Brain Tumour Detection

N Akhilesh Kumar Dutt, Siri Chandan Sai K, Neehal and MD Aman Taiyab

Department of Computer Science And Engineering B.M.S. College of Engineering Bengaluru, Karnataka, India

Mr. Basavaraj Jakkali Department of Computer Science And Engineering Visveswaraya Technological University, Belgaum Bengaluru, Karnataka, India

Abstract: Brain tumors are one of the most difficult diseases to diagnose and treat. Early detection is critical to improving patient's outcomes and reducing the risk of dementia. Various techniques are used to identify brain tumors, such as CT scans, PET scans, and MRIs. However, these processes are time consuming, expensive, and often devastating. With recent advances in machine learning, there has been increased interest in developing brain tumor detection systems that can provide more accurate, faster, and non-invasive diagnosis. In this paper, we review the latest state-ofthe-art in brain tumor detection using machine learning techniques, including deep learning, and support vector machines. We also discuss the challenges and limitations of these methods, such as data imbalance, interpretability, and generalization, and suggest potential solutions. Finally, we highlight some future directions for brain tumor detection research, such as multi-modal fusion, explainable AI, and personalized medicine.

Index Terms: Image Processing, Brain tumor, MRI, Segmentation, CNN, Keras

I. INTRODUCTION

Brain tumors are abnormal growths of cells in the brain that can be either cancerous or noncancerous. According to the American Brain Tumor Association, brain tumors account for approximately 24% of all cancers in children and 2% of all cancers in adults. Brain tumors can cause a range of symptoms, including headaches, seizures, memory loss, and difficulty with movement and coordination. Brain tumors are a significant health concern worldwide, with an estimated incidence of 11-13 per 100,00 population per year. Brain tumors are also the leading cause of cancer-related deaths in children and the second leading cause in adults under the age of 39. The diagnosis of brain tumors is challenging due to their diverse nature and location, and the symptoms they present can be non-specific.

Traditional methods for diagnosing brain tumors include imaging techniques such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET). However, these methods are often time-consuming, expensive, and require invasive procedures. Moreover, the interpretation of these imaging studies requires the expertise of a trained radiologist or neurosurgeon, which can lead to variability and potential diagnostic errors. There are better treatments that are used to reduce the growth of the tumor cells in the brain or to destroy the brain tumors such as Radiation therapy(uses high energy beams to destroy tumor cells or prevent growth),Chemotherapy(uses drugs to destroy tumor cells or to prevent growth),Targeted therapy(uses drugs that specifically target certain molecules or genetic mutations to reduce the growth of tumor cells),Immunotherapy(uses body's immune system to kill tumor cells in the brain).

Recent advancement in artificial intelligence (AI) and machine learning (ML) have paved the way for the development of automated brain tumor detection systems that can analyze medical images and provide accurate, fast, and non-invasive diagnosis. In this paper, we review the current state-of-the-art in brain tumor detection using machine learning techniques and discuss the challenges and limitations of these methods. Ultimately, the development of accurate and reliable brain tumor detection methods can lead to better patient outcomes and improved quality of life for those affected by this devastating disease.

The aim of this paper is to build an application that would help in detecting whether there are tumors in the input MRI images through the convolution neural network.



Figure 1: MRI image

II. LITERATURE SURVEY

Tonmoy Hossain et al.[1] proposed a model for segmentation and detection of brain tumor using classifiers like Support Vector Machine(SVM), K-nearest neighbour(KNN),Logistic Regression, Naive bayes and Random forest. Fuzzy C-Means algorithm is used for the image segmentation.

This model segments the Region of interest(ROI) and predicts the tumor region accurately. Among all classifiers used, SVM provides the best accuracy of 92.42%. They have proposed the second model i.e. brain tumor detection using CNN and it is implemented using keras and tensorflow. The BRATS dataset is used in both the models which consists of two classes i.e. class-0 and class-1 which represents Non-tumor and tumor MRI images. 152 training images and 68 testing images are used to achieve an accuracy of 92.82%. They have splitted the dataset to a ratio of 80:20 i.e. 174 training images and 43 testing images are used to achieve an accuracy of 97.87%. The Proposed model predicts whether the image is an tumorous or Non-tumorous MRI image.

