Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

SICAIET
2024

Research Article
Collection: Artificial Intelligence and Emerging Technologies

MRI Brain Tumor Diagnosis Using Machine Learning and Fused Optimization Scheme

Syed Ali Nawaz[1*], Rizwan Ali Shah[1], Tareef Ali Khan[2], Tanveer Aslam[1], Abdul Haseeb Wajid[1], and Mubashir Hussain Malik[3]

[1]Department of Information Technology, The Islamia University of Bahawalpur (IUB), Bahawalpur 63100, Pakistan.
[2]Department of Information Security, The Islamia University of Bahawalpur (IUB), Bahawalpur 63100, Pakistan.
[3]Department of Information Technology, Institute of Southern Punjab, Multan, Pakistan.
*Corresponding Author: Syed Ali Nawaz, E-mail: ali.nawaz@iub.edu.pk.

Academic Editor: Salman Qadri **Published:** February 01, 2024

Abstract: Brain tumors are fatal diseases that are detected following a panic, delayed, and complex process. Radiology is vast and diverse through the detection of brain tumors but examining the brain radiology images requires high skills, experience, and domain knowledge. Variations in the brain tumor tissues and similarity in the cases makes the diagnosis process more difficult. Computer-aided biomedical brain lesion identification using magnetic resonance imaging (MRI) lesses the difficulties in brain tumor detection and localization process, and it also addresses the shortage problem of skilled radiologists. In this research article, a fused optimization brain tumor diagnosis (FO-BTD) model has been proposed using brain MRI scans and machine learning classifiers. The experimental dataset comprised 200 MRIs of normal and abnormal brain tumors, either benign or malignant, collected from the Bahawal Victoria Hospital Radiology Department (BVH-RDL). A median filter was applied to reduce the noise effects and after segmenting the tumor region two ROIs of the sizes (10x10) were taken on each MRI. From each ROI 220 COM texture features were extracted, and a fused supervised feature optimization scheme gave thirty optimized features. The fused optimization comprised fisher (Fr) plus probability_of_error (POE) plus avg_correlation (AC) plus mutual_information (MI). The optimized features vector (OFV) as input to machine learning classifiers named random forest (RF) and logit-boost (LB) to classify brain MRI dataset. RF and LB classifiers gave 84.50% and 83.71%.

Keywords: MRI; Brain tumor; Machine learning; FO-BTD.

1. Introduction

Despite all the advancements, life is still threatened by deadly diseases such as, cancer, hepatitis, diabetes, cardiovascular, and Alzheimer's are few examples. Cancer can destroy any critical human organ like lungs, liver, stomach, or brain. Disastrous of cancer, just in the year 2018 are reported by the International Agency for Research on Cancer (Selvaraj, et al., 2007). The brain is composed of billions neurons and a trillion glials cells, and tumors occur when these cells proliferate abnormally, forming a mass. These abnormalities in the brain cells are identified as a neoplasm or the brain tumor [2]. Benign and malignant are two main types of brain tumors. Brain tumors are also categorized as either primary or secondary [3]. Primary brain tumors originate within the brain such as from the membranes, cranial nerves area, pituitary glands, or even within the pineal gland. Conversely, secondary brain tumors, which are more prevalent, develop when cancer cells originate from other body part and metastasize to the brain [4]. Metastatic brain tumors represent the most prevalent variety of brain tumors and are invariably malignant. The most frequent origins of these secondary brain tumors are lung and breast cancer. The manifestation of symptoms in brain tumor cases is influenced by the tumor's size, type, and specific location. Typical symptoms encompass morning headaches that alleviate throughout the day, nausea, vomiting, impaired coordination, difficulty in walking, seizures, and alterations in speech, vision, or hearing. It's important to note, though, that these symptoms can also stem from other medical issues. There are over 120 distinct types of primary brain tumors where gliomas are the most common type of the brain tumors [5].

