NeuralAgent: Brain Tumor Detection and Classification

Akshat Savaliya, Rutuja Shah, Srijha Kalyan

May 7, 2022

Abstract

The objective of this project is to explore various supervised machine learning techniques to classify brain tumors from MRI images of brain scans. Multinomial Logistic Regression, Support vector machine and various Neural Network models (Convolutional Neural Network, and MobileNet). will be leveraged to classify the types of tumors from MRI images. The trained models will then be tested on unseen MRI images to evaluate which model(s) are more suited for the classification of brain tumors which will be followed by a comparative analysis of the models.

1 Introduction

A Brain tumor is considered one of the most aggressive diseases, among children and adults. Brain tumors account for 85 to 90 percent of all primary Central Nervous System(CNS) tumors. Brain Tumors are categorized as Benign Tumors, Malignant Tumors, Pituitary tumors, etc. Proper treatment, planning, and accurate diagnostics should be implemented to improve the life expectancy of the patients. Brain tumors are detected using Magnetic Resonance Imaging (MRI) after which a specialist classifies the image to be in one of the categories mentioned above. Image classification on MRI images is an important part of medical image processing. Being able to classify correctly and efficiently is crucial because a misdiagnosis can lead to serious consequences. It is known that the biggest disadvantages of manually interpreting MRI images is that it is invasive, time-consuming, and open to sampling errors. Thus, it would be very advantageous to perform a fully automated method for the multi-classification of brain tumors using various supervised machine learning techniques. These models increase the diagnostic abilities of clinicians and radiologists to shorten the time required for a correct diagnosis [1]. In this project, we implement, evaluate, and do a comparative analysis of brain tumor classification based on MRI Imaging using three supervised machine learning techniques: Multinomial Logistic regression, Support vector machines, and Neural Networks (Convolutional Neural Network, and MobileNet).

2 Related Work

Classifying brain tumors with the use of supervised machine learning techniques has been an ongoing research for many years and a variety of techniques for preprocessing and classifying the images have been explored in various research papers. For pre-processing, methods, namely wavelet transformation, filtering, extraction of Gray level co-occurrence matrix (GLCM) for texture features, and object labeling have been commonly used. For the segmentation of the tumor in the MRI images, unsupervised learning algorithms, such as fuzzy c-means, Otsu thresholding, and others have been widely used [2]. Support vector machines (SVM) were used for automatic segmentation of tumors from MRI images with the use of fast Fourier transforms to extract features. Classification of images followed by this technique produced a test accuracy close to 98%. [3].

For classification of brain tumors, Carlo Ricciardi, et al. [10] presented an approach for classifying pituitary adenomas tumor MRIs by using multinomial logistic regression and 92% on a k-nearest neighbor with an AUC curve of 98.4%. In a study conducted by Damodaran S et al [11], Convolutional Neural networks have been used for brain tumor classification and identification. The method used in this study gained 83% accuracy for segmentation .In another study [6], Muhammad Sajjad et al suggested the classification of multi-grade tumors by applying a data augmentation technique to MRI images and then tuning it using a pre-trained VGG-19 CNN model, an augmented technique to MRI images and then tuning it using a pre-trained VGG-19 CNN model. Khwaldeh, saed et al. [8] presented a framework for classification of brain MRI images into healthy and unhealthy, and a grading system for categorizing unhealthy brain images into low and high grades, by modifying the Alex-Net CNN model which revealed 91% accuracy. Arshia Rehman et al. [7] used three different pre-trained CNN models (VGG16, AlexNet, and GoogleNet) to classify the brain tumors into pituitary, glioma, and meningioma. During this Transfer learning approach, VGG16 acquires the highest accuracy which is 98.67%. Ahmet in et al. [2] modified the pre-trained ResNet50 CNN model by removing its last 5 layers and adding 8 new layers instead and comparing its accuracy with other pre-trained models such as GoogleNet, AlexNet, and ResNet50. The updated ResNet50 model showed effective results by achieving 97.2% accuracy.

