general guideline for selecting the parameters of design procedure (even if not granted).

2.3. **Singular Value Decomposition.** Singular Value Decomposition (SVD) is one of the important tools used in digital signal and statistics data processing [17, 20]. Applying SVD on a typical $X_{\mathrm{m\times n}}$ matrix yields:

$$X=U\Sigma V^T \tag{16}$$

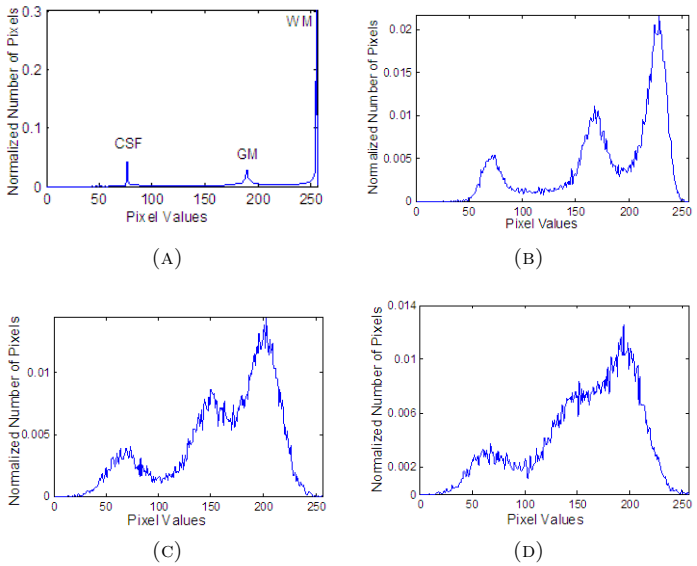


FIGURE 1. Normalized Histogram of T1-weighted Brain MRI with; A) 0% Noise and 0% INU, B) 3% Noise and 20% INU, C) 7% Noise and 20% and D) 9% Noise and 40%

Where $U_{\mathrm{m\times m}}$ and $V_{\mathrm{n\times n}}$ are singular vector matrices and $\Sigma_{\mathrm{m\times n}}$ is a diagonal matrix with the rank of r. The diagonal entries of ($\sigma_{11} > \sigma_{22} > \ \ldots > \sigma_{rr} > 0$) are equal to the singular values of X. In fact, singular values contain some information of signal energy. In a reconstruction process of X, higher singular values are more effective.

3. **Proposed Method**

Since there are three main tissues, WM, GM and CSF, in the normal brain MRI, we set the number of clusters in the FCM based algorithm to three. When FCM reaches a solution after an iterative process, the pixels are assigned to cluster with the biggest MV. In such strategies, it is expected that the extracted clusters fit to the original classes (WM, GM and CSF). Brain MRI artifacts generate an overlap between tissues. The normalized histogram [28] of a slice of T1-weighted brain MRI with and without artifacts are shown in Figure 1.

As shown in the Figure 1 (A), the pixel values of the three tissues have three separate peaks in image without artifacts. According to the Figure 1 (B), (C) and (D), there is clear overlap in gray level values of the three tissues and the overlaps increases when noise and INU levels increases. As a result, while artifact effects become further in brain MRI, the aforementioned expectation will be unattainable. In other words, a bigger MV may be assigned to the cluster that does not represent a single tissue. In the proposed method, this problem is tackled by considering the

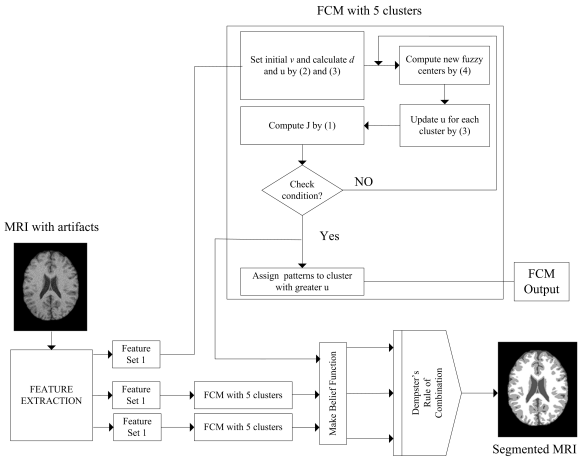


FIGURE 2. Black Diagram of the Proposed Method

main and ancillary clusters in FCM and then merging them by employing DST. The proposed method can be presented using the three main steps below:

1- Feature extraction and Feature selection.
2- Uncertain segmentation of brain tissue by considering the main and ancillary clusters.
3- Merge the ancillary clusters with the main clusters.