The process for accurately diagnosing a brain tumor involves several steps undertaken by radiologists, neurologists, and physicians, including physical examination, review of medical history, use of contrast agents, and biopsy tests [6]. The goal is to locate the abnormal tissues convincingly and the exact location, area and orientation of the abnormal tissue. Evaluation of the imaging scans, and their interpretation phase follow the physical examination and historical analysis, for the creation of digital brain images. The preferred imaging technique is Magnetic Resonance Imaging (MRI) for its superior contrast and resolution, [7]. While Computed Tomography (CT) scans are also used; however, they are not as effective as MRIs [8]. Manual brain tumor diagnosis is a complex, time-consuming and stressful task with the potential for human error, due to factors such as fatigue and information-overload, hence early and accurate diagnosis of Brain Tumor is very important and inspires a myriad of research. Also, the accurate measurement of the area of tumor is required for the targeted treatment. Machine Learning (ML) methods are significantly advancing the field of medical image analysis. Recent developments in ML have led to the creation of automated systems for diagnosing brain tumors. These systems provide vital support [2].

2. Literature Review

In a research study, an artificial neural network (ANN) model has been experimented to identify tumor in MRI images, employing Gray Level Co-occurrence Matrix (GLCM) for extracting texture features from tumors. The acceptable accuracy of 81.4% emphasizes the potential of ANN architecture and feature functions to achieve better data separation [9].

Another ANN based research study presents a method for brain tumor detection using image processing techniques including MRI images pre-processing, segmenting them using a novel method based on mean and standard deviation, two fundamental methods named Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) were applied for features extraction. The model was t-ested on 65 MRI images (37 abnormal, 28 normal), the method achieved noteworthy accuracy in distinguishing normal and abnormal tissues [10]

In another research experiment, threshold and watershed segmentation schemes were applied to isolate damaged part where SVM discriminated normal and abnormal brain tissues by giving overall classification accuracy up to 85.32% [11].

Detection and Recognition of Brain Tumor Based on DWT, PCA, and ANN is presented in another research work. The work diagnosis benign and malignant brain tumors using DWT, PCA, and ANN on MRI images. The research work introduced a novel PCA+RST hybrid feature selection method, achieving good accuracy in classifying benign and malignant tumor [12].

MRI based classification using machine learning models, particularly the support vector machine model has been designed using a subset of 32 optimized features. The research highlights the importance of feature selection and the potential impact of image quantization levels on classification accuracy in a non-invasive and fast diagnostic approach for brain lesions [13].

A variation of SVM, the least-square-support-vector-machine (LS-SVM) was incorporated by Selvaraj and his co-authors to diagnose brain tumor. The proposed model discriminated the normal brain and the tumorous brain images successfully. They used GLCM to extract the features for texture analysis of the normal brain and the tumorous brain images. The research experiment tried different classifiers including MLP, radial-basis-function, k-nearest-neighbor and LS-SVM. The LS-SVM beat the other classifiers and gave noteworthy classification accuracy [14].

Kaur and his research fellows demonstrated various image segmentation techniques. Major segmentation categories described were model-based, partial_differential_equation based, threshold-based, edge_detection based, region_growing based, clustering-based, and watershed-based. Thresholding, edge detection, region growing, and watershed seemed more promising [15].

Another novel hybrid feature selection method for MRI brain tumor classification combining Principal Component Analysis (PCA) and Rough Set Theory (RST) is presented. This approach significantly reduced the data dimensionality, focusing on the most relevant features for classification. Four classification algorithms (J48, SVM, KNN, Naive Bayes) were used for testing, with the hybrid approach achieving high classification accuracies, notably outperforming traditional methods like DWT+SVM, DWT+PCA+ANN, and others. This method demonstrates enhanced performance in distinguishing between benign and malignant tumors, highlighting its potential in aiding early and accurate brain tumor diagnosis [16].

Arsa, with other participants, proposed a brain tumor segmentation model based on Sobel operator plus thresholding. After computing the initial gradient, the threshold value was fixed, which helped to differentiate between background pixel and edge pixel. A closed contour algorithm was applied recursively to implement seeded-growing, and tumor regions were separated. The proposed model worked better than others [17].

