NEURO DISEASE DETECTION AND ANALYSIS

A PROJECT REPORT

Submitted by

PRINCE DALAL (RA2011050010093)
PRATHMESH JOSHI (RA2011050010082)

Under the Guidance of

Dr. B. PRABHU KAVIN

(Assistant Professor, Department of Data Science and Business Systems)

In partial fulfillment of the Requirements for the Degree of

BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE AND ENGINEERING

with specialization in BLOCKCHAIN TECHNOLOGY

DEPARTMENT OF DATA SCIENCE AND BUSINESS SYSTEMS FACULTY OF ENGINEERING AND TECHNOLOGY SRM INSTITUTE OF SCIENCE AND TECHNOLOGY KATTANKULATHUR – 603203

NOVEMBER 2023

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR-603203

BONAFIDE CERTIFICATE

Certified that this project report titled "**NEURO DISEASE DETECTION AND ANALYSIS"** is the bonafide work of **PRINCE DALAL [Reg No: RA2011050010093] and PRATHMESH JOSHI [Reg No: RA2011050010082]** who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion for this or any other candidate.

Dr. B. Prabhu Kavin
SUPERVISOR
Assistant Professor
Department of Data Science and Business Systems

Dr. M. Lakshmi
HEAD OF THE DEPARTMENT
Professor
Department of Data Science and Business Systems

ACKNOWLEDGEMENT

We express our humble gratitude to **Dr C. Muthamizhchelvan**, Vice-Chancellor, SRM Institute of Science and Technology, for the facilities extended for the project work and his continued support.

We extend our sincere thanks to Dean-CET, SRM Institute of Science and Technology, **Dr.T.V.Gopal**, for his invaluable support.

We wish to thank **Dr Revathi Venkataraman**, Professor & Chairperson, School of Computing, SRM Institute of Science and Technology, for her support throughout the project work.

We are incredibly grateful to our Head of the Department, **Dr M. Lakshmi,** Professor, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for her suggestions and encouragement at all the stages of the project work.

We want to convey our heartfelt thanks to our program coordinator, **Dr.S.Ganesh Kumar**, Professor, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for his inputs during the project reviews and support.

We register our immeasurable thanks to our Faculty Advisor, **Dr. Tamizhselvan C**, Assistant Professor, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for leading and helping us to complete our course.

Our inexpressible respect and thanks to my guide, **Dr. B. Prabhu Kavin**, Assistant Professor, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for providing me with an opportunity to pursue my project under her mentorship. He provided me with the freedom and support to explore the research topics of my interest. His passion for solving problems and making a difference in the world has always been inspiring.

We sincerely thank the Data Science and Business System staff and students, SRM Institute of Science and Technology, for their help during our project. Finally, we would like to thank parents, family members, and friends for their unconditional love, constant support and encouragement.

PRINCE DALAL [RA2011050010093]
PRATHMESH JOSHI [RA2011050010082]

ABSTRACT

The human brain serves as the main controller of the humanoid system. Brain cancer arises when brain tumours continue to grow, and brain tumours are caused by abnormal brain cell growth and division. In the subject of human health, computer vision is crucial since it reduces the need for human judgement to provide accurate results. CT, X-ray, and MRI scans are the most reliable and secure common magnetic resonance imaging (MRI) procedures. MRI detects even the tiniest objects. Our study focuses on the use of many approaches for brain MRI-based brain cancer diagnosis. In order to remove any present noise, we preprocessed the data in this study using the bilateral filter (BF). Convolution Neural Network (CNN) segmentation and binary thresholding are necessary for precisely determining the tumor's location. There is use of datasets for training, validation, and testing. Using our tools, we will assess whether or not the patient has a brain tumour. The final findings will be assessed using a variety of performance metrics, including sensitivity, specificity, and accuracy. The objective outcome of the recommended effort is to outperform its counterparts.

KEYWORDS: Brain tumor, Magnetic resonance imaging, Adaptive Bilateral Filter, Convolution Neural Network.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In order to perform clinical analysis and medical intervention, as well as to visually portray the operation of certain organs or tissues, medical imaging is the method and technology of producing images of the inside of the body. Medical imaging aims to identify and cure illnesses by revealing interior structures that are covered by the skin and bones. In order to detect anomalies, medical imaging also creates a database of typical anatomy and physiology.

The use of computers to handle pictures is referred to as medical imaging processing. Numerous methods and procedures, including picture acquisition, storage, display, and communication, are part of this processing. The detection and management of disorders are pursued by this procedure. In order to make it simple to identify abnormalities, this procedure generates a data bank of the organs' typical structure and function. This procedure combines radiological and organic imaging using thermal and isotope imaging, sonography, magnetic resonance imaging, and electromagnetic energy (gamma and X-rays). Numerous additional technologies are employed to capture data regarding the position and operation of the body. Compared to modulates that generate pictures, such approaches have a lot more limits.

Using a computer to alter a digital image is known as an image processing technique. This method offers several advantages, including data storage, connectivity, flexibility, and adaptation. The development of various picture scaling methods has made it possible to maintain photos effectively. This method requires several sets of rules to be applied simultaneously in the pictures. Multiple dimensions can be handled for both 2D and 3D pictures.

1.1 BRAIN ANATOMY

One of the most prevalent and, consequently, deadly brain disorders that has destroyed many lives worldwide is the brain tumour. Cancer is a disease of the brain in which cancer cells grow within the tissues of the brain. A recent study on cancer claims that over a lakh people worldwide receive a brain tumour diagnosis each year.

Figures for brain tumour patients indicate unsatisfactory outcomes despite consistent attempts to address the difficulties associated with the tumours. In response to this, researchers are focusing on computer vision to get a deeper comprehension of the first phases of tumour development and how to combat them with cutting-edge therapeutic approaches.
Computed tomography (CT) scans and magnetic resonance imaging (MR) are the two most used methods to identify whether a tumour is present and where it is located in order to make treatment options. Because of its portability and increased capacity to provide high-definition pictures of diseased tissues, these two scans are still widely employed. Currently, a number of other treatments are available for tumours, including radiation therapy, chemotherapy, and surgery. The size, nature, and grade of the tumor shown in the MR picture are only a few of the numerous parameters that influence the therapy choice. Whether or whether cancer has spread to other parts of the body is also its responsibility.

Accurate identification of the specific type of brain disorder is crucial for treatment planning in order to reduce misdiagnoses. The accuracy is frequently haphazard while using computeraided diagnostic (CAD) tools. The primary goal of computer vision is to reduce the amount of time it takes for doctors to interpret images and to generate a dependable result in the form of an associate estimate. These developments improve the consistency and accuracy of medical diagnosis; still, segmenting an MR picture of a tumour and its surrounding tissue is a highly challenging task. Another challenge that makes computerised brain tumour identification and segmentation challenging is the appearance of tumours in certain places within the brain image without differentiating picture intensities.

1.2 MOTIVATION FOR THE WORK

Cells growing abnormally inside the brain or central spinal canal is known as a brain tumour. Certain tumours may be malignant, thus prompt detection and treatment are necessary. People may be afflicted with brain tumours without realising the risk since the specific source and set of symptoms are unknown. Malignant (containing cancer cells) and benign (without containing cancer cells) are the two types of primary brain tumours.

Cells that were developing and dividing improperly gave rise to brain tumours. According to diagnostic medical imaging methods, it seems to be a solid mass. Primary brain tumours and metastatic brain tumours are the two forms of brain tumours. A primary brain tumour is one that originates in the brain and tends to remain there, whereas a metastatic brain tumour is one that spreads across the brain after originating in another part of the body.

The location, size, and kind of a brain tumour determine its symptoms. It happens when the tumour exerts pressure on the surrounding cells while compressing them. Furthermore, it also happens when a tumour obstructs the fluid that circulates throughout the brain. Headache, nausea, vomiting, and difficulty balance and walking are frequent symptoms. Diagnostic imaging techniques like CT and MRI can identify brain tumours. Depending on the type of site and the requirement for an inspection, both methods offer benefits in terms of detection. We have opted to utilise MRI scans in this study due to their ease of examination and ability to provide precise calcification and foreign mass placement.

The method most frequently used to image brain tumours and determine their proximity is magnetic resonance imaging (MRI). Aside from several alternative techniques, the traditional approach for classifying CT and MR images and identifying tumour cells is still mostly supported for human evaluation. The primary reason MR images are employed is that they are non-destructive and non-ionizing. High-definition images provided by MR imaging are widely used to identify brain tumours. MRI scans can be T1-weighted, T2weighted, or flair-weighted. Numerous methods exist for processing photos, including feature extraction, pre-processing, image segmentation, image enhancements, and classifiers.

1.3 PROBLEM STATEMENT

Our research focuses on the automatic identification and categorization of brain tumours. MRI or CT scans are often used to analyse the anatomy of the brain. Tumour detection in brain magnetic resonance imaging is the paper's goal. The primary purpose of brain tumour detection is to support clinical diagnosis. The goal is to create an algorithm that, by integrating many processes, ensures the existence of a tumour, offering a reliable technique for tumour identification in MR brain pictures. Filtering, erosion, dilation, threshold, and tumour outlining techniques such edge detection are the techniques used.

This study focuses on the extraction of tumours from MR brain images and their simplified depiction so that everyone can comprehend it. The aim of this work is to present certain important information to users in an easier-to-understand style, particularly for the medical personnel who are attending to the patient. Developing an algorithm to extract the tumour picture from the MR brain imaging is the goal of this effort. The final picture will be able to offer details on the tumor's location, size, and dimensions. Its border will also provide us information about the tumour that will be helpful in a variety of situations, giving the staff a stronger foundation on which to determine the best course of action for curing the tumour. Finally, we detect whether the given MR brain image has tumor or not using Convolution Neural Network.

