BRAIN TUMOR DETECTION

A Software Project Management REPORT

Submitted by

Pandya Keval Rajeshbhai

190430116092

in partial fulfilment for the award of the degree of

BACHELOR OF ENGINEERING

In

Information Technology

Shantilal Shah Engineering College, Bhavnagar

Gujarat Technological University, Ahmedabad

October, 2022

Shantilal Shah Engineering , Bhavnagar

New sidsar campus, Post: Vartej, Sidsar, Bhavnagar – 364060,

Gujarat, India

CERTIFICATE

Date: __/__/____

This is to certify that the Software Project management report submitted along with the project entitled **Brain Tumour Detection** has been carried out by **Pandya Keval Rajeshbhai(190430116092)** under my guidance in partial fulfilment for the degree of Bachelor of Engineering in **Information Technology**, 7th Semester of Gujarat Technological University, Ahmedabad during the academic year 2022-23.

Prof. S. J. Agravat — Internal Guide

Dr. M. S. Shah — Head of the Department

ACKNOWLEDGEMENT

There has been significant increase in the numbers of medical cases involving brain tumours in last few years, ranking it 10th most common form of tumours affecting children and adults. Detection and classification of most crucial and time taking task in field of medical image processing. Because of high variance of the size , shape , location of brain tumour Magnetic Resonance Imagining (MRI) is broadly used for detecting tumour and diagnosis various tissues abnormalities. Accurate investigation of the size and location of brain tumour plays an important role in the diagnosis of brain tumour

As a matter of first importance, I am extremely thankful to internal guide **Mr C. H. Makwana (Prof. In I.T. Dept.)** has directed me to achieve our task and giving their wide experience of information. We are additionally extremely appreciative to the SHANTILAL SHAH ENGINEERING COLLEGE for permitting me to do this project during my summer internship.

To wrap things up, we might want to recognize and thanks in huge measures to all our own kindred companion and guides for their help.

Pandya Keval Rajesbhai

(190430110692)

ABSTRACT

Brain is the most complex structure of the human body and brain tumour is one among the lethal variety of cancer. A brain tumour is a collection or mass of abnormal cells in brain. Any growth of such cells inside the brain which is enclosed by a very rigid skull can cause major pain and problems. Brain tumours is Malignant (cancerous) or Benign (noncancerous). When benign or malignant increases, it can increase the pressure inside our skull. It can damage the brain, and may even be fatal. Brain tumours are classified in two types: (i) Primary Tumour (ii) Secondary Tumour. Primary brain tumour arises in our brain. Primary tumours often Benign or Malignant. The secondary brain tumour also referred as metastatic brain tumour. It arises when cancer cells spread to our brain from other organs like lungs, breasts, skin or kidneys. Secondary brain tumours are always malignant. Benign tumours don't spread from one part of our body to another.

To minimize fatal consequences, an accurate tumour detection and classification of brain tumour is crucial for a treatment plan. In diagnosing brain tumours imaging plays a very important role. The medical field needs fast, automated, efficient and reliable technique to detect tumour like brain tumour. Detection plays very important role in treatment. If proper detection of tumour is possible then doctors keep a patient out of danger. Various image processing techniques are used

We can use a Deep Learning architectures CNN (Convolution Neural Network) generally known as NN (Neural Network) and Transfer learning for detect the brain tumour and classify the types of tumour.

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

SYMBOLS FOR USE CASE DIAGRAM:

SYMBOLS FOR SYSTEM FLOW DIAGRAM:

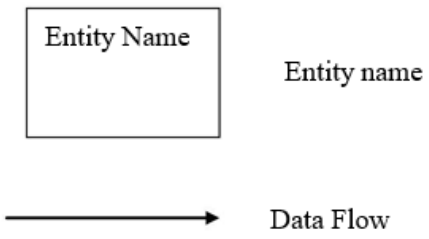

TABLE OF CONTENTS

==========

1.0 INTRODUCTION

==========

1.1 PROJECT SUMMARY

1.2 PURPOSE

1.3 OBJECTIVE

1.4 SCOPE

1.5 TECHNOLOGY AND LITERATURE REVIEW

==========

1.1 PROJECT SUMMARY

Table 1.1 project summary

Project Title	Brain Tumor detection and classification
Aim	Our main aim is to provide better method to detect and classify brain tumor
Project Category	Deep learning and machine learning method
Language	Python
Tools	Jupyter notebook, spyder IDE
Framework	Tensor flow, Keras , Numpy
Duration	2 weeks(20 June to 8 July)

1.2 PURPOSE

The purpose behind this project is to provide better method to classify and detect brain tumor and its type in order to help doctor to provide the proper treatment regarding brain tumor.

Goal of the project:

- To detect the brain tumor is percent or not.
- If it is present then classify type of the tumor

1.3 OBJECTIVE

The main objective of the project is to detect the brain tumor properly and classify its type and give information regarding it to the doctor so the doctor can do proper treatment towards it

1.4 SCOPE

What can it do?

i. It can detect the brain tumor is present or not in the given image of 2-D slice of MRI (Magnetic Resonance Imaging).

ii. It can classify the tumor type of brain if the tumor is present in the brain.

What Can't it do?

i. Convert 3-D MRI images to 2-D slices.

ii. Detection and classification on direct 3-D MRI data set

iii. Brain tumor segmentation.

iv.

1.5 TECHNOLOGY AND LIRTATURE REVIEW

Python:-
Python is a high-level, interpreted, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation.

Python is dynamically-typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly procedural), object-oriented and functional programming. It is often described as a "batteries included" language due to its comprehensive standard library.

Tensor flow:-
Tensor Flow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks.

