Glioblastoma brain tumor segmentation and survival prediction using U Net

Prabhav Srivastava

Dr Dileep Kumar Yadav

School of Computer Science, and Engineering Galgotias University, Greater Noida, India
imprabhav9451@gmail.com

Vanshika Misra

Dr Dileep Kumar Yadav

School of Computer Science, and Engineering Galgotias University, Greater Noida, India
vanshikamisra73@gmail.com

Dr Dileep Kumar Yadav

School of Computer Science, and Engineering Galgotias University, Greater Noida, India
dileep.yadav@gmail.com

Abstract

Magnetic resonance (MR) tissue isolation is a vital stage in giving particular testing of glioma hostility also, treatment reaction. Isolating magnetic resonance (MR) tissue is a useful procedure in allowing exact testing of glioma hostility and treatment reaction. This strategy was used to separate the troublesome assignment of isolating numerous fragments into discrete parallel detachment issues inside each development. Various strategies and it were utilized to lose difficulties set up each gathering. After manual/self-loader growth division, the forecast of amount imaging properties for example, volume-based measures has reflect additional action in the cycle of pictures base picture-based cancer game plan over privilege utilizing clinical protocol, for example, huge anterior posterior, cross over, and sub-par unrivaled cancer degree on a emotionally picked fragment. Such phenotyping could get to the next level careful and helpful preparation by taking into consideration the assessment of reflected organic cycles. The incredible variety of look and type of mind cancers, including sub-locales, makes division tricky, which can be exacerbated by the interaction of sanction mistakes like development or potentially field
inhomogeneity.

Keywords—Glioblastoma, Segmentation, U-Net, Magnetic Resonance Imaging.

1: Introduction

Brain tumor causes abnormal functioning of brain cells which affects the nervous system in a disastrous manner. It has various stages of serious issues with Brain that may lead to decrease the general life span of a normal human being whose brain is tumorized. Brain Tumor [1] is also of 2 types, i.e., Cancerous and Non-Cancerous. The motion inside the skull can affect the stage of Brain Tumor. If it is not detected at an early stage, it may lead to brain damage and can have life threatening

affects for a human being. According to research, around 308,102 people all over the world were diagnosed with Brain Tumor out of which 251,329 people died because of primary cancerous brain and central nervous system tumor that is the reason we have decided to focus on prediction of early-stage brain tumor. MRI (Magnetic Resonance Imaging) [1] scans are used to detect such disease and according to the results, doctor advices the most suitable treatment for the same. These treatments are done based on size, shape, stage, and part of brain where tumor isdetected.

In this project, we are focusing on detection of brain tumor using various Deep Learning techniques. There are various deep learning techniques which could help in building this project such as Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Classic Neural Network, Recurrent Neural Network (RNN) etc. but we are focusing to work using the CNN model to get the most efficient result with high accuracy and less time consuming. The basic purpose of using the CNN model is to get the excellent performance to classify the tumors.

1.1   Purpose of the Project:

The main purpose of this project was to build a CNN model that would classify if subject has a tumor or not base on MRI scan.

In this project, we are going to discuss the early prediction of Brain Tumor using Deep Learning We'll be creating a CNN model (Convolutional Neural Network) for the same. Based on some images, we are going to classify the brain tumor in order to cure or prevent the seriousness of increasing of dangerous stage. Through this project, you will get familiar with the exploratory data analysis and feature engineering techniques that need to be applied to process data.

Machine learning and Data Science are one of the fastest-growing technological fields. This field results in amazing changes in the medical field, production, robotics etc. The main reason for the advancement in this field is the increase in the computational power and availability of large amounts of data. In Data Science, this data is analyzed and made suitable for creating machine learning models and products. And Deep Learning is a subpart of Machine Learning.

1.2   Target beneficiary:

The motivation for using deep learning in the Brain tumor classification is to improve the accuracy and performance and create better strategies in order to prevent from increasing to a dangerous state.We used Convolutional Neural Network. With the help of this DL Network we predicted the probability of a person having brain tumour which are in the form of Boolean results i.e. Yes (1) or No (0).