As in the proposed method the results of FCM are merged by DST of evidence, this method is called FCMMDS. The black diagram of FCMMDS is shown in Figure 2.

As shown in Figure 2, In the FCMMDS, Each FCM with one of the selected feature set is considered as a body of evidence. Moreover, Dempster's rule of combination is utilized to combine the outputs of FCM and reduce the level of uncertainty.

3.1. **Feature Extraction and Feature Selection.** In this study, the main factors for selecting features are simplicity, containing effective information and also robustness against artifact invasion. In view of this strategy, the features extracted in the analysis are: pixel intensity (f_1), mean (f_2), standard deviation (f_3) largest singular value (f_4) and median (f_5) of the neighboring pixels. It is clear that the pixel intensity is the most important feature in BMS which contains precious information. The other features prepare significant spatial information and also f_1, f_2 and f_4 give a good robustness against noises. It is worthy to mention that noise has smaller impact on the largest singular value in comparison with other singular values.

For calculation of the mean, standard deviation and largest singular value of a prototype pixel, a 3×3 neighborhood is considered. The spatial features are formalized as f_2, f_3 and f_4 in (17) to (19).

$$f_2 = \frac{\sum_i^n x_i}{n} = E(x) \tag{17}$$

$$f_3 = \sqrt{\frac{\sum_i^n (x_i - E(x))^2}{n}} \tag{18}$$

Where x_i is a neighborhood pixel value, n=8 is the number of neighborhood pixels and E(x) is the mean or expected value. The largest singular value can be calculated as:

$$SVD(L) = [U] \underbrace{\begin{bmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix}}_{\Sigma} [V]^T \Rightarrow f_4 = \sigma_{11} \tag{19}$$

σ_{11} is the largest singular value of L which contains significant information of the matrix and can be used as an important feature in BMS.
For feature selection, there are some cases for combination of the extracted features. In this study three appropriate feature sets, $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$ are used as follows:

$$\mathbf{F}_1 = [f_1, f_2],\ \mathbf{F}_2 = [f_1, f_4],\ \mathbf{F}_3 = [f_1, f_3, f_5] \tag{20}$$

As the pixel intensity is the most important feature in the BMS, it employed as an important part in the construction of feature sets.

3.2. **Uncertain Segmentation.** The normalized histogram of T1-weighted brain MRI (with 5% noise and 20% INU) and also the associated tissues are shown separately in Figure 3.

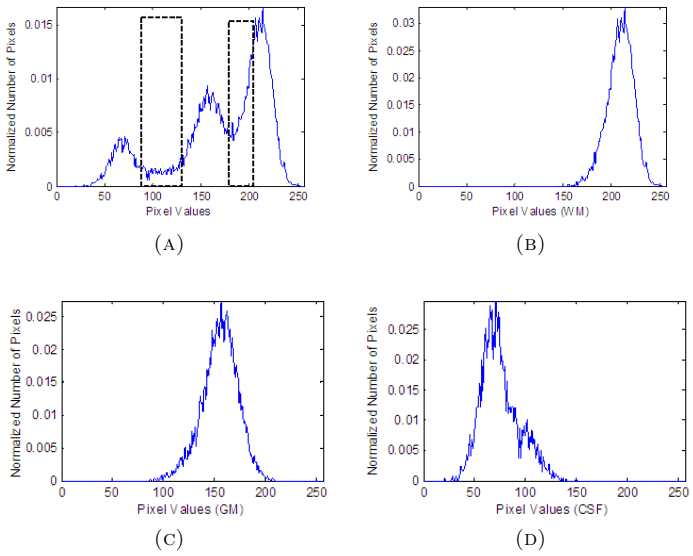


FIGURE 3. Normalized Histograms of T1-weighted Brain MRI with 5% Noise and 20% INU; A) All Tissues; B) WM; C) GM and D)CSF.

