Brain Tumor Detection & Classification Using Machine Learning

[1]Rintu Joseph
Department of Computer Science & Engineering
Mangalam College of Engineering,
Ettumanoor, India
Rintu712@gmail.com

[2]Mr. Sanoj C Chacko
Department of Computer Science
Mangalam College of Engineering,
Ettumanoor, India
sanoj.chacko@mangalam.in

ABSTRACT

The phrase "brain tumor" simply refers to the growth of aberrant brain tissues. The quickest and most common method of locating a brain tumor is through Magnetic Resonance Imaging (MRI) scans. A brain tumor may be lethal if it is not detected in the early stages of growth. As a result, it is essential to spot and diagnose tumors while looking at brain imaging. The challenge of identifying tumors from MRI scans is one that takes a lot of effort, money, and time. This study suggests a deep learning-based approach that may identify brain tumors from MRI images considerably more quickly and accurately, enabling doctors to treat patients right away. Additionally, radiologists can make decisions regarding the best therapeutic approaches to use more quickly. The trained model will also be used to predict the presence of brain tumors, automating the process, and saving time and labor.

Keywords: Brain Tumor; Machine Learning; MRI Images; Convolutional Neural Network.

I. INTRODUCTION

Brain tumor detection using machine learning is a relatively new field of study that is gaining more attention as time goes on. This type of technology is being used to detect, diagnose, and treat various types of tumors in the brain. A type of artificial intelligence called machine learning analyses data to forecast outcomes and make judgements. The use of diverse algorithms and approaches to recognize and categorize various tumor forms is necessary for the use of machine learning to the identification of brain tumors.

This type of technology can provide more accurate results and a better understanding of the underlying causes of brain tumors. Additionally, it can help doctors and researchers better understand the progression of the disease.

Supervised and unsupervised learning are two forms of machine learning algorithms utilized for brain tumor identification. Supervised learning systems use tagged data, such as medical images, to identify and classify different types of cancer. On the other hand, unsupervised learning algorithms employ unlabeled data to find patterns and trends in the data. Convolutional neural networks (CNNs) are one of the most used supervised learning techniques for finding brain tumors. This type of algorithm is used to detect tumors by analyzing medical images. CNNs are trained using labeled data, such as MRI scans, to identify and classify different types of tumors. These methods can also be used to track the evolution of tumor size and morphology over time.

Unsupervised learning algorithms can also be used for brain tumor detection. These algorithms are used to identify patterns and trends in the data without the use of labels. Common unsupervised learning algorithms used in brain tumor detection include clustering algorithms, such as k-means, and anomaly detection algorithms. Clustering algorithms are used to group similar tumors together, while anomaly detection algorithms are used to identify outliers or anomalies in the data.

Deep learning can be utilized for brain tumor identification in addition to supervised and unsupervised learning. These algorithms are designed to find patterns and trends in data that may be difficult to find using conventional machine learning methods. Because doing so tracks the progression of the disease, deep learning algorithms can be used to discover and classify different tumor types.

Overall, machine learning has become an important tool for brain tumor detection. This type of technology can provide more accurate diagnosis and treatment of tumors, as a better understanding of the underlying causes of brain tumors. Additionally, machine learning algorithms can be used to detect changes in tumor size and shape over time. With the help of machine learning, doctors and researchers can make better decisions when it comes to diagnosing and treating brain tumors.

II. RELATED WORKS

The Fuzzy segmentation technique (FCM) was used in [1] to distinguish between brain areas with tumors and those without them. A multilevel discrete wavelet transform (DWT) was also used to extract wavelet characteristics. Deep neural networks (DNNs) were included in order to accurately classify brain tumors. Comparative minimum optimization (SMO), linear discriminant analysis (LDA), and KNN classifier techniques were used to evaluate this strategy. In the DNN-based brain tumor classification analysis, the accuracy rate was 96.97 percent. However, the performance was quite subpar, and the intricacy was really high.