1.4 SCOPE

Our goal is to create an automated system that can improve, classify, and segment brain tumours. Healthcare professionals, including neurosurgeons, can use the system. It is anticipated that the system, which combines computer vision, pattern analysis, and image processing techniques, would increase brain tumour screening's effectiveness, efficiency, and sensitivity. Medical imaging projects' main objective is to minimise mistake while extracting relevant and reliable information from these pictures. The creation of supplemental tools that can aid in early diagnosis or tumour identification and location monitoring is made possible by the appropriate combination and parameterization of the stages.

CHAPTER 2

LITERATURE SURVEY

Robustness and accuracy of prediction algorithms are critical in medical diagnosis since the outcome is critical to patient care. Prediction is made using a variety of widely used classification and clustering techniques. To reduce an image's representation into a meaningful picture and facilitate analysis, medical images are clustered. The goal of a number of clustering and classification algorithms is to improve the diagnostic process's prediction accuracy in identifying anomalies.

We give a succinct overview of the various clustering techniques that have been put out between 2002 and 2018 in the literature review.We have read through twenty-five publications, each of which takes a different tack when it comes to segmentation in one or more parameters.

1. **Dr. M. Karnan and A. Sivaramakrishnan International Journal of Advanced Research In Computer and Communication Engineering, Vol. 2, Issue 4, April 2013, "A Novel Based Approach for Extraction Of Brain Tumour In MRI Images Using Soft Computing Techniques."**

A. Sivaramakrishnan proposed a creative and effective method for locating the area around a brain tumour based on an image that was processed using histogram equalisation and the fuzzy approach grouping algorithm. Principal factor analysis is used to lower the wavelet coefficient's extent, which results in the disintegration of pictures. The predicted FCM clustering algorithm's results correctly extracted the tumour region from the MR pictures.

2. **"Improved Edge Detection Algorithm for Brain Tumour Segmentation," Asra Aslam, Ekram Khan, M.M. Sufyan Beg, Procedia Computer Science, Volume 58, 2015, pp. 430–437, ISSN 1877-0509**.

   A brain-tumor segmentation method utilising increased edge detection, mostly based on Sobel feature detection, has been reported by M. M. Sufyan. The work they have given links the Sobel technique with binary thresholding operation, and it uses a secure contour procedure to dig different extents. Following that procedure, intensity values are used to remove cancer cells from the resulting image.

3. **Image Segmentation by Clustering Methods: Performance Analysis, by B. Sathaya and R. Manavalan, International Journal of Computer Applications (0975–8887), Volume 29–No.11, September 2011**.

   Different clustering methods, including K-means, improvised K-means, C-means, and improvised C-means algorithms, were presented by Sathya. In their study, an experimental analysis for large datasets made up of individual photos was given. They conducted many parametric tests to examine the implications that were found..

4. **Elchouemi, A., Singh, A.K., Prasad, P.W.C., Alsadoon, Abeer, and Devkota, B. (2018). Image Segmentation for Mathematical Morphological Reconstruction-Based Early Stage Brain Tumour Detection. 125, 115–123, Procedia Computer Science, 10.1016/j.procs.2017.12.017**.

According to B. Devkota et al. [4], aberrant tissues can be identified using morphological procedures and computer-aided detection (CAD). The morphological opening and closure procedures are the most effective segmentation technique available since they remove tumour regions with the fewest errors in the shortest amount of processing time.

CHAPTER 3

SYSTEM ARCHITECTURE AND WORK-FLOW

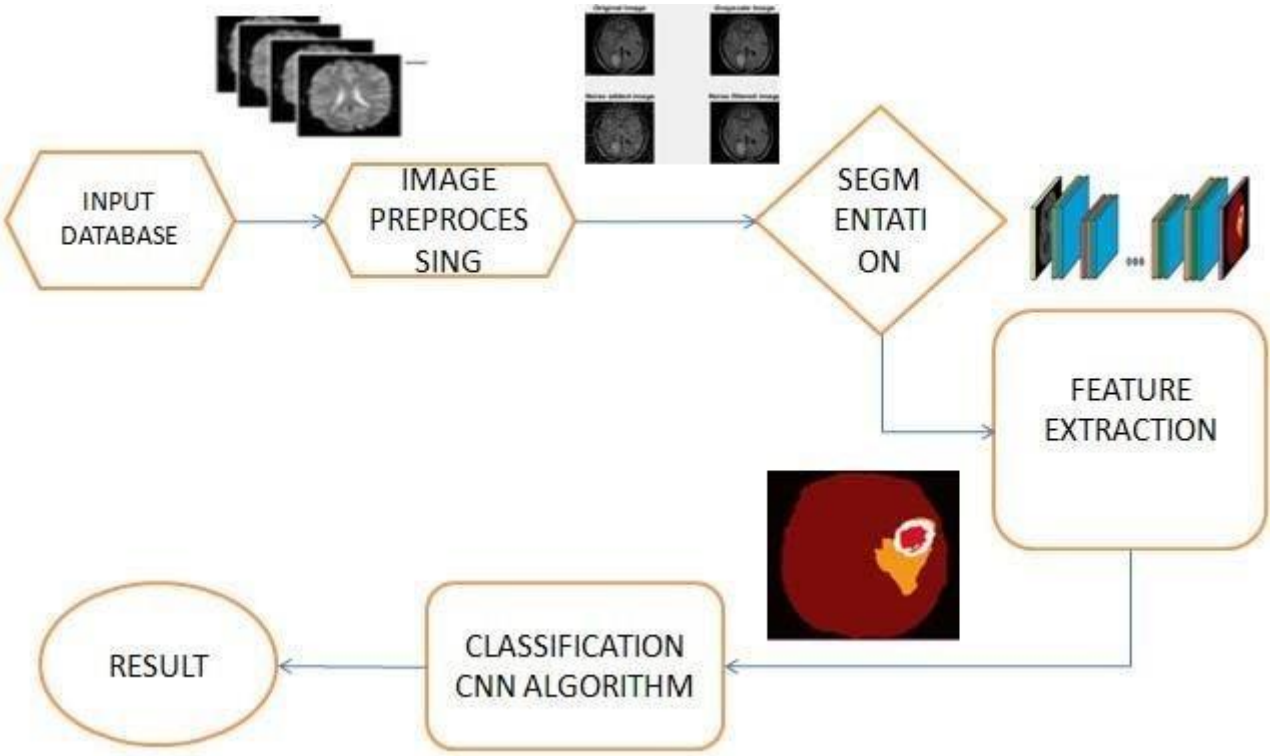

Fig 3.1 System Architecture

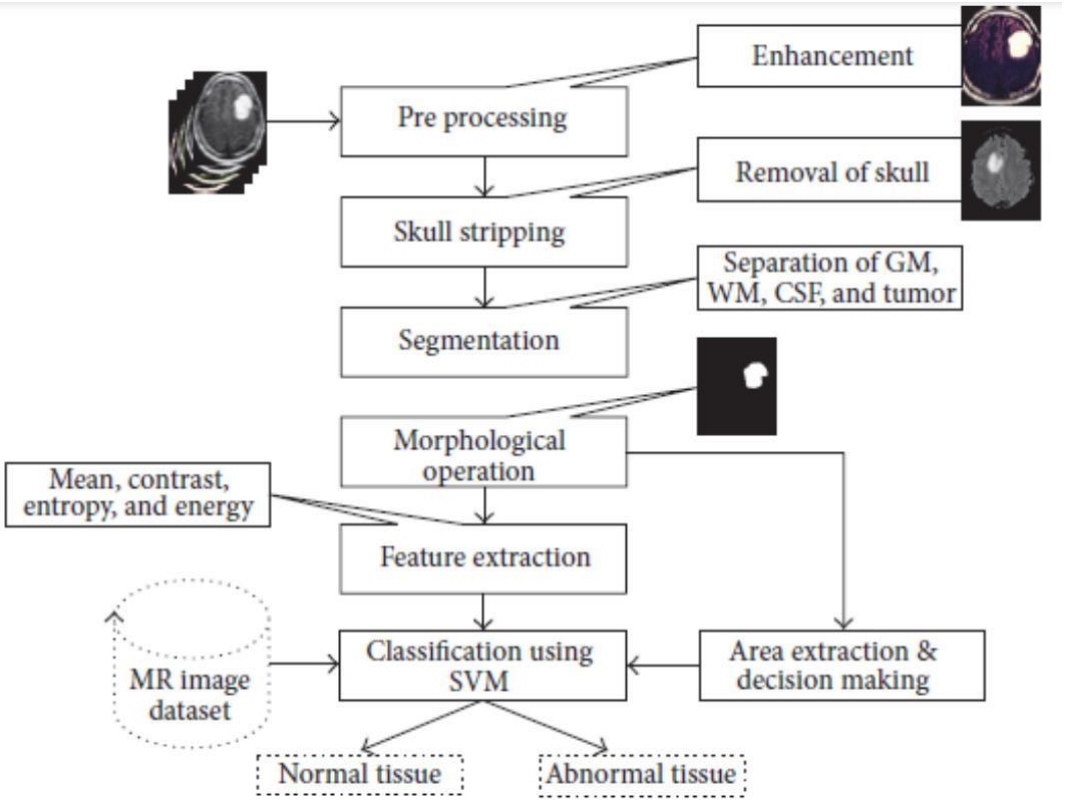

Fig 3.2 Existing work flow of brain tumor detection

• In the first stage, there is a computer-based procedures to detect tumor blocks and classify the type of tumor using Artificial Neural Network Algorithm for MRI images of different patients.

