3D MRI Based Brain Tumor Detection and Classification Using Multi Model Algorithms and Prediction by Modal Value

B.Jefferson[1*], R.S.Shanmugasundaram[2]

[1]Research Scholar, Vinayaka Mission's Research Foundation, Salem, Tamil Nadu, India.
[2]Professor, Vinayaka Mission's KirupanandaVariyar Engineering College,VMRF, Salem, Tamil Nadu, India.
*jeffson_72@yahoo.com

ABSTRACT

Magnetic Resonance Images(MRI) are essential means and have a vital role in diagnosing Braintumors. Even though Machine Learning algorithms give accuracy to some extent, still accurate prediction cannot be achieved in the medical field for the classification of Brain tumors. In this paper, Multiple algorithms are used to improve the prediction accuracy. 3D MRI images are used since they are able to offer more features than 2D images. Segmentation is carried out using watershed algorithm. The GLCM (Grey Level Co-occurrence Matrix) is used for feature extraction of the segmented region. In this paper modal values of three algorithms are used. In a dataset, a value that occurs number of times is termed as the mode or modal value. Three Machine Learning algorithms (SVM, KNN,CNN) are used and the Modal Value Prediction (MVP) is the Final Prediction achieved based on Modal value.

Keywords
Brain tumor MRI; 3D MRI; Segmentation; Watershed; Classification; Mode; GLCM.

Introduction

The brain tumor [1] is said to be an abnormal growth in the cells of brain region which is of two types: Cancerous tumor (Malignant) and non-cancerous tumor (Benign). Malignant tumors [2] may diffuse to other organs or may diffuse to brain cells from other parts of the body which can be termed as metastatic tumors.

Even though the MRI scan image supports in detecting tumors, there may be a chance of wrong prediction due to improper impressions of the image, noise and distortion. In the detection and prediction of brain tumors, accuracy plays a major role. This paper deals with the accurate prediction by considering the modal value among the prediction values from more than one Machine learning algorithms (SVM, KNN, CNN).

This following session in this paper describes the methods of the existing research of the same domain, detection and classification of brain tumors in section II, section-III describes the proposed method and architecture for the MVP prediction, and finally, section IV provides the results and analysis for the proposed method.

Related Works

Hiba Mzoughietal[3] introduced an architecture which merges the two features: local and global features with decreased weights. This architecture uses a pre-process method with normalization of intensity and ACE (Adaptive-Contrast-Enhancement) of image data to overcome the data heterogeneity. Data Augmentation is used to have a productive training in a deep 3D network.

Hassan Khotanlou[4] proposed a method which combines the contour-based and region-based paradigms. An enhanced method of segmentation was proposed, that relies on the approximate finding of symmetry plane. To have an initial detection of the tumor, two techniques are included. First method uses a fuzzy technique for classification applied to hyper-intense type of tumors and the next one represents a symmetric analysis applied to all type of tumor. The approximate symmetry plane is first calculated and then symmetry analysis is carried to find out the deviated regions.

Solmaz Abbasi[5], proposed a model to detect tumors in 3D MRI Brain images. In the pre-processing step, the histogram matching and the bias field correction are carried out. Then the ROI is located and separated. The HOG-

TOP ie, orientation gradient histogram and LBP-TOP ie, Local Binary Pattern in 3 planes(orthogonal) are considered as the features. The segmentation of the tumor is carried out using random forest algorithm.

Saif Dawood Salman Al-Shaikhli[6] approach combines features associated with texture and topological features to learn and create a dictionary. For multi-class tumor classification, Sparse coding along with dictionary learning are proposed that gives an effective result. Linear svm and sparse coding are used to classify the multi-class tumor.