Tensor Flow was developed by the Google Brain team for internal Google use in research and production. The initial version was released under the Apache License 2.0 in 2015. Google released the updated version of Tensor Flow, named Tensor Flow 2.0, in September 2019

Tensor Flow can be used in a wide variety of programming languages, most notably Python, as well as JavaScript, C++, and Java. This flexibility lends itself to a range of applications in many different sectors.

Keras:-
Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the Tensor Flow library.

Numpy:-
Numpy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays

Deep Learning:-

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.

Deep-learning architectures such as deep neural networks, deep belief networks, deep reinforcement learning, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis, climate science, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.

Convolution Neural Network(CNN):-

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of artificial neural network (ANN), most commonly applied to analyze visual imagery. CNNs are also known as Shift

Invariant or Space Invariant Artificial Neural Networks (SIANN), based on the shared-weight architecture of the convolution kernels or filters that slide along input features and provide translation responses known as feature maps

2

==========

2.0 SYSTEM ANALYSIS

2.1 STUDY OF CURRENT SYSTEM.
2.2 PROBLEM AND WEAKNESS OF CURRENT SYSTEM.
2.3 REQUIREMENT OF NEW SYSTEM.
2.4 ACTIVITY /PROCESS OF NEW SYSTEM.
2.5 FEATURES OF NEW SYSTEMS.

==========

2.1 STUDY OF THE CURRENT SYSTEM.

- There has been significant increase in the numbers of medical cases involving brain tumours in last few years, ranking it 10^{th} most common form of tumours affecting children and adults.
- Detection and classification of most crucial and time taking task in field of medical image processing. Because of high variance of the size , shape , location of brain tumour Magnetic Resonance Imagining (MRI) is broadly used for detecting tumour and diagnosis various tissues abnormalities.
- The Current system can cause the error in the prediction of the brain tumor due to issues are told before.
- So we need the system that deals with this types of issues and give us

2.2 PROBLEM AND WEAKNESS OF CURRENT SYSTEM

- Poor image processing technique.
- Manual inspections are needed.
- Time consuming.
- Late treatments due to delay in result

2.3 REQUIREMENT OF NEW SYSTEM

To overcome the weakness of the current system like time consuming, late Treatment, poor image Processing and prediction. We need different tumor detection and classification algorithm.

2.4 SYSTEM FEASIBILITY

2.4.1 Does the system Contribute to the Overall Objective of The Organization?

Yes, the purposed system contributes to overall objective of the Organization

As it follows and maintain the motives of organization as to make early result.

So the doctor can start the early treatment.

2.4.2 Can the System Be Implemented Using the Current Technology and Within the Given Cost and Schedule Constraints ?

Yes. The system can be implemented using the current technology and within the given cost and schedule constraints.

2.5 ACTIVITY/PROCESS OF NEW /PROPOSED SYSTEM

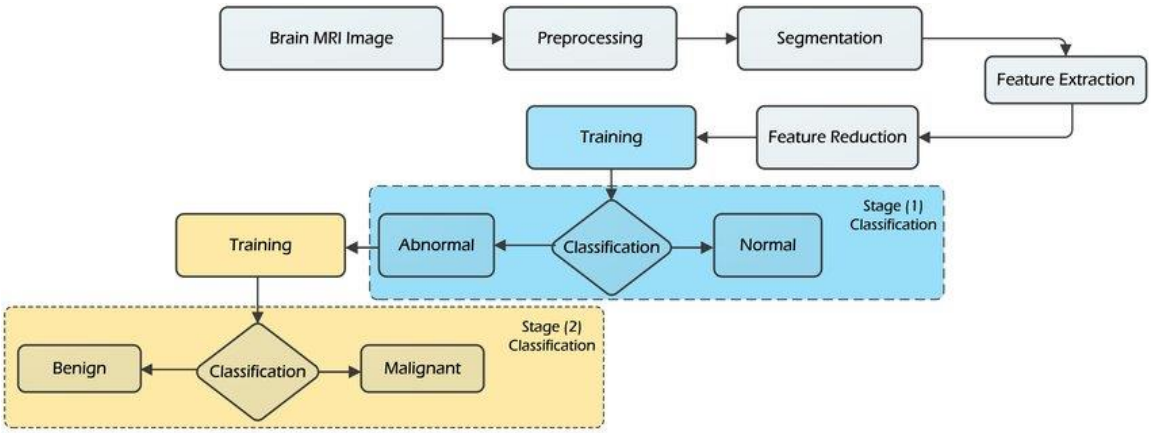

Fig 2.5.1 activity process chart

2.6 FEATURES OF NEW SYSTEM

- **Features of the new system.**

I. Time saving.

II. Good image processing techniques.

III. Accurate Prediction.

IV. Early and accurate result so the doctor can start treatment in early stages

V. Easy to use.

VI. Deals with issue like different brain size, abnormality , different tissue size etc.

VII. Fast results.

3

==========

3.0 SYSTEM DESIGN

3.1 SYSTEM DESIGN & METHODOLOGY.
3.2 DESIGN DIAGRAMS.

==========

3.1 SYSTEM DESIGN & METHODOLOGY

- I am using Incremental Methodology in development process. Incremental Model is a process of software development where requirements divided into multiple standalone modules of the software development cycle. In this model, each module goes through the requirements, design, implementation and testing phases. Every subsequent release of the module adds function to the previous release. The process continues until the complete system achieved.