• The second stage involves the use of different image processing techniques such as histogram equalization, image segmentation, image enhancement, morphological operations and feature extraction are used for brain tumor detection in the MRI images for the cancer-affected patients.

• This work is introduced one automatic brain tumor detection method to increase the accuracy and decrease the diagnosis time.

Image preprocessing: This method uses scanned images from MRIs, which contain noise. Thus, cleaning up the input image of noise is our first goal. As stated in the system flow, high pass filters are used for preprocessing and noise reduction.

Segmentation: The simplest method for segmenting images based on regions is region growth. The approach in question is categorised as pixel-based picture segmentation as well since it necessitates the first seed point selection.

Morphological operation: The border regions of the brain pictures are extracted using the morphological operation. This technique works only with binary pictures since it simply rearranges the relative order of pixel values—not their mathematical values. The fundamental process of morphology is dilation and erosion. Dilation is the addition of pixels to an object's boundary region, whereas erosion is the removal of pixels from the same boundary region.

Feature Extraction: The pictures' edges are detected by the application of feature extraction. It is the process of gathering more complex picture details including contrast, colour, texture, and shape. • **Connected component labelling**: Each collection of connected pixels with the same gray-level values is given the same distinct area name after the connected components of an image have been identified.

Tumour Identification: During this stage, we are extracting characteristics from a collection of previously acquired brain MRIs. A knowledge base is produced in order to compare.

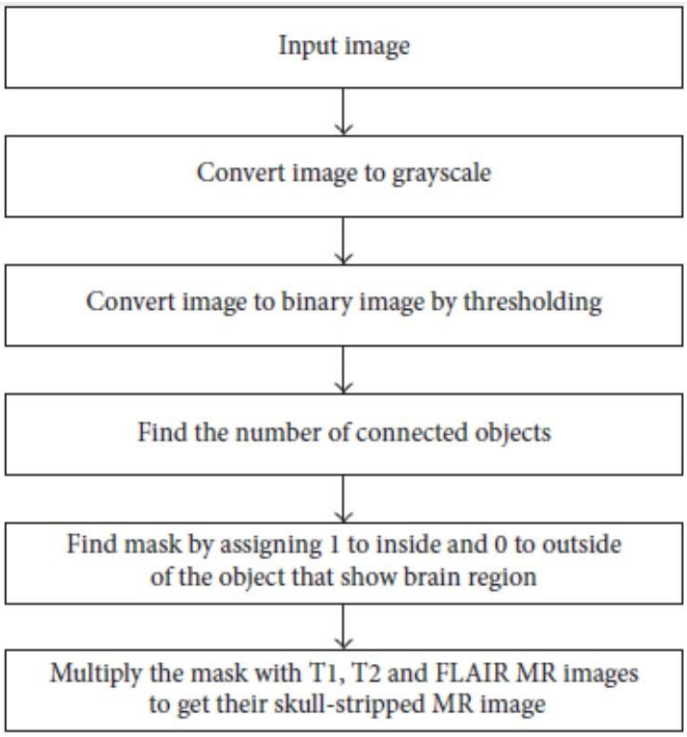

Fig 3.3 Steps used in skull stripping algorithm

a) In the first step we can take image as input. In the image we used tumor in the image and only fat and water tissues in the images.

b)In the second step convert image to grayscale
>Signal to noise
>Complexity of the code
>Learning image processing
>Difficulty of visualization

c)Then we convert image to binary image by thresholding.
Thresholding is the simplest method of image segmentation and the most common way to convert a grayscale image to binary image. In thresholding we select threshold value and then gray level value below the selected threshold value is classified as 0.and equal and greater then the threshold value are classified as 1.
d)Find the number of connected object

e)Find mask by assigning 1 to inside and 0 to outside of the object that show brain region.

f)Multiply the mask with T1,T2 and FLAIR MR images to get their skull stripped MR image

g)T1 & T2: weighted MRI
FLAIR: fluid attenuated inversion recovery weighted MRI.
Types of MRI images

T1: one tissue type is bright-FAT T2: two tissue types are bright-FAT and water

CHAPTER 4

METHODOLOGY

4.1 CONVOLUTIONAL NEURAL NETWORK

Convolutional neural networks, often known as CNNs or Conv Nets, are Deep Learning algorithms that, given an input picture, can distinguish between distinct objects and attributes by assigning each one a weight and bias that may be learned. A Conv Net requires a lot less pre-processing than other classification techniques. While filters are manually designed in more archaic approaches, Conv Nets may learn these properties and filters with sufficient training. Conv Net architecture was influenced by the way the visual cortex is organised, and it is comparable to the connection structure of neurons in the human brain. Only in a small area of the visual field known as the Receptive Field do individual neurons react to inputs. Such fields are grouped together and overlap to cover the whole visual field. Extraction of high-level characteristics from the input picture, such as edges, is the aim of the convolution operation. One Convolutional Layer is not the only place where Conv Nets may be used. Typically, Low-Level characteristics like edges, colour, gradient direction, etc. are captured by the first Cavalier. As more layers are added, the architecture also adjusts to the High-Level characteristics, providing us with a network that comprehends the photos in the dataset holistically, just as humans would.

4.2 MAGNETIC RESONANCE IMAGING (MRI)

Analysing and researching the human anatomy is possible with the use of the MRI diagnostic instrument. photos of medical patients taken over a range of electromagnetic spectrum frequencies. Each modality is appropriate for a particular purpose due to the abundance of sensors available and the physics underlying them. Magnetic fields that are around 10,000 times greater than the earth's magnetic field are used in magnetic resonance imaging (MRI) to create images. Compared to CT or ultrasound, the MRI yields more detailed pictures. A significant soft-tissue contrast is also provided by the MRI on maps of anatomical features. A magnetic resonance imaging scanner is used to examine the magnetic resonance of hydrogen (1H) nuclei in water and lipid.

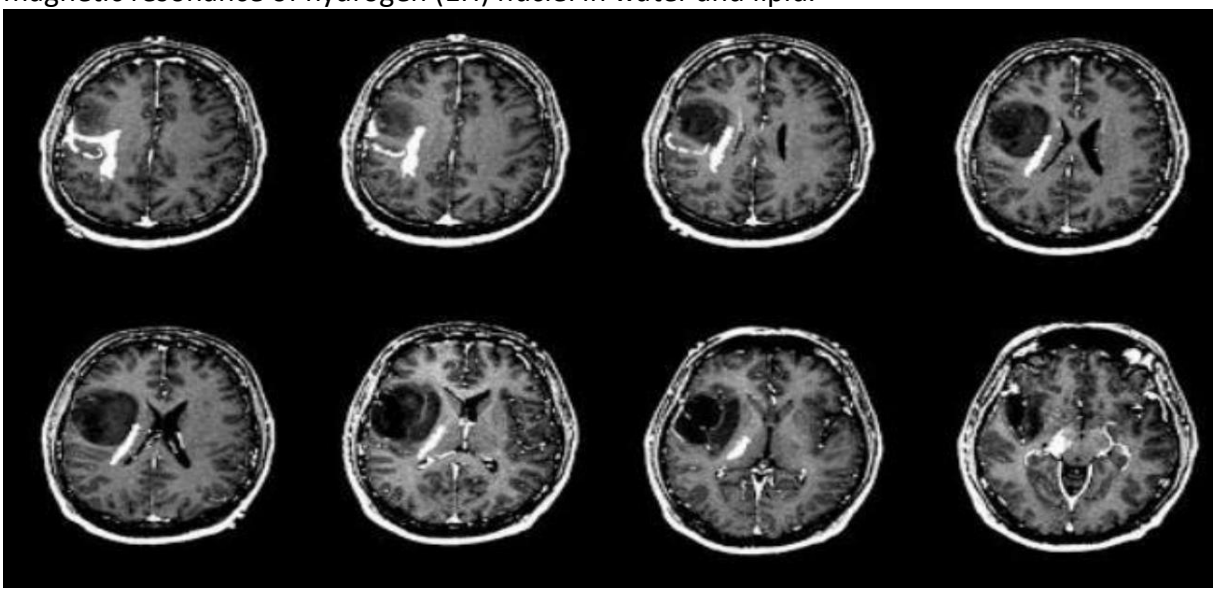

Fig 4.1 MRI of Human Brain

CHAPTER 5

CHALLENGES IN TUMOR CLASSIFICATION

Tumour identification is an extremely difficult process. Tumour location, form, and structure differ greatly between patients, making segmentation an extremely difficult process. The brain slices from many individuals are depicted in the picture below, which amply illustrates how the tumours vary from one another. It is evident that each of the eight individuals and photos below has a different tumour site. Even worse, there are differences in the intra-tumoral structure and morphology across the eight patients/images. As the pictures below show, the tumour may really be in more than one area. This does, in fact, illustrate how intricate automated segmentation is.