Three peaks in the histogram, Figure 3, represent the three main tissues in the brain. As shown in the Figure 3 (A), the GM tissue has overlap with both WM and CSF tissues, while WM and CSF have not overlap. Consequently in the FCMMDS, we assign five clusters in which all pixels between GM and CSF, also GM and WM are considered in two ancillary clusters. Therefore the five considered clusters can be categorized in two categories; three original classes of brain tissues (WM, GM and CSF), and the joint classes represented by the ancillary clusters (CSF-GM and GM-WM). By applying the FCM on all pixels with aforementioned strategy, all pixels will have five MVs for considered clusters.
In the next step, the five available clusters are converted to three classes by employing Dempster's rule of combination.

3.3. **Merging the Clusters.** As mentioned above, after FCM reach a solution, there are five MVs for each pixel in which those are represent the degree assignment of pixel to five considered clusters. In the third step of the FCMMDS, the five extracted MVs considered as BBA where the frame of discernment contains three individual classes as {CSF, GM , WM} and the focal elements of that is {{GM}, {WM}, {CSF}, {GM-WM}, {GM-CSF}}. By sorting the center of clusters and referring to the Figure 1, it is evident that the minor, middle and major centers are associated with the CSF, GM and WM respectively. Moreover the center between minor and middle represents the CSF-GM and the one of between middle and major represents GM-WM.
Suppose that for a prototype pixel the last two steps give us the five MVs for five sorted centers as follows:

F_1	$u_{\mathrm{CSF}} = 0.05$	$u_{\mathrm{CSF-GM}} = 0.15$	$u_{\mathrm{GM}} = 0.25$	$\mathrm{u}_{\mathrm{GM-WM}} = 0.35$	$u_{\mathrm{WM}} = 0.20$
F_2	$u_{\mathrm{CSF}} = 0.15$	$u_{\mathrm{CSF-GM}} = 0.05$	$u_{\mathrm{GM}} = 0.20$	$\mathrm{u}_{\mathrm{GM-WM}} = 0.32$	$u_{\mathrm{WM}} = 0.28$
F_3	$u_{\mathrm{CSF}} = 0.10$	$u_{\mathrm{CSF-GM}} = 0.40$	$u_{\mathrm{GM}} = 0.25$	$\mathrm{u}_{\mathrm{GM-WM}} = 0.10$	$u_{\mathrm{WM}} = 0.15$

Where u represents the MV of the pixel to the available cluster. According to the FCM results on the F_1 and F_2 the MVs show that there is a big concern about assigning the pixel to the one the GM or WM tissues, because the MV of the GM-WM is bigger than the other MVs. Furthermore, analyzing the figures for F_3 shows that assigning the pixel to CSF or GM is the main problem ($\mathrm{u}_{\mathrm{CSF-GM}} = 0.40$). After combining the results of FCM by (8) the following figures are achieved as: Finally, according to the combined mass values and by using the pignistic

$$\mu_{\mathrm{CSF}} = 0.04, \quad \mu_{\mathrm{CSF-GM}} = 0.01, \quad \mu_{\mathrm{GM}} = 0.75, \quad \mu_{\mathrm{GM-WM}} = 0.03, \quad \mu_{\mathrm{WM}} = 0.17$$

probability by (9), the prototype pixel can be assigned to GM.