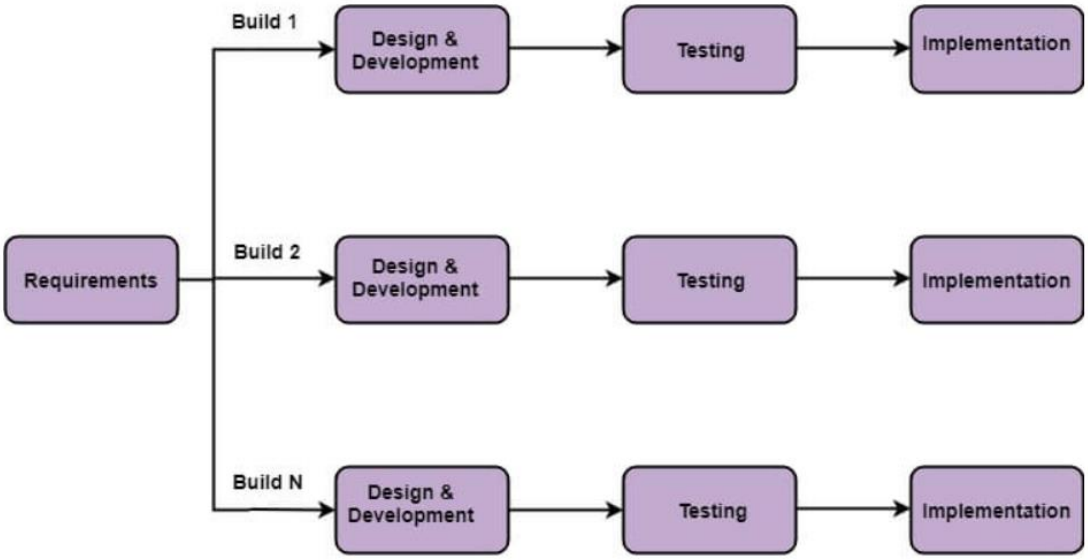

Fig: Incremental Model

Fig 3.1 Incremental Model

- **Incremental Methodology & Process**
- The various phases of incremental model are as follows:
- **1. Requirement analysis:** In the first phase of the incremental model, the product analysis expertise identifies the requirements. And the system functional requirements are understood by the requirement analysis team. To develop the software under the incremental model, this phase performs a crucial role.

- **2. Design & Development:** In this phase of the Incremental model of SDLC, the design of the system functionality and the development method are finished with success. When software develops new practicality, the incremental model uses style and development phase.

- **3. Testing:** In the incremental model, the testing phase checks the performance of each existing function as well as additional functionality. In the testing phase, the various methods are used to test the behavior of each task.

- **4. Implementation:** Implementation phase enables the coding phase of the development system. It involves the final coding that design in the designing and development phase and tests the functionality in the testing phase. After completion of this phase, the number of the product working is enhanced and upgraded up to the final system product.

3.2 DESIGN DIAGRAMS

3.2.1 FLOW CHART

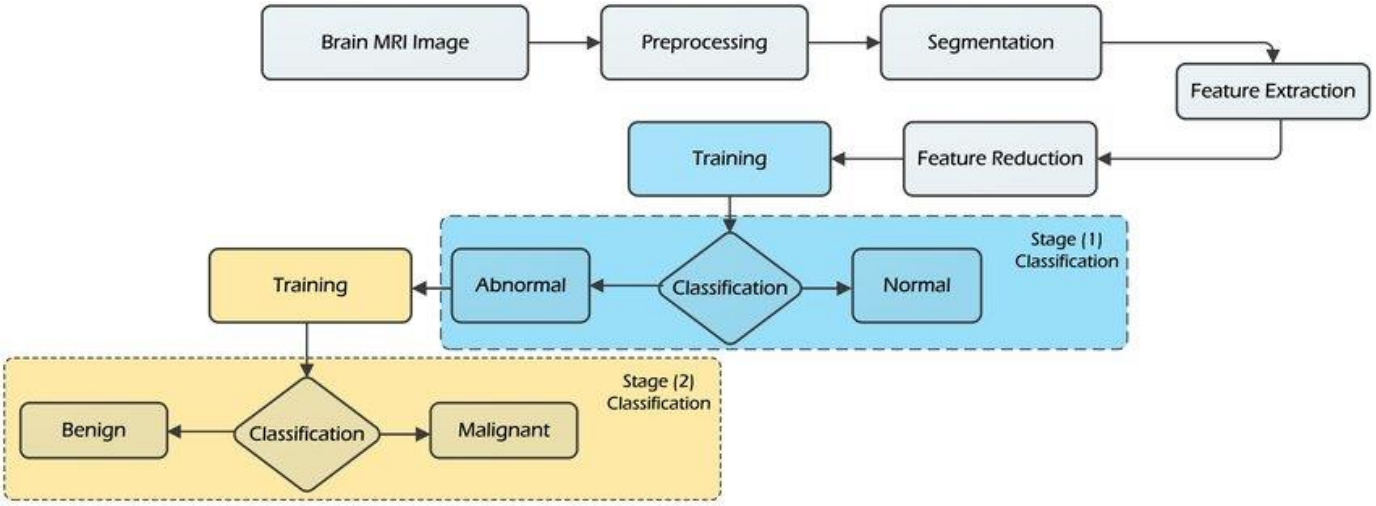

Fig 3.2.1 Flow chart

- As you can see the flow of the project starts from a MRI images . Which then preprocessed and segmented then goes for the feature extractions and reductions then these features are used to train the first model which is the tumor detection model then the same images is goes for the training for the classification model of type of the tumor