1.3   Project Scope:

Our aim is to detect the brain tumor is or not from the CNN model we have build based on the input images. the early detection of brain tumor will prevent the patient from its dangerous stage,later on we will apply the model two or three different datasets.

*   Input:
A collection of images of brain tumor. These images are of different sizes with varied lighting conditions.
*   Output:
The primary goal is to extract meaningful and accurate information from these images with the least error possible. The first step is pre-processing which is used for reduction of impulsive noises and image resizing. The next step is feature extraction, mser features and etc. The last step is model building, A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take this input image to differentiate one from the other. the evaluation of result is done on the basia of different parameters like accuracy, precision, recall values etc.

2   ) Project Description:

2.1   Reference algorithm: CNN

Convolutional neural network (CNN) is one of the most popular and used of DL networks Because of CNN, DL is very popular nowadays. The main advantage of CNN compared to its predecessors is that it automatically detects the significant features without any human supervision which made it the most used.

2.2   Data:

First of all we want the data which we are going to pre-process and using that we will train our model. So we collected the data Brain Tumor Detection CNN v1.0‖vgg16 given by Ruslan Klymentiev from the website named Kaggle. They include information like images. There were two folders named as Yes and No. In the Yes folder, it consists of 259 images of Tumor whereas in the No folder,100 images where no tumor exists. No quality data means no

quality results. This is the basic line behind the success of any algorithm. Because before using any algorithm we have to process the data in advance if we expect the best results from it. There are many potential problems with data such as incorrect values, inconsistencies, imperfections etc. Firstly we have collected the data then we have imported the data after that we have extracted the data then splitted it into training and testing and validation after that we have pre- processed the data then we have build the model, after that we have predicted the result that the images has tumor or not. The most crucial libraries for data analysis are Numpy and Matplotlib and other libraries used are tqdm, Imutils, Plotly, Tensorflow. We have used numpy for array size of data and matplolib for plotting the graph and confusion matrix.

2.3   SWOT Analysis:

●   Strength:
As an image is one of the strongest way to analyze the functioning and defects, that is why it is able to be constructed from any object that we pull from an image. A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other. The network learns to extract features while training, we just feed the image to the network, it helps the model to generate more accurate data. We are collecting datasets from Kaggle, that's a website which allows the user to access multiple datasets that are required. we use the pre-trained model that has been already trained on large datasets and extract the ,mser features and etc, from the images and use them for our tasks.

●   Weakness:

The images if not uploaded correctly may give the incorrect results. The dataset used to predict the result should be preprocessed properly.

●   Opportunities:
Predicting the Tumor will help us to know the result faster.

●   Threats:
An incorrectly predicted Tumor may result in putting the human life in danger.

2.4   Project Features:

a.   Collecting Dataset:
First of all we want the data which we are going to pre-process and using that we will train our model. So we collected the brain tumor dataset from Kaggle. There are Two fifty files and two folders in this dataset :- a) Yes b) No

b.   Data import and pre-processing:
No quality data mean there will be no quality results. This is the basic line behind the success of any algorithm. Because before using any algorithm we have to process the data in advance if we expect the best results from it. There are many potential problems with data such as incorrect values, inconsistencies, imperfections etc. that's why we need to filter the correct data without getting any kind of errors in it.

There are many pre-processing steps to deal with these issues such as data cleaning, data consolidation, data modification, data reduction, data segmentation etc. The most crucial libraries for data analysis are Numpy and Matplotlib and other libraries used are tqdm, Imutils, Plotly, Tensorflow. We have used numpy for array size of data and matplolib for plotting the graph and confusion matrix

Data Pre-processing: It improves the standard of the brain tumor MR images.Pre-processing techniques used are as follows:-
• 		cv2.erode() method is used to perform erosion on the image.
• 		Conversion of RGB color to gray-scale.
• 		The dilation operation is performed by using the cv2. dilate() method.
• 		Gaussian Blur() method of cv2 library to blur an image.
• 		Interpolation is the way the extra pixels in the new image is calculated.cv.draw: Contours function is used. It can also be used to draw any shape providedyou have its boundary points.

