A Comparative Study on Tissue Classification of Brain MR Images Using DIPY, SPM, and FSL Frameworks

Iman Azinkhah[1,2*], Mahdi Sadeghi[1,2]

1. Fintech in Medicine Research Center, Iran University of Medical Sciences, Tehran, Iran
2. Medical physics department, Faculty of medicine, Iran university of medical sciences, Tehran, Iran

ARTICLE INFO	ABSTRACT
Article type: Original Paper	**Introduction:** Image classification is a disputable part of image processing, especially brain tissue classification in Magnetic Resonance images because brain tissue signals and contrasts are close to each other. More accurate classification of MR images of the brain is crucial for the diagnosis of CSF, gray, and white matter.
	Material and Methods: A Bayesian model was used to classify 20 brain MR images using the criteria of DIPY, SPM, and FSL. Two distinct DICE and Jaccard coefficients were used to assess the similarity of all classified images.
	Results: SPM classification was more accurate than DIPY and FSL in categorizing cerebrospinal fluid. The DICE and JACCARD coefficients for the SPM classification were 97.48 ± 0.28 and 92.68 ± 0.94, respectively. The DICE and Jaccard coefficients for white matter were 95.64 ± 0.23 and 86.18 ± 1.64, respectively, while the coefficients for gray matter were 93.66 ± 0.76 and 83.62 ± 1.92 for the DIPY. **Conclusion:** The DIPY python library was able to better cluster GM and WM regions according to the results obtained.

Introduction

The most common brain tumor is reported to be glioblastoma in men, and the most non-typical brain tumor is meningioma in women. [1]. The differentiation of these brain tumors is one of the most important challenges in determining the fate of patients regarding treatment. Although the incidence of malignant brain tumors such as glioblastoma has decreased by approximately 0.8% per year, there is still an increase in its prevalence among young people [2]. The need to increase the accuracy of diagnosis in the shortest possible time is increasing, despite the significant advances in diagnostic methods for brain tumors by CT scan, PET scan, MRI, and nuclear medicine. To meet this need, the use of image-processing systems is expanding day by day, and the forefront of these image-processing systems is tissue classification techniques. On the other hand, the close contrast of brain tissues in MR images creates a more serious challenge in brain image classification. This classification mainly includes gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) in the brain. One of the most important challenges to performing optimal image processing is to remove the effects of noise and blurring at the border of the images [3]. Different image classification methods have been introduced, including atlas-based, surface-based, intensity-based, and hybrid methods. [4] The

Intensity-based method consists of thresholding, clustering, region-growing, and classification, which, using known data, allow the image to be segmented into several specific features, of which the most important is usually the intensity. The nearest neighbor classifier [5] is one of the simplest classifiers in this field and was first used by Warfield et al. [6] in brain image classification. Still, this procedure is not entirely automatic and requires training data due to the supervised nature of this classifier. In this regard, a fully automatic procedure was proposed by Cocosco et al. [7] based on sample selection with a strong algorithm. However, the most common method is the Bayesian classifier [8], in which unknown variables can be estimated automatically by creating a relationship between a set of features and variable classes. The Bayesian approach was used in different software, including Statisitical Parametric Mapping (SPM) [9], FMRIB Software Library (FSL) [10], and Diffusion Imaging in Python (DIPY) [11].

Regarding brain MR images, there are various challenges such as close contrast of brain tissues and motion artifacts, so many techniques can have different effects in image classification, In this study, the capability and accuracy of brain MR image classification using these frameworks were explored.

Corresponding Author: Tel: +989125367209; Email: imanazinkhah@gmail.com

Materials and Methods

Patients

Twenty MR images of real patients (all the patients without any history or problems) were selected. The Ethics Committee of the Iran University of Medical Sciences (IUMS) approved this study under ethics approval IR.IUMS.FMD.REC.1400.339, with the patient's written informed consent. The priority of brain T1-Weighted images was T1 MPRAGE protocol, FOV was 192 x 256, ISO images with 1x1x1 mm, TR=550ms, and TE=20ms. Figure 1 shows a sample of brain MR images with the mentioned parameters.