3 Technical approach and methodology

3.1 Dataset

A public dataset brain tumor classification dataset [4] from Kaggle is used for this project which contains around 2200 images for training and 300 images for testing.Like for dataset: https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri Below is a description of a dataset in terms of the types of classes the dataset has and the number of images per class:

Types of Tumor	Training/Testing	No of Images
Meningioma Tumor	Training	822
No Tumor	Training	395
Pituitary Tumor	Training	827
Meningioma Tumor	Testing	115
No Tumor	Testing	105
Pituitary Tumor	Testing	74

Table 1: Dataset description.

3.2 Pre-processing

The training dataset has 3 classes of brain tumors having around 2200 images. The first step of pre-processing involves resizing the images and then cropping the part using contours to remove the unnecessary pixels from the boundary of the MRI brain Images. Next, we normalize the images so that they are all compared on the same scale. These images are of different contrast and to further process the images, normalized contrast is done on images. So images with low contrast are improved by increasing the contrast by relative proportions. Moreover, vertical flips and horizontal flops have been applied on the images to introduce randomization, and Gaussian Blur has been applied to remove noise in images. The purpose of pre-processing is to reduce the computation time and noise from Images so that we can better classify the images. These preprocessing steps are used for all models to maintain consistency.

3.3 Support vector machines and Logistic Regression

The Support vector machines and Logistic regression are supervised learning algorithms that have been widely used for classification tasks in multiple domains. The block diagram of the proposed method is shown in fig1. As shown in the figure, our approach consists of 3 major stages. 1) Use of DWT-SGLDM to extract features; 2) use K-fold stratified cross validation to prevent overfitting 3) Use SVM and Multi-class Logistic Regression to construct the classifier.

3.3.1 Feature Extraction

Before feeding the data to support vector machines and logistic regression models, it is important to perform feature extraction first. Feature extraction calculates features on the basis of which image can be easily classified as normal or abnormal one. Feature extraction is the process to represent raw

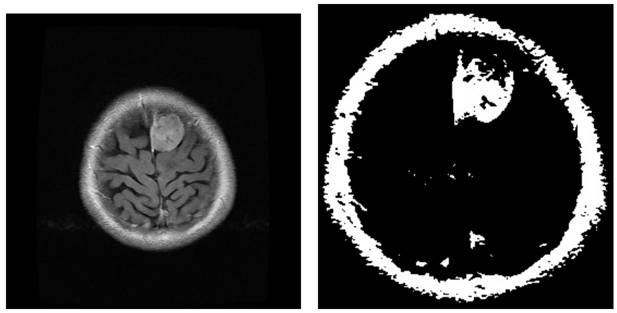


Figure 1: Result of K means (Meningioma Tumor)

images to facilitate decision making such as pattern classification. Features will be extracted from the tumor regions from MRI images. Feature extraction involves reducing the amount of data required to describe a large set of data accurately. Features are used as inputs to classifiers that assign them to the class that they represent. The intention of feature extraction is to reduce the original data by measuring positive properties, or features, that discriminate one input sample from another sample. Below are the steps that are taken to extract the features:

3.3.1.1 K-Means

K-means segmentation technique is unsupervised in nature. It segments a group of data points into k numbers of clusters. The preprocessed image is segmented using the K-means algorithm in order to take out the tumor from the MRI images. Segmentation of an image is the procedure to partition or divide the image into different regions or segments. The objective of this technique is to represent the image in a more significant and purposeful form that is easier to investigate.

3.3.1.2 Otsu Thresholding

Converting a grayscale image to monochrome is a common image processing task. Otsu's method, named after its inventor Nobuyuki Otsu, is one of many binarization algorithms, it uses a data-driven approach that can adaptively find the optimal threshold to distinguish two-class data, by going through all possible threshold values (from 0 to 255), it can find the optimal threshold value of input image. This method can be applied in image segmentation and image binarization, in the current scenario we have used it for image binarization. Otsu's thresholding is primarily based on the variance of the data- within class variance and between class variance.

3.3.1.3 Discrete Wavelet Transform

The first advantage of using Wavelet Transform (WT) is that it can preserve both the time and frequency information of the signal. Another advantage of WT is that it adopts "scale" instead of traditional "frequency" as it does not produce a time-frequency view but a time-scale view of the signal. The time-scale view is a more natural and powerful way to view data. We used Daubechies-2 for the efficient representation of smoothly changing signals.

