

# Comparison of Supervised and Unsupervised Learning Algorithms for Brain Tumor Detection

Rahul Godhani<sup>1</sup>, Gunjan Gurbani<sup>1</sup>, Tushar Juman<sup>1</sup>, Bhavika Mahadik<sup>1</sup>, Vidya Zope<sup>2</sup>

B.E. , Dept. of Computer Engineering, VES Institute of Technology, Mumbai, India<sup>1</sup>

Asst. Professor, Dept. of Computer Engineering, VES Institute of Technology, Mumbai, India<sup>2</sup>

**ABSTRACT:** MRI images are used for detecting the brain tumor. Segmentation of brain images helps to identify the size, shape, and location of the tumor in the brain. In this paper, we are comparing different segmentation algorithms on the basis of the simulation result. Segmentation methods like K-means, Random Forest, Fuzzy C Logic, and Convolution Neural Network are applied to the MRI of the brain tumor. The study involves a comparison of the accuracy of various algorithms on the BraTS dataset of 290 patients. Due to a large number of competing algorithms and techniques available for detection and diagnosis of the tumor, our comparison study mainly focuses on supervised and unsupervised learning machine learning algorithms. The use of different patterns of learning gives us more precise and detailed results as compared to conventional methods of segmentation. There are mainly four different steps involved in brain tumor detection: Pre-processing, Segmentation, Feature Extraction, and Optimization. Apart from summarizing the different segmentation algorithms, this paper also comprises of evaluation of the surveyed literature which gives detail description of the technicality used in the biomedical field.

**KEYWORDS:** MRI (Magnetic Resonance Image), Brain tumor Detection, Pre-processing, Segmentation, Feature extraction.

## I. INTRODUCTION

Cancer in a human body is the result of abnormal and uncontrolled division of the cells. A brain tumor is an intracranial tumor formed when these cells occur in the form of a mass in the brain. The tumors can be classified into two types depending on the effects that they produce on the human body, they are benign and malignant. Benign tumors are non-cancerous while the malignant are proven to be cancerous. The origin of the brain tumors can be primary origin (formed and resides within the brain) or metastatic (formed within the brain and spreads to some other parts of the body). Primary benign tumors are generally slow growing and may not cause the symptoms for a long time. Also, some tumors rarely recur after the treatment. While metastatic malignant tumors grow very fast and are likely to recur even after the treatment. Brain tumor detection requires medical imaging techniques such as MRI (Magnetic Resonance Imaging), CT (Computed Tomography) scans, X-rays or PET (Positron Emission Tomography) to form the picture of brain anatomy. Nowadays, MRI is the most widely used procedure and is always preferred by neurosurgeons as they provide fine details of the smallest abnormalities in the organs and tissues. MRI technique is non-invasive which uses radio frequency pulses and a strong magnetic field that helps in forming the internal images of the brain structure with accurate details and evaluate the type, location, shape and the metabolism of the tumor. Different MRI modalities used in tumor diagnosis are: T1-weighted MRI (T1), T2-weighted MRI (T2), and Fluid Attenuated Inversion Recovery (FLAIR). In T1-weighted MRI, compartments filled with fluid (CSF) appear dark while tissues with high-fat content appear bright while in T2-weighted MRI (T2), compartments filled with fluid (CSF) appear bright while tissues with high-fat content appear dark. FLAIR is used to separate the regions filled with CSF (Cerebrospinal Fluid). The detection of a brain tumor is a comprehensive process and requires accuracy to detect the exact location of the tumor

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 1, January 2019

without affecting the healthy tissues. Thus, segmentation plays a very crucial role in determining the exact location of the tumor inside the brain. Since manual segmentation is a time-consuming and difficult process, automated segmentation methods are the current focusing trend in the areas of research and are widely used. Our paper represents a survey on the different automated methods of segmentation which are known to give the accurate results in biomedical image segmentation. These methods are based on supervised and unsupervised ways of learning. Since, the complexity of the model is high, use of different patterns of learning and applying the intelligent algorithms based on the same would be helpful in obtaining the accurate results in the segmentation process. A brief description of these methods is summarized in our paper.

## II. LITERATURE SURVEY

### A) Detection of Brain Tumor Using K-Means Clustering

Here system acquires the images and pre-processes it under a median filter. K-means clustering is used for brain tumor detection. Here dataset consists of MRI images of size 181X272. The dataset is divided into two types-testing datasets and training dataset. Thresholding is applied on images for feature extraction and at last to recognize the tumor shape and position in MR images a proper reasoning method is used.

