

How to Cite:

Praveena, S., Singh, S. P., & Kumar, B. S. (2022). Machine learning mechanism for segmentation, progressive assessment and prediction of brain tumor growth. *International Journal of Health Sciences*, 6(S2), 7696–7709. <https://doi.org/10.53730/ijhs.v6nS2.6908>

Machine learning mechanism for segmentation, progressive assessment and prediction of brain tumor growth

S. Praveena

Department of Electronics & Communications Engineering, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India
Corresponding author email: spraveena_ece@mgit.ac.in

S. P. Singh

Department of Electronics & Communications Engineering, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India

B. Suresh Kumar

Department of Electrical & Electronics Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India

Abstract---A brain tumor is one type of illness caused by clots in the brain. A Magnetic Resonance Imaging (MRI) scan can be used to see a brain tumor in detail. Because of the similarity in color, it is difficult to distinguish brain tumor tissue from normal tissue. Brain tumors are common in all age groups, early and accurate identification of tumor type is critical for determining the most successful treatment regimens for each patient's situation. In today's world, computer interpretation of medical images plays a major role in medical diagnosis. Brain tumors can be accurately diagnosed with the use of advanced technologies. Early-stage diagnoses, as well as properly quantifying the amount and intensity of sickness, which were previously a barrier for medical science, are now eliminated by recent technologies. The automated technique for brain tumor identification was used in the study. The technique includes gray-scale conversion for reducing computation requirements. Filter operation was used to eliminate unwanted noises as much as possible to assist in improved segmentation. Next brain tumor segmentation, isolates tumor tissue from surrounding edema, fat, and cerebrospinal fluid. K-means clustering is used for this segmentation. Finally, two machine learning techniques, Support Vector Machine (SVM) and Decision Tree (DT) are used to classify benign and malignant tumors after extracting features by Gray-Level Co-occurrence Matrix (GLCM). Using the performance evaluation approach, the best model out of two is discovered,

especially accuracy, precision, recall, and F1 score are deployed for evaluation. The radiologist can use this ML model for the accurate classification of brain tumors from MRI.

Keywords--bank, loan, attributes, categorical data, metrics, mobile application.

Introduction

Annually, over 190,000 people are detected with brain tumors [1] worldwide. Although the reasons for brain tumors are uncertain, there are numerous similarities among those who suffer from the disease. Anyone, regardless of age or gender, can be affected. The rapid rise in death rates among children, adults, and especially the elderly can be attributed in part to brain tumors. Because the human is composed of millions of cells, know that cells proliferate, expand, and divide to generate new cells and tissues. The tumor may develop as a result of unregulated cell development caused by external factors. Tumors of the brain are characterized as masses that form when the brain's regulatory mechanisms are broken and abnormal brain cell proliferation occurs. There is a possibility of tumors developing in the skull, which can result in increased brain pressure and a decline in overall health. The benign tumors do not migrate to other body cells, making them the less dangerous of the two types of tumors. Malignant tumors, on the other hand, are collections of cells that are carcinogenic, destructive, and more likely to spread to other tissues and cells than benign tumors [2]. A tumor, as a group of cells, is essentially a tissue that lacks the same amount of control as healthy cells and grows at an unpredictable rate. According to data from most developed countries, brain tumors kill more than 300 people each year and are anticipated to climb in the future years [3]. According to a United News of India study, brain tumors are the tenth most frequent type of cancer in India. They cite a report from the WHO-affiliated International Association of Cancer Registries (IACR) in their report, which write about. There are 28,142 new cases of brain tumors in India each year, with a high death rate, according to the study (24,003 deaths). A study by the Clinical cancer investigation journal was done in Eastern India. It looked at 130 people who had brain tumors. The study [4] looked at people who were between the ages of four and 78, with a mean age of 42.38 and a lot of men. There are more than 2,500 Indian children who get medulloblastoma each year, according to a 2016 article in The Hindu. This is a brain tumor that spreads through cerebrospinal fluid. According to this study, approximately 4,000-5,000 people are diagnosed with a brain tumor each year, with children accounting for 20% of those diagnosed. According to a study, 90% of medulloblastoma patients can be treated, managed, and possibly preserved if brain tumors are detected early. In the tumor location, there is a lower probability of death [5]. As a result, the role of radiology in studying brain tumors through imaging has expanded. Extensive research on brain tumors has been conducted, however, the results have not been definitive. A range of imaging modalities, including MRI, ultrasonography, and computed tomography, can be used to detect brain tumors (CT). Early detection of tumors can help radiologists, physicians, and expert decision-makers, leading to better patient outcomes. MRI is by far the most popular and successful method of brain imaging. If tumor

segmentation and categorization were done manually, time and human error would have to be minimized. This issue can be overcome with automated segmentation and classification. The following points are made in connection to the paper: 1) Identification of the problem statement, 2) Prior work, 3) Research steps, 4) Data capture, 5) Image Pre-process, 6) Segmentation, 7) Feature extraction, 8) Classification algorithm, 9) Results, and 10) Conclusion