These sub-bands are selected for further processing by the gray level dependence matrix (GLDM). By applying the GLDM, 11 features are computed including entropy (ENT), contrast (CON), correlation (CORR), inverse difference moment (IDM), variance (VAR), angular second moment (ASM), entropy (ENT), Homogeneity (HOMO), Dissimilarity (DTSS), Mean, Standard Variation(STD), Kurtosis and Skewness. The co-occurrence matrix allows us to compute these features.

3.3.1.4 Gray Level Co-occurrence Matrices (GLCM)

It is the process of collecting higher-level information about an image such as shape, texture, color, and contrast. In fact, texture analysis is an important parameter of human visual perception and machine learning systems. It is used effectively to improve the accuracy of the diagnosis system by selecting prominent features. It is applied to Sub-band images found from DWT to get features. This technique follows two steps for feature extraction from the medical images. In the first step, the GLCM is computed, and in the other step, the texture features based on the GLCM are calculated. Due to the intricate structure of diversified tissues such as WM, GM, and CSF in the brain MR images, extraction of relevant features is an essential task. Textural findings and analysis could improve the diagnosis, different stages of the tumor (tumor staging), and therapy response assessment. The statistics feature formula for some of the useful features is listed below.

Mean (M): The mean of an image is calculated by adding all the pixel values of an image divided by the total number of pixels in an image:

$$M = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) \tag{1}$$

Entropy (E). Entropy is calculated to characterize the randomness of the textural image and is defined as:

$$E = -\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) log f(x, y) \tag{2}$$

Contrast (C). Contrast is a measure of the intensity of a pixel and its neighbor over the image, and it is defined as:

$$C = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x - y)^2 f(x, y) \tag{3}$$

Other features like entropy (ENT), contrast (CON), correlation (CORR), inverse difference moment (IDM), variance (VAR), angular second moment (ASM), entropy (ENT), Homogeneity (HOMO), Dissimilarity (DTSS), Mean, Standard Variation(STD), Kurtosis and Skewness have been calculated. These features are then used to classify the MRI images using SVM and Multiclass Logistic Regression.

3.3.2 Hyperparameter Tuning: Stratified K-fold Cross Validation

One of the problems that occur during the classifier training is overfitting, where the error on the training set is driven to a very small value, but when new data is presented to the network the error is large. Therefore, the first advantage of using Wavelet Transform (WT) is that it can preserve both the time and frequency information of the signal. Another advantage of WT is that it adopts "scale" instead of traditional "frequency" as it does not produce a time-frequency view but a time-scale view of the signal. The time-scale view is a more natural and powerful way to view data. We used Daubechies-2 for the efficient representation of smoothly changing signals. These sub-bands are selected for further processing by the gray level co-occurrence matrix (GLCM). By applying the GLCM, 11 features are computed including entropy (ENT), contrast (CON), correlation (CORR), inverse difference moment (IDM), variance (VAR), angular second moment (ASM), entropy (ENT), Homogeneity (HOMO), Dissimilarity (DTSS), Mean, Standard Variation(STD), Kurtosis and Skewness. The co-occurrence matrix allows us to compute these features. validation is employed to avoid over-fitting. The mechanism is to create a K-fold partition of the whole dataset, repeat K times to use K-folds for training and a left fold for validation, and finally average the error rates of K experiments.

3.3.3 Pipeline

To summarize the process; To extract the features of the MRI images, the images were first cropped to remove extra unnecessary pixels from the border and pre-processed using methods described in the pre-processing section. Then, k means clustering was applied so that we can separate the tumor from the brain. Otsu thresholding was then applied to those tumor-separated images. After that DWT was applied to those output images and a sub-band was found. GLSM was then applied on those sub-bands to find out the features mentioned above which cover the texture and other information about the image. Once the feature extraction is completed, the extracted features are fed into the SVM and logistic regression models as inputs.