Another research article introduces a fully computerized system using an Artificial Neural Network (ANN) to differentiate between benign and malignant breast tumours, to characterize breast cancer using multi-fractal dimensions and backpropagation neural networks, by analysing 184 images. The research gained high accuracy in classifying breast tumours, showing precision of 82.04% [18].

The mechanism to detect tumors in various brain sectors, such as the central middle region which is aligned with our eye level, has been investigated in another research study, where the malignant brain tumor, using Naïve Bayes classification are identified. The proposed approach involves pre-processing MRI images, extracting statistical features, and applying Naïve Bayes classifier. Tested on 50 MRI images, this method achieved an 81.25% detection accuracy for the brain tumor MRIs [12].

The study introduces an advanced approach for classifying brain tumors using MRI images. It combines Gray Level Run Length Matrix (GLRLM) and Center-Symmetric Local Binary Patterns (CSLBP) for feature extraction, which are then processed by an Artificial Neural Network (ANN) for classification. The research underscores the potential for using combined statistical texture features in medical imaging, particularly for brain tumor classification. The method achieved a good classification accuracy and outperformed than several existing methods. This demonstrates its effectiveness in distinguishing between malignant and benign brain tissues [19].

The survey of the literature elaborated that most of the research work is based on the unoptimized features selection methods and thus are less reliable as irrelevant features must impact on the authenticity of the result accuracies. Moreover, many of the schemes are based on few MRI images as their data input, thus there should be enough input data to train the machine. And lastly, mostly the surveyed scheme didn't focus on a specific dataset. We have contributed by designing a novel brain tumor diagnosis model titled FO-BTD, based on the fused optimization scheme. The proposed model diagnoses the brain tumor by classifying the MRI dataset using the machine learning classifiers.

3. Materials and Methods

This section demonstrates the complete introduction to the designed FO-BTD model. Our model is incorporating the fundamental steps of knowledge discoveries from databases using machine learning classification. At first step, the brain MRI scans were collected. At second step, the collected brain MR images were pre-processed. At third step, the tumor area was marked by the expert radiologists to confirm the ground truth value. At fourth step, GLCM features were drawn out for texture analysis. At fifth step, the extracted features were optimized by applying a fused optimization scheme. At the sixth step, machine learning classifiers were deployed to diagnose brain tumor. Complete methodology model is shown in Figure # 1.