Suresha D et al. [2] uses the combination of K-Means Clustering algorithm and support vector machine(SVM)method for detecting the brain tumors from a given input MRI image. K-Means Clustering algorithm is used to cluster the pixels with similar properties to distinguish the tumor and the non-effected part of the brain and is used to find whether tumor is detected or not in the MRI input image of the brain. if tumor is detected, they have used SVM Classifier to classify whether the tumor is benign or malignant tumor. They have used less training set of input MRI image images to obtain an accuracy of 96.9%.

Digvijay Reddy et al. [3] proposed a model for brain tumor detection using K-Means Clustering algorithm, Thresholding method and morphological operations. The similar intensity regions in the input MRI images are clustered using K-Means algorithm and the mean of each cluster is calculated and the process is done until there is change in the mean. They have used K-Means algorithm to differentiate Skull and the brain region. Thresholding is done to distinguish the Tumor and non-tumor portions in the input MRI image. The Skull in the input MRI brain image is removed by using Morphological operations. The three different kaggle datasets are taken to evaluate the accuracies and the accuracies obtained are 99.9054%,99.9146% and 99.9100%.

Angona Biswas et al. [4] proposes a model which utilizes K-Means clustering algorithm and ANN for three different types of tumor classication.K-Means Clustering algorithm is a segmentation process which is used to separate the object and background from the input image.They have used Artificial neural network structure for the classification of tumors(i.e. Giloma, Meningioma, Pituitary tumors).Neural network took 70% sample for training,15% sample for validation and 15% sample for testing. For this model the data was collected from the FigShare website and dataset consists of total 563 patients brain MRI images(246 giloma,74 meningioma and 243 pitutary brain tumor MRI images).For the Proposed model they have obtained an accuracy of 95.4%,sensitivity of 94.58% and specificity of 97.83%.

Suhib Irsheidat et al.,[5] the author constructed a model based on artificial convolution neural network that detects the tumor present in brain. This model uses CNN architecture and data augmentation which help in creating 14 new images for each individual image. The Kaggle dataset is used in which they are classified into two categories label 1 (having tumor) and label 2(do not have tumor) with a total of 253 images, which are divided into 85% for training and 15% for testing. The model predicts brain tumor with an accuracy of 88.8% on test data and 96.7% on evaluation data. Sunil Kumar et al.,[6] the author has used CNN model for brain tumor detection. This model used CNN architecture and used an unique image segmentation architecture called Watershed Algorithm which help in dividing combine or touching images. The Dataset is a combination of two datasets Br35H and Central research UK, the dataset contains 3060 images in they are classified into two categories Yes (1530 images), No (1530 images). The given validation accuracy by this model is 92 %.

Khizar Abbas et al.,[7] the author has proposed LIPC based methodology for brain tumor classification and segmentation. This model used CNN architecture and features like PCA(Principal Component Analysis) which helps to visualize data in high dimension so the classification is easy for the model. The dataset used is MICCAI 2013 the dataset contains 30 images in which they are classified into two categories High-Grade tumor(20 images) and Low-Grade Tumor(10) images. The ratio of training and testing is 70:30 respectively. The achieved accuracy by the model is 0.95%(95%).

Masoumeh Siar et al.,[8] the author has used CNN model for brain tumor detection. This model used CNN architecture with three different classifiers to obtain the best accuracy between these three which are Softmax classifier, Radial Basis Function(RBF) classifier and Decision tree classifier. The dataset contains 153 patients MRI images in which they are classified into two categories normal(80) and patient(73),they have collected a total of 1892 images in which 1666 were used for training and 226 images were used for testing. The achieved accuracy's of softmax classifier is 98.67%, Radial Basis Function classifier is 97.34% and Decision tree classifier is 94.24. The author has achieved the best accuracy of 99.12% by combing CNN- Softmax and feature extraction algorithm.