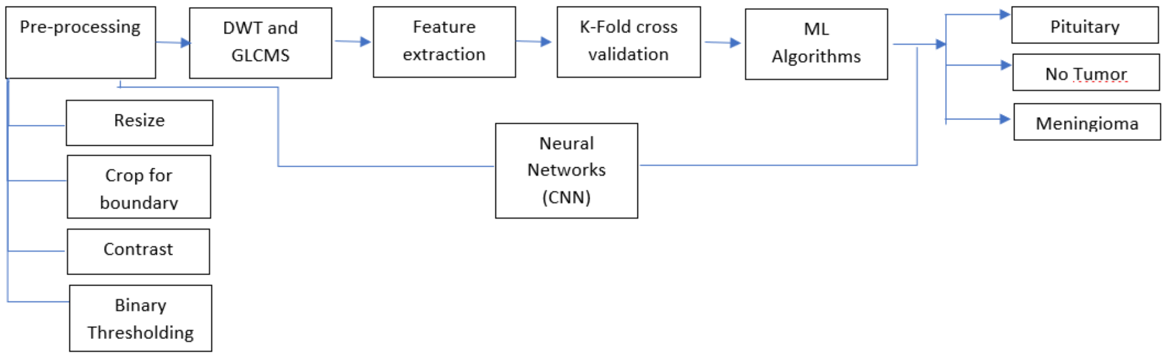


Figure 2: Pipeline of Feature Extraction of Brain MRI Images.

3.4 Neural Networks (CNN)

Deep learning models belong to a class of computing machines that are able to learn the hierarchy of features by building high-level attributes from low-level ones. The purpose of this is to automate the process of feature construction. One of the well-known deep learning models is a convolutional neural network (CNN). CNN comprises several (deep) layers of processing involving learnable operators (both linear and nonlinear) and hence, has the ability to learn and build high-level information from low-level features [2]. CNN's have been demonstrated to be effective for image classification in various domains, one of which is the classification of MRI images [3][4][5][6].

3.4.1 CNN Architecture

The CNN model implemented has 4 convolutional layers, each with a progressively increasing number of kernels 32, 64, 128, and 256 for each layer. The size of the kernel in each layer is 3x3 with stride 1 and padding 1 used. The activation function used is Relu and is applied after each layer (besides the last). We have also applied Max pooling after each layer (besides the last) of size 2x2 which reduces the width and height of the output image by a factor of 2. Furthermore, 3 fully connected components (FC) have been implemented. The flattened output of size 25x25x256 is passed to the first FC layer which then outputs 1024 features. The second layer takes 1024 features as input and outputs 64 features. And finally, the last FC layer takes 64 features as input and outputs 3 classes. Learning rate of 0.001, batch size of 64 for train and 32 for the test, and Adam optimizer have been used. The loss function used is the Cross Entropy Loss since it's an effective loss function for multi-class classification. And finally, the number of epochs used is 15, each of which outputs the train and test accuracy for that epoch.

3.4.2 Pre-trained Transfer Learning Models

3.5.1.1 Transfer Learning

In computer vision, transfer learning is usually expressed through the use of pre-trained models. A pre-trained model is a model that was trained on a large benchmark dataset to solve a problem similar to the one that we want to solve. Accordingly, due to the computational cost of training such models, it is common practice to import and use models from published literature (e.g. VGG, Inception, MobileNet).

