Detection of Brain Tumors from Magnetic Resonance Imaging by Combining Superpixel Methods and Relevance Vector Machines Classification

Ebrahim Akbari[1], Mehran Emadi[2*]
1- M.A. Student, Islamic Azad University, Mobarakeh Branch, Mobarakeh, Isfahan, Iran.
Email: Eb.ak57@gmail.com
2- Assistant Professor, Islamic Azad University Mobarakeh Branch, Mobarakeh, Isfahan, Iran.
Email: emadi.mehran49@gmail.com (Corresponding author)

Received: October 2018 Revised: November 2018 Accepted: January 2019

ABSTRACT:
The production of additional cells often forms a mass of tissue that is referred to as a tumor. Tumors can disrupt the proper functioning of the brain and even lead to the patients' death. One of the non-invasive diagnostic methods for this disease is Magnetic Resonance Imaging (MRI). The development of an automated or semi-automatic diagnostic system is required by the computer in medical treatments. Several algorithms have been used to detect a tumor, each with its own advantages and disadvantages. In the present study, an automatic method has been developed by the combination of new methods in order to find the exact area of the tumor in the MRI image. This algorithm is based on super pixel and RVM classification. The algorithm used in the super pixel method is the SLIC algorithm, which calculates for each super pixel 13 statistical characteristics and severity. Finally, an educational method introduced from the RVM classification algorithm that can detect the tumor portion from non-tumor in each brain MRI image. BRATS2012 dataset and FLAIR weights have been utilized in this study The results are compared with the results of the BRATS2012 data and The overlap coefficients of Dice, BF score, and Jaccard were 0.898, 0.697 and 0.754, respectively.

KEYWORDS: Magnetic Resonance Imaging, Super Pixel Classification, Relevance Vector Machines Classification.

1. INTRODUCTION

The brain as the central processor is one of the most important organs of the body. One of the most common diseases of the brain is the tumor. producing additional cells often form a mass of tissue that is called a tumor. The exact cause of the creation of the brain tumors is unknown. Doctors find it difficult to explain why a person has got a brain tumor and otherone hasn't [1]. One of the most important challenges facing the zoning MRI is heterogeneity in the brightness surface of image pixels. in the recommended way the superpixel method is used to deal with this problem. Simple linear iterative clustering method (SLIC) divides Images into pieces about the same size. Every piece of the picture is a Lattice Square that the size of it can be specified by the user. The initial size of the SuperPixel is considered S by the net. The geometric center of each section is considered as the superpixel center. The centers' coordinates will be updated in each iteration. The pixels are grouped according to their own space and the severity and distance. In Figure 1, an image of a brain tumor is displayed with superpixel image.

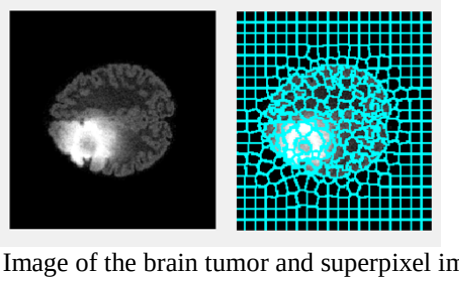

Fig. 1. Image of the brain tumor and superpixel image.

In the proposed method the RVM classification method is used to classify each superpixel into a tumor or normal tissue of the brain. In recent years the Relevance vector machine (RVM) Has attracted many scholars due to having different benefits. This method is based on the Bayes equations Related to the linear equations. Method RVM despite the simplicity of computing, is able to provide desirable results in the field of classification and regression. Among RVM applications for solving different issues, using it to solve learning issues with monitoring is often.

Ahmad lathes and colleagues Used an algorithm for diagnosing a brain tumor using the K-Means

decomposition wavelet transform and clustering to categorize the tumor.

They introduced an effective diagnostic method for brain tumor based on the wavelet-wave math morphology and K-means method. The algorithm By processing the morphology of the math and by increasing the contrast in the tumor image, it reduces the extraction process. The division and localization from suspicious areas were done using wavelet transform. Eventually, the algorithm K-Means is implemented for tumor detection [2].