4. Experimental Results

4.1. **Evaluation Criteria.** In this study, two measures have been utilized in order to evaluate the segmentation performance quantitatively. The evaluation criteria are Dice and Tanimoto similarity coefficients [11, 22] as below:

$$Dice = \frac{2\,|A_i \bigcap B_i|}{|A_i| + |B_i|} = \frac{2 \times TP}{(2 \times TP) + FP + FN} \tag{21}$$

$$Tanimoto = \frac{|A_i \bigcap B_i|}{|A_i \bigcup B_i|} = \frac{|A_i \bigcap B_i|}{|A_i| + |B_i| - |A_i \bigcap B_i|} = \frac{TP}{TP + FP + FN} \tag{22}$$

Where A_i and B_i are the reference and segmented pixels sets respectively. $|A_i|$ is the number of A_i pixels [2]. In (21) and (22) TP, TN, FP and FN are True Positive, True Negative, False Positive and False Negative respectively [22]. The greater the criteria, the better the segmentation.

4.2. **Simulated Brain MRI and Results.** T1-weighted brain MR images with slice thickness of 1 mm and volume size of $217 \times 181 \times 181$ are employed to investigate the proposed method. These images are obtained from the BrainWeb Simulated Brain Database of the McConnell Brain Imaging Centre of the Montreal Neurological Institute (MNI), McGill University (http: //www.bic.mni.mcgill.ca/brainweb). For brain images, extra non-brain tissues are removed prior to segmentation. Furthermore, the proposed algorithm compared with the three methods of FCM, BCFCM and LNLFCM [48]. In the FCM the intensity of pixels are used as feature, in the BCFCM all parameters are selected as reported in [2]. The parameters for LNLFCM are selected as: the standard deviation of the Gaussian Kernel is 30, the degree of filtering is 1000, window size of the neighborhood is 3 and the weights of all pixels in the window of searching are 7.
Figure 4 (A) shows a sample middle slice of the simulated 3D volume of MRI with 9% Rician noise and 40% INU (slice=96). The ground truth of this image is shown in the Figure 4 (B). Segmentation results of FCM, BCFCM and LNLFCM are shown at the Figures 4 (C), (D) and (E), respectively. The result of the proposed method is shown in Figures 4 (F). The Dice and Tanimoto coefficients of the corresponding images for the four methods are shown in Table 1.

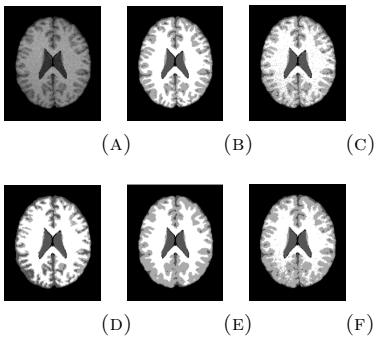


FIGURE 4. Comparison of the Segmentation Results on a Simulated MRI Brain Image: (A) Original Image with 9% Noise and 40% INU (Slice = 96); (B) Ground Truth; (C) FCM; (D) BCFCM; (E) LNLFCM and (F) FCMMDS Segmentation Results.

As shown in the Figure 4, we can visually see that the standard FCM and the BCFCM are influenced by the artifact, intensively. Also by visually evaluation of the images of Figures 4 (E) and (F), it is sense that the LNLFCM shows a better results than FDMMDS, but by referring to the ground truth, reported in Figure 4 (B), it is demonstrated that the FCMMDS has better results than LNLFCM. In fact the LNLFCM removes some small areas of different tissues in the process of artifact effect eliminating. Dice and Tanimoto coefficients of the images shown in Figure 4, reported in Table 1, confirm the finding.

	CSF		GM		WM	
	Dice	Tanimoto	Dice	Tanimoto	Dice	Tanimoto
FCM	0.897	0.813	0.826	0.704	0.897	0.813
BCFCM	0.777	0.635	0.724	0.567	0.918	0.849
LNLFCM	0.890	0.802	0.882	0.789	0.932	0.873
FCMMDS	0.929	0.867	0.888	0.799	0.930	0.870

Table 1. Different Coefficients for Different Methods in Figure 4

According to the quantitative measure concluded in Table 1, the FCM cannot represent a good result for any of the tissues. Although the BCFCM has better result then FCM in the segmentation of WM tissue, it failed in separating CSF and GM tissues. In comparison to FCM and BCFCM, the LNLFCM has the best results on all the three tissues, but the performance of our FCMMDS is clearly better. Note that to evaluate the effectiveness of segmentation method in each tissue, the quantitative criterion should be considered along with associated image. As the segmentation of upper and lower brain images are more complicated [23] and for further investigation under different artifact, the following evaluation are experimented.