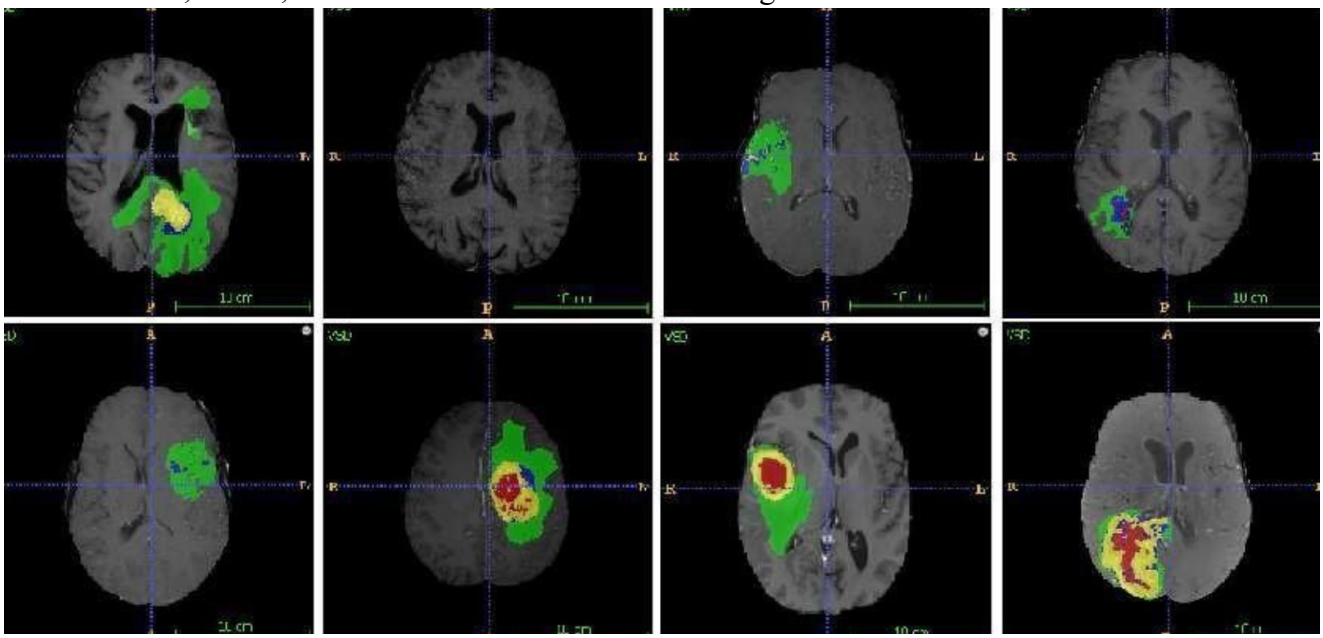

Fig 5.1 Tumor Classification

CHAPTER 6

HARDWARE AND SOFTWARE REQUIREMENTS

6.1 Hardware Requirements -

- Hard Disk : 500GB and Above
- RAM : 4GB and Above
- Processor : I3 and Above

6.2 Software Requirements –

- Operating System : Windows 7, 8, 10 (64 bit)
- Software : Python
- Tools : Anaconda (Jupyter Note Book IDE)

6.3 Software Description –

PYTHON

Python is an open-source, free computer language. Consequently, all it takes to begin using Python is a single installation. Not to mention, you may provide the community your own code. Another language that works across platforms is Python. What does this mean, then? Python may be installed and used on a variety of operating systems.
An excellent tool for visualization is Python. It offers amazing visuals using libraries like seaborn, bokeh, and Matplotlib.

PANDAS

Pandas is a well-liked Python data science library, and for good reason—among many other things, it provides strong, expressive, and adaptable data structures that facilitate data manipulation and analysis. One such structure is the Data Frame. Wes McKinney created the sophisticated data manipulation program known as Pandas. The Data Frame is its primary data structure, and it is based on the Num py package. Using rows of observations and columns of variables, data frames let you store and work with tabular data..

CHAPTER 7

UML, ACTIVITY AND SEQUENCE DIAGRAMS

7.1 UML DIAGRAM

A common, standardised modelling language used in software development is called Unified Modelling Language (UML). The asset management team is in charge of overseeing and developing this standard. A variety of graphical markup techniques are included in UML to help create visual models of software-intensive systems. It is employed in the definition, representation, alteration, building, and documentation of the artefacts of software-intensive, object-oriented systems under development. A universal modelling language that has been standardised in the software development industry is called Unified Modelling Language (UML). The asset management team is in charge of overseeing and developing this standard. It is employed in the definition, visualisation, construction, modification, and documentation of the artefacts of highly developed object-oriented software systems.

Utilising use case diagrams, one may visually represent the features offered by the system according to its users, their objectives, and any interdependencies among its use cases. There are two components to the use case diagram:
1.A use case is a set of steps that allow an object to be measured and represented as a horizontal ellipse.
2. Participants: People who engage in one or more interactions with the system as participants might be organisations, individuals, or external systems..

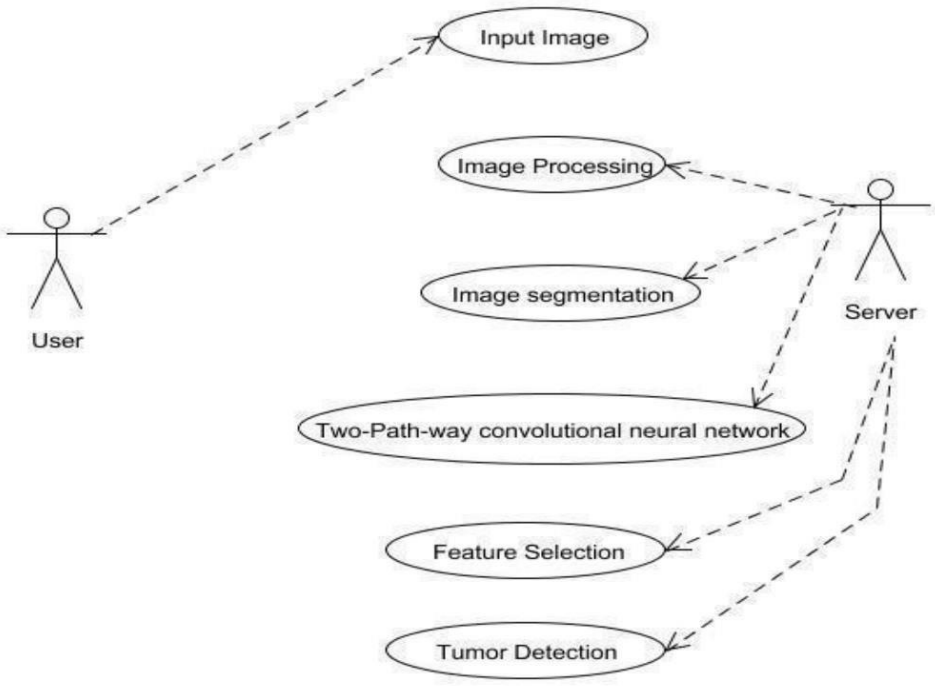

Fig 7.1 UML Representation

7.2 ACTIVITY DIAGRAM

An activity diagram is a graphical workflow model that allows for choice, iteration, and concurrency in a series of sequential tasks and actions. An activity diagram illustrates the total control flow.

The primary forms of shapes are:

Decisions are represented by diamonds, and activities are represented by rounded rectangles.
> Bars indicate when a concurrent activity begins or ends.
>A black circle denotes the workflow's beginning.
>A circle with an encircling signifies the workflow's conclusion..

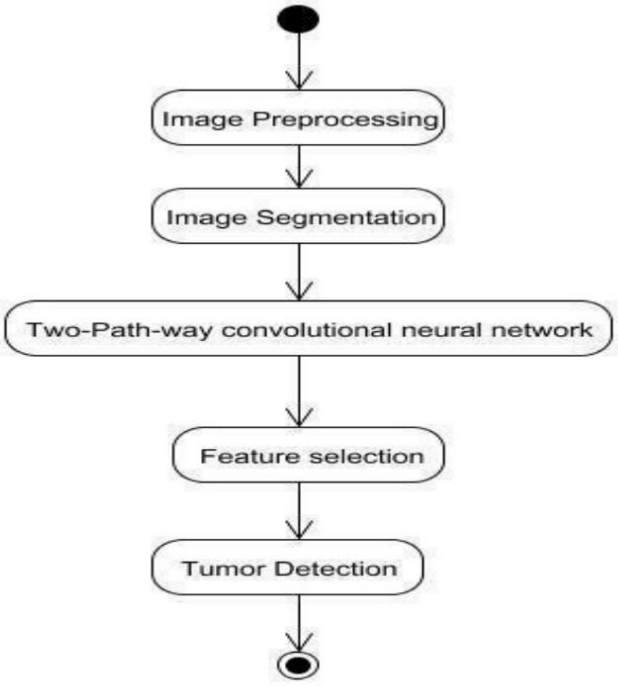

Fig 7.2 Activity Diagram

7.3 SEQUENCE DIAGRAM

An interaction diagram that displays the order and manner in which the processes interact is called a sequence diagram. This is the process of creating message sequence diagrams, also known as event diagrams, sequence diagrams, and event scenarios..