Jefferson, Shanmugasundaram[7] proposed a Model in which the features are extracted from GLCM implementing pyradiomics of python. Also using 3d-CNN, from the final layer, the features are extracted and both the features are fed in to KNN for further prediction that gives more accuracy compared to other models.
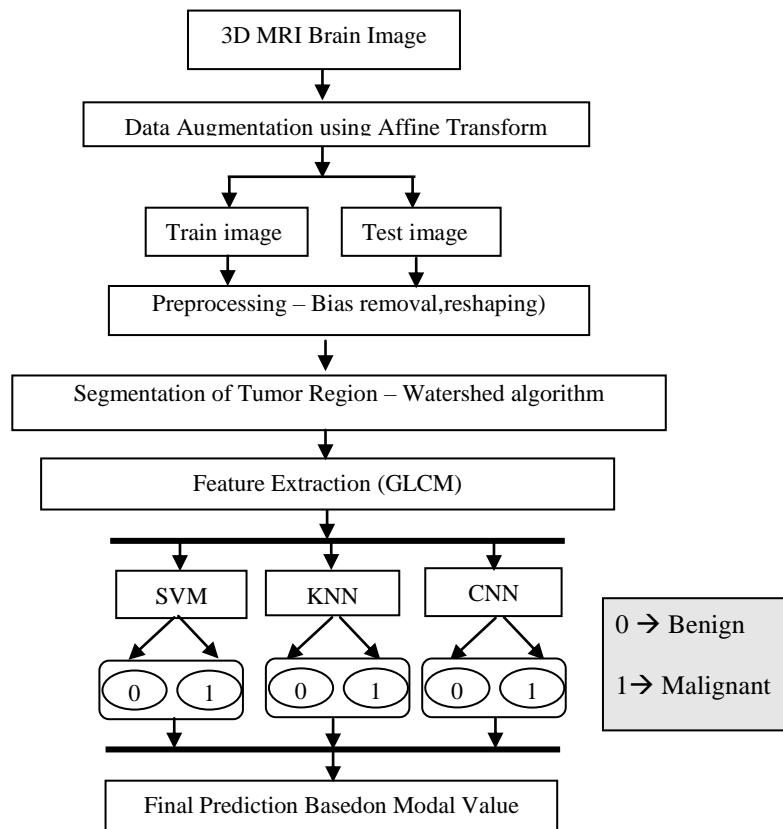
Proposed Method

Architecture of the Proposed Model

Fig-1 Model

Figure 1. Architecture Diagram

Data Set

The Dataset consists of 300 images of a 3D NIFTI format of dimension 155*240*240. After the augmentation using affine transformation, the size of the dataset is increased to 900. In which 465 images are benign and 435 images are

malignant. The dataset is divided in to Test dataset and Train dataset set. 70 percent for train set and 30 percent for test set.

Table 1. Dataset Splitting

Tumor Type	Training set	Testing set
Benign (465 images)	325	140
Malignant (435 images)	305	130
Total	630	270

Pre-Processing

In this phase, the three-dimensional images are made to fit to the Python Program. The 3D images are resized and the distortion caused by non-uniform intensity of the magnetic field during MRI scan is removed using N4ITK bias field removal algorithm. Noises are removed by the Median filter with 3x3x3 filters implemented in python. The input image dimension after resizing is 120*120*77.
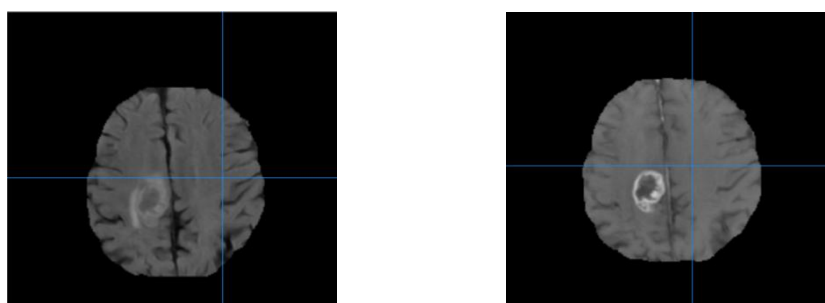

Figure 2. Sample Image after Pre-Processing

Segmentation

Segmentation is done by using one of the image segmentation techniques such as Watershed segmentation [8,9] algorithm. It is implemented in Python. SciPy, Scikit-image and OpenCV packages are used to segment the tumor region.
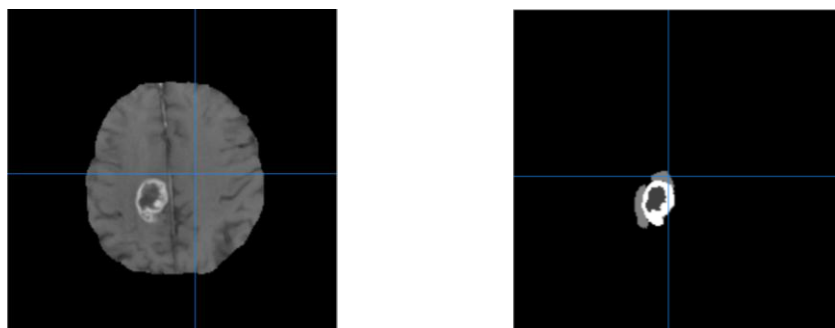

Figure 3. Sample Image after Segmentation

Feature Extraction

The features of the segmented image are extracted a feature extraction algorithm like grey level special dependence matrix method, which is commonly known as GLCM or Grey Level Co-occurrence Matrix. Pyradiomics of Python is used to get the features of GLCM.