3.2.2 USE CASE DIAGRAM

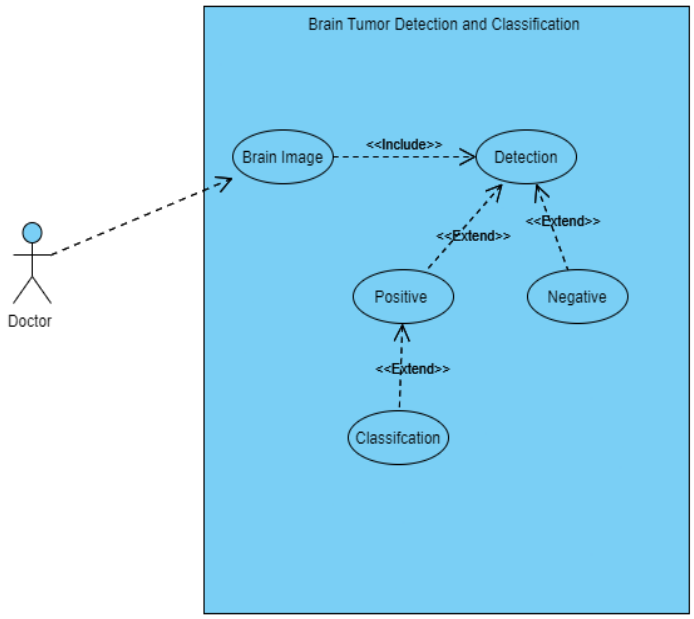

Fig 3.2.2 Use Case Diagram

3.2.3 DFD Level 0

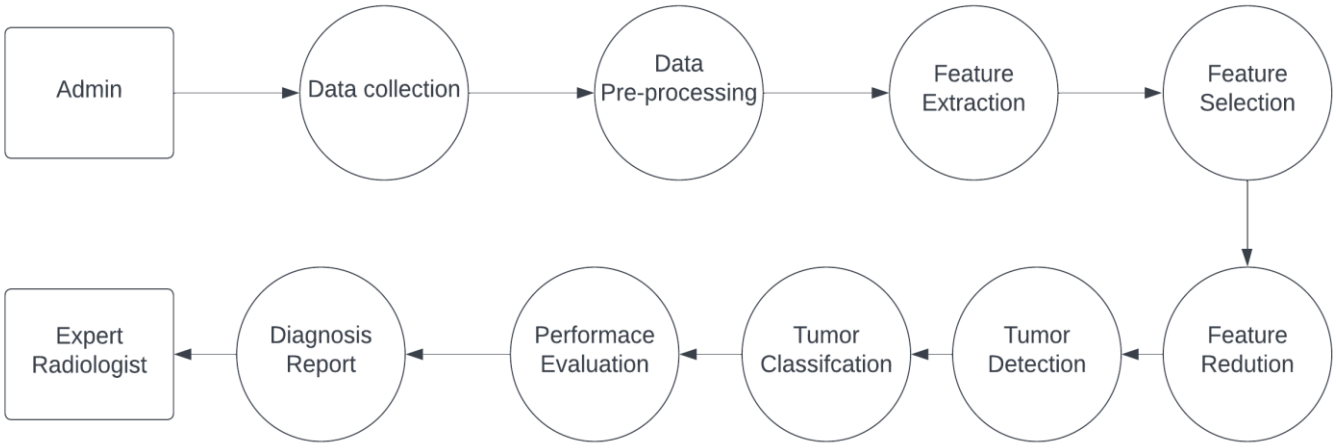

Fig 3.2.3 DFD Level 0

3.2.4 DFD Level 1

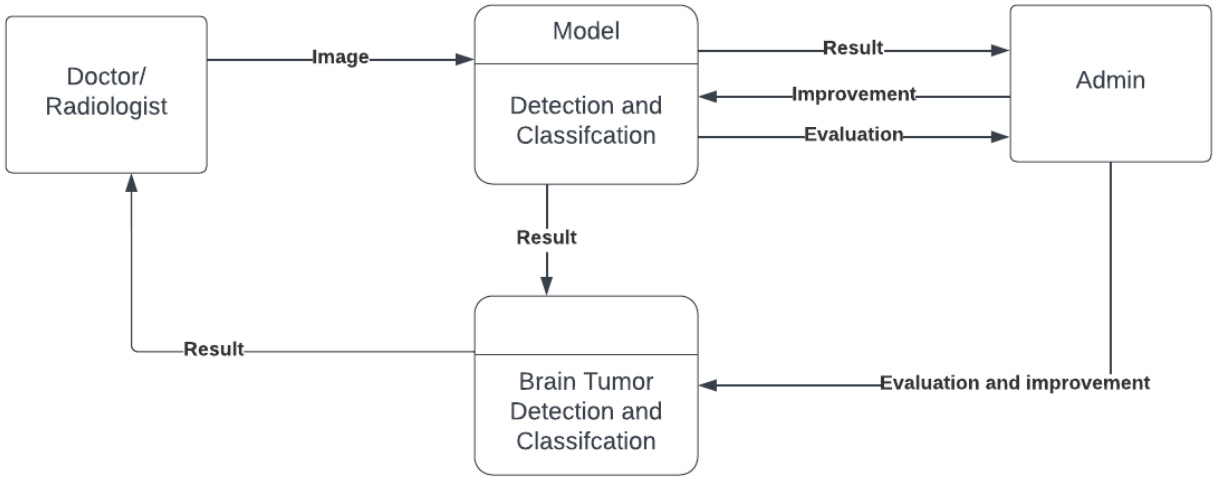

Fig 3.2.4 DFD Level 1

==========

4.0 MODEL DESIGN

4.1 MODEL ALGORITHM.
4.2 MODEL ARCHITURE.
4.3 MODEL ACCURACY.

==========

4.1 MODEL ALGORITHM

- Here we are using the deep learning algorithm for both the classification and detection of the algorithm.
- The Convolution neural network, max pooling layer, fully connected layers are used in creation of the both models.
- The Relu activation function is mainly used in all the layers accept the last layers. In the detection model the “Sigmoid” activation function is used to do the binary classification to figure out is the tumour is present or not. While in the classification model the softmax activation function is used to classify the types of tumour.
- The Convolution neural network is used to extract the features from images.
- Max pooling layer is used to Pool out maximum of features.
- The work of the activation function is to active the neuron based on the output of the previous neuron.
- The Output of the fully connected layer is the array of probability.
- The sigmoid function convert the array into the class (Either yes or no).
- The softmax also convert array into the class but it works for the multiple-classes.
- The adam optimization method is used for the Gradient decent in the back-propagation in order to reduce the loss and improve the accuracy.
- The binary cross entropy loss is used for the detection model and categorical cross entropy is used for the Classification Model.
- Here I also used the concept of the transfer learning in the classification model in order to reduce the loss and improve the accuracy