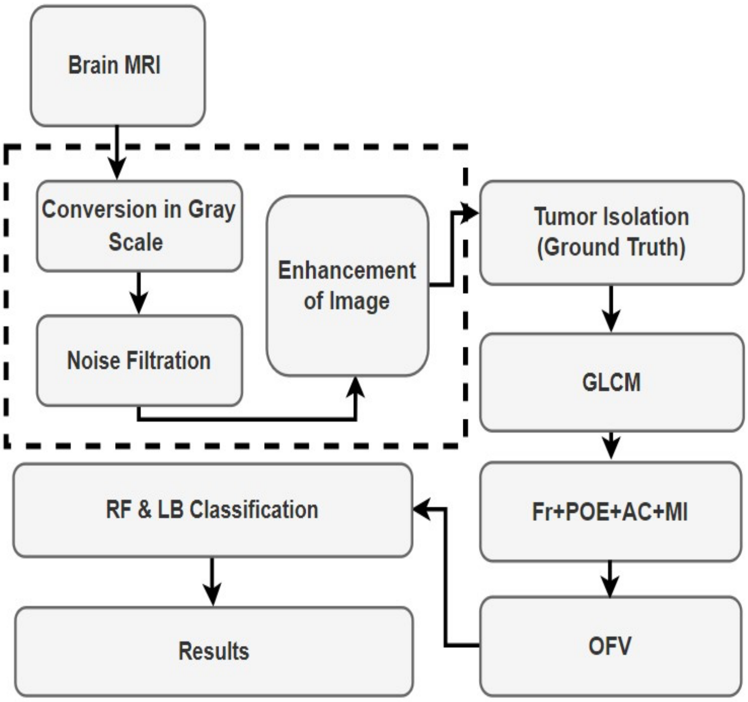

Figure 1. Methodology Diagram of Fused Optimized Brain Tumor Diagnosis Model

3.1 MRI Dataset Collection

The brain MRIs digital scans dataset for this research study was collected from the BVH-RDL[20] . A total of 200 brain MRIs were gathered where 100 MR scans were of normal brain and the 100 MR scans of tumorous brain. The tumorous MR scans were taken of different brain tumors, which include different scans of brain glioma scans.

3.2 MRI Dataset Pre-processing

The initial stage of preparing the MRIs digital scans for brain dataset classification involved image preprocessing. Initially, the MRI digital scans were transformed into grayscale images and all the images were cropped to remove the unnecessary machine numbers as well as the patient's name and the other identifications. Next, all the MR digital scans were resized to the standard size of 512×512. To reduce speckle noise, median filter is carried out on the pixel values. Ultimately, we obtain normalized, enhanced, and smooth MRI digital scans of the brain dataset. The enhanced sample brain MRIs digital scans of the normal and abnormal slices are shown in Figure # 2.

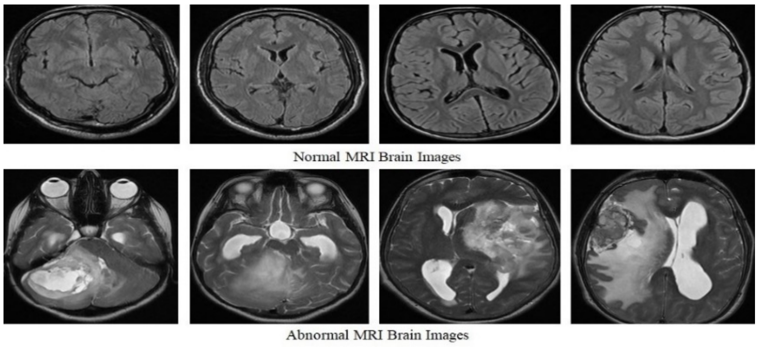

Figure 2. Sample Brain Normal and Abnormal MRIs

3.3 Feature Extraction

Feature extraction phase comprised number of key steps, which include tumor region segmentation by marking the tumor area by expert radiologist to ensure the ground truth value. When tumor areas were marked then process of taking region-of-interest (ROI) on each of the MRI digital scan was initiated. In this research experiment, a couple of equally sized ROIs of sizes 10×10 were taken on from each MRI digital scan. For texture analysis GLCM texture features were extracted from each of the ROI using MaZda 4.6 [21]. During this process, the total number of texture features were 88,000 (200×2×220). All the extracted features are described precisely.

3.4 Feature Optimization

The extracted 88,000 texture features were not sufficient for the efficient classification because a lot of unnecessary features were also becoming part of our obtained feature vector. Thus, in this step the unnecessary features were eliminated, and the features were kept in hand for further dataset classification. For efficient feature selection we designed a novel feature selection scheme by fusing multiple schemes. We fused probability-of-error, average-correlation-coefficient (POE + AC), Fisher (Fr), and the mutual information (MI). Fisher coefficient technique applies an indexing method to select the most discriminated features. POE takes the probabilistic approach to determine ratio of improperly classified features between the total number features. The AC computes the sums and averages of the old and new selections of the features and the coefficient of correlation. MI uses ranking of features and densities of the corresponding probabilities of the multiple random variables to select the most critical features. The fused optimization scheme gave the fused optimized features vector (FOV), containing the thirty most critical features which are shown in the Table # 1.