Hajar Cherguif et al.,[9] proposed a new architecture based on U-Net to solve the problem of brain tumor segmentation. It is a encoder-decoder based network architecture (consists of down sampling and up-sampling path). Architecture divided in to 3 parts- Contracting Part, The bottleneck part and the expansion part. Unlike the conventional U-Net their architecture uses 3 convolution layers per block to achieve more efficient brain tumor segmentation. Their method provided more efficient and robust segmentation compared to other approaches and showed maximum dice similarity coefficient of 0.81 for the dataset.

Chirodip Lodh Choudhury et al.,[10] proposed a model consisting of 3-layered CNN Architecture, which is later connected to Fully Connected Neural Networks. They proposed a new system based on CNN, which separate between the Brain MRI images to mark them as tumorous or not. The model is having CNN with 3 layers and requires very few steps of pre-processing to produce the results. The purpose of their research is to highlight the importance of diagnostic machine learning applications and predictive treatment. The model achieved the accuracy of 96.08%, with f-score of 97.3

Zheshu Jia et al.,[11] proposed a Fully Automatic Heterogeneous Segmentation using Support Vector Machine (FAHS-SVM) for brain tumor identification and segmentation. To identify tumor regions by combining intrinsic image structure hierarchy and statistical classification information. The proposed method can achieve promising tumor segmentation. Their experimental results indicate that the method proposed will help to identify the exact location of the brain tumor accurately and quickly. The proposed method is, therefore, critical for MR imagery brain tumor detection. The experimental results showed 98.51% of the accuracy of the proposed technology.

Aryan Sagar Methil [12] proposed a novel method involving image processing techniques for image manipulation which would aid our CNN model to classify tumor and non-tumor images better. Image Processing techniques helped us solve the illumination issues and brought the tumor into focus. Data augmentation was used to reduce the chances of overfitting, as it artificially expands the size of a training dataset, thus bringing out an improvement in the performance and the ability of the model to generalize. The system recorded an adequate accuracy of 97.94% with an excellent training recall of 98.55 % and validation recall of 99.73%.

Rasel Ahmmed et al. [13] proposed a model which mainly focuses on brain tumor detection using a robust integration method which is the integration of template based K-Means and modified Fuzzy C-Means(TKFCM) clustering algorithm that reduces operators and equipment error. Here, K-Means algorithm is to emphasized initial segmentation through the proper selection of template. Modified Fuzzy C Means technique is used to minimize the intensity homogeneities. This Modified Fuzzy C Means algorithm produces more homogeneous regions than Normal FCM algorithm and it is relatively less sensitive to noise. The Dataset of 30 brain tumor images is considered. The dataset consists of complex brain tumor images which is very difficult for the common man to identify the tumor so easily. The average sensitivity, specificity and accuracy obtained are 96.67%,100% and 97.1%.



Figure 2 : System architecture of brain tumour detection

Dr A Jagan [14] proposed a model for the detection of brain tumor by using Proposed Augmentation method. The aim of this method is to improve the segmentation accuracy, sensitivity and specifity. They have used anisotropic filter to develop the brain MR image edges for improved segmentation results. The Proposed method has been executed on ten different datasets of Fluid-Attenuated Inversion Recovery(FLAIR) 3D Brain MRI images. The Performance of the Proposed model have also been compared to EM method and Fuzzy C Means Clustering proposed method.The average accuracy of the model(97.59%) is greater than the EM method(95.60%) and Fuzzy C Means Clustering method(96.44%). The average sensitivity of the proposed model(91.99%) is greater than the EM method(91.63%) and Fuzzy C Means Clustering method(90.91%).

The average specifity of the proposed model(97.58%) is also greater than EM method(95.22%) and Fuzzy C Means Clustering method(96.79%).