Data Augmentation:
Techniques used to increase the amount of data by adding slightly modified copies of already existing data or newly created synthetic data from existing data.
 a) 		Zoom in, zoom out.
 b) 		Left shift, right shift.

c.   Training and Testing Splitting the data set:

• 		Testing data:
The Convolutional Neural Network (CNN or ConvNet), Its built-in convolutional layer reduces the high dimensionality of images without losing its information algorithm is trained using this data. The data scientist gives the algorithm with input data that matches to an expected output. The model evaluates the data repeatedly in order to gain a better understanding of its behaviour and then modifies itself to accomplish the goal.

• 		Training data:
     A huge dataset used to train a Convolutional Neural Network is referred to as trainingdata. Prediction models that apply deep learning algorithms are taught how to extract features that are important to certain goals using training data.

2.5   Design diagrams:
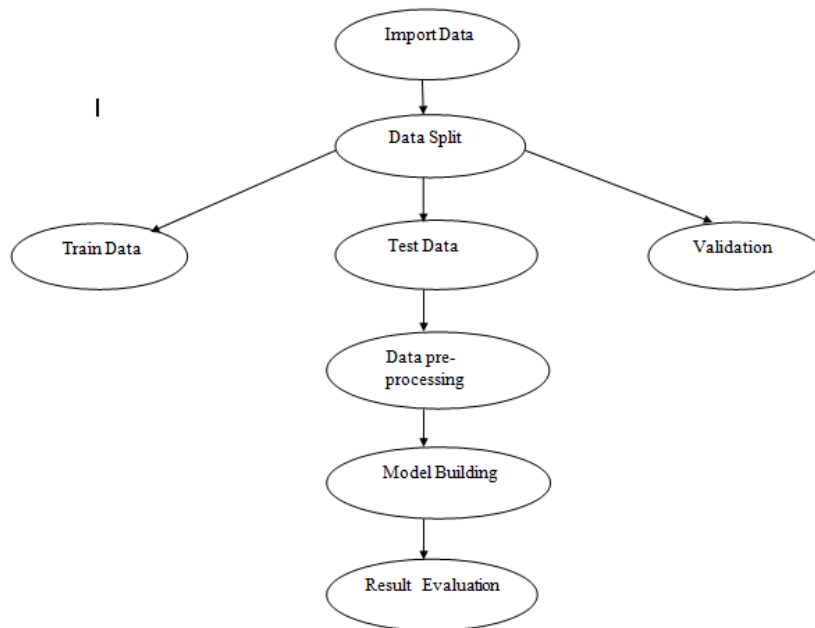
Design Diagram of Brain Tumor Classification
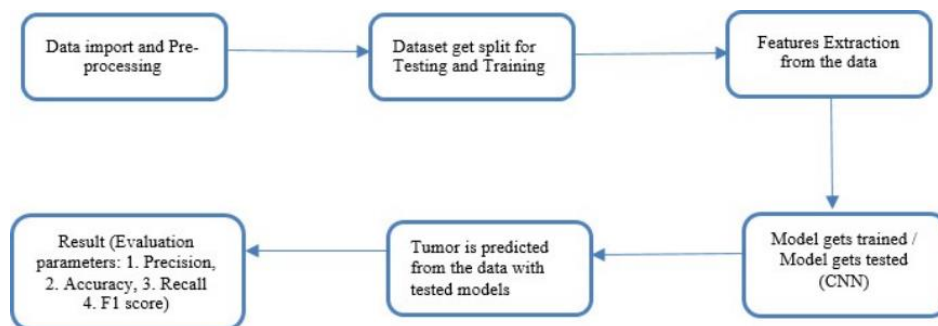
Assumption and Dependencies

- The MRI images present in the dataset should be correct.
- The CNN model created should be trained several times to get the accurate results.