4.2 MODEL ARCHITECTURE

4.2.1 DETECTION MODEL

- As you can see in the figure below , it is the architecture of the Detection model . We can see it by using the model.summray() command.
- The first layer is the input layer which takes the input form the batches here batches are in types of batch tensor. The input shape is 150X150X3 here we are using RBG 3-channels images.
- Then We have 2 convolution layers followed by two pooling layers. The size of convolution networks are 32 and 64. The stride of the max-pooling layer is 2.
- Then we have three dense layers containing neurons 128,64,1.
- The last layer contains only one neuron which is used for the binary classification.
- The last neuron has only one neuron because it has only two class to classify.
-

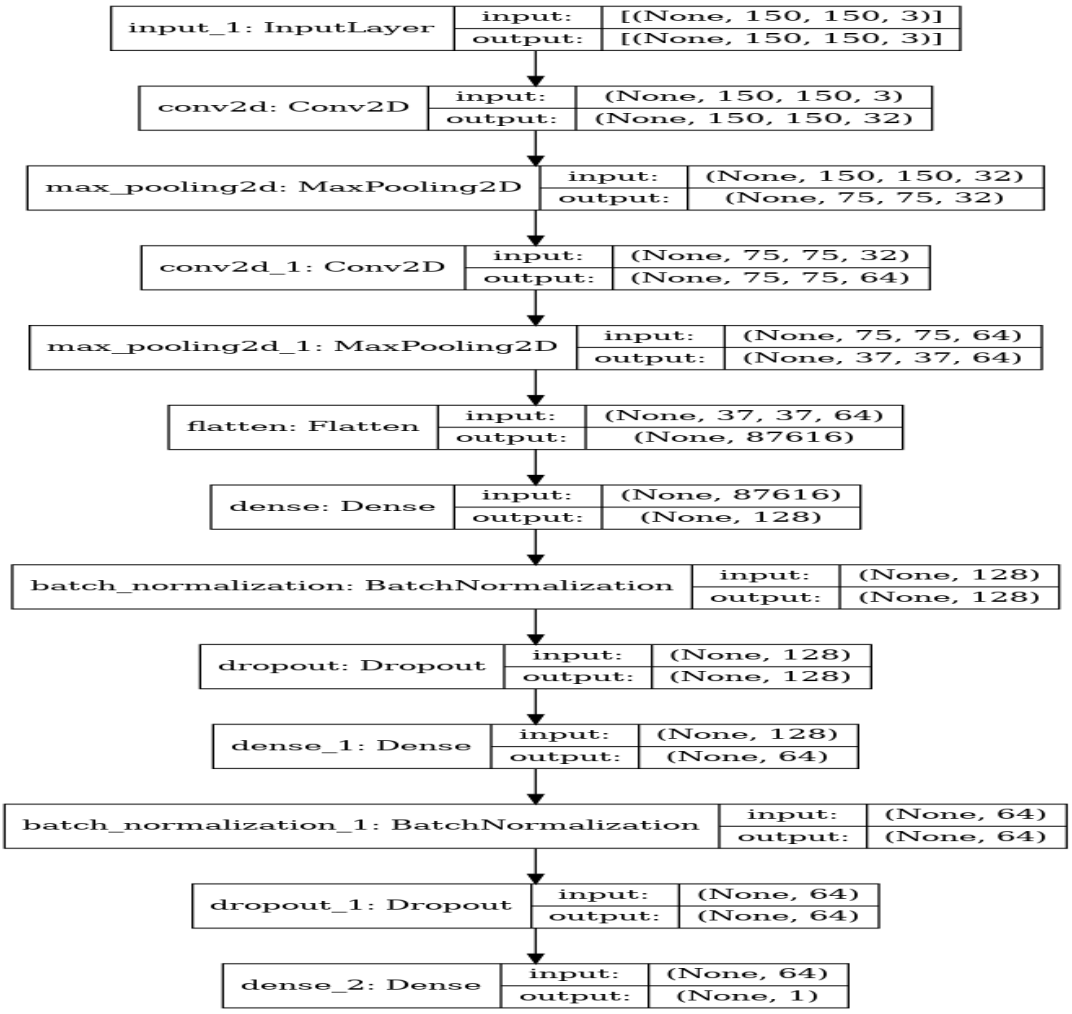

Figure 4.1 Detection Model Architecture

4.2.2 Classification Model

- The classification model has almost same architecture as the detection model.
- It takes input as 128X128X3.
- Here I am using the concept of the transfer learning mean I am using the pre-trained neural network to train the model rather than making the model for starch because it will save the time and also it will give better result.
- Here I am using pre-trained resnet-152V2 model which is trained on image net data set and has almost 152 layers and the pre-trained weights here I am using is “imagenet”.
- After that the feature extractor (Resnet-152V2) is connected to fully connected Dense layers.
- The last layer has four neurons because here I am predicting on the four class.
- All the Dense layers are using Relu activation function while last layer has softmax because of the multiclass classification.
- Here I have also used the concepts of dropout in order to reduce the over fitting

