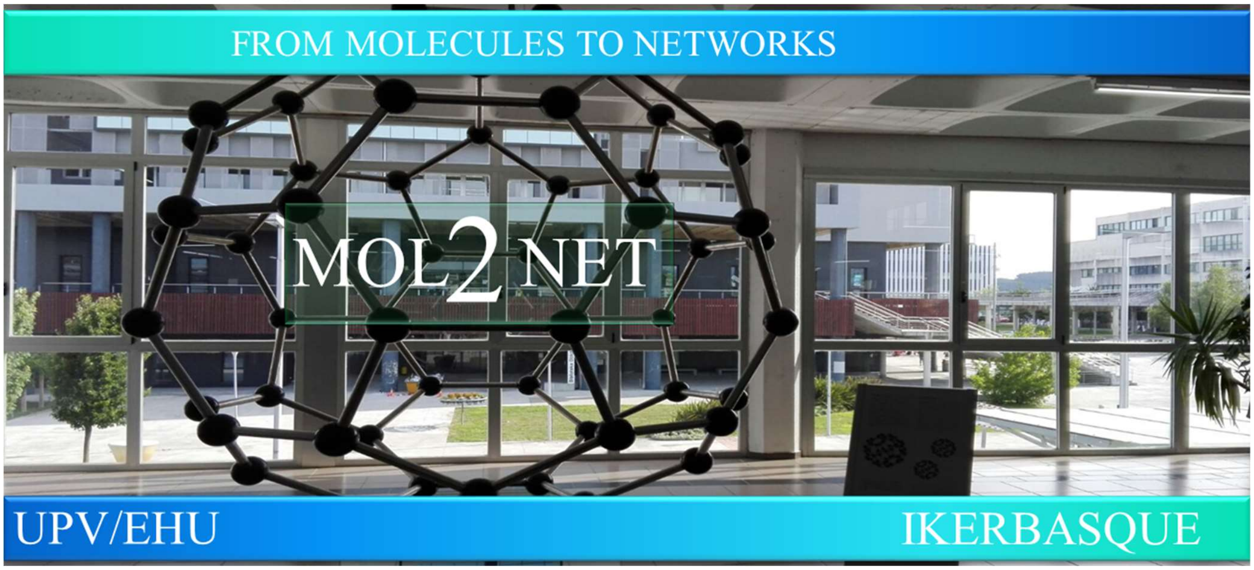
MOL2NET'22, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 8th ed.



A Machine Learning Approach for The Brain Tumor Classification Using MR Imaging

Aqib Ali [*, a], *and Samreen Naeem* [a]

