Detecting Brain Tumor Using Deep Neural Networks from MRI Images

by

A.S.M. Imamuzzaman
20141034
Redwan Islam Sakline
16201013
Sayed Rafi Junaed
17101064
Mohammad Iqbal Hossain
17101279
Dipto Das
17101135

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
June 2021

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

A.S.M. Imamuzzaman
20141034

Redwan Islam Sakline
16201013

Rafi Junaid

Sayed Rafi Junaed
17101064

M. Iqbal Hossain

Mohammad Iqbal Hossain
17101279

Dipto Das

Dipto Das
17101135

Approval

The thesis/project titled "Detection and 3D Visualization of Brain Tumor Using Machine Learning and 3D Visualization Technique" submitted by

1. A.S.M. Imamuzzaman (20141034)
2. Redwan Islam Sakline (16201013)
3. Sayed Rafi Junaed (17101064)
4. Mohammad Iqbal Hossain (17101279)
5. Dipto Das (17101135)

Of Fall, 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on August 23, 2015.

Examining Committee:
Supervisor:
(Member)

Dr. Md. Ashraful Alam , PhD
Assistant Professor
School of Data and Science
Department of Computer Science and Engineering
Brac University

Thesis Coordinator:
(Member)

Md. Golam Rabiul Alam
Associate Professor
School of Data and Science
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Mahbubul Alam Majumdar, PhD
Professor and Dean
School of Data and Science
Department of Computer Science and Engineering
Brac University

Abstract

A brain tumor is a collection of abnormal cells growth in brain. It is a neurological disease which causes great damage and affects other healthy cells of brain. It can be cancerous or non-cancerous. Nowadays, people are more concern about their health issues. So, in this thesis paper we will design and implement an efficient machine learning approach to detect brain tumor from image data. Moreover, the proposed model approaches VGG16 and ResNet50 architectural model of Convolutional Neural Network (CNN). Through this model a neurosurgeon can easily detect the brain tumor of a patient with more efficiency. Our proposed model uses MRI images, and we also make a comparison between the two architectures of CNN.

Keywords: CNN, VGG16, ResNet50, Brain Tumor

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruption.
Secondly, to our supervisor Dr. Md Ashraful Alam sir for his kind support and advice in our work. He helped us whenever we needed help.

And finally to our parents without their throughout sup-port it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

Declaration **i**

Approval **ii**

Abstract **iii**

Acknowledgment **iv**

Table of Contents **v**

List of Figures **vii**

List of Tables **viii**

Nomenclature **ix**

1 Introduction **1**
1.1 Background 1
1.2 Literature Review 2
1.3 Objective 3
1.4 Overview 3
1.4.1 Brain 3
1.4.2 Cerebrum 5
1.4.3 Cerebellum 6
1.4.4 Brain Stem 7
1.4.5 Mid Brain 7
1.4.6 Pons 8
1.4.7 Meddulla Oblongata 8
1.4.8 The Lobs of The Brain 8
1.4.9 Frontal Lobe 8
1.4.10 Parietal lobe 9
1.4.11 Occipital lobe 9
1.4.12 Temporal lobe 9
1.4.13 Brain Tumor 10
1.4.14 Primary Brain Tumor 11
1.4.15 Glomas 11
1.4.16 Meningioma 11
1.4.17 Secondary Brain Tumor 11
1.4.18 MRI 12

2 The principles of Deep Learning, Artificial Neural Networks, CNN, Image Processing, Resnet-50 and VGG16 14
2.1 Deep Learning . . . 14
2.2 Artificial Neural Networks . . . 15
2.3 Convolution Neural Network . . . 16
2.4 Convolutional . . . 17
2.5 Receptive field . . . 17
2.6 Weights . . . 17
2.7 The Operation of Convolutional Neural Networks . . . 18
2.7.1 Input image . . . 18
2.7.2 Convolution Layer . . . 18
2.7.3 Padding . . . 19
2.7.4 Non Linearity (ReLU) . . . 20
2.7.5 Pooling Layer . . . 20
2.7.6 Fully Connected Layer . . . 20
2.8 Image Processing . . . 21
2.9 Stages of Digital Image Processing . . . 22
2.9.1 Resnet-50 . . . 24
2.9.2 VGG-16 . . . 26

3 Proposed Method 28
3.1 Preparation of the Dataset . . . 28
3.2 Implementation of VGG16 and ResNet50 architecture of CNN . . . 28
3.2.1 Input Image Layer in both VGG16 and ResNet50 . . . 31
3.2.2 VGG16 and ResNet50 Convolution Later . . . 32
3.2.3 VGG16 and ResNet50 Padding . . . 32
3.2.4 VGG16 and ResNet50 Relu Layer . . . 32
3.2.5 VGG16 and ResNet50 . . . 32
3.2.6 VGG16 and ResNet50 Flatten . . . 32
3.2.7 VGG16 and ResNet50 Classification and Training . . . 33

