Ensemble Classification Using Entropy-Based Features for MRI Tissue Segmentation

Nadjet Bouchaour and Smaine Mazouzi
LICUS Lab, University 20 Aout 1955, BP 26 Route el Hadaik, 21000 Skikda, Algeria
n.bouchaour@univ-skikda.dz

Abstract. It is still hard to deal with artifacts in magnetic resonance images (MRIs), particularly when the latter are to be segmented. This paper introduces a novel feature, namely the spatial entropy of intensity that allows a pattern-based representation which enhances the MRI segmentation despite presence of high levels of noise and intensity non-uniformity (INU) within MRI data. Moreover, we bring out that ensembles of classifiers used with the proposed feature have significantly enhanced structured MRI segmentation. Thus, to conduct experiments, MRIs with different artifact levels were extracted and exploited from the Brain Web MRI database. The obtained results reveal that the proposed feature, especially when used with ensembles of classifiers has significantly enhanced the overall MRI segmentation.

Keywords: MRI tissue segmentation · Spatial entropy · Classifier boosting · Ada-Boost

1 Introduction

Pattern-based classification has wined more interest in the last decade. According to such approach, not only data are used to infer features, but also patterns that exist within data. These patterns are also used for classification and feature definition. It has been stated in several works [5] that pattern mining can help to enhance data classification, mainly with structured domains, such as object recognition and image analysis. Consequently, for several applications in the aforementioned fields, patterns can be defined as sequences of graphs in the raw data [18,26]. So defining pattern-based classification models undoubtedly allows to enhance classification, rather than using models, based solely on raw data.

Furthermore, combining classifiers and training them together with good partition of the training data, has allowed to enhance the overall classification results. Also, all the works having used the ensemble classifiers have yielded better results than those using single classifiers, regardless the weak elementary classifier that was used as constructing element of the ensemble. It has been stated in several works in the literature that combining classifiers according to
various ensemble methods allows to avoid several classical flaws in data classification. The most known flaw is the over-training problem, mainly when the volume of training data is too large. Furthermore, classifier combining allows to deals with the problem of the early convergence of the algorithm when the classification is optimization-based such as in the neural based methods [9,14].

In this work, we are interested in the classification of the MRI data using an ensemble of classifiers, namely the boosting algorithm [8]. In contrast to most of the published works, and instead of using the raw image data, taken individually at the different voxels, we consider a pattern-based representation where the neighboring voxels are aggregated to form patterns, and then use these latter as features for classification. To construct the pattern which will be used as the classification feature, we use an energy based coding, assuming that the energy defined within the neighborhood of a voxel represents well the voxel separation. Therefore, the model is based on defining within a neighborhood of a voxel an energy function represented as a spatial entropy of voxels' intensities. Such an energy function allows to consider, in addition to the image intensity, the local geometrical spatial relations that exist in the MRI data. We show that the aggregation of data according to the considered pattern, especially with energy coding, as well as the classification of these patterns with the most used ensemble of classifiers: the Adaboost, allows to significantly improve the results of classification of MRI data.

The next parts of the paper are structured as follows: firstly, in Sect. 2 we present a review of segmentation methods in medical imaging, including those based on machine learning, in addition to the main works in the literature that are close to the ours. Secondly, in Sect. 3 we provide details of our approach by presenting both the used pattern through energy coding, and the use of the Adaboost algorithm for MRI segmentation. Then, Sect. 4 is dedicated to the experimentation of our method, where we present both the used dataset and the obtained results, as well as a comparison with some methods of the literature. Finally, a concise conclusion to summarize our contribution and highlights some potential perspectives.

2 Related Work

Image segmentation is one of the most important tasks in the process of pattern recognition using visual data. It consists of subdividing the pixels/voxels of an image into distinct and homogeneous regions. There are dozens of different methods of segmentation. However, all these methods can be classified according to three main families:

1. Contour-based methods: the common principle of these methods is to detect discontinuities in visual data. These discontinuities represent the edges in the image. The detected edges are generally disjoint and open, and therefore they must be joined and closed for proper use in the subsequent recognition process.
2. Region-based methods: their principle consists in grouping the pixels/voxels of the image having the same features, in disjoint but homogeneous subsets according to a certain homogeneity criterion. These homogeneous subsets are called regions.