Table 1. Optimized Feature Set

S(0, 2) sum variance	**S(2_2) inverse_diff_momentum**	**S(0, 5) correlation**
S(1, 0) sum average	S(0_2) inverse_diff_momentum	S(0, 3) sum average
S(3, 3) entropy	S(0_3) inverse_diff_momentum	S(0, 4) correlation
S(1, 0) correlation	S(3_-3) inverse_diff_momentum	S(0,1) inverse differ momentum
S(5, -5) inverse diff momentum	S(4_-4) inverse_diff_momentum	S(2, 2) correlation
S(0, 1) angle second momentum	S(0_1) inverse_diff_momentum	S(0, 4) sum average
S(5, 5) entropy	S(0_4) inverse_diff_momentum	S(0, 4) correlate
S(1, 1) sum variance	S(5_-5) inverse_diff_momentum	S(0, 5) contrast
Skewness	S(2_-2) inverse_diff_momentum	S(0, 3) correlation
Percent 0.01%	S(0_5) inverse_diff_momentum	S(0, 3) contrast

3.5 Classification

Finally, the most optimized FOV was input the machine learning classifiers named RF and LB classifiers applying 10-fold cross-validation approach using the Weka 3.8 software [22].

3.5.1 Random Forest (RF)

RF is an excellent classification method, in which large number of decision trees are built, each utilizing distinct random selections of data and attributes. Think of each decision tree as a specialist classifier giving its own judgment on classifying the data. The final prediction is determined by first obtaining individual predictions from each tree and then selecting the most frequent outcome [23].

3.5.2 Logit Boost (LB)

Logit Boost is a boosting machine learning algorithm, best suited for binary classification problems. It is an extension of the AdaBoost algorithm, tailored specifically to handle binary classification problems. The core ideas behind LB are minimization of logistic loss, ensemble of many weak classifiers, and handling the overfitting problem. It iteratively adds weak classifiers to the model based on their previous misclassification, and the final model is a weighted combination of all the weak classifiers added during the iterations [23].

3.5.3 Performance Evaluation:

Performance measuring parameters include kappa statistics (K_Sta), receiver operating characteristic (R_O_Chrc), Total number of Instances (T_N_Ins), Time and confusion matrix. Proportions of true and false cases were described by sensitivity, specificity, and accuracy parameters. The parameters are defined by TP (true postive) rate, FP (false positive) rate [24].

4. Results

When MRI brain datasets were input to the machine learning classifiers namely, RF and LB the performance measure findings of the two classifiers are shown the Table 2. The table contains total eight columns. The first column of the table shows the names of the classifiers, and the remaining columns show the other performance measure parameters.

Table 2. Machine Learning Classifiers Performance Table

Classifiers	K_ Sta	TP Rate	FP Rate	R_O_Chrc	T_N_Ins	Time. (Sec)	Accuracy
RF	0.69	0.845	0.155	0.897	400	0.48	84.5%
LB	0.65	0.825	0.175	0.887	400	0.23	83.7093%

And the confusion matrix table for RF classifiers has been shown in the following Table 3. The table contains four columns showing the accurate classified samples and the mis-classified samples.

Table 3. Confusion Matrix for Random Forest based on 10 × 10 ROI

Classes	Normal	Abnormal	Total
Normal	186	14	200
AbNormal	48	152	200

The overall performance graph of RF and LB classifiers is shown in the Figure 3. It is abvious that RF acheied 84.5 classification accuracy, where as the LB got 83.7093 classification accuracy.

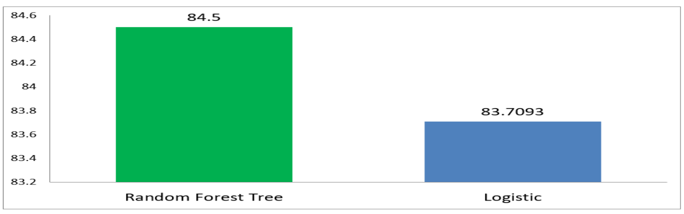

Figure 3. Sample Brain Normal and Abnormal MRIs

The above performance graph given in Figure 3, shows that RF classifier achieved 84.5% accuracy to classify the normal brain images from tumorous brain, where as the LB classifier gained 83.7093% accuracy to discriminate the normal and abnormal brain images.