### B) Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images

This paper proposes an automatic segmentation method based on Convolutional Neural Networks (CNN), exploring small 3\*3 kernels. The proposed method was validated on the BRATS 2013. For every patient in BRATS, there are four MRI sequences available: T1-weighted (T1), T1 with gadolinium-enhancing contrast (T1c), T2-weighted (T2) and FLAIR. CNN resulted in segmentation with a better delineation of the complete tumor as well as of the intratumoral structures. Using cascaded layers with small 3\*3 kernels has the advantage of maintaining the same effective receptive field of bigger kernels while reducing the number of weights, and allowing more non-linear transformations on the data.

### C) Brain Segmentation using Fuzzy C means clustering to detect tumor Region

This paper proposes an approach for segmentation of brain tumor in MRI images using fuzzy c-mean clustering. If there is any noise present in the MR image it is removed before the fuzzy C-means process is applied. In the FCM algorithm, without labels assign pixels to fuzzy clusters. With varying degree of membership in FCM, pixels belong to multiple clusters. In terms of accuracy, the several iterations are considered. Fuzzy C means need to do a full inverse distance weighting. Performance is unlimited as FCM can be used in the variety of clusters and can handle uncertainty. It gives better results in cases where data is incomplete or uncertain. It gives a better result for overlapped dataset but here we need a prior specification of the number of clusters.

### D) Brain Tumor Segmentation and Classification using Random Forests Algorithm

In this paper, brain tumor segmentation and classification are performed into three parts, Complete Tumor, Tumor Core and Enhancing Tumor. Random Decision Forest (RDF) algorithm is used. For segmentation BRATS Dataset is used which consists of T1, T1c, T2 and Flair MRI images. In this algorithm, binary decision trees are build depending on two processes. In the first process, the bootstrap set is obtained by randomly sampling the training set. And in the next process, randomization is introduced. Training sets are very expensive as it needs to cover all types of detectable objects. Here the database used limits the segmentation performance.

### E) State of the art survey on MRI brain tumor segmentation

This paper summarizes about the different techniques used in image segmentation according to their degree of human intervention in the tumor detection procedure. Manual segmentation method requires manually drawing the boundaries of the tumor regions. Semi-automatic method requires a human operator to verify the accurateness of the obtained result or to check if the segmented region is up to the mark for further processing or not. Automatic segmentation requires least human interaction in the system. Soft computing knowledge techniques that require human intelligence are generally incorporated in the detection system. Apart from summarizing the methods of segmentation

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 1, January 2019

this paper also describes the supervised and unsupervised ways of learning and currently used trends and technologies used in image segmentation. The current trends involve: Threshold-based, Region-based, Pixel classification and Model-based techniques. The first three are used in two-dimensional processing while the fourth one is mainly deployed in three-dimensional imaging.

### III.ALGORITHMS

**Table -1:** Comparison of Supervised and Unsupervised Learning Algorithms

| Supervised Learning Algorithm                           | Unsupervised Learning Algorithm                               |
|---|---|
| Input data in supervised learning is known and labelled | Input data in unsupervised learning is unknown and unlabelled |
| Very complex computation is uses                        | Less computational complexity is uses                         |
| Supervised technique gives reliable and accurate result | Unsupervised technique gives moderate accurate result         |
| Use offline analysis of data                            | Use real - time analysis of data                              |
| Number of classes are within the scope of knowledge     | Number of classes are little known                            |
| Used for forecast                                       | Used for examination  |
| Deploy training dataset                                 | Deploy just input dataset                                     |

The above table represents the comparison between Supervised and Unsupervised Learning Algorithms. These comparisons are based on computational factors, type of data, analysis and accuracy of the algorithms present in each type

#### Supervised and Unsupervised Learning Algorithm

1. K-means
2. Random Forest
3. Fuzzy C Logic
4. Convolution Neural Network

#### [A] K-means

Clustering is defined as the process in which similar objects are grouped together. Clustering is widely used in unsupervised learning. In K-means clustering, the objects are grouped together into k clusters where k is a positive integer. These k clusters are based on the similarity between the objects here in our case it is image pixels. K means groups the image pixels based on some features and similarity between the pixels. In this segmentation algorithm, the number of k clusters and their centroid points is computed randomly or by using some heuristic data. The straight line distance between each cluster centroid and the pixel is calculated. The pixel with the minimum distance is moved into the corresponding cluster. Again the new centre points are calculated by averaging the values of the pixels in the cluster.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 1, January 2019

These two steps are repeated until the values of the central pixels do not change on average. K-means is one of the most widely used clustering algorithms when image data is large. It is simple and has less computational complexity.