Sneha Grampurohit et al.,[15] the author used deep learning model's like CNN(convolution neural network) and VGG-16 for detecting the tumor in the brain through MRI images, Also used techniques like rotation in data augmentation which helps in generating images(Image Data Generator). They have taken 253 patients MRI images(2065) in which they are splitted into 3 sets Training(1445), Testing(310) and Validation(310). The total epochs are 25 from which the time consumed by CNN architecture is 5:03 mins and by VGG-16 architecture is 15:25 mins. The model achieved an accuracy of 0.9336(93.36%) for CNN architecture and in case of VGG-16 architecture the achieved accuracy is 0.9716(97.16%) where it takes more time but gets the best accuracy.

Sakshi Ahuja et al.,[16] the author used VGG-19 transfer learning-based model for brain tumor dection. The dataset is used is BraTS 2019 (challenge database) it consists of 42 patients MRI images. They divided the process into two stages, the first stage is detection phase where the model classify the MRI image into three categories: Normal, LGG and HGG and in this stage they have also used FLAIR modality which is used to analysis the parts of the brain and highlights only the tumor. In the second stage it is Segmentation phase where it provides the details like size, grading and type of tumor. The achieved accuracy by VGG-19 transfer learning-based model is 99.3%.

III. PROPOSED METHOD

a) MRI input images

The first step of this proposed system is to collect MRI images of the brain which consists of MRI images with tumors and without tumors. The whole set of all these images is considered as a Dataset. There are two MRI datasets of brain in the kaggle which were uploaded by Navoneel Chakrabarty and Ahmed Hamada. The final dataset is the combination of the above two datasets which contains each 800 tumor and normal images.



Figure 3 : Brain MRI dataset (a) Normal (b) Tumor

b) Pre processing

The main aim of this Pre-processing step is to perform different operations on the input MRI images and prepare all these images for further processing. This Pre-processing step is very much important before the Segmentation and Feature extraction steps. The dataset contains images with different dimensions. In Pre-processing step all images will be resized to a specific target size of (150,150) using The 'tensorflow.keras.preprocessing.image.module' to maintain the same dimensions for all the images. The resized images

50

100 150 200

50

can be converted into an array of numerical values using 'image.img_to_array' method.

c) Data Visualization

The Random images from the dataset which are Pre processed can be displayed to get a visual understanding of the MRI images in the dataset. The Pre processed image can be selected by using the 'matplotlib.image' module and visualized it using 'matplotlib.pyplot'.

d) Image Segmentation

The main aim of this Image segmentation step is to divide the image into multiple segments. The whole concentration will be on one particular region and that particular region will be separated from the remaining

background region. In the input MRI tumor images the main concentration will be on the tumor region so that particular region will be separated from the remaining part of the image as a part of segmentation process. Thresholding is one of the simple step for segmentation process. There are different types of thresholding techniques such as simple thresholding, Otsu's thresholding(global thresholding technique),Adaptive thresholding(local thresholding technique).

The thresholding technique used is a simple intensity-based thresholding method. This technique converts the image to grayscale using the function rgb2gray() from the skimage library and it applies a series of binary thresholding operations by setting pixel values above a certain threshold to 1 and values below the threshold to 0. The threshold values range from 0.0 to 0.7, with an increment of 0.1. This process is performed to generate multiple segmented images with different threshold levels, allowing for visual comparison and analysis of the segmentation results.



Figure 4 : (a) Input tumor image (b) Steps for generating segmented tumor region using simple intensity thresholding technique



e) Feature Extraction

In the feature extraction process the images are opened, resized, and appended to a numpy array. This occurs in the load_data() function where each image is read using the OpenCV library and then converted into RGB format, resized, and appended to a list of images. These images are then converted to a numpy array, along with their corresponding labels, and returned by the function.

f) Classification

The CNN algorithm is used for Brain tumor classification because the CNN's(Convolutional Neural Network) are mainly useful for Image classification tasks due to its advantages such as Local feature extraction, Translation invariance, Hierarchical feature learning, parameter sharing. The CNN algorithm consists of the following layers such as convolutional layers, Max Pooling layers, Flatten layers, Fully Connected layers

The first convolutional layer has 32 filters, a kernel size of (3, 3), and a ReLU activation function. The second convolutional layer also has 32 filters and a kernel size of (3, 3) with ReLU activation function. These layers are mainly used to extract features from the input images. The main aim of the Max Pooling layer is to reduce the spatial dimensions of the features and the first max pooling layer has a pool size of (2, 2) whereas The second max pooling layer also has a pool size of (2, 2). The flatten layer is useful to reshape the output of the previous layers into a 1-dimensional vector to prepare for the fully connected layers. The first fully connected layer consists of 128 neurons with a ReLU activation function and the second fully connected layer is the output layer with 2 neurons (corresponding to the two classes: tumor and normal) and a softmax activation function is used to produce class probabilities.