Karishma Sheikh and colleagues used the algorithm K-means for feature extraction from Brain cells. Before the process K-means, Noise is removed from MRI images and noise-free images are considered as an input to k-means And the tumor will be Extracted from the MRI image. Finally, Approximate reasoning is used to calculate tumor shape and calculate the tumor position [3].

ShanShen and colleagues used the method of developed fuzzy C-Mean (FCM) [3] And Artificial Neural Network [4] (ANN) to detect a tumor. The conventional FCM method is a popular distribution method for medical images, However, this is an intensity-based clustering algorithm that is not a strong method for noisy images. Although many developed FCM-based algorithms have been developed to overcome weaknesses, none of them are flawless. Compared to RFCM and other FCM-based methods, which aim to modify the complex membership function, IFCM tries to maintain all formulas, so IFCM has a better continuance than conventional FCM algorithms and simpler computing. In addition, the basic structure of the IFCM is more simple and more intelligible [4].

Selva Kumar and colleagues classified tumors into two categories: mass and malignant. K-means algorithm will be sufficient to detect mass from brain cells. Before performing K-means algorithm, noise should be eliminated from MRI images. To extract the malignant tumor's shape and its features, we use the c-mean fuzzy method to classify. Finally, an approximate logical approach is used to calculate the shape and position of the tumor. Empirically presented methods provide more precise results compared with other methods [5].

Hossein and colleagues Used an artificial network for tumor detection. They went through the following steps:

1. Data collection: First, data were collected from a total of 100 data, 50 of which were cancerous and 50 were non-cancerous.

2. Noise Removal: Used with midrange filters that have the highest signal-to-noise ratio.

3. Enhance picture quality: Extracting features can increase image quality by increasing intensity.

4. Conversion of grayscale to binary: The binary image is a sub-class of image splitting, which divides an image into parts based on the pixel values and threshold values. In this method, the value of the semi-high intensity of the threshold in full and lower half Turns Black.

5. Delete the wrong part: sometimes some parts are incorrect and should be deleted.

6. Split the image: By forming an event matrix, the gray surface of the image is formed.

This matrix is used to extract the property [6].

Rajesh and colleagues recognized brain tumors with MRI images and accurate brain segmentation. Image partitioning is mainly done using the middle filter method, K-mean clustering, and FCM classification. This method has more accurate and more effective results than other methods. In this way, a split image gets better. The system is highly scalable and user-friendly. Almost all system objectives are met and tested according to all criteria and minimize the problems caused by the manual system [7].

Yugita Sharma et al., In this paper, instead of the algorithms presented in the original space, derived from the derivative K-means algorithm based on the alternating kernel. The results of this proposed method can be used to categorize low-contrast medical images. This method has the advantages of calculating different parameters and reducing time. The validity of the new algorithm is confirmed in the process of detail of the images. Details of the proposed scheme are as follows:

1. Enter the original image as an input.
2. Make the image grayscale.
3. Find the tilt value.
4. Marking the foreground components.
5. Estimated conversion pond [8].

Sangamitera et al. First prefetched input images that eliminate the unwanted effects of the preprocessor. Then they used the split method used to divide the tumor. For better splitting results, the K-means algorithm was used to obtain a better and more precise result from the algorithm. EK-means. Subsequently, the characteristics of suspected tumor-derived images are extracted. Features such as entropy, solidarity, homogeneity, PSNR and SSIM that can be used in image categorization. For the classification of the back-propagation network, images are used as a natural image Or tumor image according to The target is 90.87% in this case [9].

Parvin et al. Used a hybrid method. For example, a tumor-based method based on tissue-based method is presented for the diagnosis and classification of the tumor. GLCM has been used as a tissue-based method to extract the characteristic of MRI images. The LS-SVM classification has been used with the MLP kernel function to categorize tumor and non-tumor images. The proposed method is more efficient and accurately split than existing methods. This method shows an accuracy of 96.63% [10].