For a step-by-step examination of patient tumor progression, a brand-new biomechanical model of tumor growth was given in [2]. It will be used to capture a major tumor impact in gliomas and distinct fringed solid tumors. To simulate tumor development, discrete and continuous techniques were merged. The suggested technique offers the potential for implicit segmentation of tumor-bearing brain pictures from atlas-based registries. Brain tissue was mostly segmented using this method. However, the computation took a long time.

In paper [3], a novel multi-feature feature (Multi FD) was utilized, and the AdaBoost classification system for the identification and segmentation of brain tumors was enhanced. Using the Multi FD feature extraction approach, the structures of brain tumor tissue were extracted. The categorization of the donated brain tissue as tumor or non-tumor was done using advanced AdaBoost.

The classification of brain voxels was detailed in Paper [4] using a Local Independent Projection (LIPC)-based classifier. The path function was also retrieved using this technique.

In [5] compares the histogram-based segmentation approach with a novel strategy for segmenting granular tumors that uses the Cellular Automata (CA) technology. For effective segmentation of brain tumors, seed selection and volume of interest (VOI) calculations were made. Incorporating tumor section segmentation was another aspect of this effort. As a result, complexity was lower, but accuracy was also lower.

ADVANTAGES

- MRI imaging can identify brain tumors.
- Because there is no human involvement, human mistake is eliminated.
- By detecting the tumor sooner, human lives can be spared.
- Intelligent artificial systems are more dependable.

DISADVANTAGES

- The model has high system requirements for proper operation.
- The dataset training process takes a long time.
- Very accurate but not entirely true.

III. MATERIALS AND METHODS

DATASET

The dataset was taken from figshare, where it had more than enough information. The main data set of images with this dataset contains MRI scans with types of tumors. The data was collected from 233 patients both male and female, aged between 18 and 96. Each scan includes T1-weighted contrast-enhanced images. The scans are labeled with 4 types of tumors such as meningioma, glioma, pituitary tumor, and images with no tumor. These can be used for research on developing tumor detection and classification algorithms.

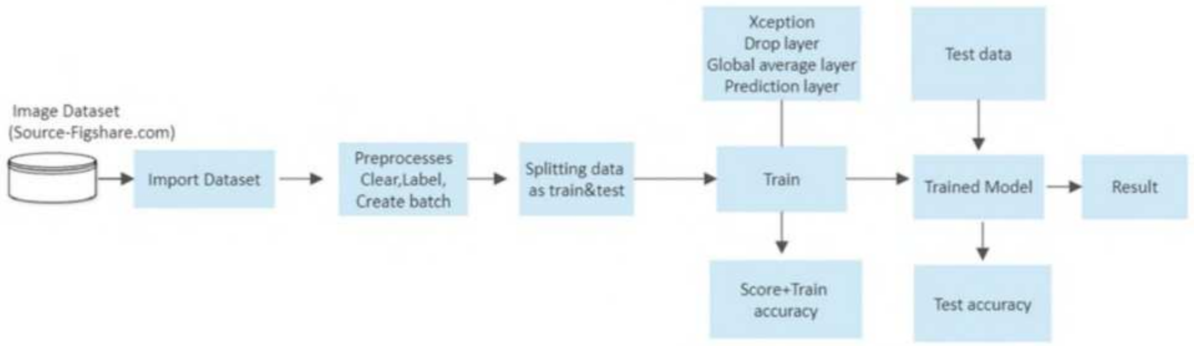
Validation set – It is the set of images that will be used during training for adjusting the parameters.

Testing set – It is the set of images that will not be involved until checking the final performance of the model.

