Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

Brain Tumour Detection Using Machine Learning

Ms. Vishaka Rani Chandramule[1], Bhagyalaxmi S R[2], Chaithanya B D[3], Diya P S[4]

[1]*Assistant Professor, Dept Of CSE, ACS College Of Engineering, Bengaluru, India*

[2] *Dept of CSE Student, ACS College Of Engineering, Bengaluru, India*

[3]*Dept of CSE Student, ACS College Of Engineering, Bengaluru, India*

[4] *Dept of CSE Student, ACS College of Engineering, Bengaluru, India*

Abstract-Diagnosing Brain Tumors Earlier - The sooner someone with cancer receives proper treatment, the more likely they are to live longer than if they waited until later to start receiving treatment. For many years, MRI has been the method of choice for detecting brain tumours with images previously scanned many times over by hundreds of thousands of doctors, clinics, and medical administrators who continue to use these same images to identify brain tumours. This creates a tremendous workload of hours (for each scan) to compare abnormal MRI scans to normal MRI scans within all the scanners currently in use, leading to the possibility of estimating the number of times these physicians have incorrectly identified scanned images due to fatigue from the repetitive motion of having to review the hundreds of thousands of scanned images for abnormal and normal results. In addition, this process has resulted in a substantial number of misidentification and clinical error associated with using MRI imaging data as it has been reviewed/evaluated manually by physicians; therefore, the development of computer algorithms and the application of "machine intelligence" as a means of assisting physicians with reading MRI's is becoming increasingly prevalent in both clinical and research environments as opposed to being completely reliant on a physician's interpretation of an image or grouping of images, now includes an algorithm assessment of thousands of previous images (abnormal & normal) scanned in order to provide statistically valid and statistically significant "probability" results for current scanning occurrences.

Key Words*:*MRI Scans, Tumors, Radiologists, Precision, Recall

1.INTRODUCTION

Brain tumors and spinal tumors are among the most difficult health issues to deal with. Some cells continue to grow too rapidly and form either large tumors or small tumors or benign tumors. The location of where the tumor is located, how big it is and what type of cell it is made up of determines the risk any individual will have. Circumstances relating to the individual will aid in treating patients and estimating the chances of recovery from a brain tumour, highlighting the need for an early diagnosis in order to execute an early intervention and limit the severity or disease progression. Doctors also diagnose brain tumours based on the use of advanced technology such as CT or MRI scans and using biopsy samples of the tumours to accurately identify the brain tumour type. CT and MRI scans will produce a high-quality image of the brain; however, many doctors prefer MRI scans due to their ability to produce the highest resolution image of the brain, differentiating between healthy and unhealthy cells.

CT scans offer advantages concerning the time it takes for a scan as CT scans produce an image of the brain faster than an MRI scan, whereas a tissue sample will be examined the same way to show what type of brain tumor it is but requires specialized equipment and skilled technician who will take several hours to analyze the sample and conclude.

1.1 Problem Statement

The vast majority of MRI scans are performed in a hospital setting, but most physician's records of the results by the radiologist are rarely accurate because physicians disagree when they look at the data, and small changes in the brain are usually missed almost every time because of the physician's excessive amount of paperwork and too many patients to administer medications and laboratory services for the same day. Therefore, the increasing demand for the services provided by the physician will hinder his performance and reduce the efficiency of the physician's resource use in the office.

1.2 Goal Of the project

The goal of this project is to develop and deploy an automated robotic device that can evaluate MRI brain scans using advanced machine learning techniques to automatically identify brain tumors.

This project will have a number of phases, including acquiring MRI scans, testing the validity of automated detection, and developing a computer model to identify tumors from MRI scans.

2. RELATED WORK

There has been a significant transformation in the way that conventional methods of machine learning have been applied to develop brain tumor detection models employing MRI imaging from the inception of the project until 2021. Feature extraction had traditionally been performed manually by three primary means: Texture analysis; Histogram intensity; and Statistical descriptors. Manual methods for extracting features have produced average to low classification results consistently (for example SVM, k-NN, etc.); however, because of the significant variability of tumor characteristics and likely variation of MRI scanner parameters from differing MRI scanners, the average and low classifications obtained from classification pipelines based on manually measured features have frequently been lower than thought achievable.