Literature Survey

The goal of the study [6] is to use a convolutional neural network (CNN) to classify brain tumors so that they can be found earlier. Three different types of CNN models are recommended for three different types of categorization tasks. At a rate of more than 95%, the first CNN model is very good at figuring out which brain cancers are real. For example, the second CNN model is 92.66% accurate at classifying brain tumors, and it can do this 92.66% of the time. When it came to brain tumors that were grade 2, grade 3, and grade 4, the third CNN model got it right 98% of the time! The grid search optimization method automatically finds and allocates all of the crucial hyperparameters of CNN to them. Using the k-means clustering algorithm, the lung capacity is broken down into parts. Lungs with tumors and lungs that aren't. Use these five statistical measures to characterize CT scans. They are all based on how much space the lungs have. The features were found using a technique called MLP, which can easily tell the difference between tumorous and non-tumorous features. The method in the article [7] worked well when using the KNN and MLP. The paper [8] says that a CNN-based hierarchical deep learning (DL) system can be used to find brain tumors. The model said that the input was either tumor or no tumor. When they looked at two different models, the presented method had an accuracy rate of 92.13% and 7.87%, compared to 7.87% and 92.13% for the other model. Whenever a tumor is observed, the state puts it into a group.

The proposed strategy was demonstrated to be a significant method for detecting brain tumors in the study [9]. A hybrid approach that combines SVM and fuzzy c-means clustering may accurately identify the tumor. In the future, a hybrid SVM algorithm will be proposed to improve accuracy while decreasing the error rate. The MRI comprises several sorts of noise. The anisotropic filter is used to improve performance. An SVM classifier is used to divide pixels into two classes. SVM with a kernel is utilized for unsupervised learning since the system is designed to function with any MRI brain image as an input. Morphological approaches are used to eliminate tumors from the segmented area. After a long wait, technology has finally been able to detect the tumor exactly [10]. Proposed work in the journal [11] with an accuracy of around 94% implies that this strategy is the best alternative for the automatic categorization of brain tumors. The proposed work is the most accurate when compared to the accuracy of other classifiers such as ANFIS, PNN, CNN, and SVM. The collection contains over 5000 MRI images that can be used to categorize tumors. Other types of tumors in the lungs and bones can be classified in the same way. The study [12] disclosed extensive Joint Intensity Matrix (JIM) features that compute the joint probability across multiparametric MRI images. These variables were also used to compare and predict gene status as well as survival outcomes in patients with lower-grade glioma (LGG). Experiments with this novel JIMs version reveal that it outperforms

the original JIMs in terms of determining gene status and survival rate of LGG patients.

Research Flow

The automated brain tumor detection research flow is shown in figure 1.

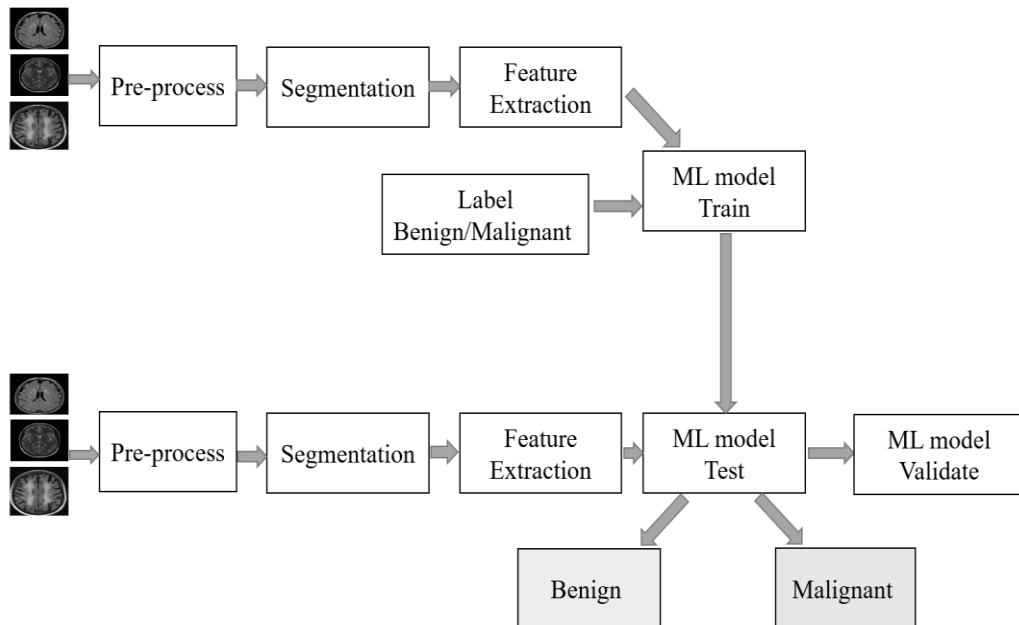


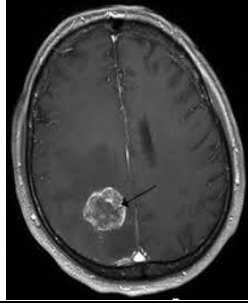
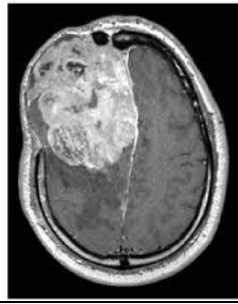
Fig. 1. Automated brain tumor detection research flow