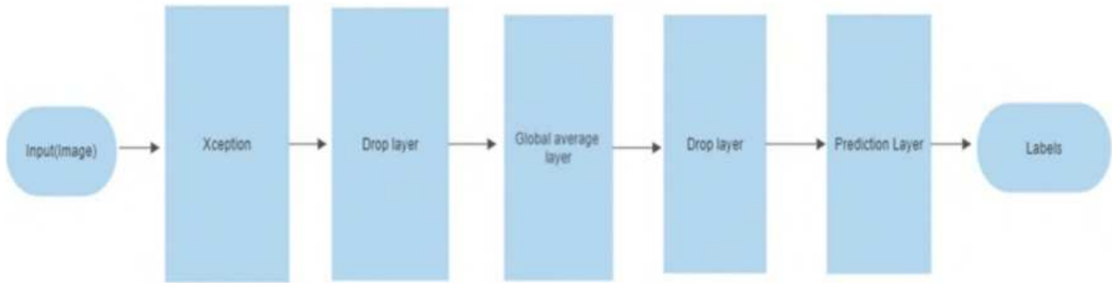
IV. EXPERIMENTAL PROCEDURE

As an initial step, the dataset was taken from figshare and stored in the drive. It gives information about different types of brain tumors. Preprocessing is an important step in any data analysis process. It involves cleaning and transforming data, so that it can be used in a machine learning algorithm or other type of analysis. This can include reshaping the image, creating the batches. In the next step, images will be labeled accordingly. Labeling is the process of classifying data into different categories by assigning it with a label. Labeling can be used to help identify the type of tumor in an image, which can then be used to make predictions or decisions. Labeling is an important part of preprocessing, as it helps to ensure that the data is organized and structured in a way that makes it easier to analyze. Further the images will be divided into batches. A batch in a CNN is a set of images that are used as input to the neural network. The batch size is the number of images that are passed to the network at a time. The larger the batch size, the more accurate the network's predictions can be, however, larger batch sizes can cause the network to take longer to train. The next step is to divide the data into train and test groups. This technique divides the data into two parts: a training set and a test or validation set. The test set is employed to assess the model's performance once it has been built using the training set. In this way, this can make sure that the model is not overfitting or underfitting the data. The advantage of CNN is that it will be able to extract the features automatically. The image contains the features such as edges, corners, etc. These features are then used to create models that can accurately detect tumors and provide insight into their characteristics. In the training model, there are multiple layers. First layer is Xception. Convolutional neural network architecture called Xception was created by Google. The model has been pre-trained on the ImageNet dataset and can be used for transfer learning in any computer vision task. Along with Xception there are few more layers such as Drop layer.

V. WORKFLOW CHART



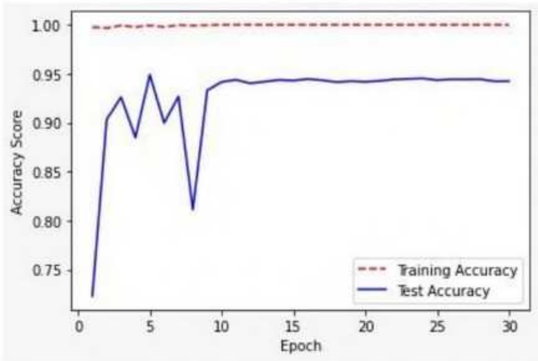
VI. STRUCTURE OF MODEL



The global average layer is a type of convolutional neural network (CNN) layer that computes the mean of all feature maps in the input volume. This layer is usually used as a regularization technique, as it forces the network to learn more general features from the input data and reduces overfitting. The prediction layer is the last layer in a neural network model. It is responsible for making predictions based on the input data. It is used to calculate the probability of a given set of inputs and outputs. The model will be trained with the 30 epochs and the accuracy will be analyzed based on the validation dataset. The model will predict the type of tumor based on the input image. Based on the input image the model returns the numpy array with the set of prediction results. The type of class with the maximum value is considered as the predicted class.