3. Methods by classification: their major asset is that they allow the learning from the labeled data, forming the ground truth. Segmentation by classification consists of assigning a label to any pixel/voxel of the image using a classifier (single or ensemble). Given that we are interested in this last family in this work, we devote the remainder of this section to introduce some methods of MRI segmentation by data classification methods.

First, classification-based segmentation methods can be further subdivided into two sub-categories:

1. Heuristic-based methods: where one or more heuristics are considered to define a pixel/voxel labeling criterion. The heuristics consider a given prior, relative to the image, to the noise, or to the distortions that the image could undergo during its acquisition. For instance, we can cite the Fuzzy C-Mean \([3]\) algorithm, where the classification prior consists in considering for the pixels/voxels at the borders of the regions and elsewhere that there exists a mixture of information, each one relating to one of the data classes. Markovian methods \([12]\) consider the prior of smoothness, where the data are considered homogeneous by parts, and any part corresponds to a homogeneous region of the image. Also, Markovian representations can express some spatial constraints that the data must respect.

2. Methods by learning: where machine learning techniques are used. The principle is to proceed by learning classifiers using labeled data, so-called training set, then use the trained classifiers to classify the data, in this case called test set. According to the latter approach, several new methods based on the combination of classifiers have emerged. The ensemble of classifiers is one of these methods, and it is based on the aggregation of a set of classifiers, called weak classifiers, trained separately according to some particular sampling, allowing on the one hand the improvement of the performance of the classification and on the other hand, dealing with some problems inherent to machine learning, where the most known is the over-training problem.

Magnetic resonance imaging (MRI) is of great importance for the establishment of correct diagnoses and thus the prescription of appropriate treatments. MRI segmentation in order to extract tissues and establish diagnostic remains an active research field \([23]\). The segmentation of an MRI consists in extracting the main tissues for which physicians and radiologists are mainly interested. These tissues are respectively CSF (Cerebrospinal Fluid), GM (Gray Matter) and WM (White Matter), for structural MRI, and also LM (Lesional Matter) for pathological MRI. Several methods for MRI segmentation have been published, starting with contour detectors, passing through region extractors, and ending with machine-learning based methods. Richard et al. \([17]\) used a distributed approach with Markovian and Bayesian categorization of MRI tissues.
The principle of their method is to segment the volume into sub volumes and then make autonomous agents cooperate to produce an overall image segmentation. The method suffer from several problems including the ad-hoc subdivision of volumes. Also, the Markovian methods are known by their minimization iterative methods that are very time consuming. By adopting the same paradigm, Scherrer et al. [20] proposed a distributed Markov model for the classification of MRI data. In their work, the authors were able to formalize the classification by using both a multi-agent system for data distribution and processing, and a Markovian representation of MRI data, allowing classification using Markovian classifiers and dealing with spatial constraints at the same time.

Several works have proposed machine-learning methods for MRI segmentation. However, few of them have used ensemble of classifiers, where mainly the unique used feature was the voxel intensity. Some authors have proceeded to feature extraction then using the obtained features with ensemble classifiers to process MRI data. Rajasree et al. have considered a fractal representation of MRI data, by using the Brownian move technique [16]. The adopted features are then used with the Adaboost algorithm to detect tumors in MRI data. Gustavo et al. have combined Genetic Algorithms (GA) and Adaboost clustering to detect the tumor area in the MRIs [13]. After a data thresholding using the GA algorithm in order to delimit the tumor area, Adaboost is trained using the obtained classification by the GA algorithm, then used to finally detect the tumor as the largest connected component in the whole image.

