Brain Tumor Classification using a Support Vector Machine

Uppala Sai Sudeep
Vignan Foundation for Science Technology and Research
Dept of CSE

Kandra Narasimha Naidu
Vignan Foundation for Science Technology and Research
Dept of IT

Pulagam Sai Girish
Vignan Foundation for Science Technology and Research
Dept of IT

Tatineni Naga Nikesh
Vignan Foundation for Science Technology and Research
Dept of CSE

Ch Sunanda
Vignan Foundation for Science Technology and Research
Dept of Biotechnology

ABSTRACT
A person's life may be protected if a brain tumor is recognized early and treated effectively. The exact diagnosis of malignancies in MRI layers becomes a meticulous effort to perform, and as a consequence, the proposed method is capable of precisely classifying the tumor. Magnetic resonance imaging (MRI) is one of the most often used methods for analyzing brain tumor pictures. There are several image classification methodologies and algorithms. The purpose of machine learning and classification algorithms is to learn automatically from training and then make accurate conclusions. This study looked at the efficacy of tumor classification algorithms for categorizing MR brain image properties.During the classification process, the statistical features of the incoming images were evaluated, and the data was carefully split into multiple categories. These data were tested using SVM (support vector machines) and Logistic Regression machine learning algorithms. With a 96 percent accuracy rate, the SVM (support vector machines) technique was demonstrated to be better than other algorithms.

Keywords
Brain Tumor, MRI (magnetic resonance imaging), PCA (Principal component analysis), SVM (Supportvector machine), LG(Logistic Regression).

1. INTRODUCTION
Computer-assisted medical clinics provide a variety of services to help sufferers in identification, treatment planning, and further procedures to address the health issues of concern. Abnormal tissue growth in and around the brain is called a brain tumor. Brain tumors are classified as LGG and HGG. To classify the tumor, various imaging techniques are very helpful. Tumors of the central nervous system (CNS) are divided into primary and metastatic (secondary), which are classified among over 150 different kinds of CNS tumors. The main tumors arise in the brain or the immediate vicinity of the brain. Metastatic tumors, on the other hand, originate elsewhere in the body and travel to the brain through circulation. In contrast to primary tumors, metastatic tumors are termed cancerous or malignant. The current gold standard for classifying brain tumors is a biopsy. However, it is frequently necessary to do a full-blown brain operation to collect a sample. MRIs, on the other hand, may be used to automatically classify brain tumors without obtaining a tumor sample, making this method more effective and less dangerous. A machine learning-based classification of brain tumors using an MRI scan is also useful in improving diagnosis and therapy planning. Because of this, research into the automated classification of MRI scans for brain tumors using machine or deep learning methods is ongoing, with encouraging findings. Many authors have come up with different ways to classify brain images to figure out what kind of tumor it is. Tumors are further grouped as meningioma, glioma, and pituitary; astrocytoma, glioblastoma, and oligodendroglioma; glioma tumor grades (I–IV); benign and malignant (I–IV); diffuse midline glioma, medulloblastoma, pilocytic astrocytoma, and ependy. Neural networks, support vector machines (SVM), K-nearest neighbors (KNN), Adaboost, and hybrid models are the most often employed classifiers in brain tumor classification. Different neural network topologies, such as feedforward, multilayer perceptron, and probabilistic, were used to build the neural network (NN). The linear, homogeneous polynomial and Gaussian radial basis function(RBF) kernels were frequently used to create the support vector machine (SVM). Classifying brain cancers from an MRI brain image has made significant progress, but there are still significant problems in doing so due to inefficient ROI detection and handcrafted feature extraction approaches for gathering descriptive information from ROIs. Due to the complexity of brain architecture and the enormous density of the brain.

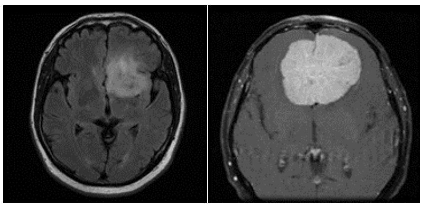
2. RELATED WORK
Zacharaki et al. [1] proposed a system to classify different grades of glioma besides a binary classification for high and low grades using SVMs and KNN. R. Ezhilarasiet al.[2] suggest utilizing a bounding box to detect the brain tumor region and forecast tumor type. The tumor is classified as malignant, benign, glial, or astrocytoma using these criteria. Brain MRI images are trained using a Support Vector Machine(SVM) from the ground up and get outstanding classification outcomes for both High-Grade Glioma (HGG) and Low-Grade Glioma (LGG). An SVM classifier is used to classify brain tumors based on depth and tumor stage. JT Kwok et al. [3] proposed discrete wavelet transformation based on the single fusion of images with all focal lengths using Support Vector Machines.