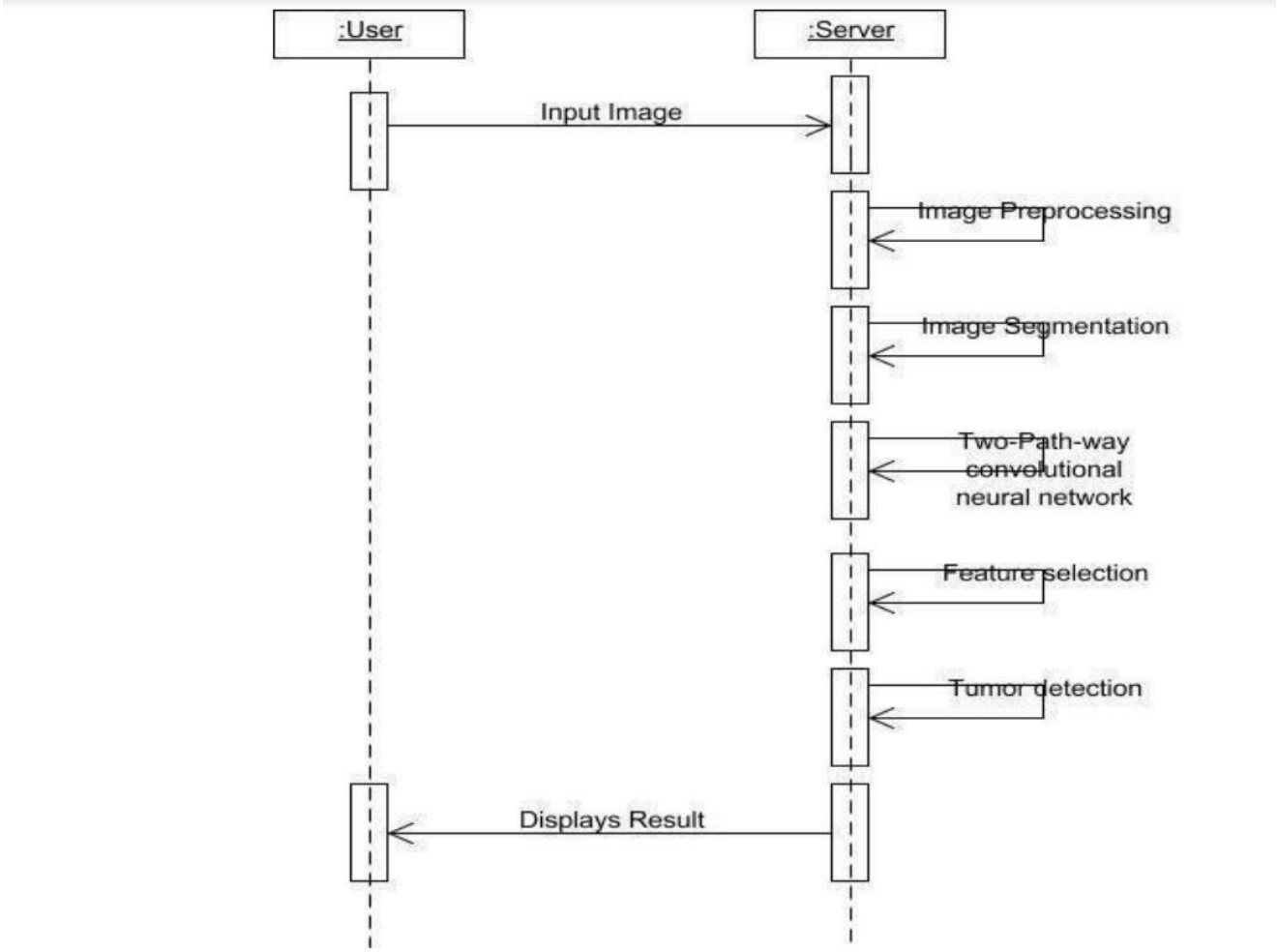

Fig 7.3 Sequence Diagram

CHAPTER 8

CODING AND IMPLEMENTATION

8.1 Setting up the Environment

```
import numpy as np  from
import tqdm import cv2
import os
import shutil
import itertools

import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelBinarizer from
sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, confusion_matrix
  import plotly.graph_objs as go
from plotly.offline import init_notebook_mode, iplot from plotly import tools

from keras.preprocessing.image import ImageDataGenerator from
keras.applications.vgg16 import VGG16, preprocess_input from keras import layers
from keras.models import Model, Sequential from keras.optimizers
import Adam, RMSprop from keras.callbacks import EarlyStopping

init_notebook_mode(connected=True)
RANDOM_SEED = 123
```

Right now all images are in one folder with yes and no subfolders. I will split the data into train, val and test folders which makes its easier to work for me. The new folder heirarchy will look as follows:

```
!apt-get install tree
clear_output() # create new
folders
!mkdir TRAIN TEST VAL TRAIN/YES TRAIN/NO TEST/YES TEST/NO VAL/YES VA
L/NO
!tree -d
```

```
.
├── TEST
│   ├── NO
│   └── YES
├── TRAIN
│   ├── NO
│   └── YES
└── VAL
    ├── NO
    └── YES

9 directories
```

```
IMG_PATH = '../input/brain-mri-images-for-brain-tumor-detection/bra in_tumor_dataset/'
# split the data by train/val/test for CLASS in
os.listdir(IMG_PATH):     if not CLASS.startswith('.'):
      IMG_NUM = len(os.listdir(IMG_PATH + CLASS))
      for (n, FILE_NAME) in enumerate(os.listdir(IMG_PATH + CLASS )):
img = IMG_PATH + CLASS + '/' + FILE_NAME
if n < 5:
            shutil.copy(img, 'TEST/' + CLASS.upper() + '/' + FI LE_NAME)
 elif n < 0.8*IMG_NUM:
            shutil.copy(img, 'TRAIN/'+ CLASS.upper() + '/' + FI LE_NAME)
else:
            shutil.copy(img, 'VAL/'+ CLASS.upper() + '/' + FILE
_NAME)
```

8.2 Data Import and Preprocessing

```
def load_data(dir_path, img_size=(100,100)):
    """
    Load resized images as np.arrays to workspace
    """     X = []     y = []     i = 0     labels = dict()     for path in
tqdm(sorted(os.listdir(dir_path))):         if not path.startswith('.'):
labels[i] = path
for file in os.listdir(dir_path + path):
            if not file.startswith('.'):
                img = cv2.imread(dir_path + path + '/' + file)
                X.append(img)
                y.append(i)
          i += 1     X = np.array(X)
y = np.array(y)
    print(f'{len(X)} images loaded from {dir_path} directory.')     return X, y, labels


def plot_confusion_matrix(cm, classes,normalize=False, c map=plt.cm.Blues):
title='Confusion matrix',
    """
    This function prints and plots the confusion matrix.
    Normalization can be applied by setting `normalize=True`.
    """
    plt.figure(figsize = (6,6))
    plt.imshow(cm, interpolation='nearest', cmap=cmap)     plt.title(title)
plt.colorbar()
    tick_marks = np.arange(len(classes))     plt.xticks(tick_marks, classes,
rotation=90)     plt.yticks(tick_marks, classes)     if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]
     thresh = cm.max() / 2.     cm = np.round(cm,2)
for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, cm[i, j],
        horizontalalignment="center",
        color="white" if cm[i, j] > thresh else "black")     plt.tight_layout()
```

```
TEST_DIR = 'TEST/'
VAL_DIR = 'VAL/'
IMG_SIZE = (224,224)


# use predefined function to load the image data into                    workspace
X_train, y_train, labels = load_data(TRAIN_DIR, IMG_SIZE)
   X_test, y_test, _ = load_data(TEST_DIR, IMG_SIZE)  X_val, y_val, _ = load_data(VAL_DIR, IMG_SIZE)
```

```
100%|██████████| 2/2 [00:00<00:00,  4.76it/s]
100%|██████████| 2/2 [00:00<00:00, 93.17it/s]
100%|██████████| 2/2 [00:00<00:00, 21.17it/s]


193 images loaded from TRAIN/ directory.
10 images loaded from TEST/ directory.
50 images loaded from VAL/ directory.
```

```
def plot_samples(X, y, labels_dict, n=50):
   """
   Creates a gridplot for desired number of images (n) from the specified set
   """    for index in range(len(labels_dict)):
      imgs = X[np.argwhere(y == index)][:n]       j = 10      i = int(n/j)

      plt.figure(figsize=(15,6))
c = 1
for img in imgs:
         plt.subplot(i,j,c) , plt.imshow(img[0]),
         plt.xticks([])  ,plt.yticks([])
c += 1
      plt.suptitle('Tumor: {}'.format(labels_dict[index]))       plt.show()
plot_samples(X_train, y_train, labels, 30)
```

```
RATIO_LIST = [] for set in (X_train, X_test, X_val):
for img in set:
        RATIO_LIST.append(img.shape[1]/img.shape[0])

plt.hist(RATIO_LIST)
plt.title('Distribution of Image Ratios') plt.xlabel('Ratio
Value') plt.ylabel('Count') plt.show()
```

```
def crop_imgs(set_name, add_pixels_value=0):
    """
    Finds the extreme points on the image and crops the rectangular out of them     """     set_new = []
for img in set_name:
        gray = cv2.cvtColor(img, cv2.COLOR_RGB2GRAY)
gray = cv2.GaussianBlur(gray, (5, 5), 0)

        # threshold the image, then perform a series of erosions +
        # dilations to remove any small regions of noise
        thresh = cv2.threshold(gray, 45, 255, cv2.THRESH_BINARY)[1]
thresh = cv2.erode(thresh, None, iterations=2)        thresh = cv2.dilate(thresh, None, iterations=2)

        # find contours in thresholded image, then grab the largest one
cnts = cv2.findContours(thresh.copy(), cv2.RETR_EXTERNAL, cv2.CH AIN_APPROX_SIMPLE)
        cnts = imutils.grab_contours(cnts)
c = max(cnts, key=cv2.contourArea)

        # find the extreme points
 extLeft = tuple(c[c[:, :, 0].argmin()][0])
extRight = tuple(c[c[:, :, 0].argmax()][0])
extTop = tuple(c[c[:, :, 1].argmin()][0])
extBot = tuple(c[c[:, :, 1].argmax()][0])

        ADD_PIXELS = add_pixels_value
        new_img = img[extTop[1]-ADD_PIXELS:extBot[1]+ADD_PIXELS, extLeft
[0]-ADD_PIXELS:extRight[0]+ADD_PIXELS].copy()
set_new.append(new_img)

    return np.array(set_new)
```

```
plt.figure(figsize=(15,6)) plt.subplot(141)
plt.imshow(img) plt.xticks([]) plt.yticks([])
plt.title('Step 1. Get the original image') plt.subplot(142)
plt.imshow(img_cnt) plt.xticks([]) plt.yticks([])
plt.title('Step 2. Find the biggest contour') plt.subplot(143)
plt.imshow(img_pnt) plt.xticks([]) plt.yticks([])
plt.title('Step 3. Find the extreme points') plt.subplot(144)
plt.imshow(new_img) plt.xticks([]) plt.yticks([])
plt.title('Step 4. Crop the image') plt.show()
```

```
# apply this for each set
X_train_crop = crop_imgs(set_name=X_train)
X_val_crop = crop_imgs(set_name=X_val)
X_test_crop = crop_imgs(set_name=X_test)
```

```
plot_samples(X_train_crop, y_train, labels, 30)
```