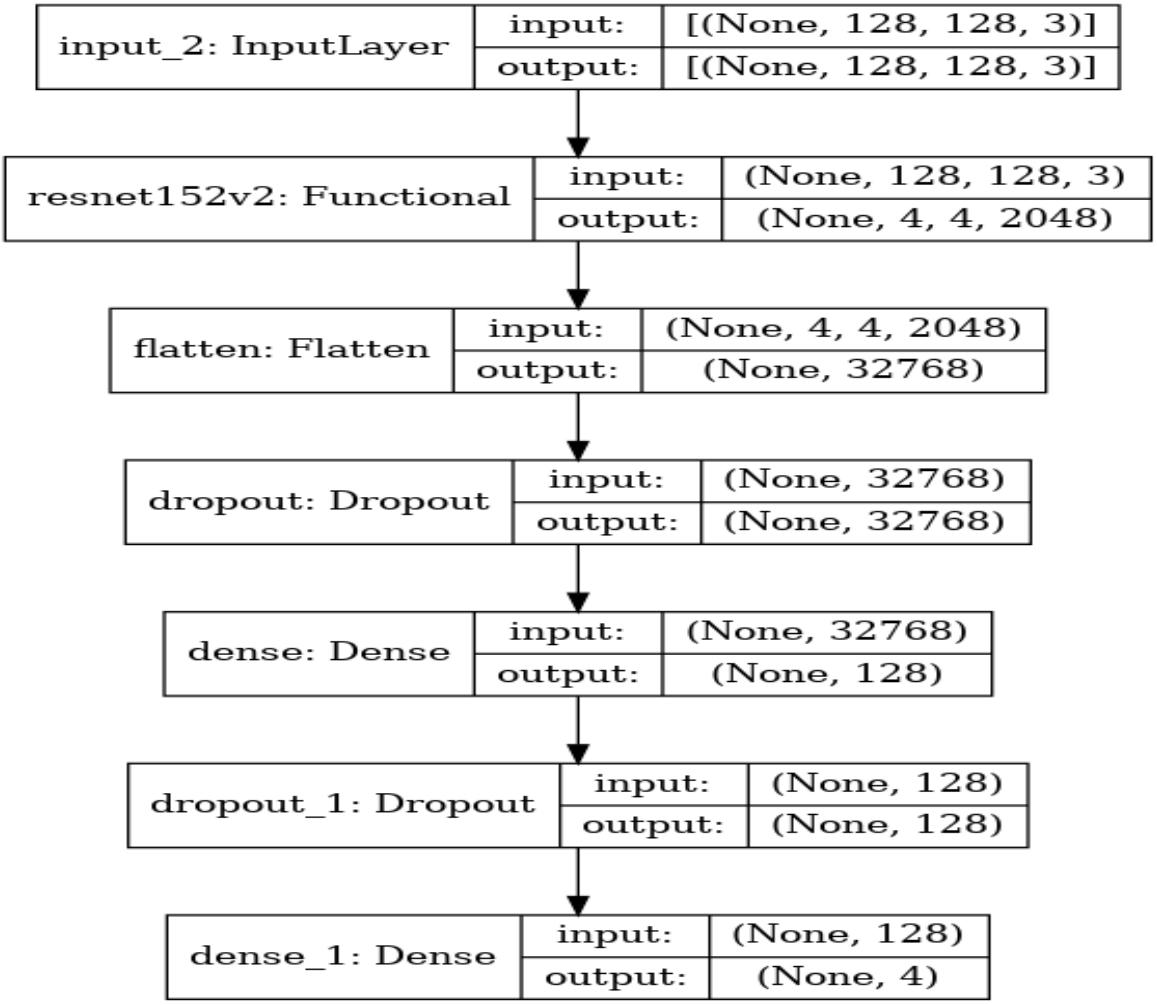

Figure 4.2 Classification Model

4.3 MODEL ACCURACY

Model	Loss Used	Training Loss	Testing Loss	Training accuracy	Testing accuracy
Detection	Binary cross	0.0000192	0.000087	99.21%	98.34%
Classification	Categorical	0.0000025	0.0000056	99.34%	99.01%