The first stage of this research is the collection of brain MRI. The image is passed through the pre-processing stage. The pre-process is very important for making the image ready for segmentation and classification. The pre-processing is used for removing noise and reducing computation. The pre-processed image is given to the K-means cluster for identifying the tumor area in MRI. After segmentation, the important features are extracted using GLCM. Next, the two ML model is trained using 80% of image data and its label. The trained models are tested by the remaining 20% of the image and asked to predict the label. The predicted result is validated for identifying the best model.

Data Acquisition

The data of the BraTS-19 is used in this research. From the beginning, brain tumor segmentation in multimodal MRI images has been a key emphasis of BraTS. The pre-operative MRI scans for BraTS 2019 are focused on gliomas, which are intrinsically diverse in appearance, form, and histology. Furthermore, BraTS'19 focuses on predicting patient overall survival using integrated assessments of radiomic features and ML approaches to determine the clinical relevance of this segmentation task. Last but not least, at BraTS'19, an experimental evaluation of tumor segmentation uncertainty will be performed. Table 1 provides a more in-depth splitting of the data.

Table 1. Image Count

Tumor Type	Benign	Malignant
Sample Image		
Total Raw data	500	500
Total Train data	400	400
Total Test data	100	100

Data Processing

Pre-processing is largely concerned with improving the quality of MR images and converting them into a format that can be used for more analysis by either a person or an ML system is required. Some of the advantages of pre-processing MR images are improved signal-to-noise ratio, enhanced image, removal of extraneous noise, smoothing the interior section of the region, and keeping its edges [13]. The phases involved in pre-processing an MRI are as follows:

Image Resize

An image processing technique is the use of an algorithm to modify an image. The versatility, scalability, storage systems, and interconnection are some of the advantages. Digital images, for example, can be used for copyright infringement because they cannot be downsized without losing quality, need massive amounts of memory, and require a faster processor for editing. This is a disadvantage of the digital image. Images may now be saved more effectively thanks to advancements in image scaling algorithms. Several criteria are used to synchronize the image at the same time. The employment of image scaling methods is necessitated by the lack of a predetermined image size. There is a chance that the tumor images obtained from the internet will contain tumors of various sizes. As a result, the image should be adjusted to keep the pixel width and height constant. In this work, the captured image has been scaled to 256*256.

Gray Conversion

Grayscale conversion is the next step in the pre-processing process. Some MRI images have a limited scale of RGB content because the red, green, and blue components have the same intensity values for each pixel, hence a single

intensity value per pixel must be defined. By converting the image's RGB content into a grayscale image with a defining intensity value, the process of filtration and enhancement is simplified [14]. Using the MATLAB program `rgb2gray`, MR images are converted to grayscale images. The brightness of this color can be expressed in decimal values ranging from 0 to 255 or binary numbers ranging from 00000000 to 11111111. The black color of a grayscale image is represented by 00000000, and the white color is represented by 11111111. This image processing approach is known as 8-bit grey scaling.

Filtering

In the field of MRI processing, de-noising has long been a hot topic of study. Approaches for reducing and removing noise are key components of processing systems. It is a technique for eliminating noise introduced during the MR imaging process. The processes of capturing, processing, and storing the MRI all have the potential to decrease image quality. Researchers continue to struggle with the difficulties of eliminating noise from original MRI because of introduces artifacts and blurs the MRI. MRI de-noising is now a significant function in medical imaging, especially in MRI. Denoising MRI can be accomplished in a variety of ways, within each set of benefits and drawbacks. This study shows that the method used is specific to the type and quantity of MRI noise. Other considerations are the MRI's de-noising efficiency, processing time, and affordability. Spatial, frequency, and wavelet domains can all be used for de-noising. Noise removal, interpolation, and re-sampling are just a few of the many purposes for which filtering in MR image processing is employed. For the best results, it is important to take into account the amount and types of noise in an image while selecting a filter [15].