3.   System Requirement

3.1   Software Requirement

Operating System   :   Windows10/8/7(64bit)
Platform   :   Google Colab
Language   :   Python

3.2   Hardware Requirements

Processor   :   Dual Core 2.7 GHz or Better
RAM   :   512 MB or Higher
Disk Space   :   512 MB

Non-functional Requirements

Non-functional requirement is a description of features, characteristics and attributes of the system as well as any constraints that may limit the boundaries of the proposed system. The non- functional requirements are essentially based on the performance, accuracy. Based on these the non-functional requirements are as follows:

- System performance: The user system should be smooth and the Usability: The system should be compatible with any PCs. It should work under anyenvironment and also under any conditions.
- Efficiency: The output should be more accurate and should have a low false positive rate.

Libraries Used

- Matplotlib:
It is a comprehensive libraryfor creating static, animated, and interactivevisualizations in Python. Matplotlib makes easy things easy and hard things possible.
Create publication quality plots.

- Tqdm:
It is used for making progress bars in python.

- Plotly:
Python plotly Library is an open-source library that can be used for data visualization and understanding data simply and easily. Plotly supports various types of plots like line charts, scatter plots, histograms, cox plots, etc.

- Imutils:

Python Package Wiki pip install imutils==0.5.4 A series of convenience functions to make basic image processing functions such as translation, rotation, resizing, skeletonization, displaying Matplotlib images, sorting contours, detecting edges, and much more easier with OpenCV and both Python 2.7 and python 3.0.

- NumPy:

NumPy is a Python library used for working with arrays.

It also has functions for working in the domain of linear algebra, Fourier transform, and matrices. NumPy was created in 2005 by Travis Oliphant. It is an open-source project and you can use it freely. NumPy stands for Numerical Python.

- Pandas:

Pandas is a Python library. Pandas is used to analyse data.

- Tensorflow:

TensorFlow makes it easy for beginners and experts to create machine learning models.

- Sk Learn:

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python.

References

[1] Tarik H., Tawfik M., Youssef D., Simohammed S., Mohammed O.J., Miloud J.E. (2020) Towards an Improved CNN Architecture for Brain Tumor Classification. In: Serrhini M., Silva C., Aljahdali S. (eds) Innovation in Information Systems and Technologies to Support Learning Research. EMENA-ISTL 2019. Learning and Analytics in Intelligent Systems, 7. Springer, Cham. https://doi.org/10.1007/978-3-030-36778-7_24

[2] D. Divyamary, S. Gopika, S. Pradeeba and M. Bhuvaneswari, Brain Tumor Detection from MRI Images using Naive Classifier, 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 620-622, doi: 10.1109/ICACCS48705.2020.9074213.

[3] Özyurt, Fatih, Eser Sert, and Derya Avcı. "An expert system for brain tumor detection: Fuzzy C-means with super resolution and convolutional neural network with extreme learning machine." Medical hypotheses 134 (2020): 109433.

[4] D´ıaz-Pernas, Francisco & Mart_nez Zarzuela, Mario & Anton- ´ Rodríguez, M´ıriam & Gonzalez-Ortega, David. (2021). A Deep ´ Learning Approach for Brain Tumor Classification and Segmentation Using a Multiscale Convolutional Neural Network. Healthcare. 9. 153. 10.3390/healthcare9020153.

[5] RAI, HARI & Chatterjee, Kalyan & Gupta, Apita & Dubey, Alok. (2020). A Novel Deep CNN Model for Classification of Brain Tumor from MR Images. 134-138. 10.1109/ICCE50343.2020.9290740.

[5] Amin, Javaria & Sharif, Muhammad & Haldorai, Anandakumar & Mussarat, Yasmin & Nayak, Ramesh. (2021). Brain tumor detection and classification using deep learning: a comprehensive survey. Complex & Intelligent Systems. 10.1007/s40747-021-00563-y

[7] Malathi, M., and P. Sinthia. "Brain Tumor Segmentation Using Convolutional Neural Network with Tensor Flow." Asian Pacific journal of cancer prevention: APJCP 20.7 (2019): 2095.

[8] Brain Tumor Detection v1.0 || CNN, VGG-16
https://www.kaggle.com/code/ruslankl/brain-tumor-detection
[1]