Recently, deep learning techniques, mainly convolutional neural networks (CNN) were widely proposed for MRI data processing. Their strong advantage is that they do not need for feature representation and extraction. In such techniques MRI data in the input are convolved to kernels in the middle layers, so features are automatically produced. Output layers classify voxels according the produced features [4,24].

As far as we know, entropy-based features for MRI processing are rare in the literature. Sarita et al. have combined probabilistic neural network and wavelet entropy for feature extraction to classify MRI data [19]. Entropy-based features were also used with optimization-based clustering, such as in the work introduced by Pham et al. [15], where authors combined fuzzy entropy clustering and multi-objective particle swarm optimization. In order to extend data representation, some authors such as Bahadure et al. [2] have transformed data from spatial domain to frequency domain, then proceed by the SVM classifier to segment such represented MRI data.

In our work, a novel representation by the spatial entropy of the voxel intensities is introduced. Contrary to the previous cited works, the proposed entropy thanks to its spatial expression, takes into account both spatial and radiometric interactions of data within the MRIs.

3 Pattern-Based Features for MRI Data Classification

In this section, we define a pattern-based feature that will be tested for brain tissue classification in MRIs. A set of weak classifiers (Support Vector Machine
(SVM) or Naïve Bayes) are combined according the boosting algorithm to be trained and to be used to label the voxels of the MRI volume. Boosting is beneficial for MRI data segmentation because it allows to avoid the over-training problem by distributing the training data over several weak classifiers. Furthermore, Adaboost algorithm allows an adaptive distribution of data which is well appropriate for MRI data, given that MRI data distribution model is not beforehand known. At first, the MR image is pre-processed using a skull-strip algorithm, namely FSL Brain Extraction Tool (BET) [10, 22], to remove the non brain tissues. In our case we have preferred to avoid the noise filtering, given that, firstly, the MRI data are usually altered on the boundaries between the different tissues by the partial volume effect, where an averaging of the intensities at these voxels aggravate such effect. Secondly, because the proposed pattern allows the noise reduction on voxel classification even without MRI data smoothing. Such characteristic of the proposed method is explained by the fact that the used entropy-based pattern around a given voxel is a combination of the whole voxels belonging to the voxel neighborhood. Before introducing details how our method proceeds to label voxels, firstly, we present in the following section the Adaboost algorithm for two-classes classifiers, and then we pursue with how such an algorithm can be generalized for problems with more than two classes.

3.1 AdaBoost for MRI Data Classification

AdaBoost is a meta algorithm for data classification. It consists of an ensemble of classifiers, called weak in this case, that adaptively boost the performance of the ensemble classifier [8]. The outputs of the latter are combined according to a set of weights obtained according to the training data, providing the final AdaBoost classifier. Adaptation in such meta algorithm refers to that fact that the weak classifiers are adjusted by the subsets of the training data that were wrongly classified at the previous iterations. The algorithm bellow is an Adaboost implementation with a binary weak classifier $h$. After the learning of the ensemble classifier using the training set $(x_1, y_1), \ldots, (x_m, y_m)$ the whole hypothesis of the ensemble classifier is expressed by the function $H(x)$, where $x$ is an instance from the test set to be labeled $y \in \{-1, +1\}$.

To deal with problems with more than two classes, lets $M$ classes, $M - 1$ binary weak classifiers are used to substitute one weak $M$-classes’ classifier. There are two well known methods that allow combining binary classifiers to build a multi-class classifier. The first method proceeds according to the principle winner takes all [21], where the $k^{th}$ classifier allows to distinguish the class $k$ from all the other classes. For the second method, a classifier is dedicated to each couple of classes [11], where its role is to distinguish the two classes for which it is dedicated. We note that the two methods are implemented in the most of the machine-learning platforms.
**Algorithm 1: Adaboost**

**Result:** H(x)

Given: \((x_1, y_1), ..., (x_m, y_m)\) where \(x_i \in X, y_i \in \{-1, +1\}\)