The FCMMDS and the other mentioned algorithm are applied on an upper, lower and another middle slice of 3D volume with various artifacts levels and the results are shown in Figure 5 and Table 2.

Figures 5 (A), (G) and (M) are the simulated MRI images with noise levels 3%, 5%, 7% and without INU, respectively. Figures 5 (B), (H) and (N) show the ground truth of the corresponding noisy images. FCMMDS outputs are shown in Figures 5 (F), (R), (L), respectively. The mean Dice and Tanimoto similarity coefficients of the Figure 5 are shown in the Table 2.

	Slice30		Slice70		Slice130	
	Dice	Tanimoto	Dice	Tanimoto	Dice	Tanimoto
FCM	0.838	0.721	0.933	0.874	0.965	0.932
BCFCM	0.550	0.379	0.834	0.715	0.532	0.362
LNLFCM	0.901	0.820	0.904	0.825	0.900	0.818
FCMMDS	0.920	0.852	0.933	0.874	0.942	0.890

Table 2. Mean of Coefficients for Different Methods in Figure 5

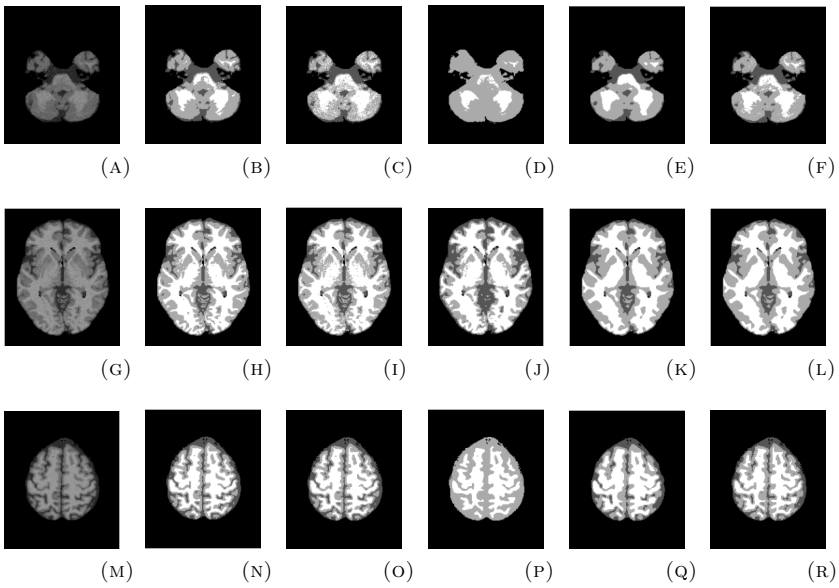


FIGURE 5. Comparison of the Segmentation Results on a Simulated MRI Brain Image: (A, G and M) Original Image with 7, 5 and 3% Noise (Slice = 30, 70, 130 Respectively); (B, H and N) Ground Truth; (C, I and O) FCM; (D, J and P) BCFCM; (E, K and Q) LNLFCM and (F, L and R) FCMMDS

According to the Table2, in lower noise levels (slice 130 with 3% noise) the FCM has better results than proposed method, but when the noise level increases dramatically (slice=30, noise 7%), the algorithm fails. By comparing the results of the proposed method, FCM, BCFCM and LNLFCM with the ground truth, reported in the Figures 5 and also the corresponding Dice and Tanimoto coefficients in the Table 2, it is evident that the FCMMDS performs better than the other three methods in this image set. These results demonstrate that FCMMDS is a robust technique for BMS.