Researchers have been working on new models of U-Net (U-Net++, Attention U-Nets) since 2022 in order to achieve greater accuracy when creating segmentation maps from medical imaging data. Part of the motivation behind developing this type of model is to be able to match the edges of a tumour more accurately than current methods, which should result in fewer false-positive diagnoses. Many of the researchers currently working in this field will begin investigating methods for providing answers regarding how and why segmentations were made using Explainable AI methods (such as Grad-CAM), as many of the results produced using these three different imaging modalities show similar quantitative results. Using these techniques will allow the researcher to understand which areas of tumour tissue are responsible for the prediction of the tumour being identified or not, thus providing additional insight into the accuracy of the tumour segmentations produced from the image data of these različnih MRI modalities (T1, T2, FLAIR).

In addition to using these explainable AI techniques, many of the researchers are now also examining the feasibility of using Multiple MRI Imaging Modalities (T1, T2, FLAIR) to augment their diagnostic striking support from automated systems and validate that their findings based on objective clinical evaluation methods will provide clinical-grade confidence in the automated systems they develop. Demonstrations and evaluations of real-world applications across various institutions will begin in 2024 and continue through 2025. In practice, deep learning solutions can be implemented using very low-resource computing and lightweight architecture to produce "light weight" solutions for deep learning applications.

3.OVERVIEW OF THE PROPOSED PROJECT

In this article, we will present a new way to analyze how the anatomy of the brain, as imaged by Magnetic Resonance Imaging (MRI) technologies, can be combined with state-of-the-art artificial neural network (ANN) technologies to generate new sets of features to analyze an MRI dataset semantically. We believe that such an analysis could improve the ability to detect small or otherwise obscure abnormalities in MRI scans that are normally missed by traditional imaging techniques.

The system will categorize and sort multiple types of unique images by their subtype (e.g., gliomas, meningiomas, pituitary and other non-brain tumors) and standard brain tissue. It will also reduce the time needed to review all images/tumor subtypes as well as increase the accuracy associated with evaluating individual images by using performance metrics generated during operation of these functions. All images must first go through specific initial processing steps that enable neural networks to operate on the images before evaluation (e.g., remove noise, resize images, balance pixels), so that ultimately, all images produced are uniform in their formatting which will assist neural networks to have been trained using the training methodology outlined above.
The performance of the system will be evaluated based on quantitative metrics generated from collecting observations made in real time (e.g., true positives, precision, recall, f1, etc.) compared to using hypothetical models

4.SYSTEM ARCHITECTURE AND METHODOLOGY

This project is founded on an AI-based method that will automatically analyse MRI images to detect brain tumours. The AI system is going to have a five-step structure where there are five separate steps that depend on each other will be undertaken. These steps will be comprised of: obtaining MRI brain images (from either publicly available MRI databases or clinical facilities), preparing the obtained MRI brain images so that they can be analysed; extracting features from the prepared MRI brain images; categorising (classifying) the MRI images as to whether or not there is a brain tumour); and visualising (showing) the results of the classification/categorisation.

The first step of the classification phase is to obtain the MRI brain images and place them into labelled datasets for supervised learning purposes (e.g. Are there brain tumours in the MRI image?). The images obtained for classification will be obtained from either publicly available databases of MRI images or from clinical facilities.

To create a standardised set of diffusion data, Parsons has established a number of different facilities for pre-processing any of the above images prior to being established in an all-diffusion dataset.

By pre-processing (through the use of image standardisation) images will correct the differences in sizes, reduce noise and normalise the distribution of sizes in all of the images; this should result in a higher degree of contrast in all of the diffuse parent images and throughout the classifying process will provide a clearer understanding of the appearance of the diffuse parent images.

Feature extraction is separated into two classes: A) Manually constructed models designed under the assumptions from traditional machine-learning rules (texture features, histogram features and statistical features). B) Produced features through the automatic aggregation of features generated by a deep neural network (CNN) built specifically for use with this application. Our methodology uses a CNN to read those hierarchical types of features via MRI scans of aggressive cancer. By recognizing patterns, and anomalies of the structure of the tumor and also the intensity difference between all the parts of the volume of the tumor. This reduces the need to manually do all of the feature extraction from the MRI before performing analysis and/or classification with those features.