The next step would be resizing images to (224,224) and applying preprocessing needed for VGG-16 model input.

```
X_train_prep = preprocess_imgs(set_name=X_train_crop, img_size=IMG_SIZE)
X_test_prep = preprocess_imgs(set_name=X_test_crop, img_size=IMG_SIZE)
X_val_prep = preprocess_imgs(set_name=X_val_crop, img_size=IMG_SIZE)
```

8.3 CNN MODEL

8.3.1 Data Augmentation

Demo

```
demo_datagen = ImageDataGenerator(
rotation_range=15,    width_shift_range=0.05,
height_shift_range=0.05,    rescale=1./255,
shear_range=0.05,
   brightness_range=[0.1, 1.5],    horizontal_flip=True,
vertical_flip=True
)
os.mkdir('preview') x =
X_train_crop[0]
x = x.reshape((1,) + x.shape)
 i = 0
for batch in demo_datagen.flow(x, batch_size=1, save_to_dir='preview', save_pr efix='aug_img',
save_format='jpg'):    i += 1    if i > 20:        break
plt.imshow(X_train_crop[0]) plt.xticks([])
plt.yticks([])
plt.title('Original Image')
plt.show()

plt.figure(figsize=(15,6)) i = 1 for img in
os.listdir('preview/'):
   img = cv2.cv2.imread('preview/' + img)    img = cv2.cvtColor(img,
cv2.COLOR_BGR2RGB)    plt.subplot(3,7,i)    plt.imshow(img)
plt.xticks([])    plt.yticks([])    i += 1    if i > 3*7:        break
plt.suptitle('Augemented Images') plt.show()
```

Apply

```
TRAIN_DIR = 'TRAIN_CROP/'
VAL_DIR = 'VAL_CROP/'

train_datagen = ImageDataGenerator( rotation_range=15,
width_shift_range=0.1,
    height_shift_range=0.1,

    shear_range=0.1,
    brightness_range=[0.5, 1.5], horizontal_flip=True,
vertical_flip=True,
    preprocessing_function=preprocess_input

)

test_datagen = ImageDataGenerator( preprocessing_function=preprocess_input
)



train_generator = train_datagen.flow_from_directory( TRAIN_DIR,
color_mode='rgb', target_size=IMG_SIZE, batch_size=32,
class_mode='binary', seed=RANDOM_SEED
)



validation_generator = test_datagen.flow_from_directory(VAL_DIR,
color_mode='rgb', target_size=IMG_SIZE, batch_size=16,
class_mode='binary', seed=RANDOM_SEED)
```

8.3.2 Model Building

```
# load base model
vgg16_weight_path = '../input/keras-pretrained-models/vgg16_weights_tf_dim_order ing_tf_kernels_notop.h5'
base_model = VGG16(
   weights=vgg16_weight_path,     include_top=False,
   input_shape=IMG_SIZE + (3,) )
```

```
NUM_CLASSES = 1

model = Sequential() model.add(base_model)
model.add(layers.Flatten()) model.add(layers.Dropout(0.5))
model.add(layers.Dense(NUM_CLASSES, activation='sigmoid'))

model.layers[0].trainable = False

model.compile(
   loss='binary_crossentropy',     optimizer=RMSprop(lr=1e-4),
metrics=['accuracy']
)
model.summary()
```

```
_________________________________________________________________ Layer (type)
Output Shape                                                      Param #
================================================================= vgg16 (Model)
(None, 7, 7, 512)                                                 14714688
_________________________________________________________________ flatten_1 (Flatten)
(None, 25088)                                                     0
_________________________________________________________________ dropout_1 (Dropout)
(None, 25088)                                                     0
_________________________________________________________________ dense_1 (Dense)
(None, 1)                                                         25089
=================================================================
Total params: 14,739,777
Trainable params: 25,089
Non-trainable params: 14,714,688
_________________________________________________________________
```

8.3.3 Model Performance

```
predictions = model.predict(X_val_prep)
predictions = [1 if x>0.5 else 0 for x in predictions]

accuracy = accuracy_score(y_val, predictions) print('Val Accuracy = %.2f' %
accuracy)

confusion_mtx = confusion_matrix(y_val, predictions)
cm = plot_confusion_matrix(confusion_mtx, classes = list(labels.items()), norm alize=False)
```

```
Test Accuracy = 1.00
```