## [B] Random Forest

Random forest is a more flexible machine learning algorithm. In this algorithm, binary decision trees are building depending on two processes. It is a supervised algorithm. Random forest builds multiple decision trees and merges them together to get a more accurate and stable prediction. In the first process, the bootstrap set is obtained by randomly sampling the training set. And in the next process, randomization is introduced. The best split is chosen by randomly choosing features from each node. Each and every tree in random forest is a weak classifier.

## [C] Fuzzy C Logic

Fuzzy clustering is the partitioning of the data in a collection of clusters which contains lacking distinction or singularity. Accuracy is calculated based on several iterations, the more the number of iterations, the more they give accuracy. In FCM fuzzy partitions are created by iterative optimization of object function, with the update of the cluster centre and membership function. Fuzzy logic is logic of fuzzy sets; a fuzzy set has potentially an infinite range of truth values between one and zero. FCM provides a better result for overlapped data points or overlapped regions which belongs to more than one or single cluster.

1. A homogeneous system has the same properties at every point. I.e. data in one cluster should be as same as possible.
2. In heterogeneity system, there is variation from one service to another i.e. data or objects from different clusters are follows dissimilarity property.

## [D] Convolution Neural Network (CNN)

CNN is a type of feed-forward artificial neural network which is inspired by the biological nervous system. CNN has an ability to adapt on how to do tasks on given data which makes it important in Brain Tumor detection due to variations in size and shape of the tumor. Convolution Neural Networks are preferred over Fully Connected networks to compute less number of weights during each layer. Hence, Convolutional layers have fewer weights to train compare to other artificial neural networks which make CNN less prone to over fitting. CNN consists of input, output, and hidden layers. Input layer consists of 3D MR image and output layer which helps in classification and tumor area. The hidden layers consist of ReLU (Rectified Linear Unit), Convolution and Pooling layers. A specific model is chosen to work on CNN. The ReLU layer contains the activation function. The convolution layer will compute the output of previous layer neurons that are connected to local regions in the input, each multiplied between their weights and a small region they are connected to in the input volume. Pooling layer helps in down sampling in dimensions. If we hoard multiple convolutional layers, the obtained features become more conceptual with the increasing depth.

## IV. PROPOSED SYSTEM

The first step in the proposed system of our project is the acquisition of the MR images followed by their enhancement or pre-processing. The pre-processed image is further segmented by applying segmentation algorithms. The last step is the extraction of the tumor characteristics such as shape, intensity, sensitivity, area calculation, etc. The flow diagram of the proposed system for our project is as shown below:

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 1, January 2019

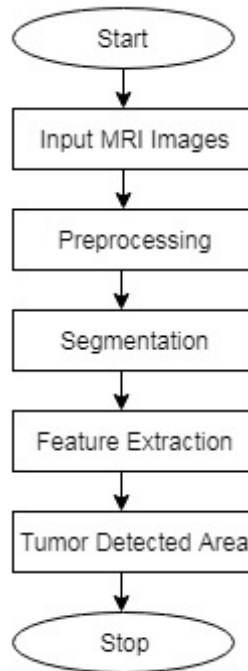


Fig 1: Steps in Tumor Detection

## [A] Pre-processing

The pre-processing step in tumor detection plays a very significant role since the input image may not be of desired quality. This step is useful when the acquired image is blurred or there is presence of some noise in the image. In order to get the exact tumor regions from the MRI, the input image is pre-processed to get an enhanced version of the image on which further operations of image processing could be performed. The first operation in image pre-processing is converting the input RGB colour image into a gray-scale image. The total image size is also reduced by eliminating the redundant and irrelevant data from the image. If there is noise present in the image, then it can be removed by applying various filters on the MR image. The median filter is effective in removing salt and pepper noise from the image whereas, a high pass filter enhances the high-frequency components in the image. Thresholding on other hand makes the all pixel values below the threshold value 0 and above the threshold value. Edge detection can be performed by applying Prewitt-Sobel mask and blurring effect can be reduced by applying a Wiener filter. The pre-processing step significantly improves the quality of the image so that it becomes suitable for segmentation operation.