[a] *College of Automation, Southeast University, Nanjing, China.*

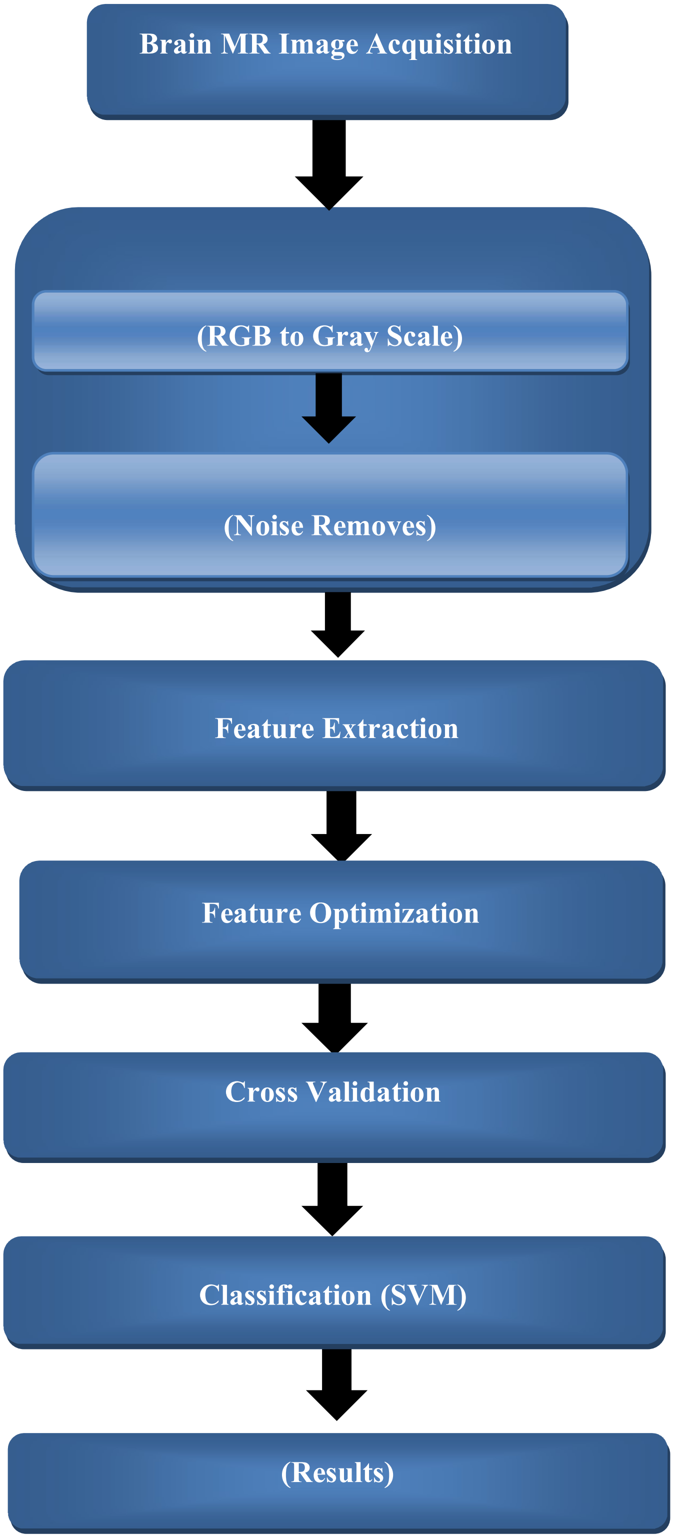
. * *Corresponding author: aqibcsit@gmail.com*

Abstract.

This research presents a novel approach for denoising, extracting, and detecting tumors on MRI images. Images obtained from an MRI scanner are helpful to medical professionals in the research and diagnosis of brain disorders and malignancies. This activity aims to assist the radiologist and the physician in obtaining a second opinion on the diagnosis. The ambiguity that existed in the characteristics of magnetic resonance (MR) images has been resolved more straightforwardly. In the paper, the magnetic resonance imaging (MRI) image that was obtained from the machine is analyzed. The study takes advantage of the data collected in real-time. A variety of noise-reduction filters are used throughout the execution of the fundamental preprocessing steps. After the image has been de-noised, it is segmented, and then feature extraction is carried out. The wavelet transform is used in order to extract the features. The wavelet transform is superior to other techniques in terms of its applicability to the MRI image feature extraction process. The characteristics are then sent to the classifier, which conducts classification via Random Forest. A comparison is made between the categorization procedure and more traditional approaches.

Keyword: MRI, Brain Tumor, Classification, Random Forest, Machine Learning.

Graphical Abstract



Introduction

Glioma is the primary brain tumor that occurs most often. Glioma accounts for 30–40% of all brain tumors in adulthood and approximately 80–90% of brain malignant tumors. Surgical resection is the most effective therapeutic approach currently available [1]. In clinical applications, automatic segmentation of brain tumors is beneficial for measuring tumor features. It may also aid clinicians in the diagnosis, treatment planning, and survival prediction of brain disorders. Magnetic resonance imaging, also known as MRI, is a technique that is commonly used in the field of radiographic imaging. Because it causes no damage, emits no ionizing radiation, and provides high contrast in the imaging of soft tissues, MRI has become the imaging method of choice for diagnosing and treating brain tumors [2]. At the moment, the industry standard for the segmentation of brain tumors. Nevertheless, it is costly, time-consuming, and open to interpretation. As a result, an MRI automated segmentation approach for brain tumors that is both quick and accurate is of utmost importance for clinical applications [3].