CHAPTER 9

OUTPUT

Data Import and Processing

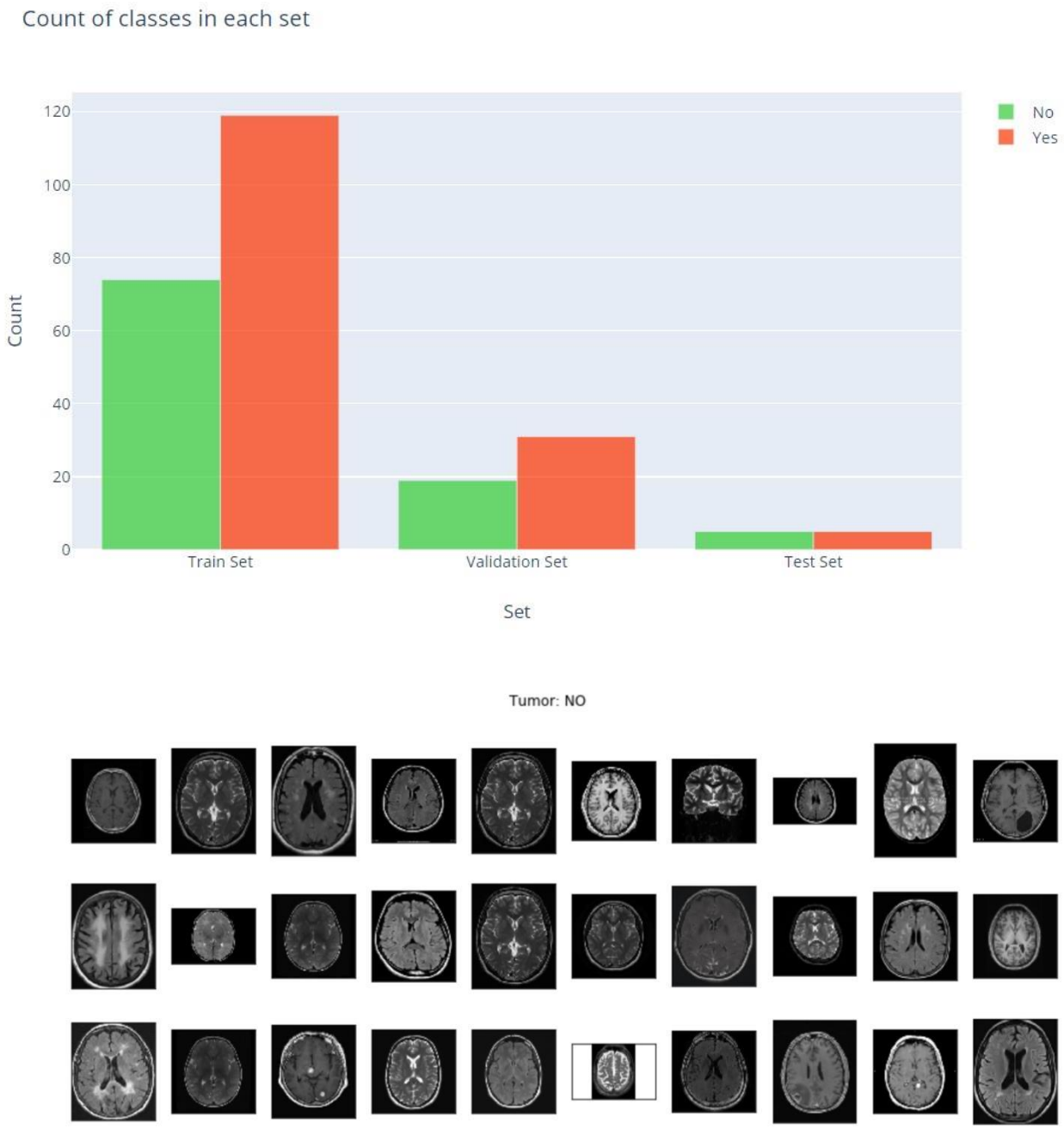

Fig 9.1 MRI of Brain without Tumor

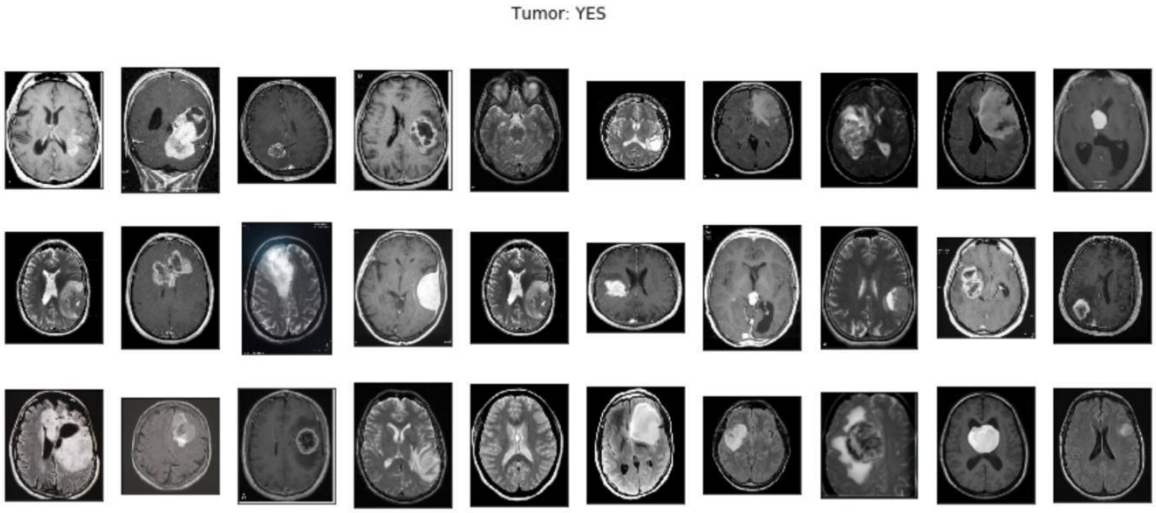

Fig 9.2 MRI of Brain with Tumor

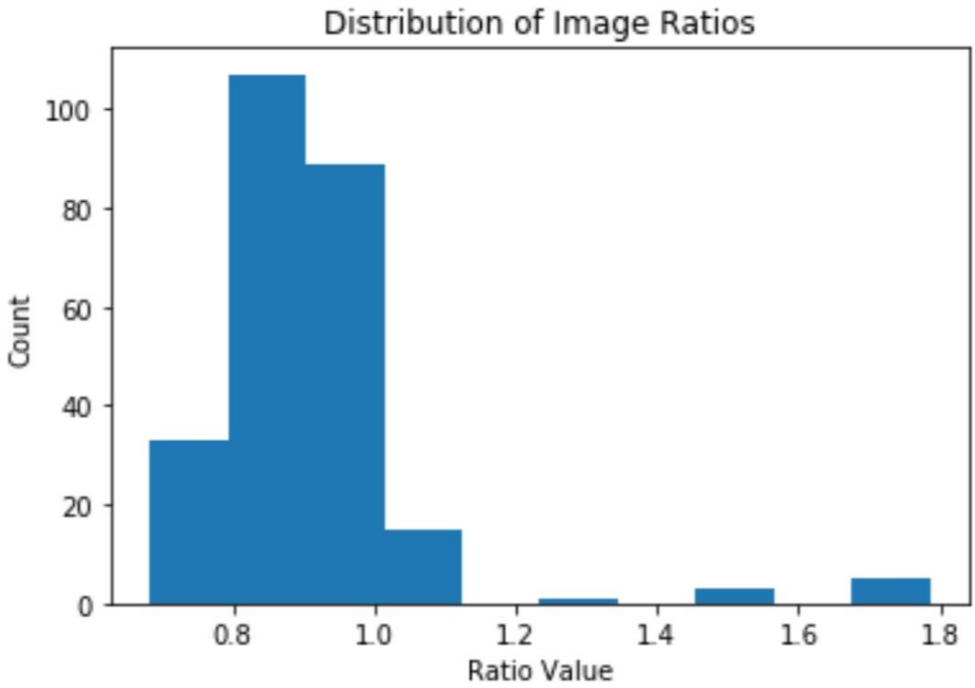

Fig 9.3 Count- Ratio Value Graph

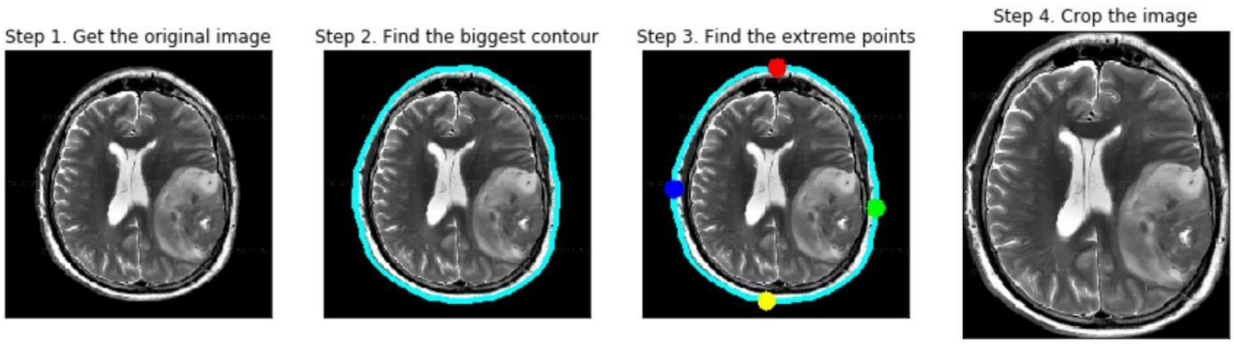

Fig 9.4 Growing Tumor Images

Data Augmentation

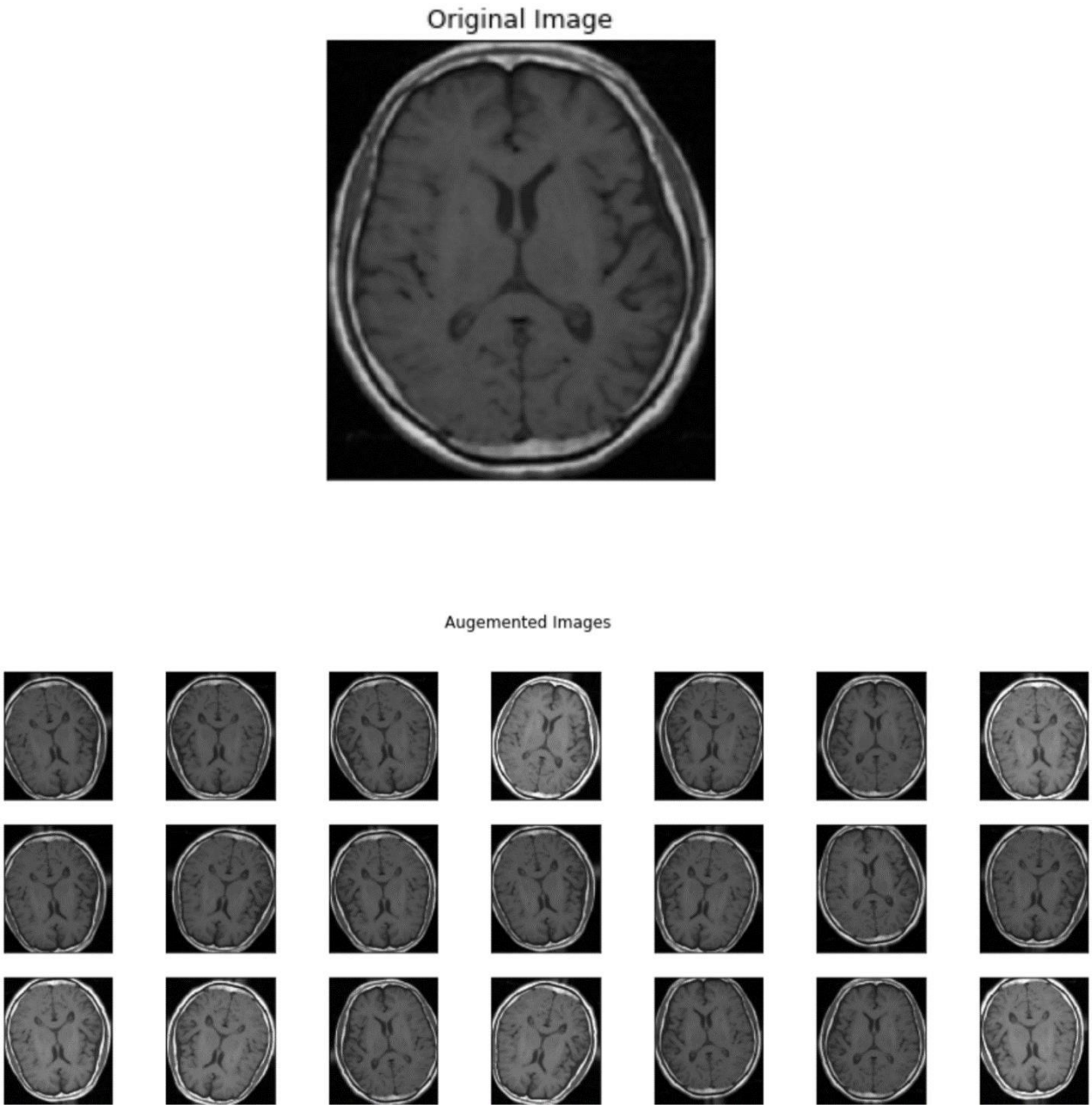

Fig 9.5 Augmentation Images

Model Performance

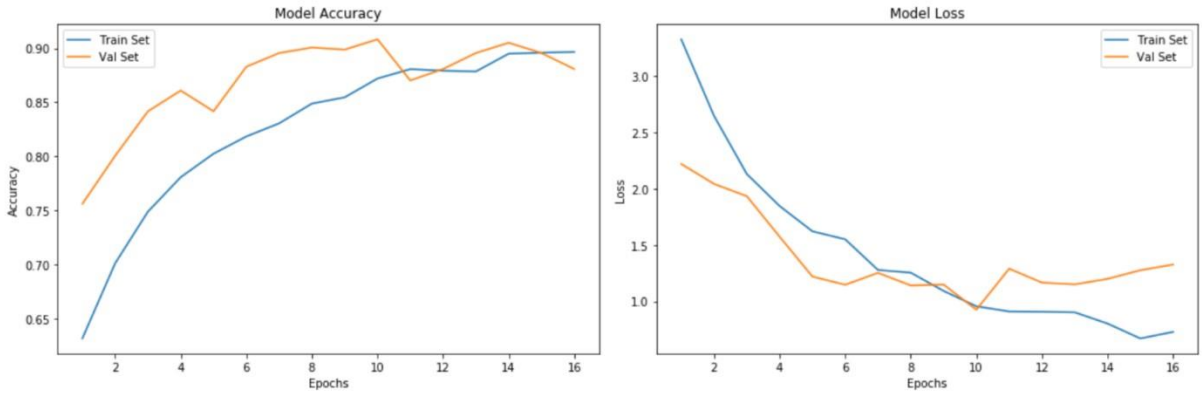

Fig 9.6 Model Performance

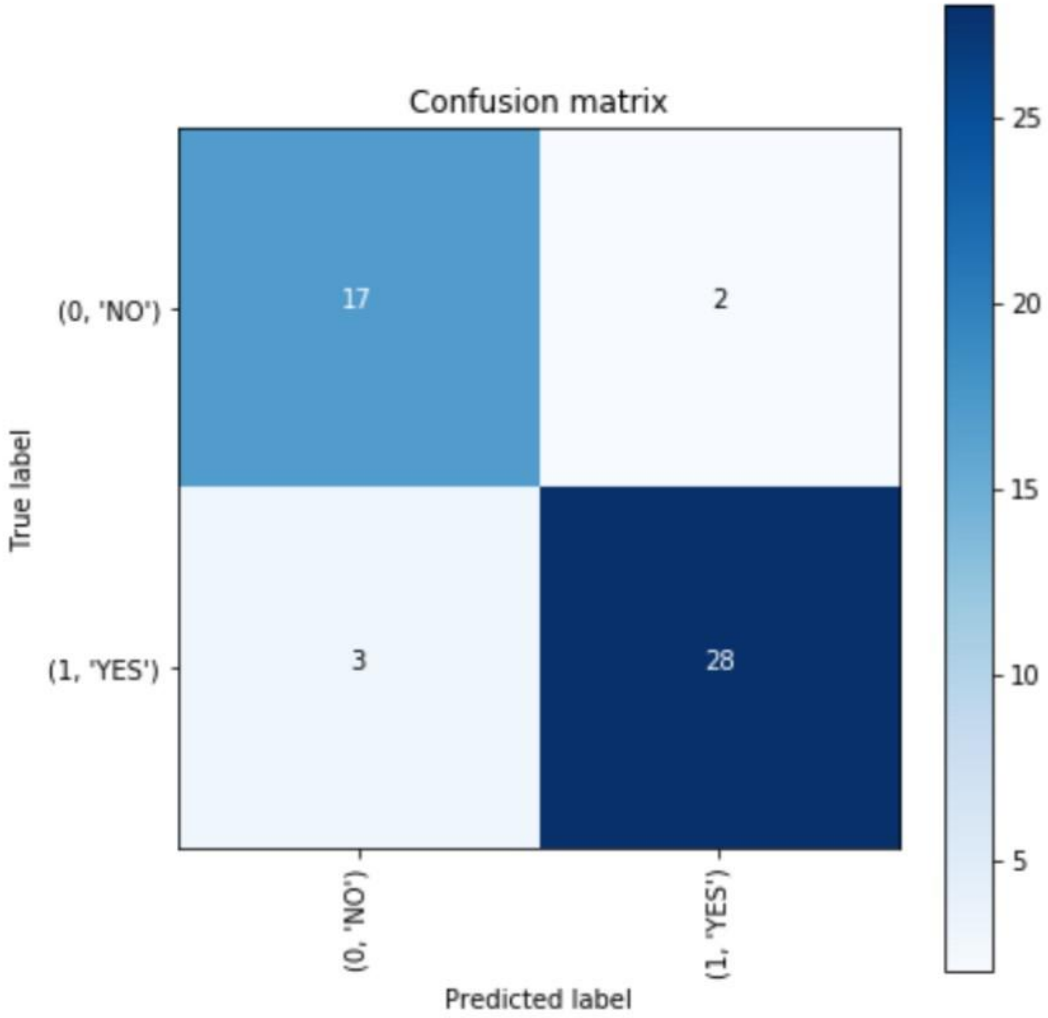

Fig 9.7 Confusion Matrix

CHAPTER 10

CONCLUSION

We focused on feature-based existing work in brain tumor identification. We examine image processing techniques such as segmentation, pre-processing, extraction of features, and classification in feature-based learning. Additionally, research CNN and VGG16, two deep learning approaches .This system detects if a tumor is there or not; if it is, the model returns yes; if not, it returns no. and the CNN and the VGG 16 Model have been compared. In terms of comparison, VGG 16 performs better than CNN. Even yet, there's always room for progress in this sector of development, thus not every activity is considered flawless. I've picked up a ton of information and understanding regarding the topic of development.

CHAPTER 11

REFERENCES

[1] "Tumor detection in MR images using one class immune feature weighted SVMs," L. Guo, L. Zhao, Y. Wu, Y. Li, G. Xu, and Q. Yan, IEEE Transactions on Magnetics, vol. 47, no. 10, pp. 3849– 3852, 2011.

[2]In the International Journal of Engineering Research and Applications, volume 3, pages 1686– 1690, 2013, R. Kumari wrote, "SVM classification an approach on detecting abnormality in brain MRI images."

[3]DICOM Example Image Collections, available at http://www.osirix-viewer.com.

[4] "Simulated Brain Database: Brain web" The website brainweb.bic.mni.mcgill.ca/cgi/brainweb1.

[5] Accessible over the internet at www.cancer.ca/~/media/CCE 10/08/15.

[6] J. C. Buckner, P. D. Brown, B. P. O'Neill, F. B. Meyer, C. J. Wetmore, J. H. Uhm, "Central nervous system tumors." In October 2007, Mayo Clinic Proceedings, Vol. 82, No. 10, pp. 1271–1286.

(7) Deepa and Singh Akansha (2016) reviewed the use of tomography in the detection of brain tumors. The International Conference on Sustainable Global Development via Computing (INDIACom)