Shiong g and colleagues used the MSFCM algorithm to partition MRI images, including superpixel methods and FCM methods. First, the image is split into several superpixels, and then a deep split is performed for areas with a larger gray variance to set the threshold. To get fuzzy members for each superpixel, the FCM algorithm for categorization uses superpixels instead of pixels. Ultimately, MRI images of the brain with superpixels are merged into the same category. The experiment demonstrates that the proposed method of the FCM method is more efficient and provides good results in the segmentation of MRI images that contain both non-uniform and non-uniform images. This makes it possible for experts to make segmentation of MRI images of the human brain with high accuracy and effectiveness compared to the FCM method and according to similarity criteria. In addition, empirical results have shown that local exploitation of the broadcast method to categorize correctly in terms of superpixels and the function of the button to measure the similarity between superpixels is very suitable for use in medical images, including the division of the dataset from sequential medical images at one time. [11].

2. PROCEDURE

In Fig. 2, the steps are shown as a flowchart

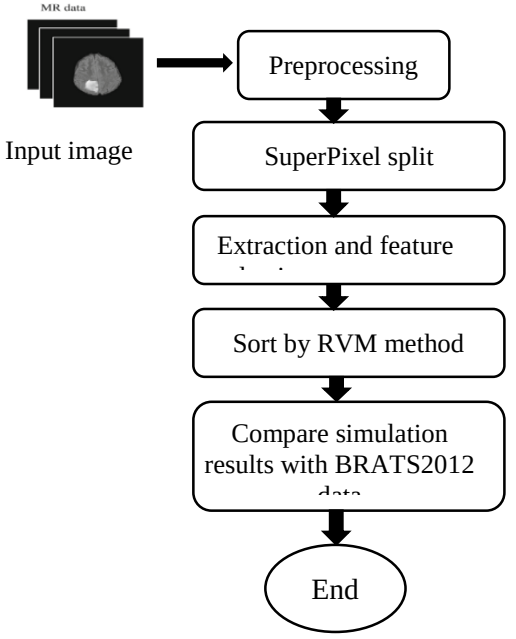

Fig. 2. Flowchart steps proposed procedure.

2.1. Input Images and Data Sets

The MRI image used in this research is the FLAIR protocol, which can have a JPG or mha extension. The images with the mha extension are from 181 images of different brain sections. In this study, the image of 110 was used for completeness. First, we convert all images taken to a size of 200 × 200 and adjust the brightness of the images to everyone. In this study, the BRATS 2012 general education curriculum focuses more on the overall assessment of the proposed method. In this section, the data is described and the division results are presented and discussed. The description of the BRATS 2012 data collection is the interpretation of the clinical educational data collection, which consists of information from several MRI centers from 35 patients with glioma (13 degrees below and 22 degrees high). For this training suite, the correctness of the foundation is provided by a trained physician. For each patient information, T1, T2, and FLAIR images are available. Data were obtained from several centers using different scanners with different field strengths (1.5T and 3T). In this study, only FLAIR images were used to evaluate the proposed method [12]. After the superpixel division, in each image, superpixels are classified into tumors and non-tumors by two methods of classification of SVM and RVM.

2.1. Preprocessing

First, the skull will be deleted from all MRI images using the FSL software. This software is a collection of analytical software designed for brain research applications. Of course, this software runs under the Linux operating system [13]. For images of the same size and size, all of the images taken are converted to a size of 200 × 200 and we set the brightness of the images for everyone.