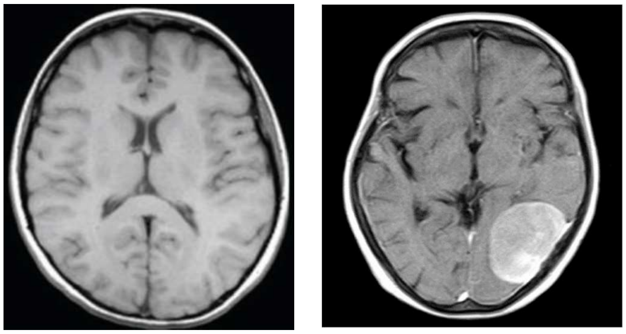
At the moment, there are primarily two categories of techniques that may be used for the automated segmentation of images depicting brain tumors. (1) Techniques of machine learning that is based on human-created characteristics. This technique uses various classifiers for the many distinct manual characteristics available, such as support vectors with spatial and intensity information. Machine [4], Gaussian mixture model with intensity features. On the other hand, these approaches involve the manual extraction of features, which is time-consuming, expensive, and prone to errors. Models that are based on manual features need to have adequate generalization. (2) Utilizing an all-encompassing deep learning methodology. This approach can produce more accurate segmentation results without requiring the construction of labor-intensive manual features. For instance, the fully convolutional neural network (called Convolutional Neural Networks, or FCN) extends image-level categorization to the pixel level using volume. It has become the pioneering work of deep learning in the application of semantic segmentation because the multilayer substitutes the fully connected layer and delivers excellent segmentation results [5]. This study was done on semantic segmentation. The fact that it is not sensitive to precise characteristics does not evaluate the interaction between pixels and pixels to their full extent, and the results of its segmentation need to be acceptable are all drawbacks of this method. U-Net network is a modification and extension of FCN. It uses cascade operation to merge deep features and shallow features; it also substitutes the summation approach in FCN, and it solves the issue of profound feature loss caused by the up-sampling process [6].

The precision of the segmentation has been significantly enhanced. An application for doing segmentation on three-dimensional medical images is known as the V-Net network [7]. The coding portion uses the residual network, which lessens the possibility of gradient disappearance or gradient explosion while simultaneously guaranteeing the steadily increasing depth of the network.

Materials and Methods

.

Medical image segmentation classifies each voxel by tissue type and anatomy. Segmentation improves medical image visualization, presents more information, and permits quantitative assessments of image structures. Segmentation is utilized for quantifying tissues, classifying brain white matter, gray matter, and cerebrospinal fluid, diagnostic imaging, and localizing anomalies, malfunctions, and diseases [8]. Neuroimaging is done before neurosurgery for therapeutic planning and decision-making. Manual, semiautomatic, and fully automated brain tumor segmentation exist. Manual segmentation involves designating the region of interest or tumor borders. Diagnosed anatomy is labeled. Manual demarcation requires complicated graphical user interfaces. Tumor area selection was also laborious. Clinical experiments involve manual demarcation. It is commonly utilized in clinical studies since tissue distinction requires human knowledge and experience. Semiautomatic brain tumor segmentation uses manual inputs and outputs. Fully automated approaches use the software in the acquiring machine to diagnose tumors without human intervention. Human intelligence-based algorithms like machine learning and decision-making modules are built in. Segmentation must be clear and understandable. This work offers automated brain tumor segmentation with little human intervention [9].



Non-Tumor **Tumor**

Figure 1: Sample Dataset

In this examination, which was carried out at the public dataset titled "Brain MRI Images for Brain Tumor Detection - Kaggle," for the sake of research, the dataset had a total of 300 samples with confirmed cases of brain cancer. These samples were given to other researchers so they may continue their work [10]. After that, the information was sorted using five different machine learning models, and then it was submitted to the test after being randomly divided into 10 different cross-validation groups. The classifier's Support vector machine (SVM) test accomplishments in this experiment were compared using Python. Models were constructed and tested to see how well they performed. Based on these features, it is possible to draw further conclusions about attributes such as the mean, standard deviation, worst-case scenario, and maximum value. Based on these results, a diagnostic class is created, which designates whether the tumor is benign or malignant. This class is denoted by the letters B (benign) and M (malignant).