5. **Conclusion**

In the proposed method, called FCMMDS, the appropriate combination of the spatial information are used as feature sets. In the FCMMDS by assigning all pixels to main and ancillary clusters, the trouble of uncertainty in pixel values is tackled. This process creates extra clusters in which in the final steps, they merged to main clusters using Dempster's rule of combination. The FCMMDS has been compared with standard FCM, BCFCM and LNLFCM algorithms in which the proposed method shows good results. It is shown that FCMMDS has satisfactory results on upper and lower brain slices which are critical in brain MRI segmentation.

Furthermore, the results of our proposed algorithm under various levels of artifacts demonstrated that FCMMDS is a robust technique for brain MRI segmentation.

References

[1] W. Abd-Almageed, A. El-Osery and C. Smith, *A fuzzy-statistical contour model for MRI segmentation and target tracking*, presented at the SPIE, Orlando, FL, USA, (2004), 25–33.

[2] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag and T. Moriarty, *A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data*, IEEE transactions on medical imaging, **21(3)** (2002), 193–199.

[3] S. P. Awate, H. Zhang, T. J. Simon and J. C. Gee, *Multivariate segmentation of brain tissues by fusion of MRI and DTI data*, presented at the Proceedings of the 2008 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Paris, France, (2008).

[4] M. Balafar, A. Ramli, M. Saripan and S. Mashohor, *Review of brain MRI image segmentation methods*, Artificial Intelligence Review, **33(3)** (2010), 261–274.

[5] M. Beynon, D. Cosker and D. Marshall, *An expert system for multi-criteria decision making using Dempster Shafer theory*, Expert Systems with Applications, **20(4)** (2001), 357–367.

[6] E. Binaghi and P. Madella, *Fuzzy DempsterShafer reasoning for rule-based classifiers*, International Journal of Intelligent Systems, **14(6)** (1999), 559-583.

[7] I. Bloch, *Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account*, Pattern Recognition Letters, **17(8)** (1996), 905–919.

[8] M. Bomans, K. H. Hohne, U. Tiede and M. Riemer, *3-D segmentation of MR images of the head for 3-D display*, IEEE transactions on medical imaging, **9(2)** (1990), 177–183.

[9] C. Brechbhler, G. Gerig and G. Szkely, *Compensation of spatial inhomogeneity in MRI based on a multi-valued image model and a parametric bias estimate*, In Visualization in Biomedical Computing, (1996), 141–146.

[10] K. S. Chuang, H. L. Tzeng, S. Chen, J. Wu and T. J. Chen, *Fuzzy c-means clustering with spatial information for image segmentation*, Computerized Medical Imaging and Graphics : the Official Journal of the Computerized Medical Imaging Society, **30(1)** (2006), 9-151.

[11] A. Demirhan and I. Gler, *Combining stationary wavelet transform and self-organizing maps for brain MR image segmentation*, Engineering Applications of Artificial Intelligence, **24(2)** (2011), 358–367.

[12] J. Ghasemi, R. Ghaderi, M. R. Karami Mollaei and A. Hojjatoleslami, *Separation of brain tissues in MRI based on multi-dimensional FCM and spatial information*, Eighth International Conference on in Fuzzy Systems and Knowledge Discovery (FSKD), (2011), 247–251.

[13] J. Ghasemi, M. R. Karami Mollaei, R. Ghaderi and A. Hojjatoleslami, *Brain tissue segmentation based on spatial information fusion by Dempster-Shafer theory*, Journal of Zhejiang University - Science C, **13(7)** (2012), 520–533.

[14] J. D. Gispert, S. Reig, J. Pascau, J. J. Vaquero, P. Garcia-Barreno and M. Desco, *Method for bias field correction of brain T1-weighted magnetic resonance images minimizing segmentation error*, Human brain mapping, **22(2)** (2004), 133–144.

[15] M. Hasanzadeh and S. Kasaei, *Multispectral Brain MRI Segmentation based on Fuzzy Classifiers and Evidence Theory*, presented at the 15th Iranian Conference on Electrical Engineering, ICEE, Tehran, Iran, 2007.