Initialisation: \(D_1(i) = \frac{1}{m}\) for \(i = 1, ..., m\);

for \(t = 1, ..., T\) do

Train weak learner using distribution \(D_t\);

Get weak hypothesis \(h_t: X \rightarrow \{-1, +1\}\);

Aim: select \(h_t\) with low weighted error:

\[\varepsilon = \Pr_{i \sim D_t}[h_t(x_i) \neq y_i];\]

Choose \(\alpha = \frac{1}{2} \ln \left(\frac{1}{1 - \varepsilon}\right)\);

Update, for \(i = 1, ..., m\):

\[D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha t y_i h_t(x_i))}{Z_t};\]

where \(Z_t\) is a normalization factor (chosen so that \(D_{t+1}\) will be a distribution).

end

Output the final hypothesis:

\[H(x) = \text{sign}(\sum_{t=1}^{T} \alpha_t h_t(x))\]

---

### 3.2 MRI Data

The MRI volume obtained after skull-striping is a set of voxels that each one can belong to one of the three remaining tissues, namely, the Cerebrospinal Fluid (CSF), the Gray Matter (GM), and the White Matter (WM). Each of them is characterized by its mean intensity and the corresponding standard-deviation \((\mu_c, \sigma_c)\), \(c \in \{CSF, GM, WM\}\). We also assume that the intensity distribution in each tissue is Gaussian (see Formula 1).

\[
f_c(x_i, \mu_c, \sigma_c) = \frac{1}{\sigma_c \sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \mu_c)^2/2\sigma_c^2}\]

(1)

where \(x_i\) is the intensity of the voxel at the location \(i\).

### 3.3 Energy Coding-Based Classification

Our proposed entropy-based pattern aims to capture interactions between the voxels belonging to a local neighborhood. Such interactions can be represented according to an energy function. So, the proposed pattern for a given voxel \(i\) in the MRI volume is a vector of three components, where each component represents the spatial entropy of the intensities of the similar voxels in the neighborhood. Such subsets of similar voxels are obtained by the \(k\)-means algorithm, applied on the voxel’s neighborhood with three classes (CSF, GW, WM) (see Eq. 2).

\[
E_c = -\sum_{D_c} P_i \times \log_2 P_i
\]

(2)
where $D_c$ denotes the set of the voxels belonging to the class $c$, and $P_i$ is the probability that the voxel belongs to the class $c$:

$$P_i = \frac{\frac{1}{\sigma_c \sqrt{2\pi}} e^{\frac{1}{2} \left( x_i - \mu_c \right)^2 / 2\sigma_c^2}}{\sum_c \frac{1}{\sigma_c \sqrt{2\pi}} e^{\frac{1}{2} \left( x_i - \mu_c \right)^2 / 2\sigma_c^2}}$$

(3)

$\mu_c, \sigma_c$ are respectively the mean and the standard-deviation of the intensities of the voxels belonging to the class $c$ and situated in the neighborhood of the voxel in question $(i)$. So, a clustering by the $k$-means algorithm is performed at the voxel neighborhood, so the three subsets of voxels and their respective couples of $(\mu_c, \sigma_c)$, $c \in \{CSF, GM, WM\}$ are obtained. We notice that for the training MRIs we do not need for clustering in voxel neighborhood because we have the ground truth that allows to know the voxels of each cluster what allows to calculate the three spacial entropies.

Fig. 1. Principle of the proposed energy-coding-based classification.

As it can be noticed on Fig. 1, the vector of features that will be used for classification by the SVM or the Naïve Byes is composed, in addition to the voxel’s intensity $x_i$, of the three spatial entropies $E_1, E_2$, and $E_3$, obtained according to the clustering of the set of the voxels forming the local neighborhood. Such a pattern captures well the interactions of the voxels, and expresses well the spatial constraints that exist within the MRI data. The entropies $E_1$, $E_2$, and $E_3$ allow to distinguish the cases where the voxel is in the neighborhood of a tissue border or not. Also, they allow distinguishing if a voxel is affected by a high deviation due to noise. Obviously, the intensity of the voxel in question is considered for classification, so the resulting class is likely that of the tissue with the mean intensity is the closest, but adjusted if needed by the interaction with voxels in the neighborhood, expressed by the introduced spatial entropies.