3.5.1.2 MobileNet

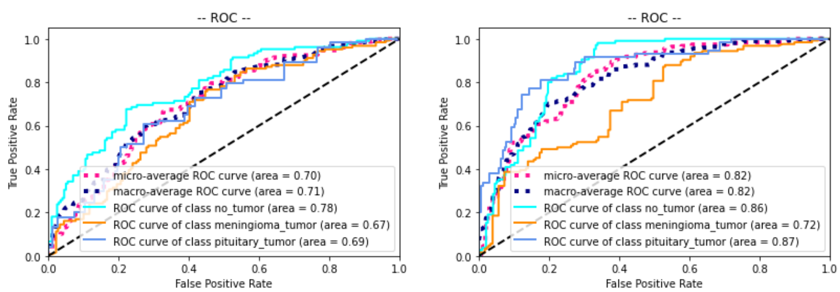
MobileNet is a type of pretrained convolutional neural network designed for mobile and embedded vision applications. They are based on a streamlined architecture that uses depthwise separable convolutions to build lightweight deep neural networks that can have low latency for mobile and embedded devices. MobileNet is a streamlined architecture that uses depthwise separable convolutions to construct lightweight deep convolutional neural networks and provides an efficient model for mobile and embedded vision applications [15]. The structure of MobileNet is based on depthwise separable filters. Depthwise separable convolution filters are composed of depthwise convolution filters and point convolution filters. The depthwise convolution filter performs a single convolution on each input channel, and the point convolution filter combines the output of depthwise convolution linearly with 1x1 convolutions. Learning rate of 0.001, batch size of 64 for train and 32 for the test, and Adam optimizer have been used. The loss function used is the Cross Entropy Loss since it's an effective loss function for multi-class classification. And finally, the number of epochs used is 15, each of which outputs the train and test accuracy for that epoch.

4 Results

4.1 Traditional Machine Learning Classifiers

For SVM and Logistic Regression, we used features calculated from DWT and GLCM to represent the MRI image shape, texture and contrast properties. Feature matrix for MRI images was generated and stored for further classification. Apart from that we use k-fold cross validation to find the best performing model on our dataset.

- In SVM, we used RBF, Linear and Polynomial kernels and found that Linear kernels are working better than others. RBF kernel is almost as efficient as linear kernel with 67% accuracy.
- Another traditional machine learning algorithm we used was Logistic Regression. We applied cross validation in order to find the best performing model for LR. We got 54.5% training accuracy and 51.3% testing accuracy. There is a large scope of improving this accuracy which is discussed in the future work section.



4.2 CNN and MobileNet

- The convolutional Neural network was able to achieve an 82.3 percentage test accuracy. Figure 3 below shows how the training loss decreases as the number of Epochs increases and figure 3 shows how the training and test accuracy increases as the number of epochs increases - which is exactly the type of output that was expected of the CNN model.
- The pretrained Mobilenet model is the state-of-art model for the Brain tumor image classification by providing an accuracy of 98% for train data and 96% for test data. The loss and accuracy graph is shown figure 4.

5 Conclusion and Future Work

We can see that classifiers based on Neural Networks(NN) are outperforming SVM and Logistic Regression with a large margin in accuracy. But we do have a scope for improving the feature extraction part. To efficiently get the features from the tumor region we need to extract out the tumor region with minimum noise like brain tissues and skull. We have used KMeans, Otsu thresholding and contour

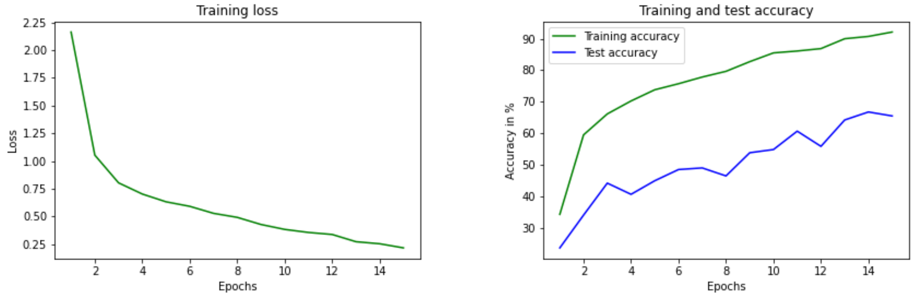


Figure 3: CNN Results

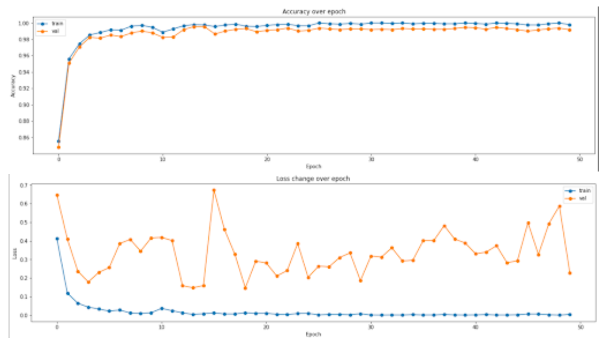


Figure 4: Mobile-net accuracy and loss graphs

Figure 5: Results from the Mobilenet prediction

Table 2: Results of Brain Tumor Classification

Models	Meningioma	No Tumor	Pituitary	Final Accuracy
SVM	61.78	63	77.8	68
Logistic Regression	86	44	13.5	51.15
CNN	93	47.29	95.23	82.3
MobileNet	97	92	98	96