5. Discussion

In this research experiment, two machine learning classifiers, namely, RF and LB were deployed to classify normal and abnormal brain MRI datasets. The performance measure parameters of RF classifier were presented in Table 2. The table shows that RF beat the LB classifier and gave overall 84.5% results accuracy whereas the LB classifier achieved 83.7093% classification results accuracy. The comparison given in Table 4 exhibits that the proposed model the FO-BTD gives the more classification accuracy among the other brain tumor diagnosis methods.

Table 4. Comparison with Other State-of-the-art Techniques

Source Citation	Extracted Features	Classifier	Overall Maximum Accuracy
[9]	Texture Features	ANN	81.4%
[18]	Texture Features	ANN	82.04%
[12]	Texture Features	Naïve Bayes	81.25%
(Selvaraj, et al., 2007)	Texture Features	SVM, MPL,	77 to 98%
Proposed CABGD Model	GLCM	PART	85%

6. Conclusions

This research study provides a novel computer-aided brain glioma identification model entitled FO-BTD has been designed. Firstly, glioma-infected patients and healthy brain MRI datasets were collected from the Radiology of Bahawal Victoria Hospital (BVH-RDL). Histogram equalization, grey-level collection, noise removal, and tumour localization according to the ground truth were the image pre-processing steps. A couple of ROIs of size 10×10 were circled on each of the MRIs, and GLCM features were extracted from each ROI for texture analysis. A composite feature selection scheme comprising Fr plus POE plus AC plus MI gave the thirty most optimal features set. The obtained features were input into three machine learning classifiers to identify glioma brain tumours. The LB classifier gave 83.71% classification accuracy and the RF classifier got 84.50%. section is not mandatory but can be added to the manuscript if the discussion is unusually long or complex.