Plotting of model accuracy

- Detection Model

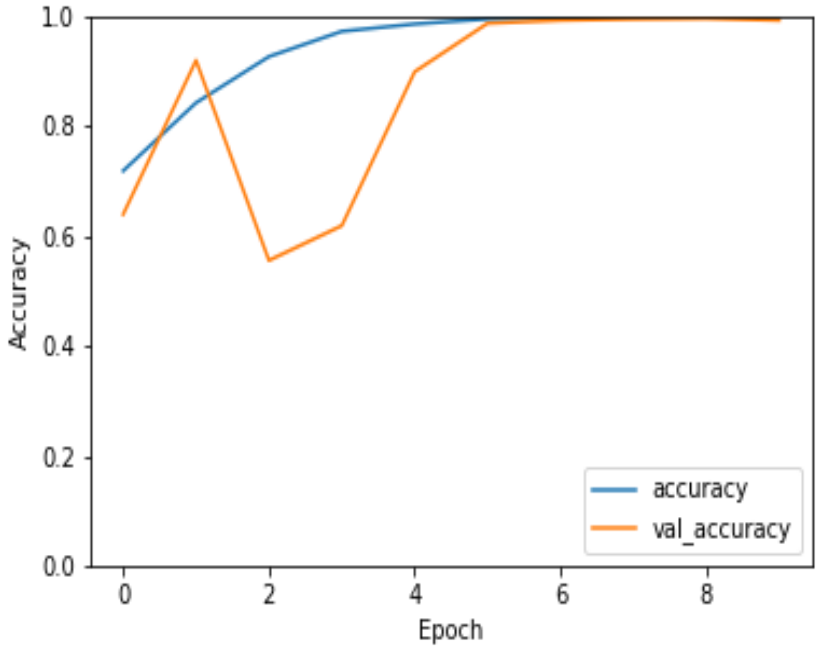

Fig 4.3 Detection Accuracy

- Classification Model

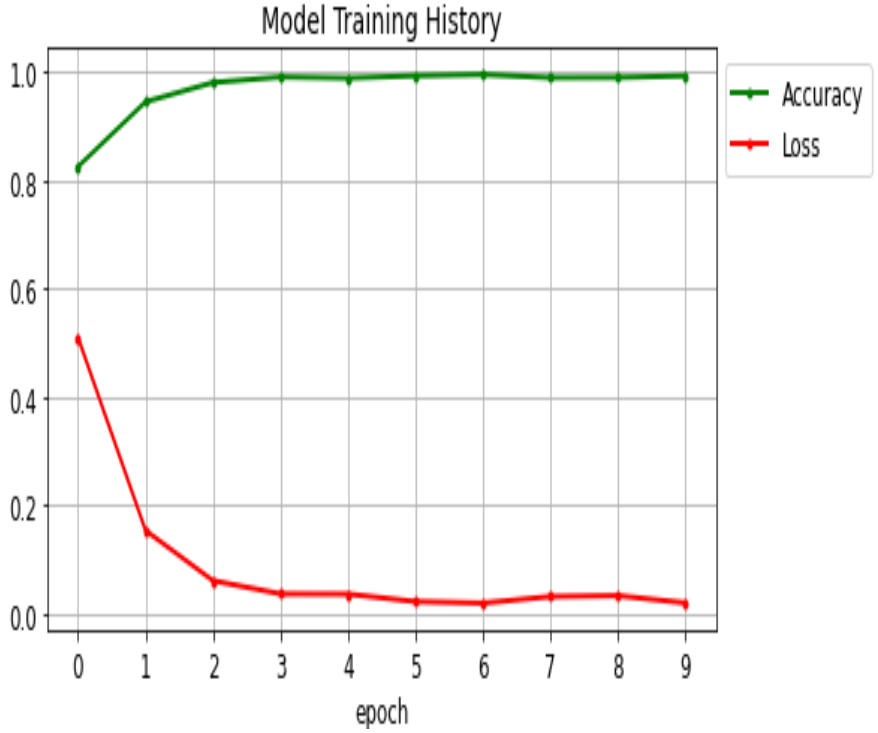

Fig 4.4 Classification Accuracy

5

==========

5.0 IMPLEMENTATION

5.1 IMPLEMENTATION STEPS

==========

5.1 IMPLEMENTATION STEPS

- There The Major steps of the Implementation are as below:-

1. Build and Collect the Dataset.

2. Train Model.

3. Evaluate Model
.
4. Model Deployment.

5.1.1 Build and Collect dataset

- The Data collection and building of the dataset is very import task to solve the problem of Machine Learning.

- As per survey , 80% of the time is spend on data collection and pre-processing in solving the machine learning problem.

- Here I am using mixture of many datasets in order to get better accuracy. Because a wise man says
"the more data , the more learning , the more learning the better results".

- The data set here I am using are from the Kaggle which is subsidiary of Google LLC, is an online community of data scientists and machine learning practitioners.

- Here I am using two types of the data first one is for tumor detection and other for the tumor classification.

- In the data pre-processing task I am doing following operations on the dataset.

i. Image resizing.

ii. Noise Removing.

iii. Train/Test partitioning (80/20).

iv. Removing un-necessary images.

5.1.1.2 Datasets which I used

5. Here I have used two types of dataset 1.Dataset containing the image about the Positive and negative image and 2nd one which contains images of the different type of brain tumor.

6. All the images are the 2-D image slices of the 3-D MRI volumes of the different subjects. Various noise removing , image enhancing techniques are applied on

images in order to give us good quality images in which all the features can extracted easily in order to get better results

- Detection Dataset:- The detection Dataset Here I have used is Brain Tumor dataset which is available on kaggle.com. This dataset contains 2 classes 1st is positive and 2nd is negative . This dataset contain about 4600 images. All the images are in .jpg format and 150X150 RGB images.

- Classification Dataset:- In the classification dataset , I have used dataset of Brain Tumor MRI Dataset which is also from kaggle.com. This dataset contains more than 7000 images. Each images are in .jpg format and 150X150 RGB images. This dataset contains four classes.

  1.glioma.
  2.meningioma.
  3.notumor .
  4.pituitary.

5.1.2 Train The model

- After collecting and pre-processing the data , next step is training the model . It is most important step of the project.
- The process of training an ML model involves providing an ML algorithm (that is, the learning algorithm) with training data to learn from. The term ML model refers to the model artifact that is created by the training process.
- Here I am creating two Models 1st one for the detection of tumor and 2nd one for the classification of the tumor.
- Here we are using the deep learning algorithm for both the classification and detection of the algorithm.
- The Convolution neural network, max pooling layer, fully connected layers are used in creation of the both models.
- The Relu activation function is mainly used in all the layers accept the last layers. In the detection model the “Sigmoid” activation function is used to do the binary classification to figure out is the tumour is present or not. While in the classification model the softmax activation function is used to classify the types of tumour.
- The Convolution neural network is used to extract the features from images.
- Max pooling layer is used to Pool out maximum of features.
- The work of the activation function is to active the neuron based on the output of the previous neuron.
- The Output of the fully connected layer is the array of probability.
- The sigmoid function convert the array into the class (Either yes or no).
- The softmax also convert array into the class but it works for the multiple-classes.
- The adam optimization method is used for the Gradient decent in the back-propagation in order to reduce the loss and improve the accuracy.
- The binary cross entropy loss is used for the detection model and categorical cross entropy is used for the Classification Model.
- Here I also used the concept of the transfer learning in the classification model in order to reduce the loss and improve the accuracy
- Here I am using a Deep Learning architectures CNN (Convolution Neural Network)

generally known as NN (Neural Network) and Transfer learning for detect the brain tumour and classify the types of tumour.