[8] Squire's Fundamentals of Radiology, Sixth Edition, R. A. Novellines, M. D. UPR, 2004.

[9] D. C. Preston (2006). Brain and Spine Magnetic Resonance Imaging (MRI) from Basics. CaseMeded at Case.Edu.

[10] American College of Radiology Glossary of MR Terms, Hendrik RE. (2005)..

IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 4, pp. 1451–1458, 2015. A. Demirhan, M. Toru, and I. Guler, "Segmentation of tumour and edoema along with healthy tissues of brain using wavelets and neural networks."

[12] Bahadure, A.K. Nilesh Bhaskarrao (2017, March 6). taken from the publication https://www.hindawi.com/journals/ijbi/2017/9749108/.

[13] Mohsin, Sajjad, Malik, and Abdullah, A. H. (2012) "Effective method of applying morphological operations to MRI skull stripping after false background detection to detect brain tumour," International Journal of Information and Education Technology, vol. 2, no. 4, pp. 335–337.

[14] On April 4, 2017, Gavale, P. M., Aher, P. V., & Wani, D. V. taken from the PDF file IRJETV4I462.pdf from https://www.irjet.net/archives/V4/i4.

[15] The article "State of the art survey on MRI brain tumour segmentation" was published in 2013 by N. Gordillo, E. Montseny, and P. Sobrevilla in Magnetic Resonance Imaging, vol. 31, no. 8, pp. 1426– 1438.

[16] On November 3, 2016, Samantaray, M. from http://ieeexplore.ieee.org/document/7727089/ was retrieved.

[17] Nandi, A. (April 11, 2016) taken from the website http://ieeexplore.ieee.org/document/7449892/

The paper "Morphology based enhancement and skull stripping of MRI brain images" was presented by C. C. Benson and V. L. Lajish in the Proceedings of the International Conference on Intelligent Computing Applications (ICICA '14), which took place in Tamilnadu, India in March 2014.

[19] S. Z. Oo and A. S. Khaing, International Journal of Research in Engineering and Technology, vol. 3, no. 3, pp. 367–374, 2014, "Brain tumour detection and segmentation using watershed segmentation and morphological operation."

[20]The article "Skull stripping magnetic resonance images brain images: region growing versus mathematical morphology" was published in the International Journal of Computer Information Systems and Industrial Management Applications in 2011. It was written by R. Roslan, N. Jamil, and R. Mahmud.