Preprocessing

Removal of non-brain tissues: The goal was to achieve brain tissues, which have the least interaction with the signals of non-brain tissues. Because signals from fat tissues and skull bones can affect the signals from brain tissues, it was necessary to remove these tissues. Here we are faced with only two types of cells: brain tissue and non-brain tissue. In this study, the FSL (the brain extraction tool) was used. The technique used in this software has been reported as favorable for adult images[12]. In this procedure, the gravity center of the brain is found and then it is moved to all the cells above it to find the boundary edge of the brain tissue.

Bias field correction: One of the major problems of MR images is intensity heterogeneity, which must be solved before any image processing in a bias field correction procedure. The heterogeneity was more vital because the used device was a 1.5 Tesla. The main cause of this inhomogeneity is the change in the sensitivity of the receiving coils and the interference between the magnetic field and the patient's body [13, 14]. This problem can cause brain white matter to be measured in another area of brain gray matter with the same intensity, thus impairing the classification of the brain because it is assumed that the intensity of a particular brain tissue is approximately uniform throughout the image. In this study, the bias field correction technique of the BrainSuit program was used [15]. In this program, it was assumed all other brain tissues were removed except Cerebrospinal fluid (CSF), white matter (WM), and Grey Matter (GM), and then a series of average values for the desired tissues, i.e., CSF, WM, and GM, were calculated. The average values were used for a partial volume measurement model that includes the effect of local heterogeneity calculations. This model was matched with the histogram of a sub-volume of the image, and thus the increase in the sub-volume relative to the intensity of the whole image was estimated. The estimated values were used to create a three-cubic B-Spline, which is an estimate of the same non-uniformity. By dividing the final image by this spline, the corrected bias field was achieved [16].

Bayesian classification

The Bayesian classifier, which is a common parametric classifier, is one of the most commonly used in probability theory and statistics. It is named after Thomas Bayings and is believed to provide information about the probability of

an event based on its possible conditions. [17] The class probabilities of unknown variables are estimated by modeling their relationship with feature sets using a Bayesian classifier. This model employs maximum posterior (MAP) estimation and Bayesian inference to approximate the output image x with the observed image, by decreasing or eliminating the posterior distribution P (xy) of possible labels X:

$$\hat{x} = \arg max P(x|y) \qquad (1)$$

In the Bayesian framework, there are three probability distributions:

(1) The prior distribution $P(\vec{x})$ consist of knowing the possible configurations before viewing the actual image.

(2) The posterior distribution of $P(x|y)$ is derived after an observation is made.

(3) The probability $P(y|x)$ is characterized by the probability of receiving an observation with respect to a set of model parameters. But Bayes' law says:

$$P(x|y) = \frac{P(y|x)\,P(x)}{P(y)} \qquad (2)$$

By attention to equation 1:

$$\hat{x} = \arg max \left(\frac{P(y|x)\,P(x)}{P(y)} \right) \qquad (3)$$

Also, considering that the probability of $P(y)$ is quite clear, then we can finally say that:

$$\hat{x} = (P(y|x)P(x)) \qquad (4)$$

In order to classify the new data, this rule ensures that every pixel is assigned to the class with the highest posterior probability. Therefore, it is need to define two other parameters according to formula 3, P(x) prior distribution and $P(y|x)$ probability model or observation model. In this study, a Gaussian distribution was used for the probability model and Markov Random Field (MRF) [18] was used for the prior distribution. The use of spatial texture information and neighboring voxels in brain MR image classification has been of great importance. The gray intensity of the surrounding voxels is responsible for the intensity associated with a particular voxel. According to this point of view, Markov random field theory serves as a model for this concept that local features of an image, in which the global features of the image follow the local interactions. MRF techniques have been highly successful in classifying brain MR images, especially in the matter of noise factor [9, 10, 19].