[16] T. Heinonen, P. Dastidar, H. Eskola, H. Frey, P. Ryymin and E. Laasonen, *Applicability of semi-automatic segmentation for volumetric analysis of brain lesions,* Journal of Medical Engineering And Technology, **22(4)** (1998), 173–178.

[17] S. K. Jha and R. D. S. Yadava, *Denoising by singular value decomposition and its application to electronic nose data processing*, IEEE Sensors Journal, **11(1)** (2011), 35–44.

[18] L. Ji and H. Yan, *An attractable snakes based on the greedy algorithm for contour extraction*, Pattern Recognition, **35(4)** (2002), 791–806.

[19] Z. X. Ji, Q. S. Sun and D. S. Xia, *A modified possibilistic fuzzy c-means clustering algorithm for bias field estimation and segmentation of brain MR image*, Computerized Medical Imaging and Graphics, **35(5)** (2011), 383–397.

[20] L. Jui-Hsiang, T. Ming-Feng, C. Lumdo and C. C. P. Chen, *Accurate and analytical statistical spatial correlation modeling based on singular value decomposition for VLSI DFM applications*, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, **29(4)** (2010), 580-589.

[21] F. Kyoomarsi, H. Khosravi, E. Eslami and M. Davoudi, *Extraction-based Text Summarization Using Fuzzy Analysisn*, Iranian Journal of Fuzzy Systems, **7(3)** (2010), 15–32.

[22] Lladó, A. Oliver, M. Cabezas, J. Freixenet, J. C. Vilanova, A. Quiles, L. Valls, L. Ramió-Torrent and À. Rovira, *Segmentation of multiple sclerosis lesions in brain MRI: a review of automated approaches*, Information Sciences, **186(1)** (2012), 164–185.

[23] A. W. Liew and H. Yan, *An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation*, IEEE transactions on medical imaging, **22(9)** (2003), 1063–1075.

[24] A. Liew and H. Yan, *Current methods in the automatic tissue segmentation of 3D magnetic resonance brain images*, Current Medical Imaging Reviews, **2(1)** (2006), 91–103.

[25] E. G. Mansoori, M. J. Zolghadri and S. D. Katebi, *Using distribution of data to enhance prformance of fuzzy classification systems*, Iranian Journal of Fuzzy Systems, **4(1)** (2007), 21–36.

[26] E. G. Mansoori, M. J. Zolghadri, S. D. Katebi, H. Mohabatkar,R. Boostani and M. H. Sadreddini, *Generating fuzzy for protein classification*, Iranian Journal of Fuzzy Systems, **5(2)** (2008), 21–33.

[27] T. McInerney and D. Terzopoulos, *Deformable models in medical image analysis: a survey*, Medical Image Analysis, **1(2)** (1996), 91–108.

[28] S. B. Mehta, S. Chaudhury, A. Bhattacharyya and A. Jena, *Handcrafted fuzzy rules for tissue classification*, Magnetic Resonance Imaging, **26(6)** (2008), 815–823.

[29] F. Moayedi, R. bostani, A. R. Kazemi, S. Katebi and E. Dashti, *Subclass Fuzzy-SVM classifier as an efficient method to enhance the mass detection in mamograms*, Iranian Journal of Fuzzy Systems, **7(1)** (2010), 15–31.

[30] W. J. Niessen, K. L. Vincken, J. Weickert, B. M. T. H. Romeny and M. A. Viergever, *Multiscale segmentation of three-dimensional MR brain images*, International Journal of Computer Vision, **31(2)** (1999), 185–202.

[31] D. L. Pham and J. L. Prince, *An adaptive fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneities*, Pattern Recognition Letters, **20(1)** (1999), 57–68.

[32] D. L. Pham and J. L. Prince, *Adaptive fuzzy segmentation of magnetic resonance images*, IEEE Transactions on Medical Imaging, **18(9)** (1999), 737-752.

[33] D. L. Pham, C. Xu and J. L. Prince, *A survey of current methods in medical image segmentation*, Annual Review of Biomedical Engineering, **2** (2000), 315–337.