M.Gurbina, M.Lascu.[4] used Wavelets Transforms such as Discrete Wavelet Transform (DWT) and Continuous Wavelet

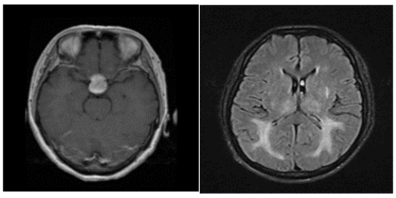
Transform (CWT) for detecting tumors and SVM for classifying tumors. El-Dahshan et al. [5] used DWT to categorize brain tumor images as normal or abnormal. Wavelet is computationally and storage-intensive. A new dimension reduction strategy PCA was used to reduce the feature vector's dimension and improve discrimination. Principal component analysis decreases data dimensionality and research costs. SVM classifies input data.Hebli Amrutha et al. [6] created a technique for detecting benign and malignant tumours in brain MRI images.

3. PROPOSED METHODOLOGY

The principal goal is to identify Meningioma, Glioma, pituitary, or no tumor from brain tumor data set as shown in Fig 1. (a),(b),(c), and (d). The authors proposed the framework for tumor classification as shown in Fig 2. The input image is initially preprocessed to prepare it for classification. The images that have been preprocessed are sent to a feature extractor. Principal Component Analysis (PCA) is used in this scenario to extract and reduce characteristics. PCA reduces the complexity of the predictor space. Classification models that are less prone to overfitting can be built by lowering dimensionality. By linearly modifying predictors and creating a new set of variables known as principal components, PCA reduces redundant dimensions.



(a) Meningioma (b) Glioma



(c) Pituitary (d) No tumor

A. *Pre-processing Input MR Image*

The MRI images, taken from multiple scanners are suffering from varying intensities and noise caused during image registration, blur edges, and so on. To address this problem, Images must be pre-processed before being fed into the proposed classification network in order to remove noise, uniformity across all images. Varying intensities are normalized using min-max normalization using eq 1., which scales intensity values to [0, 1]. The normalized intensity value is denoted by $f^{'}(x, y)$, while the image's minimum and maximum values are denoted by V_{max} , V_{min} .

$$f^{'}(x, y) = \frac{f(x,y) - V_{min}}{V_{max} - V_{min}} \quad (1)$$

We resize the normalized picture to 200x200 after normalization. Resizing MR images speeds up training and reduces memory reserve. In reality, the MRI dataset contains 512x512 MR images, while the size utilized for later classification is 200x200. Figure 1 depicts a resized image that has been adjusted (3).

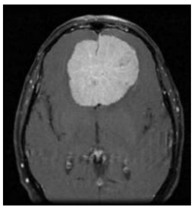


Fig 3. Normalized MR image

B. *feature extraction using PCA*

Various methods like DWT, GLCM and PCA are the existing approaches in machine learning for extracting feature maps from input data. Authors employed PCA that identifies, measures, and retains most of the data associations for classification tasks. The covariance matrix is used to detect feature correlation. The variance of different components is the diagonal of the covariance matrix. The eigenvalues and eigenvectors (principal components) of the covariance matrix are used to create the transformation matrix (TM). To reduce the data, eigenvectors with low eigenvalues are removed from the transformation matrix. As a result, PCA projects data onto eigenvector space, which has fewer dimensions than vector space.The eigenvectors are sorted in decreasing order within the transformation matrix.

C. *Classification*

Machine learning approaches identify hidden patterns with the help of learned features of the input data and are labeled in order to classify. Support Vector Machine(SVM) finds a hyperplane that classifies the data in N-dimensional space. SVM is commonly referred to as non-parametric. The Support Vector Machine (SVM) is a sophisticated machine learning system built through analytical learning. SVM categorizes MRI brain tumor training data intoeither malignant or benign. Because the data in training samples are organized as vectors, the number of rows in each vector corresponds to the number of observations. The amount of columns and medical images represents the collection of characteristics. This classifier can discriminate between Types of Tumors using practice samples. Images of bothmalignant and normal brains can be classified.