More and more people are turning to Adaptive Median Filtering (AMF) [16] as a better de-noising method than conventional median filtering (MF). The AMF performs spatial processing to determine if an MR image has been influenced by noise. Pixels are classified as noise using the AMF's comparison of each pixel to its neighboring pixels. Both, the comparative threshold and the neighborhood window size can be altered. A noisy pixel is structurally out of alignment with the pixels to which it is architecturally comparable. The noise-labeled pixels in a neighborhood are utilized to replace these troublesome pixels with a median value obtained from all of the neighboring pixels. When the adaptive MF is used, the neighborhood window size can be changed. The neighborhood window of an MF, on the other hand, remains constant throughout the process. The AMF is a better fit for this type of situation because the MF suffers from high amounts of impulsive noise. Furthermore, unlike the regular MF, the AMF conserves MRI properties namely non-impulsive noise. To deal with noise, the adaptive MF may vary the size of S_{xy} while it's operating.

AMF algorithm [17]:

Stage 1

$$\begin{aligned}
 A_1 &= z_{med} - z_{min} \\
 A_2 &= z_{med} - z_{max} \\
 \text{if } A_1 > 0 \ \& \ A_2 < 0, \text{Stage 2} \\
 \text{else, hike window size}
 \end{aligned}$$

Stage 2

$$\begin{aligned}
 B_1 &= z_{xy} - z_{min} \\
 B_2 &= z_{xy} - z_{max} \\
 \text{if } B_1 > 0 \ \&\ B_2 < 0, z_{xy} \\
 &\text{else, } z_{med}
 \end{aligned}$$

Where,

z_{med} → Gray level median

z_{min} → Gray level min

z_{max} → Gray level max

z_{xy} → Gray level @ co-ordinate (x, y)

S_{max} → S_{xy} max

Thresholding

The pixels in an image background can be distinguished from those in an anticipated target by applying a threshold to each pixel, which is accomplished with the use of a process known as thresholding. Depending on the image obtained used, pixels grey levels belonging to target and pixels assigned to the background may be significantly different from one another. When Otsu wants to use his technique of thresholding, he goes through all the probable threshold levels and counts the number of pixels that fall along both sides of the threshold, which are those in the foreground or background. Make sure that when add up all of the parts, get a value that is as low as possible in intensity. In general, the most straightforward feature to share among "pixels" in a given location is intensity. As a result, thresholding is commonly used to distinguish between light and dark zones. In this case, a greyscale image is converted to a binary image by converting all pixels below a particular threshold value (TV) to "0" and all pixels above the TV to "1" [18]. The creation of a "criterion function" that results in a dissociation between areas is the foundation of optimal thresholding. When it comes to intensity, an objective criteria function is made. The function that maximizes to choose the threshold. Otsu's thresholding chooses the TV that has the least difference between black and white pixels in different classes. It is used in qualitative statistics to measure how statistically different two things are. In the first step of image segmentation, there is a set of criteria for maximizing a class variance. This criterion is similar to the criterion for maximizing the conventional among variance, but it is not the same. Following that, a recursive method is created to quickly find the best threshold. It uses the new requirement to do this. Many things can do with images to make them more monochrome. A grayscale image is reduced to a binary image using Otsu's method.

Image Segmentation

Image segmentation is a critical step in analyzing MRI data for brain tumors. In the aforementioned situation, the human expert manually segments the tumor. This manual segmentation is a time-consuming, arduous operation that frequently involves lengthy procedures, and the results are highly dependent on human ability. Furthermore, these findings differ from expert to expert and are rarely replicated by the same expert. Although the segmentation method has been researched for a long period, it is still in need of improvement, especially in

the medical image field. This is partly because the present segmentation approach [19] has the following flaws: (1) There is no universally accepted gold method of describing nodules or tumor margins due to adherence between the tumor and adjacent tissues, and defining tumor margins is highly arbitrary. Because of the large range of possible subjective interpretations, consistent segmentation results are difficult to get. This complicates the ML-based segmentation method since the training set's segmentation labels are not always correct. (2) Physical segmentation of tumors takes a while for physicians. The tumor must be marked layer by layer, as larger nodules are typically scattered across numerous slices. Furthermore, modeling nodules is challenging because it lack typical geometric features. (3) The tumor margin will be vague due to the influence of adhesions and other variables, affecting segmentation. (4) The ensuing feature extraction technique, which depends on stable and reproducible derived features, requires consistent tumor segmentation. Despite the problems and limits associated with the segmentation method discussed previously, scientists are continuously refining the technique for specific circumstances. As a result, automatic segmentation and reproducible segmentation methods are in high demand. MRI segmentation allows for improved classification of brain tumor subtypes, which aids in diagnosis. It enables precise delineation, which is critical in radiation or surgical planning [20].