After extraction has occurred the classifier shall be used to classify the image as tumourous or non-tumourous and whether or not the image quality can be classified. The way in which the tumour classifier uses a different extraction algorithm means that different extraction algorithms are going to be utilized to classify whether or not the image contains a tumour and how (or if) accurately the extraction algorithms classify the tumour in the patient image. Classification will utilize the following extraction methods such as an Annex neural net (a feed forward neural network) or Support Vector Machine (SVM). The data set to train or fit the tumour classifier as well as to evaluate the classifier will be the set of data which has been labelled with both positive and negative classification markers. Using backpropagation the best fit classifier will have been determined. The three metrics (accuracy, precision, and recall) will be determined through the use of a confusion matrix to compare the output of the classifier against the original lesion data set.

In summary, the system can distinguish different types of diagnoses as well as make explicit predictions regarding patient prognosis. The system will also provide visual tools (e.g., heatmaps; activation maps) to aid in the interpretation of the results and improve the ability of health care professionals to interpret those results.

The entire processing chain of software is integrated into one standardised way of analysing MRI brain images by means of re-inventing the systematic methods of processing MRI brain imaging data, which includes enhancing the condition of the image, localising the brain tumour, learning what the tumour looks like, and classifying the features of the tumour, so that the user has a single platform for detecting the pathology of brain tumours.

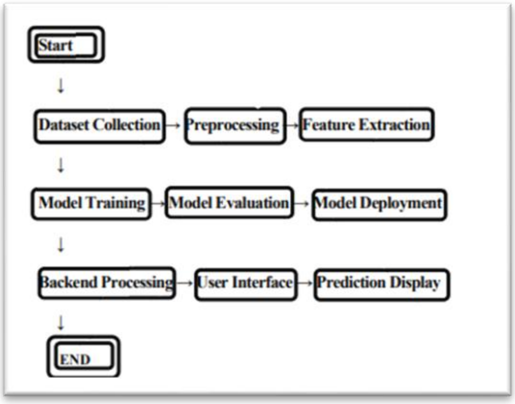

Fig 4.1 Methodology diagram

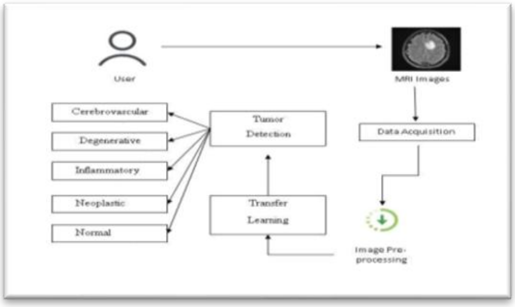

Fig 4.2: System architecture

5.PERFOMANCE MATRIX

The approach for evaluating the performance of a brain tumour detection system will involve creating confusion matrices. These matrices are an often used statistical method that will allow us to evaluate the effectiveness of a system's ability to predict over time based on the output from the brain tumour detection system being developed. Confusion matrices provide a concise summary of predicted results from a classification technique and are broken down into four categories; True positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

To illustrate this using the image of the miracle of MRI where a brain tumour has been identified as having been detected by the proposed model, the true positives will be the images of brain tumours (that is, an MRI image containing a brain tumour) that had been successfully identified by the proposed model. Conversely, the true negatives will be the number of normally functioning MRI pictures (i.e., images containing no brain tumour) that had been successfully identified as 'normal' by the proposed model.

An accurate assessment of the System’s overall performance is required to accurately evaluate performance, and in order to find this information, we must estimate the total number (or total count) of properly detected MRI images (properly classified) versus all detected/classified MRI images.

The accuracy metric may provide a good estimate of how well the proposed Brain Tumour System performs; however, it should not be used as an overall measure of performance due to the diversity of the types of cases and unequal distributions of those types of cases within the database. Therefore, there will be additional performance metrics used to provide as complete of an evaluation as possible, regarding both the exercise and test phases of the trial, for the proposed Brain Tumour System.

6.EXPERIMENTAL SETUP AND RESULT ANALYSIS

As part of creating a brain tumor detection system we created a number of controlled experiments that were designed to measure how well an algorithm could detect and identify different types of tumors based on many computer files (images) taken of the brains of patients pre-surgery, and then to measure the accuracy of the algorithm in identifying the tumors versus the non-tumors taken from the same patients pre-surgery. We generated a data set of MRI images of both tumors and non-tumors by using patient MRI's of their brains on MRI's of patients with the same diagnosis (for example; a person diagnosed with a tumor). Each of the images in the data set (half-tumor, half-non-tumor) had an assigned label indicating if a tumor was present or not (for example; a tumor was present in the MRI). We created both training (using the first half of each data set) and testing (using second half) sets for the algorithm from the image data sets according to the same set of standards in order to ensure a fair evaluation of how well the algorithm works with data sets. We defined a ratio of the number of images in the training set to the number of images in the testing set and used this standard definition for both image data sets.