2.3. SuperPixel Classification

The Simple Linear Iterative Clustering (SLIC) divides images into pieces with similar dimensions and sizes. Each image is a square grid whose size can be specified by the user. The primary size of the network is considered for this Super Pixel S. The geometric center of each section is considered to be the superpixel center. The coordinates of these centers are updated every time. Pixels are grouped according to their space and severity and distance. The spatial distance $\boldsymbol{d_s}$ between the ith and the jth pixel is calculated from the following equation:

$$d_s = \sqrt{(x_j - y_i)^2 + (y_j - y_i)^2} \quad (1)$$

*X and y are spati*al coordinates of pixels. The intensity of the distance $\boldsymbol{d_c}$ between two pixels is defined as:

$$d_c = \sqrt{(I_j - I_i)^2} \quad (2)$$

That is I_i and I_j the values of the intensity are normal ith and jth pixels. The overall distance measurement is a combination of spatial distances and intensity, and then D is calculated as follows:

$$D = \sqrt{{d_c}^2 + (\frac{d_s}{s})^2 m^2} \quad (3)$$

Where m is the compression coefficient that shows the flexibility of the superpixel boundaries. The high values of m representing more compact parts and lower values of m create more flexible boundaries. There are various types of features that can be considered to teach a strong classification for the diagnosis and division of the brain tumor, as SuperPixel suggests in this study. Brain tissue has complex structures, so the severity characteristics are not sufficient to divide the tumor, so complementary features are used to improve the accuracy of the division.

2.4. Relevance Vector Machine Classification (RVM)

In recent years, the interface carriers have been considered by many researchers due to different benefits. This method is based on the linear equations of the bait equation. RVM, while the simplicity of computation, can provide favorable results in terms of classification and regression. Among RVM applications for solving different issues, its use in solving supervised learning issues is of interest. Figure (3) shows a two-dimensional feature space divided into two parts.

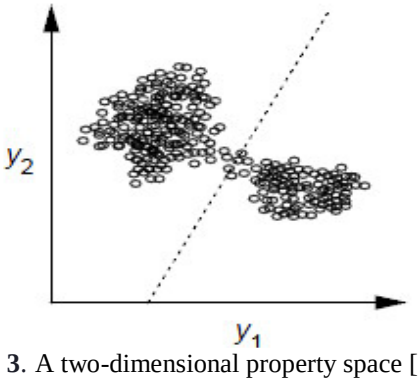

Fig. 3. A two-dimensional property space [14].

All pixels with their corresponding attributes on the left are in the same category and the rest of the pixels are in the other category. One of the most commonly used parametric classifiers is the classification of the most probable or bias. This categorizer assumes that the brightness of the pixels is an independent sample of a mixture of a series of statistical distributions, usually Gaussian. Learning data is obtained by identifying the representative of each of the components and then estimating it based on which RVM method is based on the same method.

3. IMPLEMENTING THE ALGORITHM

In superpixel division, images are allocated to roughly regular pieces with almost similar brightness values and brightness. These pieces are based on the two parameters s and m, which must first be obtained from the optimal values. For each superpixel, the characteristics, mean, standard deviation, variance, mean absolute deviation, mean absolute deviation, variance coefficient, roughness, mean and severity states, central moments, range, range between scale and entropy are calculated, after selecting the feature to find The most important feature of the RVM classification method is to divide the MRI image into a tumor and non-tumor.

3.1. Obtaining Optimal Values of s

One of the influential parameters in superpixel segmentation is the area of each superpixel (s). The size of the superpixel can make calculations faster and may provide enough information to extract features such as stable texture properties. However, a large superpixel size may include more than one class of pixels, which results in the calculation of the wrong properties (such as small areas of the clot) Blood or bleeding) and is not suitable for small lesions. While a small superpixel size is more likely to consist of just one class of pixels and is preferable to a smaller fragmentation, superpixels, however, may not have enough pixels to measure stable properties, and the time to produce small parts is too high. So, you should find an optimal value for s that can achieve the best overlap result and, in order to do this in a statistical society, obtain 20 images of the MRI of the tumor with different s values and calculate the mean of the overlapping coefficients that are in the table. (1). According to the table, if the number of superpixels is 400, the best result is obtained. The results are presented as diagrams in Fig. 4.

Table 1. The Relationship of Superpixel S Size with Overlapping Criteria.