[34] S. Prima, N. Ayache, T. Barrick and N. Roberts, *Maximum likelihood estimation of the bias field in MR brain images: investigating different modelings of the imaging process*, presented at the Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention, Utrecht, The Netherlands, 2001.

[35] S. Ramathilagam, R. Pandiyarajan, A. Sathya, R. Devi and S. R. Kannan, *Modified fuzzy c-means algorithm for segmentation of T1-T2-weighted brain MRI*, Journal of Computational and Applied Mathematics, **235(6)** (2011), 1578–1586.

[36] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, Princeton., 1976.

[37] S. Shen, W. Sandham, M. Granat and A. Sterr, *MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization*, IEEE Transactions on Information Technology in Biomedicine : a Publication of the Ieee Engineering in Medicine and Biology Society, **9(3)** (2005), 459–67.

[38] A. Simmons, P. S. Tofts, G. J. Barker and S. R. Arridge, *Sources of intensity nonuniformity in spin echo images at 1.5 T*, Magnetic Resonance in Medicine : official journal of the Society of Magnetic Resonance in Medicine, **32(1)** (1994), 121–8.

[39] M. Y. Siyal and L. Yu, *An intelligent modified fuzzy c-means based algorithm for bias estimation and segmentation of brain MRI*, Pattern Recognition Letters, **26(13)** (2005), 2052-2062.
[40] J. G. Sled, A. P. Zijdenbos and A. C. Evans, *A nonparametric method for automatic correction of intensity nonuniformity in MRI data*, IEEE Transactions On Medical Imaging, **17(1)** (1998), 87–97.
[41] P. Smets and R. Kennes, *The transferable belief model*, Artificial Intelligence, **66 (2)** (1994), 191–234.
[42] M. Styner, C. Brechbuhler, G. Szekely and G. Gerig, *Parametric estimate of intensity inhomogeneities applied to MRI*, IEEE Transactions on Medical Imaging, **19(3)** (2000), 153–165.
[43] M. Tabassian, R. Ghaderi and R. Ebrahimpour, *Combination of multiple diverse classifiers using belief functions for handling data with imperfect labels*, Expert Systems with Applications, **39(2)** (2012), 1698–1707.
[44] M. Tabassian, R. Ghaderi and R. Ebrahimpour, *Knitted fabric defect classification for uncertain labels based on Dempster-Shafer theory of evidence*, Expert Systems with Applications, **38(5)** (2011), 5259–5267.
[45] L. Tzu-Chao, *Switching-based filter based on Dempsters combination rule for image processing*, Information Sciences, **180(24)** (2010), 4892–4908.
[46] v, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, 1981.
[47] F. Valente, *Multi-stream speech recognition based on DempsterShafer combination rule*, Speech Communication, **52(3)** (2010), 213–222.
[48] J. Wang, J. Kong, Y. Lu, M. Qi and B. Zhang, *A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints*, Computerized Medical Imaging and Graphics : the Official Journal of the Computerized Medical Imaging Society, **32(8)** (2008), 685–698.
[49] R. R. Yager, J. Kacprzyk and M. Fedrizzi, *Advances in the Dempster-Shafer theory of evidence*, New York ; Chichester: Wiley, 1994.
[50] D. Q. Zhang and S. C. Chen, *A novel kernelized fuzzy c-means algorithm with application in medical image segmentation*, Artificial Intelligence in Medicine, **32(1)** (2004), 37-50.

Jamal Ghasemi*, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran
E-mail address: j.ghasemi@umz.ac.ir

Mohamad Reza Karami Mollaei, Faculty of Electrical and Computer Engeniering, Babol University of Technology, P.O.Box 484, Babol, Iran
E-mail address: mkarami@nit.ac.ir

Reza Ghaderi, Shahid Beheshti University, Tehran, Iran
E-mail address: r_ghaderi@sbu.ac.ir

Ali Hojjatoleslami, School of computing, University of Kent, Canterbury,CT2 7PT UK
E-mail address: S.A.Hojjatoleslami@kent.ac.uk

*Corresponding author