4 Result Analysis 34

5 Conclusion and Future Work 37
5.1 Concluding Remark . . . 37
5.2 Future Works . . . 37

Bibliography 39

List of Figures

1.1 Human Brain 5
1.2 Brain Tumor 11
1.3 MRI Scanner 13

2.1 Difference between Deep Learning and Machine Learning 15
2.2 An Artificial Feedforward Neural Network 16
2.3 Neural network with many convolutional layers 18
2.4 RGB 3 Channels. 19
2.5 Convoluting a 5x5x1 image with a 3x3x1 kernel to get a 3x3x1 convolved feature. 19
2.6 Types of Pooling. 21
2.7 Steps in digital image processing. 22
2.8 Steps in digital image processing. 25
2.9 Steps in digital image processing. 25
2.10 Steps in digital image processing. 26
2.11 Steps in Digital Image Processing. 26
2.12 Steps in digital image processing. 27

3.1 Workflow of the Proposed Model 29
3.2 Implementation of VGG16 30
3.3 Implementation of ResNet50 31

4.1 VGG-16 Model 34
4.2 ResNet-50 Model 35

List of Tables

4.1 Model performance 35
4.2 Results from VGG-16 and Resnet-50 36

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

CNS Central Anxious Framework

CSF Cerebrospinal Fluid

CT Computed tomography

GM Grey matter

GMM Gaussian Mixture Model

MRI Magnetic resonance imaging

MTR Magnetic Resonance Tomography

NMR Nuclear magnetic resonance

$NMRI$ Nuclear Magnetic Resonance Imaging

PET Positron emanation tomography

SOM Selforganization map

VTA Ventral tegmental vicinity

WM White matter

Chapter 1

Introduction

A brain tumor is occurred by the developing of unusual cell on the surface of the brain. It produces various kinds of indications based on which part of the brain is affected. These signs includes migraine, seizures, visionary problem and mental problem. A migraine is basically more painful with morning vomiting. More specific side effects are inconvenience in speaking and strolling in sensation. Tumors that originate in the brain, known as primary brain tumors, may be generally classified as those that grow within the brain and are called for the cell from whence they originated. Some tumors are not malignant. Some instances may be damaging and inconvenient. Lymphoma, Cystic Oligodendroglioma, and Ependymoma as well as Anaplastic astrocytoma are all types of tumor derived from neural tissue (which is histologically benign but continues to be dangerous). Tumors that originate in other regions of the body may be able to spread to the brain, resulting in tumors or brain tumors. Cancers that spread to the brain often develop secondary brain tumors as well as melanomas in the skin. To examine the brain structure completely imaging modalities such as computed tomography (CT), positron emanation tomography (PET) and attractive reverberation imaging are often used. When compared to PET and CT, MRI is thought to be far superior. It allows for a careful investigation of brain life systems without the need for physical movement or ionization radioactivity. Tumor areas are clearly visible using MRI, allowing for a much more effective treatment method. MR imaging represents a useful perspective on the human body. As a consequence, MRI will advance in medical science by one line. Unsupervised approaches like fuzzy c-means and self organization maps are used to categorize MR images (SOM). Many MR imaging investigations rely on both supervised and unsupervised methods for classifying malignant or benign MR images. In this research, we apply supervised machine learning approaches[3] .

1.1 Background

The medical imaging technology of magnetic resonance imaging (MRI), Nuclear Magnetic Resonance Imaging (NMRI) or Magnetic Resonance Tomography (MRT) is used in radiology to visualize within body structures. MRI provides a thorough analysis of the human body when compared to other medical imaging techniques[2] . On the other hand, MRI pictures do not emit any hazardous radiation. The soft tissue of the human body may be adequately contracted using MRI. Because the brain is made up of soft tissue, MRI is utilized to detect any brain disorders such as cancer and tumors. T1 and T2 weighted pictures, FLAIR and MRS pictures are two forms of MRI imaging. Grey matter (GM), White matter (WM), and Cerebrospinal Fluid (CSF) are three different types of MRI imaging (CSF). White matter is brighter than gray matter in a T1 weighted picture. Gray matter, on the other hand, is lighter than White matter in a T2 weighted picture [8].

1.2 Literature Review

The following have been cited by more than a few brain tumor identification and segmentation techniques. A novel produced by Cui (2018) on automated segmentation focused primarily on the Convolutional Neural Network of cascaded deep learning. It has two networks of sub. Localization Network (TLN) and an intra tumour classification network (ITCN). Using the tumour localization network, the tumour region is isolated from the MRI brain slice. ITCN helps to label the defined tumour area into a couple of sub-regions. The dataset used was Brain MRI images for Brain Tumor Detection dataset. The evaluation can be performed through sensitivity [18].