Classification of brain T1-Weighted MR images

DIPY: The DIPY python library was used to classify brain tissues in T1-Weighted MR images into CSF, WM, and GM. This library will implement all three main features of this study, i.e. using Bayesian theory, MRF model for prior distribution, and Gaussian distribution for a probability model. To run DIPY, we will need two other libraries, Matplotlib [20], and numpy [21]. As mentioned earlier, BIAS field correction and removal of non-brain tissues were performed on all MR images before the implementation of the classification code.
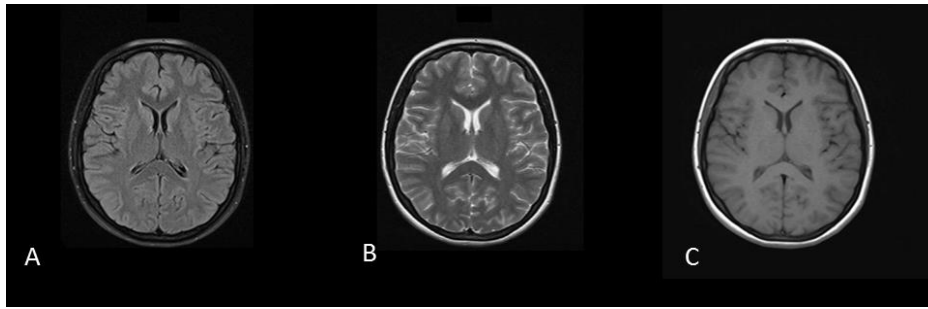
Figure 1. A sample of brain MR images taken from patients with parameters: TR=550ms, TE=20ms, FOV=192x156, and voxel size= 1x1x1. (A) FLAIR, (B) T2 Weighted, and (C) T1-weighted image*s*
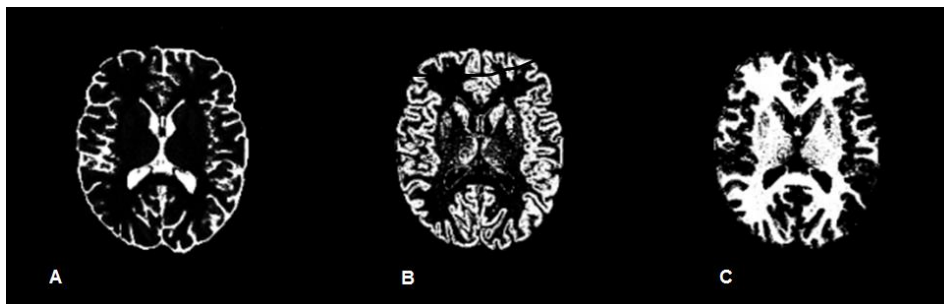

Figure 2. Brain MR images are classified into (A) cerebrospinal fluid, (B) gray matter, and (C) white matter by the DIPY Python package

All brain T1-Weigthed images of patients were classified and probability maps of each determined tissue were obtained. Figure 2 demonstrates the brain T1-Weighted MR images classified into CSF, WM, and GM by the DIPY python library.

SPM: The SPM (Statistical Parametric Mapping) program is dedicated software for functional and statistical analysis of neuroimaging data. In this study, SPM8 under the MATLAB program was used. A single model in all classifications even in pre-processing used in SPM. The estimation of the prior probability of each tissue was done using the Bayesian method. Voxels with more tissue probability were selected as the members of that tissue class. SPM8 routinely classified brain MR images into three groups including CSF, gray, and white matter.

FSL: The FSL (FMRIB's Software Library) program was developed by the Oxford University Department of Functional Brain Studies for the functional and statistical analysis of brain data. The FSL program has several toolboxes; in this study, the special toolbox for brain classification, namely FAST was used. The most important features of FAST-FSL are the correction of non-uniformities in intensities caused by the magnetic field in the MR image and the use of a Hidden Markov Random Field (HMRF) model. These two processes repeat to reach a desired convergence and finally, brain MR images are classified into CSF, GM, and WM.

Evaluation Methods

It is necessary to have a specific standard to evaluate the success rate of a brain MR image classification technique so that one can be assured of the success rate

of a method. One of the most important helping quantities in this field is the investigation of the amount of spatial overlap of images [22]. One of the simplest indicators is the DICE similarity coefficient [23], which is also known as the specific agreement ratio [24]. The DICE similarity coefficient compares the amount of spatial overlap between two segments or two binary images. This quantity is calculated according to equation 5:

$$Dice\ Coefficient = \frac{2|img_{segment} \cap img_{ref}|}{|img_{segment}| + |img_{ref}|} \quad (5)$$

The value of a DICE ranges from 0 to 1, 0 indicating no spatial overlap between two sets of binary classification results, and 1 indicating complete overlap. The second quantity measured to investigate the amount of spatial overlap is the Jaccard coefficient [25]. According to formula 6, the range of Jaccrad quantity is variable between 0 (no overlap) and 1 (complete overlap).

$$Jaccard\ Coefficient = \frac{|img_{segment} \cap img_{ref}|}{|img_{segment} \cup img_{ref}|} \quad (6)$$

Here, a comparison of the spatial overlap between the results obtained by DIPY, SPM, and FAST-FSL and the results of manual classification has been made. In this way, the labeled images (labeled manually) of each tissue were used as a reference, and the results of the classification approach were converted into binary images with voxel resolution and image dimensions similar to the reference image.