We also created a standardized training set to help the model learn the characteristics that tumors have to have in order to be identified by the model.MRI images were standardized based on size so they all contain the same type of data as well as each other; this also reduced noise from the images.

Various techniques (rotation, flipping, scaling) were used to create more diverse data in order to reduce the likelihood that the model will overfit due to too little variation in training samples. The Model has been constructed using either machine learning or deep learning methodologies and trained through several epochs while optimally adjusting the hyper-parameters (learning rate and batch size). Once the model has been trained, the model will have its parameters fine-tuned so that the package identified as tumour or non-tumour contains as few errors as possible through the use of back-propagation.

The model learned to distinguish between brain tissue with and without tumours and how to use the features developed to do so. The model was evaluated against four metrics (Accuracy, Precision, Recall, & F-Score). In addition to the metrics, the results were then evaluated with a confusion matrix illustrating the number of true positives (actual tumours successfully identified), true negatives (patients who do not have a tumour correctly classified), false positives (instances where tumour patients were incorrectly classified), and false negatives (instances where non-tumour patients were incorrectly classified).

The results are all in high confidence (>90% prediction accuracy) for the classifier system. In addition, theThere were a small number of false positives, which resulted in a low precision score and produced an overall F1 score that demonstrates similar or balanced performance across several categories when comparing the general outcome(s) of the model as a whole. These results therefore demonstrate that the automated system can provide clinically valid and reliable data to assess the clinical validity of the images of brain tumours by evaluating how they were classified.

7.CONCLUSION

This research article describes a highly functional and responsive framework designed to facilitate the automated detection of brain cancer tumours via magnetic resonance imaging (MRI) scans using machine learning algorithms. The framework will help produce consistent results given the variance seen in the different types of medical imaging by creating a pre-processed image of MRI scans that has been standardised prior to being processed; creating an adaptive training system based on the characteristics of the extracted feature set; and providing assistance with the classification of the resulting data based on how successfully the images were analysed and classified.

The validation testing results display evidence of consistent, continuous system performance over different independent tests, which lends strong support to the generalizability of this model across other sample test sets. Furthermore, the validation testing produced high sensitivity results that support the use of this trained model for successful identification of cases where tumours are involved in diagnosis.

Clinical screenings for tumors could utilize a trained framework to provide additional assistance, as the development of new mechanisms that create minimal rates of false positive finding throughout all testing phases. The overall performance of a trained framework is consistently high for all evaluated metrics, regardless of the distribution used in training for any specific evaluation.

Adding additional research sites and gathering multi-site MRI datasets to expand the amount of available data for future research will provide new directions in this area. New methods of regularly collecting multi-site MRI datasets will enable looking at new ways of improving data quality.

The release of scalable, clear and accessible AI-based tools, in conjunction with the completion of AI-based tools to provide the support (decision) visualisation of these tools, will create further value through the integration of Explainable AI approaches with those tools.

This new technology will make it easier for this framework to be used in real-world healthcare situations than in only research-based settings from now on. This paper describes an evidence base to support the area of the technical efficacy of this framework, the diagnostic reliability of this framework and the adaptability of this framework for diagnosing brain tumors. Therefore, this framework will provide a decision-making tool used as cognitive support for making high-confidence decisions in the interpretation of CT scan images (as obtained by radiologists).

REFERENCES

[1] “PBVit A Vision Setup Using Image Patches”, 2025 – by Pratik Kumar Chauhan, alongside Deepak Kumar Verma, with contributions from GhanshyamG. Tejani.

[2] “Content-Based Brain Resonance Image Retrieval”, 2025-by Bedriye Doga, Hursit Burak Mutlu, U. Rajendra Acharya.

[3] “Brain Tumour Detection using Intelligence Technology”, 2024- by Shubangi Solanki, Uday Pratap Sign, Siddarth Sign.

[4] “Mutual Information MapReduce and Minimum Quadrabgle Classification for tumour Big Data”, 2023-manikandan Ramchandran, Rizwan Patan, ambeshwar Kumar, Sohel Hosseini, Amir H.Gandomi.

[5] “Pattern Descriptor Orientation and Map Algorithm_Based Brain Pathology Detection”, 2021 - B. Deep, M.Murugappan, G. Sumithra, Mufti Mahmud.