Out of the 600 data points, 150 fall into the category of benign, and 150 falls into the category of malignant. The total accuracy of the categorizing work done by the system may be determined by comparing the test classes to the classes generated by the system and calculating the percentage of it that was predicted correctly by both sets. Within the categorized dataset, there is a comprehensive study of the following four probable outcomes: The term "true positive" refers to the instance in which the very first positive sample is accurately identified as positive (TP) [11]. Conversely, a false negative (FN) occurs when the first positive sample is incorrectly categorized as unfavorable while the initial negative sample is detrimental. In this scenario, the FN is the more serious of the two results. It is referred to as a true negative (TN) when correctly identified. When the initial sample that should have been negative is incorrectly identified as positive, it is referred to as a false positive (FP). The confusion matrix is a matrix that displays all of these different conditions [12].

Results and Discussion

- Support vector machine (SVM)
- Time taken to build the model: 0.13 seconds
- Test mode: Cross Validation 10

Table 1: SVM Classifier Summary

Total Number of Instances	300	
Correctly Classified Instances	295	**98.3333 %**
Incorrectly Classified Instances	5	**1.6667 %**
Kappa statistic	0.9667	
Mean absolute error	0.0268	
Root mean squared error	0.1089	
Relative absolute error	5.3547 %	
Root relative squared error	21.7721 %	

Table 2: SVM Classifier Detailed Accuracy

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	Class
0.967	0	1	0.967	0.983	0.967	0.998	**Tumor**
1	0.033	0.968	1	0.984	0.967	0.998	**Non-Tumor**
0.983	0.017	0.984	0.983	0.983	0.967	0.998	**Weighted Avg.**