Clustering is a technique for categorizing a collection of data into a predefined number of categories. K-means is among the most often used techniques for this. It separates a big amount of data into k categories using k-means clustering. It categorizes a set of data by dividing it into k different groups. Two elements contain the K-means method. The first step identifies the k centroid. In the second step, each point is assigned to the cluster having the nearest centroid to the data point. The Euclidean technique would be used to ascertain the distance between the closest centroid and the origin. Furthermore, the centroid of every cluster is computed to guarantee that the cluster's points are associated in the minimum possible Euclidean distance. The division's clusters are made up of components and their centres. Each cluster is connected by a line that travels through its centre. This line is referred to as a "centroid." To summarise, K-means is an iterative algorithm that seeks to minimize the total distances among all items and their cluster centroid. Consider an image with a resolution of $x * y$ that must be clustered into k number of clusters. The cluster centers are represented by the input pixels $p(x, y)$ and c_k , respectively.

The K-means clustering algorithm is as follows:

1. Initialize the cluster size k and the center.
2. Calculate the Euclidean distance d between the image's center and each pixel in the image using the relation given below for each pixel.

$$d = ||p(x, y) - c_k|| \quad [1]$$

3. Allocate each pixel to the closest point based on the distance d .
4. Once allocating all pixels, recalculate the new center using the equation shown below.

$$c_k = \frac{1}{k} \sum_{y \in c_k} \sum_{x \in c_k} p(x, y) \quad [2]$$

5. Repeat the process till the tolerance level is met.
6. Resize and modify the cluster pixels to create an image.

K-means has the advantage of being easy to implement, but it also has certain drawbacks. The final clustering findings' quality is dependent on the arbitrary selection of the initial centroid. As a result, the results will differ if the initial centroid is chosen at random. As a result, the starting center will be carefully selected so that achieve the desired segmentation. When creating a K-means clustering algorithm, another term to bear in mind is computational complexity. It is determined by the number of data items, clusters, and iterations. Figure 2 shows the MRI in each step like raw image, gray image, filtered image, threshold image, and segmented image.

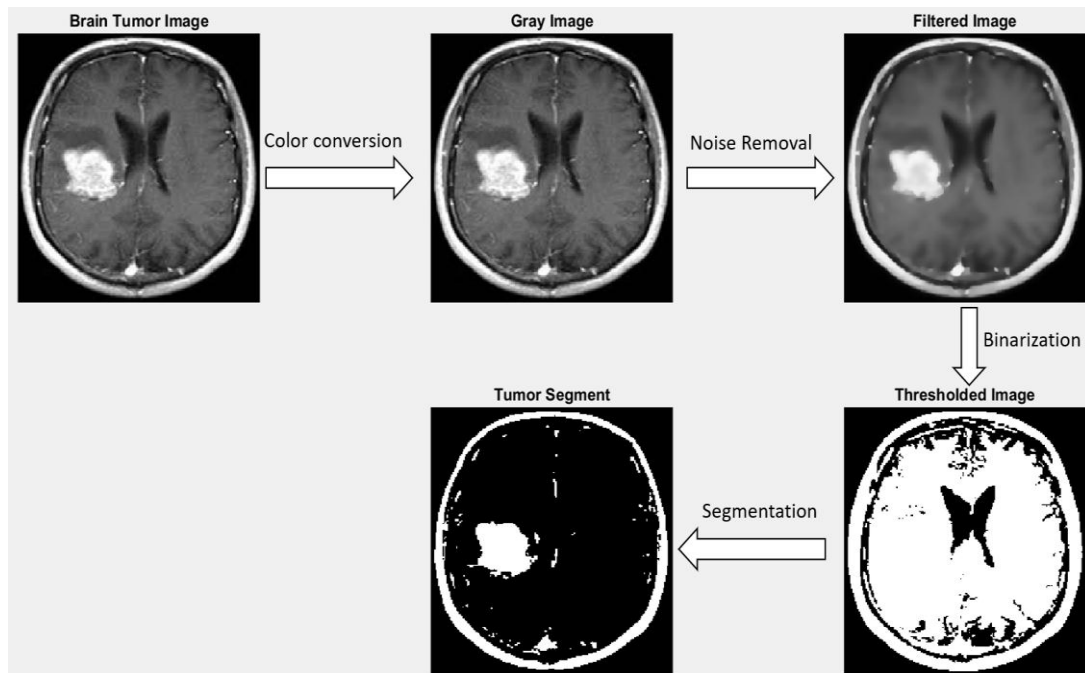


Fig. 2. Raw image to segmentation

Feature Extraction

Extracting image features is a very important part of the study. Textural features that were investigated and demonstrated to work can help people tell the difference between tumors and regular brain tissue when they look at a brain scan. The GLCM feature is premised on the reality that two grey levels that are next to each other happen to be in the same image. By multiplying the quantity of grey - levels in the MRI by their ROI dimension, the feature is determined by calculating. The GLCM looks at the space between pixels to figure out the texture. A different name for it is called the gray-level spatial dependence matrix, or GSD matrix. This method captures an image's higher-level details, like its shape, texture, color, and contrast, by taking a picture. Texture analysis highly depends on people's visual acuity and ML mechanisms. It's being used to enhance the accuracy of the detection method by choosing features that are important to the process. The GLCM is being used to look for features in medical images, and they are used to do this. A two-step process. First, the GLCM is measured. In the next

phase, the GLCM is estimated with texture features. These images can be used to study the complex structure of brain tissues, and it takes a lot of work to figure out which properties are important. The MRI shows 12 different features. Adding more features makes the classifier's calculations more complicated, so the number of features grows.