3.1 EVALUATION METRICS

Accuracy is used to assess the model's performance. TP (True Positive) shows the instances in which actual output and expected output differ, and TN (True Negative) denotes circumstances where the output is positive. FN (False Negative) shows the cases in which the actual output is true and the predicted output is incorrect. The term FP (False Positive) refers to situations in which the forecast is right but the actual outcome is incorrect.

$$\text{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} (2)$$

4. EXPERIMENTAL RESULTS

The findings obtained using the suggested approach are reported in this section. The Google Colab environment is used in the experiments in this paper. We used the Kaggle 2018 training dataset, which included four different types of tumor images. In an 80:20 split, the whole data is divided into training and testing data. Following validation, there are 2304 train data samples and 576 test data samples. After doing All the Pre-processing and Feature Extraction of the input data it is given to the Classification. In this paper, we have compared the Training, Testing, and Accuracy scores by Two Methods namely Support Vector Machine and Logistic Regression.

Table (2) illustrates the Scores.

Table 2: Performance evaluation of various models

Reference	Model Name	Data Set	Test Accuracy
[1]Zacharaki et	SVM	Real Human brain MRI data	85
[2]R.Ezhilarasi et al	R-CNN	Brats 2017	88
[4]M.Gurbina, M.Lascu	SVM and Wavelet-Transforms	Brats 2018	92
[5]EL-sydedA,El-Dahshan	Hybrid Techniques	Real Human Brain MRI Data	92
[6]AmruthaHelbi	SVM	Collected images from Clinical MRI centers	90
[7] K. Sudharani	K-NN	Brats 2015	95
[8]Parvathy	PCA	CE-MRI	87.6
[9] Yu-Dong Zhang	PCA and Kernel SVM	Collected from Harvard Medical School	95
	SVM	Clinical MRI Dataset	96

Table 1: List of Hyper Parameters

Model	SVM
Data Set	Clinical MRI Dataset
Image Size	256x256
Normalization Techniques	Min-Max Normalization
Metrics	Accuracy

5. CONCLUSION

We suggested a system for brain tumor classification based on dimension reduction in this research. Because the feature size is large, we apply PCA to reduce calculation time. In the classification challenge, the SVM classifier obtains 96% accuracy. The SVM method is effective in classifying brain tumors. The suggested method is tested on many Brain Tumor Images, and the results are the best and most effective.

6. REFERENCES

[1] Zacharaki et al.Classification of Brain Tumor Type and Grade Using MRI Texture and Shape in a Machine Learning Scheme,2009 Dec;62(6):1609-18.DOI: 10.1002/mrm.22147.

[2] R. Ezhilarasi et al.Tumor Detection in the Brain using Faster R-CNN,August 2018 Conference: 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), DOI:10.1109/I-SMAC.2018.8653705

[3] Li, Shutao, JT-Y. Kwok, IW-H. Tsang and Yaonan Wang. "Fusing images with different focuses using support vector machines." IEEE Transactions on neural networks15, no. 6 (2004): 1555-1561.DOI:https://doi.org/10.1109/TNN.2004.837780

[4] M.Gurbina, M.Lascu,"Tumor Detection and Classification of MRI images using Different Wavelets Transforms and Support Vector Machines".Published 1 july 2019 42ndInternationalConference on Telecommunications and signal processing.DOI:10.1109/TSP.2019.8769040

[5] EL-syded A. El-Dahshan," Classification using deep learning neural networks for brain tumors.Future computing and informatics Journal." https://doi.org/10.1016/j.fcij.2017.12.001

[6] AMRUTA HEBLI. Dr. SUDHA GUPTA," Brain Tumor Prediction and Classification using Support Vector Machine". 2017 International Conference on Advances in Computing, Communication and Control (ICAC3).DOI: 10.1109/ICAC3.2017.8318767

[7] K. Sudharani, T. C. Sarma and K. Satya Rasad, "Intelligent Brain Tumor lesion classification and identification from MRI images using a K-NN technique," 2015 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kumaracoil, 2015, pp. 777-780. DOI: 10.1109/ICCICCT.2015.7475384

[8] Parvathy Jyothi. Robert A Singh," Classification of Tumors from Brain MR Slices using PCA for Discriminative Models". International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181 Vol. 10 Issue 12, December-2021.

[9] Yu-Dong Zhang, LenanWu."AnMr Brain Images Classifier via Principal Component Analysis and Kernel Support Vector Machine".January 2012Progress In Electromagnetics Research 130:369-388 DOI:10.2528/PIER12061410

[10] Varuna Shree, N., Kumar, T.N.R. Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. Brain Inf. 5, 23–30 (2018) DOI:10.1007/s40708-017-00755.

[11] Vinotha, K., 2014. "Brain Tumor Detection and Classification Using Histogram Equalization and Fuzzy Support Vector Machine Approach," International Journal of Engineering and Computer Science ISSN23197242 3(5): 5823-5827.