4 Experimentation and Evaluation

The experimentation of the proposed pattern has been done using MRI volumes from the well known database brain web [7]. This database provide a
large set of MRI volumes with their ground truth labeling, enabling authors to use machine-learning based methods, and quantitatively evaluate the proposed methods. Furthermore MRIs can be obtained according to various levels of artifacts, namely noise, and INU. All MRIs are $181 \times 217 \times 181$ voxels of size. In this work, they are considered only MRIs with T1 modality. As in most of works in the literature of the field, we have used 70% of data for training and 30% for testing.

4.1 Parameter Selection

Only one parameter is to be set in the proposed method. It consists of the number of iterations of the Adaboost algorithm, which corresponds also to the number of weak classifiers that will be combined to classify the MRI voxels. This parameter is quantitatively adjusted by testing a set of values and then we choose that from which there is no more significant enhancement of classification accuracy, expressed by the Dice index (see subsection bellow). Finally, we want to point out that the size of the neighborhood for entropy calculation is fixed at $3 \times 3$.

4.2 Performance Evaluation

Two main indexes are usually computed to evaluate and compare segmentation methods based on classification and clustering. They are namely Jaccard and Dice indexes. Based on true positive ($TP$), true negative ($TN$), and false positive ($FP$) labeling instances, Jaccard coefficient is expressed as follows:

$$Jaccard = \frac{TP}{TP + TN + FP}$$  \hspace{1cm} (4)

Dice coefficient can be expressed as:

$$Dice = \frac{2TP}{2TP + TN + FP}$$  \hspace{1cm} (5)

We opted for Dice index for two main reasons: the first one is that it is the most cited in the literature and the second one is that it has been considered by works with which we compare ours.

4.3 Experimental Results

As it has been cited above, we have experimented the proposed features by using two different weak classifiers, namely the SVM classifier and Naïve Bayes classifier. Indeed, we have started our experimentation by testing several classifier models and we have noticed that the retained classifiers yielded score better. So, the introduced results corresponds to the use of two sets of weak classifiers, combined each one according the AdaBoost algorithm. We have considered two different classifiers, and we have compared our results to those of previous works,
in order to show that the improvement of the results is not due to the classifier itself, but to the proposed features, namely the spatial entropy.

First, we introduce in Figs. 2, 3, and 4 respectively the MRI obtained with 1% noise level and 0% INU level, the three brain tissues using SVM as weak classifier, and the three brain tissues using Naïve Bayes as weak classifier. We can visually notice, by comparing with the MRI image, that the obtained tissues are well delimited.

Table 1 shows the segmentation results according to the Dice index using a the SVM classifier as a weak classifier with the voxel intensity and its spatial entropies as the classification features.
Fig. 4. MRI segmentation results with noise level set to 1% and INU set to 0%. (a), (b), (c) are respectively the White matter, the Gray matter and the CSF for the case where the weak classifier is a Naïve Bayes.

Table 1. Segmentation results according to the Dice index for the different MRIs and the different brain matters (WM, GM and CSF). The classification features by the Support Vector Machine are the voxel intensity and its spatial entropies.