Overlapping metrics				
Jaccard	BF Score	Dice		
0.220	0.001	0.399	200	Superpixel number
0.555	0.156	0.650	300	
0.849	0.915	0.701	400	
0.787	0.771	0.740	500	
0.787	0.764	0.764	600	
0.741	0.712	0.784	700	
0.767	0.712	0.729	800	
0.765	0.701	0721.	900	
0.738	0.967	0.712	1000	

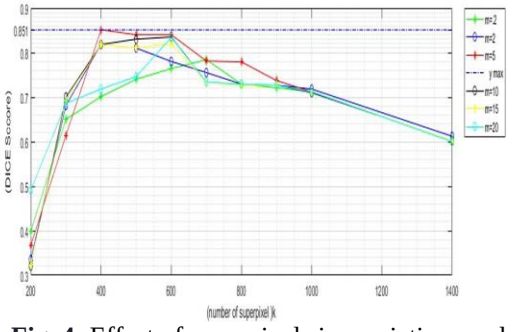

Fig. 4. Effect of superpixel size variations and compression coefficients on the overlap coefficient.

3.2. Obtaining Optimal Values of m

Another important parameter in the superpixel division step is the compression factor (m). The greater part of this parameter results in harder partitions that are more stable and usually have less noise (such as holes and pixels with a low density), however, the splitting may not follow very well the boundaries of the texture, especially when there are clear and clear boundaries Not have While lower compression values lead to more flexible and precise boundaries, splittings may produce separate and separate pixels. It may also produce very limited superpixels that are not suitable for tissue analysis. In this study, the compression coefficient, in addition to the visual examination, is achieved by changing this parameter and measuring the overlapping criteria to an optimal value of m =5 this has good flexibility for different images.

To compare this method with BRATS data, a study published in Ref. [15] can be cited due to the fact that this study is based on the double classification (ie, the active tumor core against the normal tissue of the brain) based on the FLAIR protocol, compared to The current methods published in BRATS data are difficult. It is important that the proposed method is applied to BRATS data, data coming from several centers and using different scanners.

To investigate the effect of the m-factor, defined in the equation, on superpixel boundaries, different values are applied from 0 to 20, and the results are presented in Table (2). The compression coefficient is obtained at tighter boundaries of m = 20 while at a very flexible boundary m = 0. But the changes and disruptions of the superpixel shape increase.

In order to obtain the optimum value of m, although it has reached the value of 5 in the diagram of the figure (3), however, the number of super pixels is 400 and repeats with different values of m, the simulation is repeated and the results are given in table (2). By examining the results and diagrams of the figure (5) and visual inspection of the boundaries in the superpixel area, m = 5 has been selected that contains coherent boundaries. Therefore, the compression coefficient is considered in all experiments; and for measuring the performance of superpixel division, the Dice, BF score and Jaccard coefficients have been used.

Table 2. The relationship between the size of the compression coefficient and the overlapping criteria.

Overlapping metrics				
Jaccard	BF Score	Dice		
0.556	0.609	0.635	0.2	The size of the compression coefficient m
0.612	0.729	0.714	0.8	
0.780	0.868	0.818	2	
0.849	0.915	0.845	5	
0.821	0.901	0.838	10	
0.788	0.844	822.	15	
0.749	0.726	0.818	20	

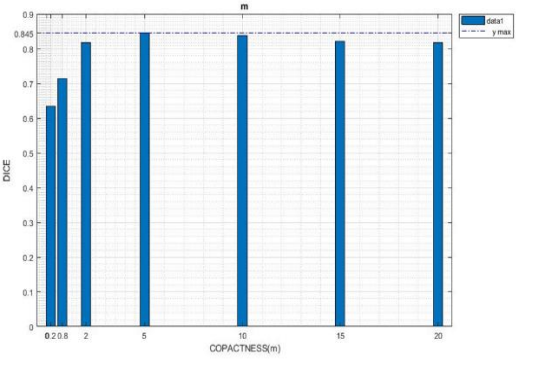

Fig. 5. Chart of the effect of compression coefficient variations on the DICE overlap factor.