Results

The comparison of DICE and Jaccard coefficients in the brain classification T1-Weigthed images by DIPY, SPM, and FAST-FSL is shown in Table 1. CSF classification by SPM was higher and better than other classification software packages in both coefficients (DICE = 97.48 ± 0.28, Jaccard = 92.68 ± 0.94). But DIPY has classified GM and WM more thoroughly than SPM and FAST-FSL. So, the DICE coefficient of WM and GM region for DIPY was 95.64 ± 0.23 and 93.66 ± 0.76.

By looking at the diagram in Figure 3, the bar graph of DICE and Jaccard coefficients for the classification of the three regions of CSF, WM, and GM can be better compared. In all measurements, the DICE coefficient was greater than the Jaccard coefficient.

Table 1. DICE and Jaccard coefficients in different classification software packages for CSF, WM, and GM

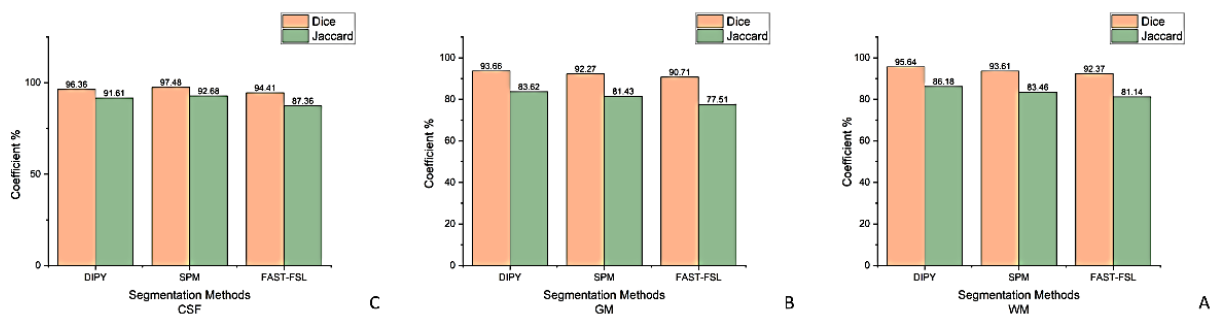| software packages	Evaluation Method	CSF (Mean ± SD %)	WM (Mean ± SD %)	GM (Mean ± SD %)
DIPY (Bayesian)	DICE	96.36 ± 0.97	95.64 ± 0.23	93.66 ± 0.76
	Jaccard	91.61 ± 1.25	86.18 ± 1.64	83.62 ± 1.92
SPM	DICE	97.48 ± 0.28	93.61 ± 0.81	92.27 ± 0.26
	Jaccard	92.68 ± 0.94	83.46 ± 1.91	81.43 ± 1.62
FAST-FSL	DICE	94.41 ± 0.83	92.37 ± 0.63	90.71 ± 0.48
	Jaccard	87.36 ± 1.11	81.14 ± 1.42	77.51 ± 1.93

Figure 3. Bar chart of DICE and Jaccard coefficients in different classification software packages for CSF, WM, and GM

Discussion

The classification of brain tissues is very helpful for the studies of Alzheimer's disease [26], epilepsy [27], and diseases related to the cortical brain tissues. Because manual methods are not accurate, various techniques have been proposed for the classification of patients' brain MR images. However, providing these techniques is still considered a big challenge. Brain MR image classification methods can be considered a subset of classification methods based on image intensity [4].

The purpose of this method is to use data with similar labels for image classification. These labels can be for intensity or other features of the image. The classification methods themselves can be supervised or unsupervised. Supervised methods require trained images, which means they must be done manually, which is out of the scope of this paper. One of the supervised methods is the Bayesian parametric classifier. The class probability of known variables can be approximated by a Bayesian estimator, which then models the possible relationships between underlying features and class variable names. This study aimed to classify brain images based on Bayesian techniques; the differences between the classifications performed by other software packages used in this study were also desirable. In some areas of the brain, it seems that the

classification of CSF has been done better than that of GM and WM, and it seems that there is a need to conduct more studies in this field. One of the challenging points of classification is that usually there is no spatial modeling; for this reason, by choosing an approach like MRF in this study, neighborhood, and geometric information was also included in a classified approach. Among the classification methods that are commonly used in medical imaging are cluster algorithms, one of them is the C-Means fuzzy method, which was used to compare this technique. The most important feature of cluster methods is self-training using available data, which is considered an unsupervised method [28]. One of the most important drawbacks of this method seems to be the unfavorable performance of this technique in noisy images with non-uniformity.