|   | WM | GM | CSF |
|---|----|----|-----|
|   | N INU 1 | 3 | 5 | 7 | 1 | 3 | 5 | 7 | 1 | 3 | 5 | 7 |
| 0% | 99.17 | 98.10 | 96.60 | 94.08 | 98.23 | 96.21 | 93.53 | 89.25 | 98.46 | 97.09 | 95.60 | 93.63 |
| 20% | 98.40 | 97.70 | 96.26 | 94.30 | 96.86 | 95.51 | 93.05 | 89.47 | 97.78 | 96.85 | 95.59 | 93.73 |
| 40% | 97.02 | 96.55 | 94.95 | 92.75 | 94.59 | 93.67 | 90.80 | 87.02 | 96.92 | 96.29 | 94.69 | 92.99 |
| 60% | 93.58 | 91.79 | 89.48 | 87.3 | 86.37 | 83.78 | 79.41 | 73.97 | 87.94 | 87.07 | 85.32 | 81.04 |
| 90% | 90.40 | 89.08 | 87.42 | 85.78 | 83.05 | 80.17 | 76.66 | 71.44 | 87.89 | 86.32 | 84.04 | 80.38 |

Table 2. Segmentation results according to the Dice index for the different MRIs and the different brain matters (WM, GM and CSF). The Adaboost of the Naïve Bayes classifier with the spatial entropies of the voxel and its intensity as the classification features.

|   | WM | GM | CSF |
|---|----|----|-----|
|   | N INU 1 | 3 | 5 | 7 | 1 | 3 | 5 | 7 | 1 | 3 | 5 | 7 |
| 0% | 98.90 | 97.53 | 95.17 | 91.48 | 97.41 | 94.88 | 90.30 | 82.45 | 97.36 | 96.05 | 94.18 | 91.96 |
| 20% | 98.08 | 97.14 | 94.32 | 91.55 | 95.94 | 94.15 | 88.98 | 82.57 | 96.72 | 95.73 | 94.45 | 92.15 |
| 40% | 96.19 | 94.52 | 92.30 | 89.67 | 92.56 | 89.54 | 84.63 | 78.68 | 95.39 | 94.78 | 92.95 | 91.72 |
| 60% | 90.09 | 88.58 | 87.08 | 86.24 | 77.73 | 73.94 | 69.88 | 68.08 | 86.29 | 85.47 | 83.97 | 80.68 |
| 90% | 86.18 | 85.70 | 85.24 | 84.89 | 67.67 | 66.91 | 65.69 | 65.31 | 83.97 | 84.05 | 82.16 | 79.96 |

Table 2 shows the segmentation results according to the Dice index using a Naïve Bayes as a weak classifier with the voxel intensity and its spatial entropies as the classification features.

According the results introduced in Table 1 and Table 2, we notice the high scores of segmentation when the spatial entropy is used as a classification feature.
This improvement can be explained by the capture of the interactions between neighboring voxels. These interactions are expressed thanks to an energy, formulated as a spatial entropy. A voxel is not classified solely according to its value (gray level) but according to the strength of its interactions with its neighboring voxels.

According the introduced results, we can notice that the method has well scored, and it presents a strong robustness against noise and INU, where Dice index was not drastically fall with high levels of these two artifacts. For instance, for white matter, the variation of the Dice index is from 98.40 to 96.26 with the SVM classifier for a variation in noise from 1% to 5%, with INU fixed to 20%. We can notice the same robustness against the INU. The variation is from 99.17 to 93.58 for the INU varying from 0% to 60% with a noise level set at 1%. With the Naïve Byes classifier, the method scores also well, and presents also good robustness against noise and INU. For CSF, the variation of the Dice index is from 97.78 to 95.59 with the SVM classifier for a variation in noise from 1% to 5%, with INU fixed to 20%. Also, the same robustness is noticed against the INU. The variation is from 98.46 to 96.92 for the INU varying from 0% to 40% with a noise level set at 1%.

Figures 5 and 6 introduce the variations of the Dice index according the number of iterations of AdaBoost algorithm, respectively in the two cases: using the SVM, or using the Naïve Bayes as weak classifier. We can notice that Dice index uniformly grows with the SVM classifier for the three tissues, and becomes stable from the 9th iteration. It is the same case for the Naïve Bayes classifier, however the latter become stable little early that the SVM one, namely from the 8th iteration. According to the two cases, we ca, conclude that the Adaboost algorithm does not require many iterations (less than 10 iterations) to be trained. Such a low number of iterations allows us to conclude that the classification of voxels according the proposed methods is too fast and can be envisaged for realtime applications.