The dataset is divided into training and testing data. The training data contains 1600 images(800 tumor and normal images) and the testing data contains 320 images(160 tumor and normal images). The model is compiled using the Adam optimizer(adapts the learning rate during training) and trained by using 1600 training images and labels evaluated using the test data, and predictions are made on new images to classify them as either tumor or normal. The training process is repeated for a total of 10 epochs. The training process and model architecture is shown below.

Model: "sequential"		
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 148, 148, 32)	896
max_pooling2d (MaxPooling2D)	(None, 74, 74, 32)	0
conv2d_1 (Conv2D)	(None, 72, 72, 32)	9248
max_pooling2d_1 (MaxPooling 2D)	(None, 36, 36, 32)	0
flatten (Flatten)	(None, 41472)	0
dense (Dense)	(None, 128)	5308544
dense_1 (Dense)	(None, 2)	258
Total params: 5,318,946		

Trainable params: 5,318,946 Non-trainable params: 0

Figure 5 : Model Architecture

Epoch 1/10 10/10 [] - 245 2s/step - loss: 1.7404 - accuracy: 0.5625 - val_loss: 0.7981 - val_accuracy: 0. 588	
Epoch 2/10 10/10 [
Epoch 3/18 10/10 [====================================	
Epoch 4/10 10/10 [====================================	
5000 Boylo [
5425 Boch 6/10 10/10[==================================	
940/ Booch 7/10 18/10 [====================================	
Gpch 8/10 18/10 [==========] - 24s 2s/step - loss: 0.0557 - accuracy: 0.9867 - val_loss: 0.1020 - val_accuracy: 0. 9680	
Epoch 9/10 10/10 [====================================	
Epoch 10/10 10/10 [========] - 22s 2s/step - loss: 0.0162 - accuracy: 0.0904 - val_loss: 0.1137 - val_accuracy: 0. 9594	

Figure 6 : Training Process for 10 epochs

IV. RESULTS AND DISCUSSIONS

Figure 7 : Tumor and Normal image results



The Objective of the proposed system is to detect whether there is tumor or not in the input MRI image. There are two different datasets uploaded by Navoneel Chakrabarty and Ahmed Hamada. The final dataset is the combination of both the above datasets, So there are 1600 images (800- Tumor, 800- Normal) and the final dataset is divided into train and testing data based on the test ratio 0.2. The training data consists of 1600 images (800- Tumor, 800- Normal) and the testing data consists of 320 images (160- Tumor, 160- Normal).



Figure 8 : Bar Graph for Testing and Training images

After completing the training process, the variation in the results of training accuracy vs validation accuracy and the training loss vs validation loss is shown in the below graph



Figure 9 : (a) Train Accuracy vs Validation Accuracy (b) Train Loss vs Validation Loss



Figure 10: Selecting an MRI image for detecting tumor



Figure 12 : Confusion matrix nal

V. CONCLUSION AND FUTURE SCOPE

This paper presents a deep learning method for detecting brain tumours. Early detection of cancer can help with timely and effective treatment. This model used CNN algorithm for better detection of tumors. A simple intensity based thresholding technique is used to highlight the tumor region and separate the remaining part of the image. We obtained an accuracy of 99.06% by using CNN algorithm.

In terms of future scope implementing more advanced image processing techniques, incorporating data augmentation to increase the size and diversity of the dataset, and deploying the model to a web application where the patient can upload their MRI image, the application will predict whether the tumor is benign or malignant.