Table 3: Confusion Matrix result using SVM Classifier

Classified as	**A**	**B**
A = Tumor	145	5
B = non-Tumor	0	150

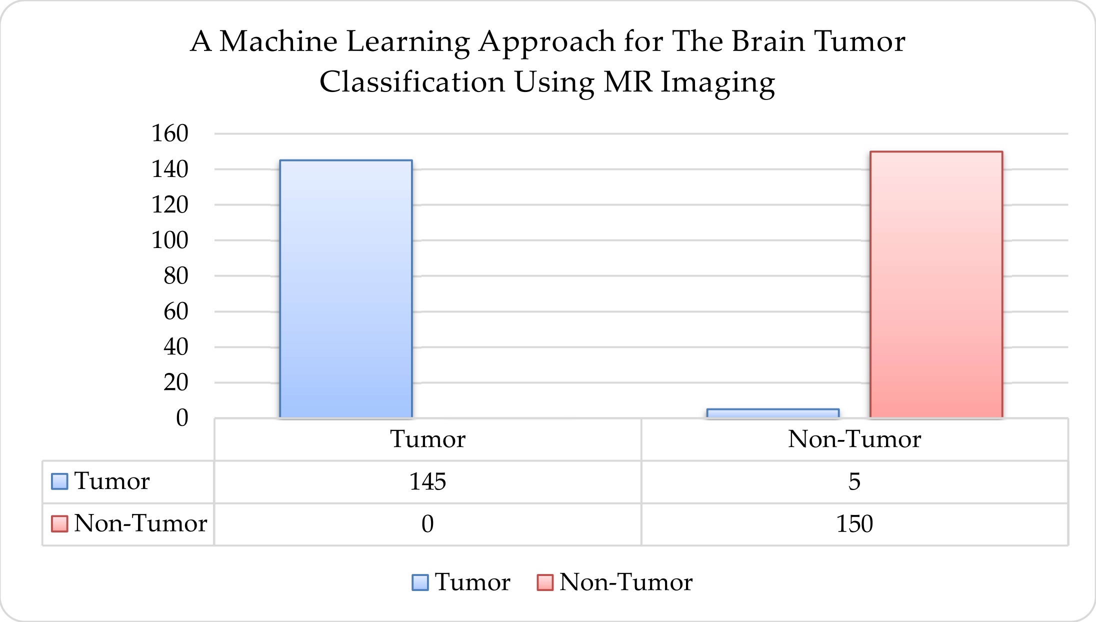


Figure 2: Accuracy of Dataset using SVM Classifier

Conclusions

Within the scope of this study, we have performed preprocessing on the MRI images and feature extraction. One of the most critical components of this work is the compilation of the database information. A further benefit of this study is that it makes use of both real-time Brain MR images as well as those that have been simulated. Second, a comprehensive preprocessing method is used to get rid of the undesired sounds. A high percentage of successful completions of this stage has ensured that the system as a whole will provide accurate results.

Last but not least, an ideal feature set is derived from these brain MR images. This feature set is highly significant for the performance optimization of the automated system. Because the convergence rate is also one of the performance indicators of this work, the number of features used in this work is manageable to prevent any computational complexity from occurring. In the not-too-distant future, the classification algorithm will be rewritten to use the SVM approach that was presented.

References

[1]. Ayadi, W., Elhamzi, W., Charfi, I., & Atri, M. (2021). Deep CNN for brain tumor classification. Neural Processing Letters, 53(1), 671-700.

[2]. Xie, Y., Zaccagna, F., Rundo, L., Testa, C., Agati, R., Lodi, R., ... & Tonon, C. (2022). Convolutional neural network techniques for brain tumor classification (from 2015 to 2022): Review, challenges, and future perspectives. Diagnostics, 12(8), 1850.

[3]. Sharif, M. I., Khan, M. A., Alhussein, M., Aurangzeb, K., & Raza, M. (2022). A decision support system for multimodal brain tumor classification using deep learning. Complex & Intelligent Systems, 8(4), 3007-3020.

[4]. Kumar, S., & Mankame, D. P. (2020). Optimization driven deep convolution neural network for brain tumor classification. Biocybernetics and Biomedical Engineering, 40(3), 1190-1204.

[5]. Ayadi, W., Charfi, I., Elhamzi, W., & Atri, M. (2020). Brain tumor classification based on hybrid approach. The Visual Computer, 1-11.

[6]. Raja, P. S. (2020). Brain tumor classification using a hybrid deep autoencoder with Bayesian fuzzy clustering-based segmentation approach. Biocybernetics and Biomedical Engineering, 40(1), 440-453.

[7]. Alqudah, A. M., Alquraan, H., Qasmieh, I. A., Alqudah, A., & Al-Sharu, W. (2020). Brain tumor classification using deep learning technique--a comparison between cropped, uncropped, and segmented lesion images with different sizes. arXiv preprint arXiv:2001.08844.

[8]. Khan, H. A., Jue, W., Mushtaq, M., & Mushtaq, M. U. (2020). Brain tumor classification in MRI image using convolutional neural network. Math. Biosci. Eng, 17(5), 6203-6216.

[9]. Guan, Y., Aamir, M., Rahman, Z., Ali, A., Abro, W. A., Dayo, Z. A., ... & Aamir, M. (2021). A framework for efficient brain tumor classification using MRI images. Math. Biosci. Eng, 18, 5790-5815.

[10]. Naeem, S., Ali, A., Qadri, S., Khan Mashwani, W., Tairan, N., Shah, H., ... & Anam, S. (2020). Machine-learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images. Applied Sciences, 10(9), 3134.

[11]. Ali, A., Qadri, S., Khan Mashwani, W., Kumam, W., Kumam, P., Naeem, S., ... & Sulaiman, M. (2020). Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image. Entropy, 22(5), 567.

[12]. ALI, A.; NAEEM, S.; ZUBAIR, M. Machine Learning Based Classification of Chronic Kidney Disease Using CT Scan Images, in Proceedings of the MOL2NET'22, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 8th ed., 26–31 December 2022, MDPI: Basel, Switzerland, doi:10.3390/mol2net-08-13903.