References

1. F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA Cancer J Clin, vol. 68, no. 6, pp. 394–424, 2018, doi: 10.3322/caac.21492.
2. S. A. Nawaz, D. M. Khan, and S. Qadri, "Brain Tumor Classification Based on Hybrid Optimized Multi-features Analysis Using Magnetic Resonance Imaging Dataset," Applied Artificial Intelligence, vol. 36, no. 1, 2022, doi: 10.1080/08839514.2022.2031824.
3. T. Sinha, "Tumors: Benign and Malignant," Cancer Ther Oncol Int J, vol. 10, no. 3, pp. 1–3, 2018, doi: 10.19080/ctoij.2018.10.555790.
4. E. Bianconi et al., "An estimation of the number of cells in the human body," Ann Hum Biol, vol. 40, no. 6, pp. 463–471, 2013, doi: 10.3109/03014460.2013.807878.
5. L. D.N. et al., "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary," Acta Neuropathol, vol. 131, no. 6, pp. 803–820, 2016, [Online]. Available: http://link.springer.de/link/service/journals/00401/index.htm%5Cnhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed18a&NEWS=N&AN=610343744
6. M. J Strong and J. Garces, "Brain Tumors: Epidemiology and Current Trends in Treatment," Journal of Brain Tumors & Neurooncology, vol. 01, no. 01, 2016, doi: 10.4172/2475-3203.1000102.
7. S. Bauer, R. Wiest, L. P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," Phys Med Biol, vol. 58, no. 13, pp. 1–44, 2013, doi: 10.1088/0031-9155/58/13/R97.
8. G. Perkins,A. Liu, "Primary Brain Tumors in Adults: Diagnosis and Treatment - American Family Physician," Am Fam Physician, vol. 93, no. 3, pp. 211–218, 2016, [Online]. Available: www.aafp.org/afp
9. S. Jain, "Brain Cancer Classification Using GLCM Based Feature Extraction in Artificial Neural Network," International Journal of Computer Science & Engineering Technology, vol. 4, no. 7, pp. 966–970, 2013.
10. Z. Faisal and N. K. El Abbadi, "Detection and recognition of brain tumor based on DWT, PCA and ANN," Indonesian Journal of Electrical Engineering and Computer Science, vol. 18, no. 1, pp. 56–63, 2019, doi: 10.11591/ijeecs.v18.i1.pp56-63.
11. S. K. Hatcholli Seere and K. Karibasappa, "Threshold Segmentation and Watershed Segmentation Algorithm for Brain Tumor Detection using Support Vector Machine," European Journal of Engineering Research and Science, vol. 5, no. 4, pp. 516–519, 2020, doi: 10.24018/ejers.2020.5.4.1902.
12. H. Rajni, B. Soujanya, V. K. Reddy, and S. M. Vinutha, "Brain Tumor Detection using Naïve Bayes Classification," no. June, pp. 6851–6854, 2020.
13. R. Ortiz-Ramón, S. Ruiz-España, E. Mollá-Olmos, and D. Moratal, "Glioblastomas and brain metastases differentiation following an MRI texture analysis-based radiomics approach," Physica Medica, vol. 76, no. November 2019, pp. 44–54, 2020, doi: 10.1016/j.ejmp.2020.06.016.
14. H. Selvaraj, S. T. Selvi, D. Selvathi, and L. Gewali, "Brain mri slices classification using least squares support vector machine," IC-MED International Journal of Intelligent Computing in Medical Sciences and Image Processing, vol. 1, no. 1, pp. 21–33, 2007, doi: 10.1080/1931308X.2007.10644134.
15. D. Kaur and Y. Kaur, "International Journal of Computer Science and Mobile Computing Various Image Segmentation Techniques: A Review," International Journal of Computer Science and Mobile Computing, vol. 3, no. 5, pp. 809–814, 2014, [Online]. Available: www.ijcsmc.com
16. G. Alabbas, F. F. K. Hussain, and A. Hadi, "Hybrid Approach Based Feature Selection for MRI Brain Tumor Classification," no. September 2022, pp. 0–19, 2020.
17. A. Aslam, E. Khan, and M. M. S. Beg, "Improved Edge Detection Algorithm for Brain Tumor Segmentation," Procedia Comput Sci, vol. 58, pp. 430–437, 2015, doi: 10.1016/j.procs.2015.08.057.
18. P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule Networks for Brain Tumor Classification Based on MRI Images and Coarse Tumor Boundaries," in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, May 2019, pp. 1368–1372. doi: 10.1109/ICASSP.2019.8683759.

19. M. Mudda, R. Manjunath, and N. Krishnamurthy, "Brain Tumor Classification Using Enhanced Statistical Texture Features," IETE J Res, pp. 1–12, 2020, doi: 10.1080/03772063.2020.1775501.
20. QMC-BVH, "Bahawal Victoria Hospital - Quaid-e-Azam Medical College," Quaid-e-Azam Medical College - Bahawal Victoria Hospital. 2020. Accessed: Jun. 05, 2020. [Online]. Available: https://www.qamc.edu.pk/bahawalvictoriahospital/
21. M. Strzelecki, P. Szczypinski, A. Materka, and A. Klepaczko, "A software tool for automatic classification and segmentation of 2D/3D medical images," Nucl Instrum Methods Phys Res A, vol. 702, pp. 137–140, 2013, doi: 10.1016/j.nima.2012.09.006.
22. R. R. Bouckaert, E. Frank, R. Kirkby, P. Reutemann, A. Seewald, and D. Scuse, "WEKA Manual for Version 3-7-2," 2002, [Online]. Available: http://netcologne.dl.sourceforge.net/project/weka/documentation/3.7.x/WekaManual-3-7-12.pdf
23. I. H. Witten, E. Frank, and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 3rd ed. Elsevier, 2011. doi: 10.1016/C2009-0-19715-5.
24. S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," vol. 35, no. 5, pp. 1240–1251, 2016.