In order to show the effectiveness of the proposed features, we introduce in Table 3 a comparison between the obtained MRI data classification results with SVM as weak classifier and those of some well cited works from the literature. We have considered MRI with 20% INU and different noise levels, and we compared results for WM and GM tissues.

Table 3. Comparison results according Dice index with some work from the literature.

| Noise method                    | WM  | GM  |
|---------------------------------|-----|-----|
| Fast%                           | 97  | 95  |
| SMP5%                           | 94  | 94  |
| NL-FCM%                         | 94  | 91  |
| Boosted spatial entropy (with SVM)% | 98  | 98  |
We can notice from the previous table that our method scores nearly better than all the cited methods. They are more close to Fast [25], considered the best, compared to the others [1,6]. Such results confirm that the proposed features when they are used with an ensemble classifier, such as Adaboost, are well appropriate to represent local patterns in MRIs data, especially when the latter are to be segmented.

4.4 Results Analysis and Discussion

According to the different results introduced below, we notice the strong improvement in segmentation results when spatial entropy is used as classification features, in addition to intensities of voxels. The use of the spatial entropy of the voxel with its intensity has allowed a large increase in the values of the Dice index, even with high levels of noise and INU. This can be explained by
the fact that the proposed spatial entropy, in addition to its ability to consider the voxel and its neighborhood, and as it has an energy nature, it expresses the interaction force between the voxels in the MRIs. Thus, a voxel is not classified solely according to its value, but also according to the force of interaction of the voxel with its neighborhood. Taking into account the spatial entropy has shown better robustness against noise and INU, especially when the levels of these artifacts are high. Such a result can be explained by the widening of the voxel’s field of interaction beyond its local neighborhood.

5 Conclusion

In this paper, we have introduced a new pattern-based set of features of MRI data and how it was used with an ensemble classifier for voxel classification. The introduced features consist, in addition to the voxel’s intensity, of the spatial entropies
of intensities of similar voxels in the voxel neighborhood. Such entropies are interesting to capture the interaction between neighboring voxels, so that they allow the latter to be better classified. Indeed, labeling a voxel in MRI does not depend only on its intensity but also on its interaction with the neighborhood. With the AdaBoost algorithm, two weak classifiers were tested, namely the SVM and the Naïve Bayes classifier. The experimentation of the proposed method, by varying the different artifact levels, showed a strong improvement in the classification results in both cases. Furthermore the Adaboost algorithm converges in less than 10 iterations, which make the training stage too fast. In future work, the proposed features will be tested with different classifiers including deep ones, and more optimisation could be applied to the AdaBoost algorithm, by considering priors on brain MRI data.