Tumor Classification

Data mining frequently employs systems that generate classifiers. Classification algorithms in data mining are capable of handling huge amounts of data. It could be used to make assumptions about the names of categorical classes, characterize understanding, and classify new data with training datasets and corresponding labels. When it came to ML classification techniques, this study focused on the DT and SVM algorithms.

DT

DT has extensive applications in a variety of disciplines, including ML, image processing, and pattern identification. Because each test in the series compares a numeric property to a TV, DT is an efficient and effective tool for connecting fundamental tests. When use the DT, essentially employing a divide-and-conquer algorithm that works from the top down. Other than the root, all nodes are evaluated to see if they have any qualities that can be checked in the training sample set. Following that, the training sample set is subdivided into several subsets based on the outcomes of the testing. Each subset represents a new leaf node, and the prior process repeats for every new leaf node to guarantee that looping proceeds to achieve the predetermined termination state [21].

The process of building a DT is separated into two stages: tree construction and tree pruning. Tree construction is the first stage of the procedure. During the tree-building step, a subset of the training sample is picked and a DT is formed using the breadth-first recursive approach till every leaf node corresponds to the same category. The second stage is classification. Secondly, there is the pruning phase, which checks the produced DT and corrects any errors before cutting the DT and introducing nodes till the proper DT is constructed. Pruning, a recursive operation that ends in a DT, decreases the effect of noise on categorization accuracy by reducing the number of misclassifications. Generally, the bigger the amount of information gained, the greater the amount of "purity enhancement" that may be produced by partitioning the sample using attributes. This implies that choosing attributes for DT partitioning based on information gain is a valid choice.

The information entropy is commonly employed to estimate the amount of information, data entropy as w .

$$Q_L(w) = \frac{1}{2} \sum_{i,j} P_{i,j} (w^T x_i - w^T x_j)^2 \quad [3]$$

The following equation is used for determining the gain rate:

$$m_i(g) = \frac{fit_i - worst(g)}{best(g) - worst(g)} \quad [4]$$

SVM

SVMs are supervised teaching methods that use data to look for patterns and gain knowledge from them. SVMs are often employed for categorization and regression. An SVM training function produces a hyperplane for two-class problems, which assigns new data to one of 2 classifications on either phase of a hyperplane. This makes a non-probabilistic binary classifier [22]. When using an SVM model, think of observational data as space points. This allows to group them into different groups based on the highest distance to the nearby assessment data point of any category. People then put the new things they learn into a class based on which side of the divide they are on. As long as the problem is not differentiable in the input space, Soft-margin SVM can't find a powerful separating hyperplane that reduces the sum of miscategorized data points and works well when it doesn't work well. With a kernel, the data can be moved into a higher-D space called a kernel, where they can be separated in a straight line. Instead of having to solve a high-order partitioning hypersurface in the input space, a linear hyperplane can be found in the kernel space to separate the many classes associated with the categorization question. As compared to learning a nonlinear surface, trying to switch to kernel space. This makes it a good choice. This means that the kernel or Gram matrix must be evaluated on all pairs of data points as positive and semidefinite.

$$K(x, u) = \sum_r \varphi_r(x) \varphi_r(u) \quad [5]$$

where $\varphi(x)$ belongs to the Hilbert space

The Gaussian radial basis function (RBF) kernel is used in this work

$$K(x, u) = \exp \left[-\frac{\|x-u\|}{\sigma} \right] \quad [6]$$

Results

The tumor's location can be detected with pinpoint accuracy by analyzing 1000 MRI data. In this study, the BraTS-19 dataset was used to train and evaluate the model. The collected image is divided into two parts: 80% and 20%. The image has been pre-processed, including grey conversion, noise reduction, and binarization. Binarization is accomplished using the threshold method. An AMF is used to minimize noise. The tumor is then segmented from the original image using the K-means cluster. GLCM is an application for extracting features from an MRI image. In the second phase, SVM and DT classifiers are employed to classify the images. Finally, the optimum model is determined using metrics such as accuracy, recall, precision, and f1-score. Table 2 displays the model and metric values. Table 2 also includes the formula used to calculate the metrics. In the formula, TP and TN indicate the model's correct estimations of malignant as well as a benign brain tumors. The FP is the value of a false estimate of a benign tumor, while the FN represents the inverse of the FP . In the formula, T represents the sum of all four elements.