1264

Classification and Prediction

With the help of the extracted features, the segmented tumor region is then classified as benign or malignant. This process is done by three unique machine learning algorithms, which are SVM, KNN and CNN.

KNN Classifier
KNN is a multiclass classifier which uses distance metrics. The distance of one test observation of test dataset from all the observations of the training dataset is calculated and the nearest neighbours of them are determined. This will be done for each and every test observation and thus finds the similarities in data. The distance is calculated using Euclidian distance as shown below:

$$EucDist(a,b) = \sqrt{\sum_{i=1}^{n}(ai - bi)^2}$$

KNN is implemented in python in which the extracted features are given as input. For that scikit-learn KneighborsClassifer package is used to implement the algorithm. For better accuracy, the K value (number of neighbours) is initialized to 1(K=1). The algorithm gives 241 correct predictions and 29 wrong predictions among 270 test datasets. The detail of Confusion Matrix (CM) is

Table 2.KNN-Confusion Matrix

Test Images = 270	Actual Benign	Actual Malignant
Predicted Benign	126 [True Benign]	15 [False Benign]
Predicted Malignant	14 [False Malignant]	115 [True Malignant]

$$Accuracy = \frac{TB + TM}{TB + TM + FB + FM} X\ 100 = 89.2$$

SVM Classifier
Since SVM [10] works as a binary classifier based on supervised learning, it is proven to perform classification with more accuracy. The classifier is supposed to classify 2 classes: benign or malignant. The RBF (Radial Basis Function) is used as a kernel. The lower-dimensional input space is transformed into a higher dimensional space by this kernel. The parameter gamma ranges from zero to one. (0 to 1). A perfect value for Gamma is assumed as 0.1, which will be the default value for gamma. The gamma value needs to be specified manually in the classification algorithm. SVM is implemented using class: sklearn.svm.SVC() in python and got an accuracy of 92.5 percent.

Table 3.SVM-Confusion Matrix

Test Images = 270	Actual Benign	Actual Malignant
Predicted Benign	130 [True Benign]	8 [False Benign]
Predicted Malignant	12 [False Malignant]	120 [True Malignant]

3DCNN Classifier
The performance of CNN increases according to the number of training data. The input image size is 120x120x77. The model consists of three convolution layers. The number of filters is in increasing order 8,16,32 for each layer. The kernel size is 3x3x3 which is fixed for all the filters. Maxpooling layer is applied after the first convolution layer and third convolution layer with maxpool_size 2x2x2. Relu activation is used in association with all the convolution

layers. All the data are flattened and send to the fully connected layer where softmax activation function is used to get the final prediction. KerasTensorflow of library of python is used to implement the model.
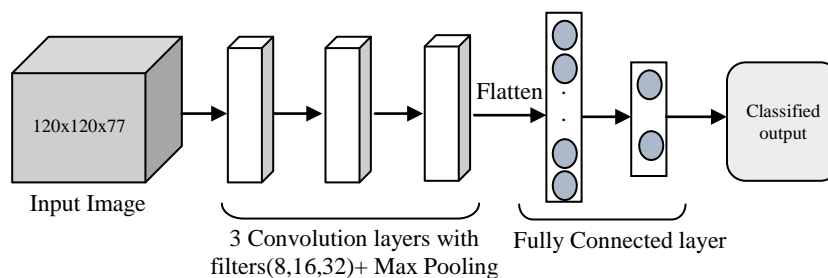

Figure 4. 3DCNN

Table 4.3DCNN-Confusion Matrix

Test Images = 270	Actual Benign	Actual Malignant
Predicted Benign	133 [True Benign]	5 [False Benign]
Predicted Malignant	7 [False Malignant]	125 [True Malignant]

The accuracy for the model is calculated as 95.5 Percentage.

For MVP (Model Value Prediction), the modal value (Mode) is calculated by considering the appearance of data value that occurs many times. The values predicted by the three classifiers are grouped and the value that occurs frequently will be the final prediction value which is the Modal Value Prediction (MVP). MVP is the value that most frequently occurs among the three predicted values by KNN, SVM and CNN. If two or more prediction is correct then the final MVP have a correct prediction.