So, in this research, the number of super pixels is 400 and the compression coefficient is 5. After implementing the superpixel algorithm, for each superpixel, 13 properties must be extracted, stored in a 13 × 400 matrix, each row representing the characteristics of a superpixel. These features include contrast, solidarity, homogeneity, moderate, standard deviation, entropy, variance, flatness, corrosion, roughness, RMS and IDM.

According to these 13 attributes, the image is converted to gray, and by pivotalizing the pixels of the tumor, one and non-tumor pixels are assigned a zero number, resulting in the extraction of the tumor as a black and white image. Then, the labeled pixels match the original image, and the area around the tumor area is marked with a blue border and shown in one shape.

3.3. Evaluation Parameters

In this research, three parameters have been used to evaluate the research method. According to the definitions in Table (3), the parameters can be defined as follows.

Table 3. Introduction of evaluation parameters.

Description	Parameter
The number of pixels per image is 255 (the presence of a tumor in both pictures)	a
The number of pixels that vary in their two images (wrong in detecting a tumor)	b
The number of pixels in each of their photos is 0 (no tumor in both pictures)	c
The number of pixels that have a tumor	TP(True Posetive)
Pixel numbers that have a tumor	TN(True Negative)
The number of pixels that do not have a tumor but the tumor is detected	FP(False Posetive)
The number of pixels that have a tumor but no tumor is detected	FN(False Negative)

Formulas for estimating similarity coefficients based on the parameters of Table 3 are as follows [16]:

$$Dice = \frac{2*TP}{(2*TP+FP+FN)} \quad (4)$$

$$jaccard = \frac{TP}{(TP+FP+FN)} \quad (5)$$

$$Recall = \left(\frac{TP}{(TP+FP)}\right) \quad (6)$$

$$Precision = \left(\frac{TP}{FP}\right) \quad (7)$$

$$BF = \frac{2*Recall*Precision}{(Recall+Precision)} \quad (8)$$

In order to better compare the proposed algorithm, in addition to the RVM method, the SVM classification method is also applied to the data. In Table 4, the mean of overlap coefficients for the diagnosis of a brain tumor for 35 patients with cancer is presented by SVM and RVM methods.

Table 4. Average values of overlapping parameters.

	DICE		**BFscore**		**Jaccard**	
	SVM	**RVM**	**SVM**	**RVM**	**SVM**	**RVM**
The average value	0.795	0.898	0.612	0.752	0.657	0.784

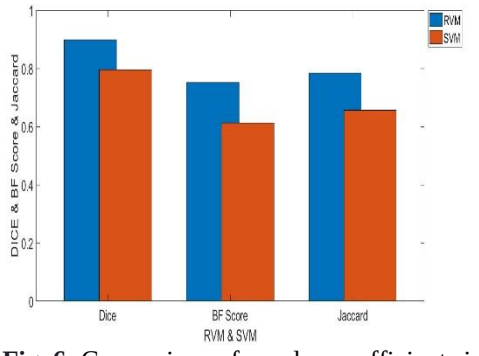

Fig. 6. Comparison of overlap coefficients in both RVM and SVM methods.

In all samples, three overlapping parameters are measured. Figures (6), (7) and (8) show the results of the comparison of these three parameters for the SVM and RVM algorithms. In Fig. 4, the results are presented as diagrams.

In Figs. 7, 8 and 9, the results of tumor detection for 35 MRI images are plotted using SVM and RVM as diagrams and the average overlapping coefficients are displayed.