The accuracy value obtained on test data by SVM is 90% and by DT is 90%. The recall of both model (SVM & DT) is 90.38%, 88.46%. The highest precision of 95.91% is given by SVM and DT gives the lowest precision of 92%. Finally, the F1 score is more in SVM when compared to DT. The table is values are plotted using the line graph and it is shown in figure 3. The figure depicts that SVM gives a superior response in all metrics categories than the DT. The orange 3D line shows the SVM, and blue depicts the DT results.

Table 1. Model Comparison Using Metrics

MODEL	ACCURACY	RECALL	PRECISION	F1 SCORE
Formula	$\frac{TP + TN}{T}$	$\frac{TP}{TP + FN}$	$\frac{TP}{TP + FP}$	$\frac{2TP}{2TP + FP + FN}$
DT	90	88.4615	92	90.1961
SVM	93	90.3846	95.9184	93.0693

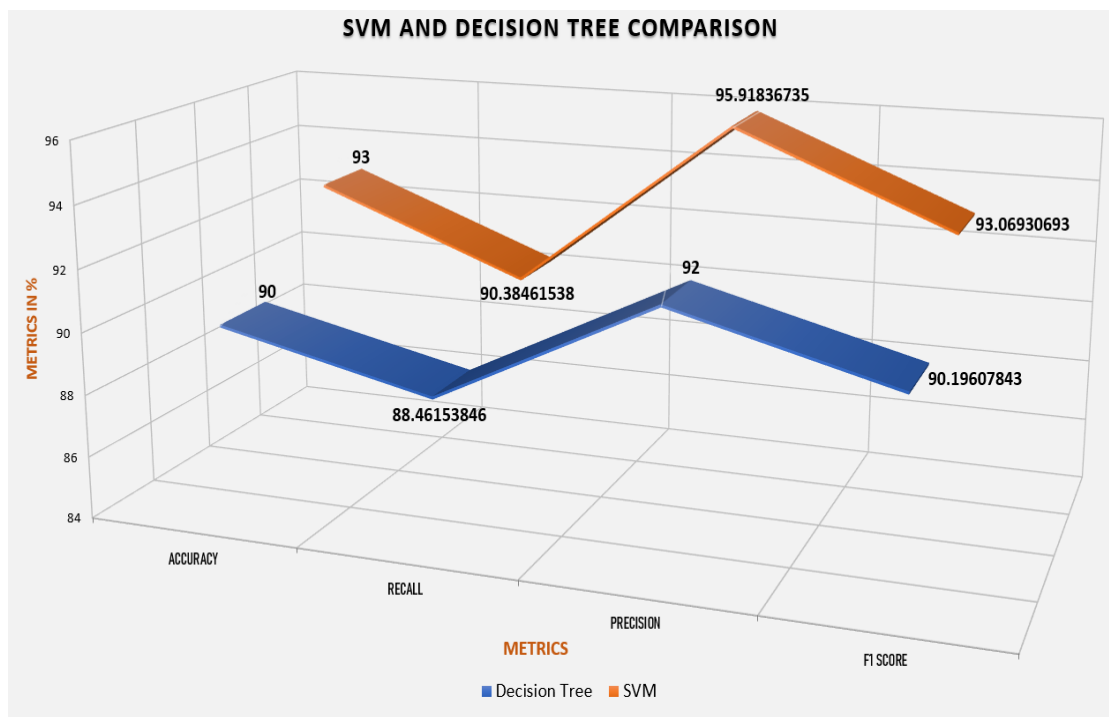


Fig. 3. Comparison of SVM and DT

Conclusion

Brain tumors must be discovered and treated immediately since they constitute a significant health risk to the human race. A total of 1000 BraTS-19 MRI data were used to detect a brain tumor. The data is processed using the threshold approach to binarize and reduce image noise. Next processing, the following two procedures are required: segmentation and classification. Because of the intricacy of the MRI and the deficiency of anatomical models that correctly describe the numerous deformations in each component, medical image segmentation is a difficult

problem. K-means clustering techniques with good accuracy and computation speed are used for segmentation. This paper proposes a brain tumor partitioning approach that requires little to no human interaction. The major goal of this proposed technique is to allow human specialists or neurosurgeons to identify patients in a short amount of time. Tumors are identified as benign or malignant using ML-based on their appearance. This work seeks to detect brain tumors using MRI and two distinct techniques, SVM and DT. Both approaches have a high accuracy rate in detecting tumors. SVM is exceedingly promising in the analysis of brain tumors, with an accuracy rate of 93%. Tumors found in other parts of the body can be spotted and analyzed using the same method. Different classifiers and optimization methodologies could be employed in the future to increase performance by accurate identification of segmentation and extraction method.