Results and Discussion

The test dataset (270 images) is used to test the trained classifiers (KNN, SVM and CNN), and the prediction results are observed for each test image. Table-1 shows the prediction result of each test image for three classifiers.

Table 5. Final Prediction by MVP

Image	Values As per Labelled Image (Actual Result)	KNN Classifier	SVMClassifier	CNN Classifier	Final Prediction (MVP)
Brain_mri_01	1	1	1	1	1
Brain_mri_02	0	1	0	0	0
Brain_mri_03	0	0	1	0	0
Brain_mri_04	1	0	1	1	1
Brain_mri_05	0	0	0	0	0
Brain_mri_06	1	1	1	1	1
Brain_mri_07	0	0	0	0	0
Brain_mri_08	1	1	1	1	1

1266

Brain_mri_09	0	0	0	0	0
Brain_mri_10	0	0	0	0	0
..	..	..	..	..	..
..	..	..	..	..	..
Brain_mri_270	1	1	0	1	1
0 ->Benign, 1-> Malignant					

Among 270 test dataset, 29 images are wrongly predicted by KNN, 20 are wrongly predicted by SVM and 12 are predicted wrongly by CNN. So, the individual accuracy level is 89.2, 92.5 and 95.5 respectively. Final Prediction (MVP) is done as shown in table-5.

Table 6. Accuracy Analysis

Classifier	Correct Prediction	Wrong Prediction	Accuracy
KNN	241	29	89.2
SVM	250	20	92.5
3DCNN	258	12	95.5
Final Prediction by MVP	263	7	97.4

Among the three classification algorithms, if two or more predict wrongly then the final MVP will be wrong. In our, case KNN and SVM both predicted the same four cases wrongly and KNN and CNN both predicted the same three cases wrongly and totally MVP predict Seven cases wrongly (4 due to KNN and SVM and 3 due to KNN and CNN).
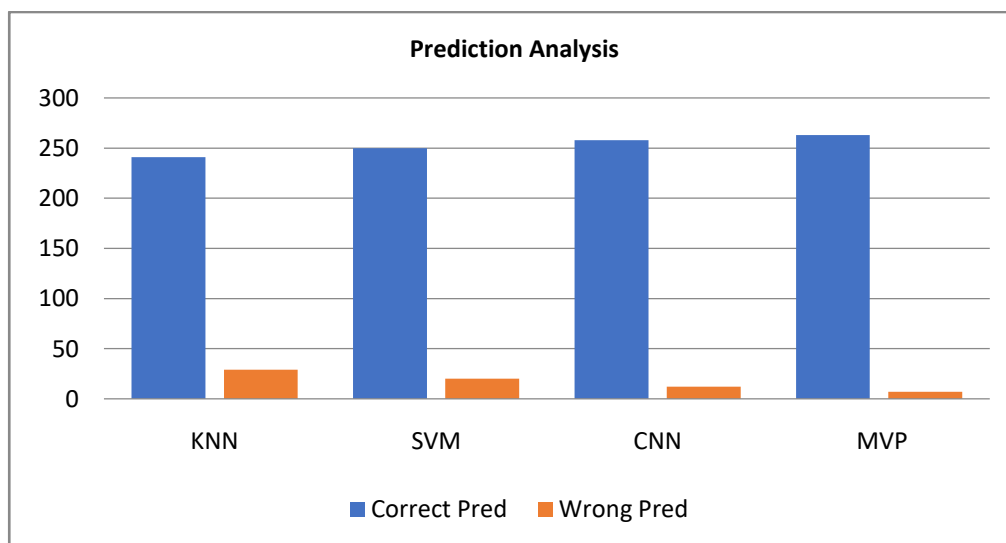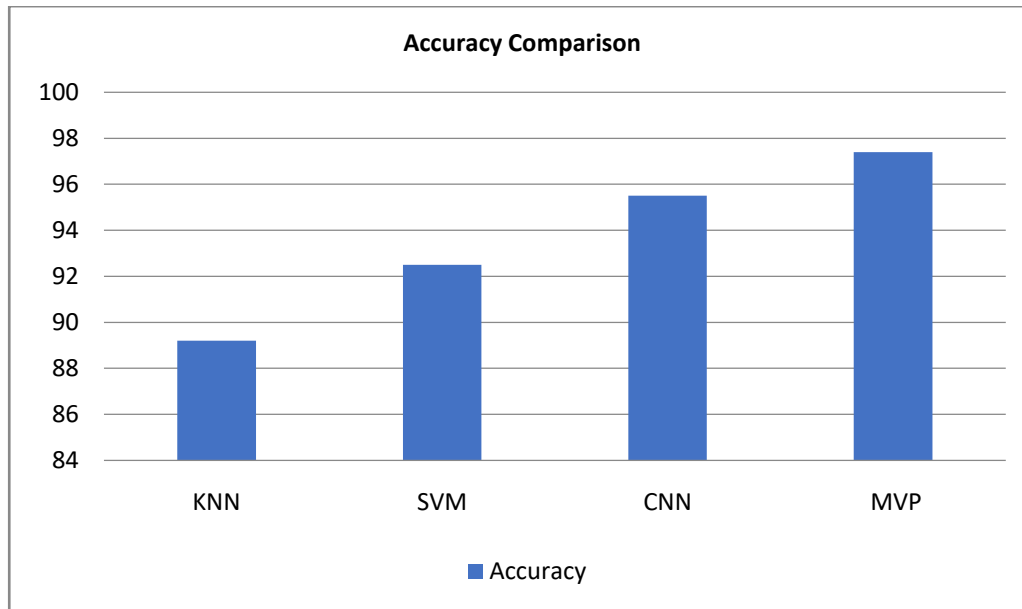