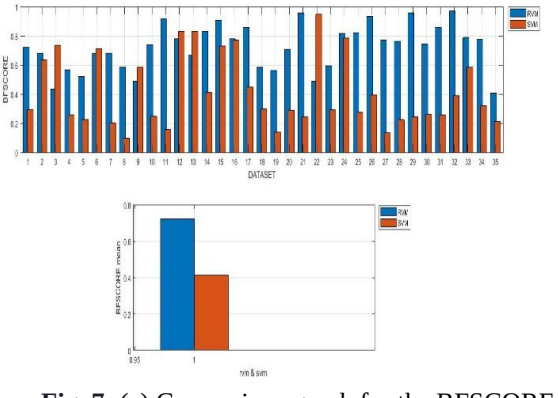

Fig. 7. (a) Comparison graph for the BFSCORE parameter for all samples (b) The mean of the BFSCORE parameter.

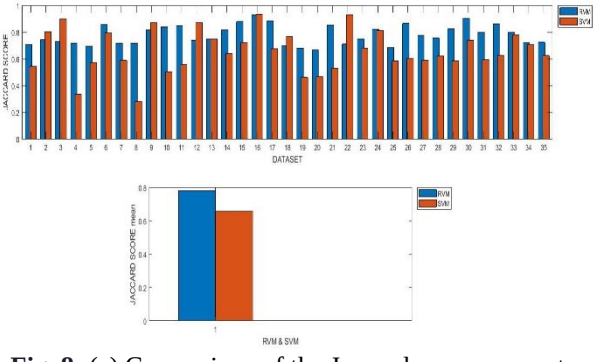

Fig. 8. (a) Comparison of the Jaccard score parameter for all samples (b) The average parameter of Jaccard score.

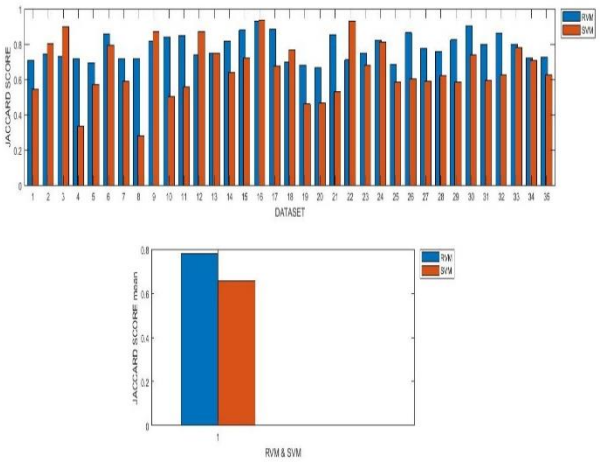

Fig. 9. (a) Charts for comparing the Jaccard score for all samples (b) The average parameter of the Jaccard score.

In Fig. 10, images of tumor extraction are shown in three manual methods, SVM and RVM.

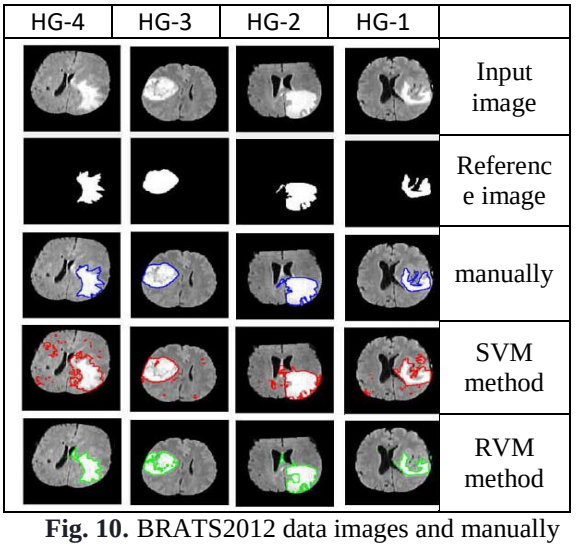

Fig. 10. BRATS2012 data images and manually subdivided images, SVM and RV.