REFERENCE

1. T. Hossain, F. S. Shishir, M. Ashraf, M. A. Al Nasim and F. Muhammad Shah, "Brain Tumor Detection Using Convolutional Neural Network," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934561.

2. D. Suresha, N. Jagadisha, H. S. Shrisha and K. S. Kaushik, "Detection of Brain Tumor Using Image Processing," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 844-848, doi: 10.1109/ICCMC48092.2020.ICCMC-000156.

3. D. Reddy, Dheeraj, Kiran, V. Bhavana and H. K. Krishnappa, "Brain Tumor Detection Using Image Segmentation Techniques," 2018 International Conference on Communication and Signal Processing (ICCSP), 2018, pp. 0018-0022, doi: 10.1109/ICCSP.2018.8524235

4. A. Biswas and M. S. Islam, "Brain Tumor Types Classification using K-means Clustering and ANN Approach," 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), 2021, pp. 654-658, doi: 10.1109/ICREST51555.2021.9331115

5. S. Irsheidat and R. Duwairi, "Brain Tumor Detection Using Artificial Convolutional Neural Networks," 2020 11th International Conference on Information and Communication Systems (ICICS), 2020, pp. 197-203, doi: 10.1109/ICICS49469.2020.239522.

6. S. Kumar, R. Dhir and N. Chaurasia, "Brain Tumor Detection Analysis Using CNN: A Review," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS),2021,pp. 1061-1067, doi: 10.1109/ICAIS50930.2021.9395920.

7. K. Abbas, P. W. Khan, K. T. Ahmed and W. -C. Song, "Automatic Brain Tumor Detection in Medical Imaging using Machine Learning," 2019 International Conference on Information and Communication Technology Convergence (ICTC), 2019, pp. 531-536, doi: 10.1109/ICTC46691.2019.8939748.

8. M. Siar and M. Teshnehlab, "Brain Tumor Detection Using Deep Neural Network and Machine Learning Algorithm," 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), 2019, pp. 363-368, doi: 10.1109/ICCKE48569.2019.8964846.

FOR

- M. Siar and M. Teshnehlab, "Brain Tumor Detection UsingDeep Neural Network and Machine Learning Algorithm," 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), 2019, pp. 363-368, doi: 10.1109/ICCKE48569.2019.89648
- C. L. Choudhury, C. Mahanty, R. Kumar and B. K. Mishra, "Brain Tumor Detection and Classification Using Convolutional Neural Network and Deep Neural Network," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), 2020, pp. 1-4, doi: 10.1109/ICCSEA49143.2020.9132874.
- Z. Jia and D. Chen, "Brain Tumor Identification and Classification of MRI images using deep learning techniques," in IEEE Access, doi: 10.1109/ACCESS.2020.3016319.
- A. S. Methil, "Brain Tumor Detection using Deep Learning and Image Processing," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 100-108, doi: 10.1109/ICAIS50930.2021.9395823.
- 13. R. Ahmmed and M. F. Hossain, "Tumor detection in brain MRI image using template based K-means and Fuzzy C-means clustering algorithm," 2016 International Conference on Computer Communication and Informatics (ICCCI), 2016, pp. 1-6, doi: 10.1109/ICCCI.2016.7479972.
- A. Jagan, "A New Approach for Segmentation and Detection of Brain Tumor in 3D Brain MR Imaging," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 1230-1235, doi: 10.1109/ICECA.2018.8474874.
- S. Grampurohit, V. Shalavadi, V. R. Dhotargavi, M. Kudari and S. Jolad, "Brain Tumor Detection Using Deep Learning Models," 2020 IEEE India Council International Subsections Conference (INDISCON), 2020, pp. 129-134, doi: 10.1109/INDISCON50162.2020.00037.
- 16. S. Ahuja, B. K. Panigrahi and T. Gandhi, "Transfer Learning Based Brain Tumor Detection and Segmentation using Superpixel Technique," 2020 International Conference on Contemporary Computing and Applications (IC3A), 2020, pp. 244- 249, doi: 10.1109/IC3A48958.2020.233306

OPEN ACCESS JOURNAL