Figure 5. Prediction Analysis

Figure 6. Accuracy Comparison

Conclusion

Each Classification algorithms will have misclassification rate according to the performance of the algorithm and their associated algorithms. The misclassification rate may change between classification algorithms. The image misclassified by one algorithm may be correctly classified by other. Here more than two algorithms are used to classify the brain tumor and the results are analysed for 270 test images. The mode value of the predicted result (MVP) for each image is calculated and analysed. An improved accuracy of 97.4 percent is achieved as shown in fig-5. This can be extended by involving more odd number of classifiers and calculate MVP to get more precise and accurate results.

References

[1] Stefan Bauer, A survey of MRI-based medical image analysis for brain tumor studies, *Physics In Medicine and Biology*, http://dx.doi.org/10.1088/0031-9155/58/13/R97.

[2] B. Jefferson, R. S. Shanmugasundaram, Assessment on Brain Tumor Detection Techniques in Hyperintense MR Images, *IJRTE*, ISSN: 2277-3878, Volume-8 Issue-5, January 2020

[3] Hiba Mzoughi1,2 & Ines Njeh1, Deep Multi-Scale 3D Convolutional Neural Network (CNN) for MRI Gliomas Brain Tumor Classification, *Journal of Digital Imaging*, https://doi.org/10.1007/s10278-020-00347-9.

[4] Hassan Khotanlou, 3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models. *Fuzzy Sets and Systems, Elsevier- 2009, 160 (10), ff10.1016/j.fss.2008.11.016ff. ffhal-01251278f*

[5] Solmaz Abbasi, Detection of Brain Tumor in 3D-MRI Images using Local Binary Patterns and Histogram Orientation Gradient, *Neuro computing,* *https://doi.org/10.1016/j.neucom.2016.09.051*.

[6] Saif Dawood Salman Al-Shaikh, Brain tumor classification using sparse coding and dictionary learning, *2014-IEEE International Conference on Image Processing (ICIP)*, 978-1-4799-5751-4/14/$31.00 ©2014 IEEE

[7] B. Jefferson, R. S. Shanmugasundaram, Brain Tumor Classification in 3D-MRI using Features from Radiomics and 3D-CNN combined with KNN Classifier, *IJEET Vol.12, Issue 2, Feb 2021, DOI: 10.34218/IJEET.12.2.2021.017*

[8] J. Sijbers,P.Scheunders, Watershed-Based Segmentation Of 3D MR Data For Volume Quantization, *Magnetic Resonance Imaging*, Vol. 15, No. 6, pp. 679-688. 1997 Elsevier Science Inc.

[9] Anton S. Kornilov, An Overview of Watershed Algorithm Implementations in Open Source Libraries, J. Imaging 2018, 4, 123; doi:10.3390/jimaging4100123

[10]   R.Ahmmed, Classification of tumors and its stages in brain MRI using SVM and ANN,2017, ECCE, doi:10.1109/ECACE.2017.7912909.