4. CONCLUSION

In this paper, the integration of superpixel segmentation with RVM classification algorithm is used to improve the diagnosis of cancerous brain tumors. Given that the pixels of tissue in the brain do not have the same gray level in MRI images, the superpixel method is used to arrange the image so that it can perform a proper classification using the extracted properties of each superpixel. In this research, the SLIC algorithm is used for the superpixel method. There are two effective parameters in this algorithm, one of which is the superpixel size (S), or the number of superpixels of an image (K) and the compression coefficient (m). These two parameters, based on the experiment conducted in the study, were obtained for K values 400 and for m values 5. For each superpixel, 13 statistical characteristics and severity are calculated, which are used in the RVM classification algorithm to classify superpixels into tumors and non-tomorrows. For a better evaluation, the RVM algorithm was implemented for BRATS2012 data. According to the results, the overlapping parameters of Dice, BFsccore and Jaccard in the RVM algorithm were 0.8898, 0.752, and 0.7484, respectively, and 795 for the SVM algorithm respectively 0, 0, 612, and 657/0.

REFERENCES

[1] V. L. and S. P., **"Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations,"** IEEE Transactions on Pattern Analysis & Machine Intelligence, pp. 583-598, 1991.

[2] A. kharrat, **"Detection of Brain Tumor in Medical Images,"** in 2009 3rd International Conference on Signals, Circuits and Systems (SCS), 2009.

[3] S. K., S. V. and T. S., **"Clustering based Segmentation Approach to Detect Brain Tumor from MRI Scan,"** International Journal of Computer Applications, p. 118, 2015.

[4] S., Shen, **"MRI Fuzzy Segmentation of Brain Tissue Using Neighborhood Attraction with Neural-Network Optimization,"** IEEE transactions on information technology in biomedicine, pp. 459-467, 2005.9(3).

[5] J. Selvakumar and A. Lakshmi, **"Brain Tumor Segmentation and its Area Calculation in Brain MR Images using K-mean Clustering and Fuzzy C-mean Algorithm,"** in IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012), 2012.

[6] H. M.A., **"Lung Cancer Detection Using Artificial Neural Network & Fuzzy Clustering,"** International Journal of Advanced Research in Computer and Communication Engineering, 2015.4.(3).

[7] V. Rajesh, **"Brain Tumor Segmentation and its Area Calculation in Brain MR Images using K-Mean Clustering and Fuzzy C-Mean Algorithm.,"** Fuzzy Systems, pp. 103-107, 2015.

[8] Y. Sharma and p. Kaur, **"Detection and Extraction of Brain Tumor from MRI Images using k-Means Clustering and Watershed Algorithms,"** International Journal of Computer Science Trends and Technology, pp. 8-32, 2015. 3(5).

[9] P. Sangamithraa and S. Govindaraju, **"Lung Tumour Detection and Classification using EK-Mean Clustering,"** in International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET). 2016., 2016.

[10] G. Praveen and A. Agrawal**, "Hybrid Approach for Brain Tumor Detection and Classification in Magnetic Resonance Images,"** in 2015

Communication, Control and Intelligent Systems (CCIS), 2015.

[11] S. Ji, **"A New Multistage Medical Segmentation Method Based On Superpixel And Fuzzy Clustering,"** in Computational and mathematical methods in medicine 2014, 2014.

[12] BRATS, "**The Virtual Skeleton Database Project**," [Online]. Available: https://www.smir.ch/BRATS/Start2012. [Accessed 2 1 2018].

[13] M. Soltaninejad and G. Yang**, "Automated Brain Tumour Detection and Segmentation using Superpixel-Based Extremely Randomized Trees in FLAIR MRI,"** International Journal of Computer Assisted Radiology and Surgery, vol. 12, no. 2, pp. 183-203.

[14] C. DL Pham and. P. J.L. , **"A Survey of Current Methods in Medical Image Segmentation,"** *Annual Review of Biomedical Engineering*, vol. 2, 2000.

[15] M. e. a. B. H., **"The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS),"** *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993-2024, 2015.

[16] A. Islam, S. M. S. Reza and K. M. Iftekharu, **"Multifractal Texture Estimation for Detection and Segmentation of Brain Tumors,"** *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, vol. 60, no. 11, 2013.