- Here I am using the tensor flow frame work for creation of the model. And also using libraries like OS, Numpy , Pandas and keras for implementation.
- The programming language I used here is Python 3.8.9.
- Each model is trained for 20-30 epochs using Adam optimizer and learning rate of 0.00001.
- The loss function used in Detection model is binary cross entropy loss and in classification model loss function is used sparse_categorical_crossentropy loss.
- The accuracy matrix is used as evaluation matrix for both model.

5.1.3 MODEL EVALUATION

- The Model evaluation is based on the two losses
- I . Binary cross entropy :- Binary cross entropy compares each of the predicted probabilities to actual class output which can be either 0 or 1. It then calculates the score that penalizes the probabilities based on the distance from the expected value. That means how close or far from the actual value.
- Equation of binary cross entropy is as below:-

$$\text{Log loss} = \frac{1}{N} \sum_{i=1}^{N} -\left(y_i * \log(p_i) + (1-y_i) * \log(1-p_i) \right)$$

Figure 5.1 Binary cross entropy

- II. Sparse categorical cross entropy:- It is Used as a loss function for multi-class classification model where the output label is assigned integer value (0, 1, 2, 3…).
- It is mostly used to classify the categorical data.
- It works on principle of the one hot encoding.

$$-\sum_{c=1}^{M} y_{o,c} \log(p_{o,c})$$

Figure 5.2 Sparse Categorical cross entropy.

- In the backpropogation of the model , this both losses are used in order to improve accuracy and decrease the loss.
- The Adam optimizer at learning rate(0.01) is used as gradient decent method.
- The dropout and data augmentation is also used in order to avoid over fitting.

- The accuracy table of model is as below.

Model	Loss Used	Training Loss	Testing Loss	Training accuracy	Testing accuracy
Detection	Binary cross	0.0000192	0.000087	99.21%	98.34%
Classification	Categorical	0.0000025	0.0000056	99.34%	99.01%

6

==========

6.0 CONCLUSION

6.1 CONCLUSION.
6.2 OVERALL ANALYSIS OF PROJECT.
6.3 PROBLEM ENCOUNTERED AND SOLUTIONS.
6.4 SUMMARY OF PROJECT.
6.5 LIMITATION AND FUTURE ENHANCEMENT.

==========

6.1 CONCLUSION.

- Here I have created the computerized method for detection and identification of the brain tumor using deep learning concepts.
- The input images which are slices of 3-D MRI volumes are pre-processed using an adaptive bilateral filtering technique for the elimination of noises that are present inside the original image.
- The Convolution neural network is used to extract features from image in order to detect and classify the images.
- Here I have also used concept of the transfer learning in order to increase accuracy.
- The both propose models have accuracy almost as 99% on both training and testing set.
- To avoid over fitting, we need to ensure we have plenty of testing and validation of data i.e. dataset is not generalized. This is solved by Data Augmentation and Dropout.
- The main goal is to help the doctor to detect the tumor in early stage so doctor can start treatment and cancer can prevent in the early stage of development

6.2 OVERALL ANAYLSIS OF PROJECT

- Since the first day of the project, I started learning a lot of new things from the respected internal guide.
- First of all, the first week was all about project selection. So, discussed about various domains and technology with the guide.
- Then, Selected one of the projects from the suggested projects by guide. Then started requirement gathering and analysis about the domain.
- Then I stared the planning about the project like which types of tasks are needed to be completed. And which task required how much of the time. Then I stared diagrams to get better idea of the project.
- And likewise continued my work on further modules and learned many new things while developing the project by self also from the internal guide.

- Had some difficulties in between in development part but the internal guide was very supportive and always up for any kind of help.

- So my overall experience for this project was very good and I learn a lots of things form our guide.

6.3 PROBLEM ENCOUNTERED AND POSSIBLE SOLUTION

Problem: Find the proper dataset for the problem.

Solution: I asked my internal guide about this problem and he suggested me to find out the dataset on kaggle.com.

Problem: Needed the higher GPUS and RAM to run the code

Solution: I search for the this problem on the web and find out to run the code on the cloud platforms.

6.4 SUMMARY OF THE PROJECT

- As We know this project about the brain tumor detection and classification on the human brain . The main motive is to help the doctor to detect the tumor in early stage so he can start the treatment earlier so it can increase the chances of patient.

6.5 LIMITATION AND FUTURE ENHANCEMENT.

Limitations:

1. Can’t detect on 3-D volumes.
2. Need 2-D slices.
3. Need some medium specification to run.
4. Some time may be slower

Future Enhancement:

1. Run on 3-D volume directly.
2. Segmentation and region analysis

REFERENCES

1. Brain tumor detection and classification using machine learning: a comprehensive survey
2. Keras Documentation
3. Tensor flow Documentation.
4. Image Processing techniques using tensor flow documentation
5. Convolution Neural Network.
6. Transfer learning.
7. Data Augmentation.
8. Loss functions.
9. Brain tumor information.
10. Brain Tumor Reserch paper
11. Dataset from kaggle
12. Python libraries documentation