Steps for pre-processing are as follows:

- 1) Firstly, the acquired MR image is converted from an RGB colour image to a gray scale image. A gray scale image often reduces the computational time and space complexity to process the image.
- 2) The second step in the pre-processing stage is to remove the noise present in the MRI. A 3x3 noise removing filter is applied on magnetic resonance image (MRI) in order to remove the noise. The type of filter which is used depends on the type of noise present in the image such as a median filter is excellent in reducing salt and pepper noise.
- 3) Thirdly, edge enhancement is performed. This mainly done using a high pass filter. For this purpose, the high pass filter mask is used.

## [B] Segmentation

Image Segmentation involves segmenting the image i.e. partitioning of an image into a set of pixels. The image segmentation results in contours extracted from the image and a set of regions that cover the whole image together. All the pixels in an area are similar in intensity, texture, or colour. The goal of image segmentation is to change the representation of a digital image into something that is easier to analyse and also gives meaningful outcome.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 1, January 2019

The different approaches of segmentation are :( i) first; find limits between regions based on discontinuities in intensities.

(ii) Second, define thresholds based on the distribution of pixels.

(iii) Third, directly find the regions.

## [C] Feature Extraction

Feature extraction is a comprehensive understanding of the image features. This process extracts crucial information about organ under diagnosis. In this process, texture, shape, contrast, and colour etc. is extracted. This process is used to simplify the analysis, and to reduce the time consumed to find the type of tumor. It is mainly used for reducing the complexity in classifying the characteristics of an image. It is therefore used to increase the factor of accuracy.

In this process, the following features are extracted:

- Type
- Location
- Contrast
- Intensity

## V. CONCLUSION

In this paper, a survey on various brain tumor detection algorithms is conducted. A comparative study is made on the various algorithms. At first, the various methods, which are being currently used in MRI image processing are broadly studied. This involved studying the available algorithms, use of these algorithms in various fields, etc. Based on that survey, this paper was written mentioning the various techniques in use. A brief description of each algorithm is also provided. Also, a detailed description of processes involved in brain tumor detection technique is given. Segmentation is the most indicative and expressive part of the detection process. Computational time required for each algorithm will also be considered. As the identification of a tumor is an intricate and tactful task, here much more importance is given for accurate and reliable result. Hence a detailed methodology that highlights a new prospect for developing a sturdier image segmentation technique is much sought.

## REFERENCES

- [1] Sérgio Pereira, Adriano Pinto, Victor Alves, and Carlos A.Silva Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO.5, MAY 2016.
- [2] A. Florence, Dr. J.G.R Sathiaselalan Fuzzy c-means Clustering Algorithm for Brain Tumor Segmentation International Journals of Advanced Research in Computer Science and Software Engineering ISSN: 2277-128X (Volume-7, Issue-6).
- [3] Najeebullah Shah, Sheikh Ziauddin, Ahmad R. Shahid Brain Tumor Segmentation and Classification using Cascaded Random Decision Forests 017 14th International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI-CON).
- [4] Vipin Y. Borole, Sunil S. Nimbhore, Dr. Seema S. Kawthekar Image Processing Techniques for Brain Tumor Detection: A Review International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) ISSN 2278-6856.
- [5] Kaus, M.R., Warfield, S.K., Nabavi, A., Black, P.M., Jolesz, F.A. and Kikinis, R., 2001. Automated segmentation of MR images of brain tumors. Radiology, 218(2), pp.586-591.
- [6] Ashwini A. Mandwe, AnisaAnjum Detection of Brain Tumor Using K-Means Clustering International Journal of Science and Research (IJSR) ISSN (Online): a. 2319-7064.
- [7] Sanjeev Thakur, Luxit Kapoor A Survey on Brain Tumor Detection Using Image Processing Techniques 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence10.1109/CONFLUENCE.2017.7943218.
- [8] DevendraSomwanshi, Pratima Sharma, Deepika Joshi, Ashutosh Kumar An efficient Brain Tumor Detection from MRI Images using Entropy Measures IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2016), December 23-25, 2016.
- [9] Miss. Shrutika Santosh, Prof. Swati Kulkarni, Prof. AkshataRaut Implementation of Image Processing for Detection of Brain Tumors Proceedings of the IEEE 2017 International Conference on Computing Methodologies and Communication 10.1109/ICCMC.2017.8282559.
- [10] BhushanPawar, Siddhi Ganbote, SnehalShitole, MansiSarode, RupaliPandharpatte Optimizing Problem of Brain Tumor Detection Using Image Processing e-ISSN: 2395 -0056.
- [11] AMRUTA PRAMOD HEBLI, SUDHA GUPTA BRAIN TUMOR DETECTION USING IMAGE PROCESSING: A SURVEY ISSN: 2347-6982 Volume-5, Issue-1, Jan.-2017.