VII. Results and Discussions

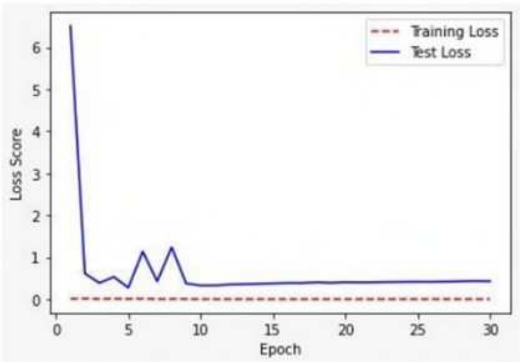
The results of using convolutional neural networks (CNNs) for brain tumor detection have been promising. The use of CNNs has been shown to be effective for classifying brain tumors into different types. CNN was used to accurately classify these different types of brain tumors with an accuracy of up to 100% only during training period and test accuracy is up to 95%. These results suggest that CNNs can be a useful tool for automated brain tumor detection and classification. Further research is needed to increase the accuracy will exist all the time. Additionally, more research is needed to determine the best approach for using CNNs for brain tumor detection and classification. The number of times a whole dataset is submitted to and received by the neural network is called an epoch.



Accuracy Score

Here following inferences made from below accuracy graph.

- Training accuracy was maintained 95- 100% throughout.
- Validation accuracy increased with the number of epochs but there was a fluctuation between 4-10 epochs.
- As the epochs crosses 10 the validation accuracy becomes stable resulting with 94-95% accuracy.
- The fluctuation might be because of noise in the data, the type of model being used, the number of parameters, the type of optimization algorithm being used, the learning rate, and the amount of regularization etc.



Loss Score

The above graph displays the loss score during the training of a model.

- Loss is a number that symbolizes the total of the model's errors.
- The attrition score with the retraining set of photos stays very close to 0 to 0.5%.
- With the validation data the model's loss score fluctuates between the epochs 4-12 and thereafter it fluctuates with very less variations.
- The loss score oscillates less as the number of epochs rises

VIII. CONCLUSION

The model is based on the machine learning algorithm CNN (Convolutional Neural Network). It helps to predict just by reducing and resizing the image without losing any important information that will be used for predicting. The created model achieves an accuracy of 97.79% when applied to the training set and an accuracy of 82.86% when applied to the validation set. The loss gradually starts decreasing with the increase in the number of epochs. The model loss is very less when applied to the training set whereas it is high when applied to the validation set. In future, different datasets would be applied to this model, to further increase the overall accuracy.

REFERENCES

[1] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. ElHorbaty and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," Future Computing and Informatics Journal, pp. 68-71, 2018.

[2] S. Bauer, C. May, D. Dionysiou, G. Stamatakos, P. Buchler and M. Reyes, "Multiscale modeling for Image Analysis of Brain Tumor Studies," IEEE Transactions on Biomedical Engineering, vol. 59, no. 1, pp. 25-29, 2012.

[3] Islam, S. M. Reza and K. M. Iftekharuddin, "Multifractal texture estimation for detection and segmentation of brain tumors," IEEE Transactions on Biomedical Engineering, pp. 3204-3215, 2013.

[4] M. Huang, W. Yang, Y. Wu, J. Jiang, W. Chen and Q. Feng, "Brain tumor segmentation based on local independent projection-based classification," IEEE Transactions on Biomedical Engineering, pp. 2633-2645, 2014.

[5] Hamamci, N. Kucuk, K. Karaman, K. Engin and G. Unal, "Tumor-cut: Segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications," IEEE Transactions on Medical Imaging, pp. 790-804, 2012.

[6] H. B. Menze, A. Jakab, S. Bauer, K. Farahani and K. Justin, "The Multimodal Brain Tumor Image Segmentation Benchmark," IEEE Transactions on Medical Imaging, pp. 1993-2024, 2014.

[7] Jin Liu, Min Li, Jianxin Wang, Fangxiang Wu, Tianming Liu and Yi Pan, "A survey of MRIbased brain tumor segmentation methods," Tsinghua Science and Technology, pp. 578-595, 2014