References

1. Ashburner, J., Friston, K.J.: Unified segmentation (2005)
2. Bahadure, N.B., Ray, A.K., Thethi, H.P.: Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. Int. J. Biomed. Imaging 9749108:1–9749108:12 (2017). https://doi.org/10.1155/2017/9749108
3. Bezdek, J., Ehrlich, R., Full, W.E.: FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. 10, 191–203 (1984)
4. de Brébisson, A., Montana, G.: Deep neural networks for anatomical brain segmentation. CoRR abs/1502.02445 (2015). http://arxiv.org/abs/1502.02445
5. Bringmann, B., Nijssen, S., Zimmermann, A.: Pattern-based classification: a unifying perspective. CoRR abs/1111.6191 (2011). http://arxiv.org/abs/1111.6191
6. Caldairou, B., Passat, N., Habas, P., Studholme, C., Rousseau, F.: A non-local fuzzy segmentation method: application to brain MRI. Pattern Recogn. 44(9), 1916–1927 (2011). https://doi.org/10.1016/j.patcog.2010.06.006, https://hal.archives-ouvertes.fr/hal-00476587
7. Cocosco, C., Kolokian, V., Kwan, R.S., Evans, A.: Simulated brain database home-page. https://brainweb.bic.mni.mcgill.ca/brainweb. Accessed 03 June 2020
8. Freund, Y., Schapire, R.: A short introduction to boosting (1999)
9. Hornik, K.: Approximation capabilities of multilayer feedforward networks. Neural Networks 4(2), 251–257 (1991). https://doi.org/10.1016/0893-6080(91)90009-T
10. Jenkinson, M.: Bet2 : MR-based estimation of brain, skull and scalp surfaces. In: Eleventh Annual Meeting of the Organization for Human Brain Mapping (2005). https://ci.nii.ac.jp/naid/10030066593/en/
11. Knerr, S., Personnaz, L., Dreyfus, G.: Single-layer learning revisited: a stepwise procedure for building and training a neural network. In: Fogelman Soulié, F., Hérault, J. (eds.) Neurocomputing: Algorithms, Architectures and Applications. NATO ASI Series, vol. F68, pp. 41–50. Springer-Verlag, Heidelberg (1990). https://doi.org/10.1007/978-3-642-76153-9_5
12. Li, S.: Markov random field modeling in image analysis. In: Advances in Pattern Recognition (2009)
13. Oliveira., G.C., Varoto., R., Jr., A.C.: Brain tumor segmentation in magnetic resonance images using genetic algorithm clustering and adaboost classifier. In: Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies. BIOIMAGING: BIOIMAGING, vol. 2, pp. 77–82. INSTICC, SciTePress (2018). https://doi.org/10.5220/0006534900770082

14. Park, J., Sandberg, I.W.: Approximation and radial-basis-function networks. Neural Comput. 5(2), 305–316 (1993). https://doi.org/10.1162/neco.1993.5.2.305

15. Pham, T.X., Siarry, P., Oulhadj, H.: A multi-objective optimization approach for brain MRI segmentation using fuzzy entropy clustering and region-based active contour methods. Magn. Resonance Imaging 61, 41–65 (2019)

16. Rajasree, R., Columbus, C.C.: Brain tumour image segmentation and classification system based on the modified adaboost classifier. Int. J. Appl. Eng. Res. 10(14) (2015)

17. Richard, N., Dojat, M., Garbay, C.: Distributed markovian segmentation: application to MR brain scans. Pattern Recogn. 40(12), 3467–3480 (2007). https://doi.org/10.1016/j.patcog.2007.03.019

18. Roma, A.A., et al.: Invasive endocervical adenocarcinoma: a new pattern-based classification system with important clinical significance. Am. J. Surg. Pathol. 39(5), 667–672 (2015). https://doi.org/10.1097/pas.0000000000000402

19. Saritha, M., Paul Joseph, K., Mathew, A.T.: Classification of mri brain images using combined wavelet entropy based spider web plots and probabilistic neural network. Pattern Recogn. Lett. 34(16), 2151–2156 (2013). https://doi.org/10.1016/j.patrec.2013.08.017

20. Scherrer, B., Forbes, F., Garbay, C., Dojat, M.: Distributed local MRF models for tissue and structure brain segmentation. IEEE Trans. Med. Imaging 28(8), 1278–1295 (2009)

21. Schölkopf, B., Burges, C., Vapnik, V.: Extracting support data for a given task. In: KDD (1995)

22. Smith, S.: Fast robust automated brain extraction. Human Brain Mapp. 17 (2002)

23. Yamanakkanavar, N., Choi, J.Y., Lee, B.: MRI segmentation and classification of human brain using deep learning for diagnosis of Alzheimer’s disease: a survey. Sensors 20(11), 3243 (2020). https://doi.org/10.3390/s20113243

24. Zhang, W., et al.: Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. NeuroImage 108, 214–224 (2015)

25. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57 (2001)

26. Zhou, C., Cule, B., Goethals, B.: Pattern based sequence classification. IEEE Trans. Knowl. Data Eng. 28, 1285–1298 (2016)