Table 3: Performance Metrics

Models	Precision	Recall	F1-score
SVM	0.67	0.68	0.68
Logistic Regression	0.59	0.48	0.51
CNN	0.82	0.80	0.83
MobileNet	0.96	0.95	0.95

removal to grab the tumor part of the MRI image. There are other ways like using fuzzy c-mean and other thresholding algorithms, also Deep Neural networks, to help improve the efficiency of separating tumor regions (ROI). So there is a scope of exploring these methods apart from the one we used and see if it increases the accuracy or not. Apart from that we can increase the number of features extracted including informative features that can capture tumor shape and texture.

We see that using pretrained models helps to significantly improve the accuracy and the performance of classifying types of tumors from brain MRI images to 96% The CNN model was able to get a classification accuracy of 82.3%. Although this is an improvement to the accuracy of the SVM and logistic regression models, there is still scope for improvement for the accuracy of the CNN model. One thing that can be improved is the pre-processing of images. Techniques that amplify the region of the tumor could be applied such that the tumor is highlighted, leading to a better feature extraction in the convolutional layers and hence, a better overall accuracy.

6 GitHub Link

https://github.com/AkshatSavaliya98/ML_Brain_Tumor_Anaylsis_Akshat_Rutuja_Srijha

7 References

[1] V.P.Gladis Pushpa Rathi1 and Dr.S.Palani. Brain Tumor MRI Image classification with feature selection and extraction using linear discriminant analysis.

[2] Mamta Mittal, Lalit Mohan Goyal, Sumit Kaur, Iqbaldeep Kaur, Amit Verma, D. Jude Hemanth, Deep learning based enhanced tumor segmentation approach for MR brain images, Applied Soft Computing, Volume 78, 2019, Pages 346-354, ISSN 1568-4946.

[3] Xiaohong W. Gao, Rui Hui, Zengmin Tian, Classification of CT brain images based on deep learning networks, Computer Methods and Programs in Biomedicine, Volume 138, 2017, Pages 49-56, ISSN 0169-2607.

[4] Justin S. Paul, Andrew J. Plassard, Bennett A. Landman, Daniel Fabbri, "Deep learning for brain tumor classification," Proc. SPIE 10137, Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging, 1013710 (13 March 2017)

[5] Hassan Ali Khan, Wu Jue, Muhammad Mushtaq, Muhammad Umer Mushtaq. Brain tumor classification in MRI image using convolutional neural network[J]. Mathematical Biosciences and Engineering, 2020, 17(5): 6203-6216. doi: 10.3934/mbe.2020328

[6] El Boustani, A., Aatila, M., El Bachari, E., El Oirrak, A. (2020). MRI Brain Images Classification Using Convolutional Neural Networks. In: Ezziyyani, M. (eds) Advanced Intelligent Systems for Sustainable Development (AI2SD'2019). AI2SD 2019. Advances in Intelligent Systems and Computing, vol 1105. Springer, Cham. https://doi.org/10.1007/978-3-030-36674-$2_3$2

[7] Conference Proceedings 2296, 020023 (2020); https://doi.org/10.1063/5.0030978 Published Online: 16 November 2020

[8] A. Kharrat, M. B. Halima and M. Ben Ayed, "MRI brain tumor classification using Support Vector Machines and meta-heuristic method," 2015 15th International Conference on Intelligent Systems Design and Applications (ISDA), 2015, pp. 446-451, doi: 10.1109/ISDA.2015.7489271.

[9] 32. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," IEEE Transactions on Systems, Man and Cybernetics, vol. 3, no. 6, pp. 610–621, 1973

[10] Damodharan S., Raghavan D.Combining tissue segmentation and neural network for brain