In general, the SPM program has different practical goals compared to the other two programs which emphasize the field of examining the performance of brain data. The FSL and DIPY follow a more general goal in the field of brain image classification. However, it seems that the SPM8 package can better classify than CSF according to the use of an optimized registration model, the use of several data spectra, and the further development of tissue probability estimation maps. But

DIPY, in addition to being easier to run in the Python environment, has been more successful in classifying white matter and gray matter. As mentioned earlier, DIPY uses an MRF to model the background probability of brain tissue, which of course is also used in SPM8. Therefore, it seems that the most important feature of DIPY compared to other classification programs in this study was to find more optimal solutions using maximum and repeated conditional commands.

Finally, it should be said that the MRF technique can work well for the classification of brain MR images that are accompanied by intensity heterogeneity, but one of its limitations is the need for the appropriate selection of parameters controlling the strength of spatial interactions so that improper selection can lead to overly smooth division and as a result, lose important structural details. [29] Also, it should be said that similar to the result of the Zamanpour et al. [30] study, we can choose the appropriate classification software packages according to the purpose of image analysis.

Conclusion

Classification of MR images is a very active field in the medical field. It can be said that a single method does not answer all needs in the field of classification of MR images because each of them has strengths in one area and weaknesses in another. Like the technique used in this study, SPM uses a Bayesian classifier, but the CSF classification is better according to the DICE and Jaccard coefficients obtained by SPM. However, the classification of the other two regions GM and WM completely by the method used by DIPY in this study has been more suitable than the other two software packages.

Acknowledgment

References

1. Miller KD, Ostrom QT, Kruchko C, Patil N, Tihan T, Cioffi G, et al. Brain and other central nervous system tumor statistics, 2021. CA: a cancer journal for clinicians. 2021 Sep;71(5):381-406. doi:10.3322/caac.21693

2. Lin D, Wang M, Chen Y, Gong J, Chen L, Shi X, et al. Trends in intracranial glioma incidence and mortality in the United States, 1975-2018. Frontiers in oncology. 2021 Nov 1;11:748061. doi:10.3389/fonc.2021.748061

3. Razavi SE, Khodadadi H. Segmentation of Magnetic Resonance Brain Imaging Based on Graph Theory. Iranian Journal of Medical Physics. 2020 Jan 1;17(1):48-57. doi:10.22038/ijmp.2019.35600.1447

4. Despotović I, Goossens B, Philips W. MRI segmentation of the human brain: challenges, methods, and applications. Computational and mathematical methods in medicine. 2015;2015. doi:10.1155/2015/450341

5. Duda RO. PE hart and DG Stork, pattern classification. Address: Wiley-Interscience. 2001.

6. Warfield SK, Kaus M, Jolesz FA, Kikinis R. Adaptive, template moderated, spatially varying statistical classification. Medical image analysis. 2000 Mar 1;4(1):43-55. doi:10.1016/S1361-8415(00)00003-7

7. Cocosco CA, Zijdenbos AP, Evans AC. A fully automatic and robust brain MRI tissue classification method. Medical image analysis. 2003 Dec 1;7(4):513-27. doi:10.1016/S1361-8415(03)00037-9

8. Wells WM, Grimson WE, Kikinis R, Jolesz FA. Adaptive segmentation of MRI data. IEEE transactions on medical imaging. 1996 Aug;15(4):429-42. doi:10.1109/42.511747

9. Ashburner J, Friston KJ. Unified segmentation. neuroimage. 2005 Jul 1;26(3):839-51. doi:10.1016/j.neuroimage.2005.02.018

10. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE transactions on medical imaging. 2001 Jan;20(1):45-57. doi:10.1109/42.906424

11. Garyfallidis E, Brett M, Amirbekian B, Rokem A, Van Der Walt S, Descoteaux M, et al. Dipy, a library for the analysis of diffusion MRI data. Frontiers in neuroinformatics. 2014 Feb 21;8:8. doi:10.3389/fninf.2014.00008

12. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp. 2002;17(3):143-155. doi:10.1002/hbm.10062

13. Collins CM, Liu W, Schreiber W, Yang QX, Smith MB. Central brightening due to constructive interference with, without, and despite dielectric resonance. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2005 Feb;21(2):192-6. doi:10.1002/jmri.20245

14. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE transactions on medical imaging. 1998 Feb;17(1):87-97. doi:10.1109/42.668698

15. Shattuck DW, Leahy RM. BrainSuite: an automated cortical surface identification tool. Medical image analysis. 2002 Jun 1;6(2):129-42. doi:10.1016/S1361-8415(02)00054-3

16. Shattuck DW, Sandor-Leahy SR, Schaper KA, Rottenberg DA, Leahy RM. Magnetic resonance image tissue classification using a partial volume model. NeuroImage. 2001 May 1;13(5):856-76. doi:10.1006/nimg.2000.0730

17. Theodoridis S. Machine learning. Chapter 7—Classification: a tour of the classics. doi:10.1016/b978-0-12-818803-3.00016-7

18. Li SZ. Markov random field modeling in image analysis. Springer Science & Business Media; 2009 Apr 3. doi:10.1007/978-4-431-66933-3

19. Pham DL, Xu C, Prince JL. Current methods in medical image segmentation. Annual review of biomedical engineering. 2000 Aug;2(1):315-37. doi:10.1146/annurev.bioeng.2.1.315

20. Hunter JD. Matplotlib: A 2D graphics environment. Comput Sci Eng. 2007;9(3):90-95. doi:10.1109/MCSE.2007.55

21. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature. 2020;585(7825):357-362. doi:10.1038/s41586-020-2649-2

22. Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ, Wells III WM, Jolesz FA, Kikinis R. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. Academic radiology. 2004 Feb 1;11(2):178-89. doi:10.1016/S1076-6332(03)00671-8

23. Yao AD, Cheng DL, Pan I, Kitamura F. Deep learning in neuroradiology: a systematic review of current algorithms and approaches for the new wave of imaging technology. Radiology: Artificial Intelligence. 2020 Mar 4;2(2):e190026. doi:10.1148/ryai.2020190026

24. Fleiss JL, Levin B, Paik MC (2003) The measurement of interrater agreement. In: Shewart WA, Wilks SS (eds). Statistical methods for rates and proportions 2003. John Wiley & Sons, p 598–626. doi:10.1002/0471445428.ch18

25. Crum WR, Camara O, Hill DL. Generalized overlap measures for evaluation and validation in medical image analysis. IEEE transactions on medical imaging. 2006 Oct 30;25(11):1451-61. doi:10.1109/TMI.2006.880587

26. DeCarli C, Maisog J, Murphy DG, Teichberg D, Rapoport SI, Horwitz B. Method for quantification of brain, ventricular, and subarachnoid CSF volumes from MR images. Journal of computer assisted tomography. 1992 Mar 1;16(2):274-84. doi:10.1097/00004728-199203000-00018

27. Gilmore RL, Childress MD, Leonard C, Quisling R, Roper S, Eisenschenk S, et al. Hippocampal volumetrics differentiate patients with temporal lobe epilepsy and extratemporal lobe epilepsy. Archives of neurology. 1995 Aug 1;52(8):819-24. doi:10.1001/archneur.1995.00540320103017

28. Wadhwa A, Bhardwaj A, Verma VS. A review on brain tumor segmentation of MRI images. Magnetic resonance imaging. 2019 Sep 1;61:247-59. doi:10.1016/j.mri.2019.05.043

29. Dubey RB, Hanmandlu M, Gupta SK, Gupta SK. The brain MR image segmentation techniques and use of diagnostic packages. Academic Radiology. 2010 May 1;17(5):658-71. doi:10.1016/j.acra.2009.12.017

30. Zamanpour SA, Ganji Z, Bigham B, Zemorshidi F, Zare H. Evaluation and Comparison of Automatic Brain Segmentation Methods Based On the Gold Standard Method. Iranian Journal of Medical Physics/Majallah-I Fīzīk-I Pizishkī-i Īrān. 2023 Jul 1;20(4). doi:10.22038/ijmp.2022.66025.2134.