References

1. Arif, Muhammad & Ajesh, F. & Shamsudheen, Shermin & Geman, Oana & Izdrui, Diana-Roxana & Vicoveanu, Dragos, "Brain Tumor Detection and Classification by MRI Using Biologically Inspired Orthogonal Wavelet Transform and Deep Learning Techniques", *Journal of Healthcare Engineering*, vol. 2022, pp. 1-18, 2022.
2. Shibuya, M, "Welcoming the New WHO Classification of Pituitary Tumors 2017: Revolution in TTF-1-Positive Posterior Pituitary Tumors", *Brain Tumor Pathology*, vol. 35, issue. 2, pp. 62-70, 2018
3. Himaja Byale, Dr. Lingaraju G M and Shekar Sivasubramanian, "Automatic Segmentation and Classification of Brain Tumor using Machine Learning Techniques", *International Journal of Applied Engineering Research*, vol. 13, no. 14, pp. 11686-11692, 2018
4. M. A. Dorairangaswamy, "A novel invisible and blind watermarking scheme for copyright protection of digital images," *IJCSNS International Journal of Computer Science and Network Security*, vol. 9, no. 4, 2009.
5. Arti Tiwari, Shilpa Srivastava, Millie Pant, "Brain Tumor Segmentation and Classification from Magnetic Resonance Images: Review of selected methods from 2014 to 2019", *Pattern Recognition Letters*, vol. 131, 2019
6. Irmak, E, "Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework", *Iran J Sci Technol Trans Electr Eng*, vol. 45, pp. 1015-1036, 2021.
7. S. Potghan, R. Rajamenakshi, and A. Bhise, "Multi-Layer Perceptron Based Lung Tumor Classification," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 499-502, 2018.
8. Khan, Abdul & Abbas, Sagheer & Khan, Muhammad & Farooq, Umer & Khan, Wasim & Siddiqui, Shahan & Ahmad, Aiesha, "Intelligent Model for Brain Tumor Identification Using Deep Learning", *Applied Computational Intelligence and Soft Computing*, pp. 1-10, 2022.
9. Parveen and A. Singh, "Detection of a brain tumor in MRI images, using a combination of fuzzy c-means and SVM," 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN), pp. 98-102, 2015.
10. M. H. O. Rashid, M. A. Mamun, M. A. Hossain, and M. P. Uddin, "Brain Tumor Detection Using Anisotropic Filtering, SVM Classifier and

- Morphological Operation from MR Images," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), pp. 1-4, 2018.
11. B. Jabber, K. Rajesh, D. Haritha, C. Z. Basha and S. N. Parveen, "An Intelligent System for Classification of Brain Tumors with GLCM and Back Propagation Neural Network," 2020 4th International Conference on Electronics, Communication, and Aerospace Technology (ICECA), pp. 21-25, 2020.
 12. A. Chaddad, C. Desrosiers, B. Abdulkarim, and T. Niazi, "Predicting the Gene Status and Survival Outcome of Lower Grade Glioma Patients with Multimodal MRI Features," *IEEE Access*, vol. 7, pp. 75976-75984, 2019
 13. A. Demirhan, M. Toru, and I. Guler, "Segmentation of tumor and edema along with healthy tissues of the brain using wavelets and neural networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1451-1458, 2015.
 14. Kanan, Christopher & Cottrell, Garrison, "Color-to-Grayscale: Does the Method Matter in Image Recognition?", *PLOS ONE*, vol. 7, pp. e29740, 2012.
 15. Patel K, Mewada H, "A review on different image de-noising methods", *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 2, issue. 1, pp. 155-159, 2014
 16. L Lin, X Meng, X Liang, "Reduction of impulse noise in MRI images using the block-based adaptive median filter", 2013 IEEE International Conference on Medical Imaging Physics and Engineering (ICMIPE), pp.132-134, 2013
 17. Ali, Hanafy. "MRI Medical Image Denoising by Fundamental Filters" In *High-Resolution Neuroimaging: Basic Physical Principles and Clinical Applications*, IntechOpen, 2018
 18. Bangare, Sunil & Dubal, Amruta & Bangare, Pallavi & Patil, Suhas., "Reviewing Otsu's Method for Image Thresholding", *International Journal of Applied Engineering Research*, vol. 10, pp. 21777-21783, 2015.
 19. Jie Tian and Di Dong and Zhenyu Liu and Jingwei Wei, "Radiomics and its Clinical Application, Chapter 1 - Introduction", *Academic Press, The Elsevier and MICCAI Society Book Series*, pp. 1-18, 2021
 20. Kunapuli, Seshadri & Bhallamudi, Praveen, "A review of deep learning models for medical diagnosis", *Machine learning, big data, and IoT for Medical Informatics*, pp. 389-404, 2021
 21. M. S. Yorgun and R. B. Rood, "A decision tree algorithm for the investigation of model biases related to dynamical cores and physical parameterizations," *Journal of Advances in Modeling Earth Systems*, vol. 8, no. 4, pp. 1769-1785, 2016.
 22. Awad, M., Khanna, R, "Support Vector Machines for Classification", In *Efficient Learning Machines*, pp. 39-66, 2015