Khawaldeh (2018) given machine learning method for medical image classification and segmentation. For the classification of brain clinical images, the method uses Convolutional networks. The brain tumour is classified into low and excessive grades [20]. Haveri (2017) defined segmentation of brain tumour using deep neural networks to classify tumour into low and high grades. Dong (2017) proposed non-invasive magnetic resonance methods as a diagnostic tool to discover brain tumour barring ionizing radiation.The network used in this task was the deep convolutional network based on u-net.This offers the most favorable results for the central regions of tumors [13].

To discover the gliomas-based brain tumour, Hussain (2017) suggested a segmentation algorithm. It uses a deep convolutional neural network algorithm to detect an irregularly formed tumor. Therefore, with proper brain tumor segmentation, the survival chances of the affected person are improved. By adding max-out and drop-out layers in the patch processing, the trouble of overfitting is eliminated. In addition, this proposed algorithm uses a pre-processing approach to put off the unnecessary noise and post-processing helps to eliminate small false positives using morphological data set used to be BRATS from 2013 [14].

Chinmayi (2017) presented an approach using Bhattacharya coefficient for MRI brain tumour segmentation and classification. With the use of the anisotropic diffusion filter, the unwanted skull parts have been eliminated. It uses deep learning CNN to train the MRI brain tumour image. Finally, the consequences of the proposed method compared in terms of accuracy. The findings will allow the radiologist to define the size and location of a tumor. Kamnitsas (2017) developed a brain lesion segmentation that is a challenging mission performed using the convolutional three-D network. 3D CNN is a high-quality approach that provides sufficient segmentation without raising the cost of computing and the wide range of training parameters [15].

Isil (2016) stated that one of the difficult tasks in the medical field is brain tumour segmentation. Using early brain tumor prognosis, the patient's lifespan is prolonged. Manual segmentation of brain tumours for huge amounts of data is a painstaking process. Automatic segmentation now incorporates deep learning techniques for segmentation for a few days. For a massive amount of MRI-based photograph data, this provides effective segmentation. The paper analyzed the present-day deep learning approaches. Wang (2017) endorsed the state-of-the-art performance for automated medical segmentation generated by the Convolutional Neural Network. But it failed to produce robust results for medical use. The problem is corrected by using the new interactive segmentation system focused on deep learning. The proposed method allows the CNN model adaptive to a particular test image that can also be controlled or unsupervised [17].

1.3 Objective

To combat this swelling, several improvements have been made, propelling clinical research to new heights. The paper's main claim is that MR images can reveal tumor evidence in the cerebrum. The primary goal of brain tumor identification is to aid in the clinical conclusion so that professionals can recognize the tumor without difficulty. Currently, the methods used by nervous system specialists for research are not completely error-free, implying that manual division is unquestionably not a good idea. The study proposes an AI-based strategy for the division of cerebrum images and tumor identification using a variety of grouping methodologies that improve the presentation, minimize the complexity, and chip away at continuing data. The goal of this exploratory project is to classify and divide brain tumors. Nowadays, the majority of health-care systems rely on clinical image processing to separate data about the human body and assess almost all illnesses. This exam paper explains how to recognize and highlight tumors of the cerebrum that are influenced by extraction. As long as our suggested technology allows us to correctly identify the tumor and provides an easier approach to the brain tumor, we agree that our approach will be more exact than inspecting the brain tumor. In this way, we see that when this assessment comes to be realized in physical form, the latter will likewise become highly restricted, with more precision feasible.

1.4 Overview

1.4.1 Brain

The structure of human framework is made up of the spinal cord and the neuron cells that forms the human brain. The cerebrum, brainstem, and cerebellum are all parts of the brain. It is in charge of the majority of the body's workouts, preparing, joining, and planning the information it receives from the sense organs and making decisions on the information provided to the rest of the body. The brain is protected and confined by the cranial bones of the skull. The brain, a surprising three-pound organ that regulates all bodily components, decodes information from the outside world and exemplifies the center of the brain and soul. The cerebrum speaks to a variety of things including insight, creativity, emotions, and memory. The cerebrum, cerebellum and brainstem make up the brain. The brain collects information from us in the form of location, fragrance, touch, taste, and hearing - all of which are frequently present at the same time. Later, it pulls all these collected data in an understandable sequence and stores in our memory cell. Our thoughts, memory, and speech as well as the growth of our arms and legs and the capabilities of other internal organs are all under the control of the cerebrum. The central anxious framework is formed by the cerebrum and spinal cord (CNS). The PNS is formed of spinal nerves, which branch out from the spinal cord, and cranial nerves, which branch out from the brain. The cerebrum, cerebellum, and brainstem are the three main components of the brain [21] .

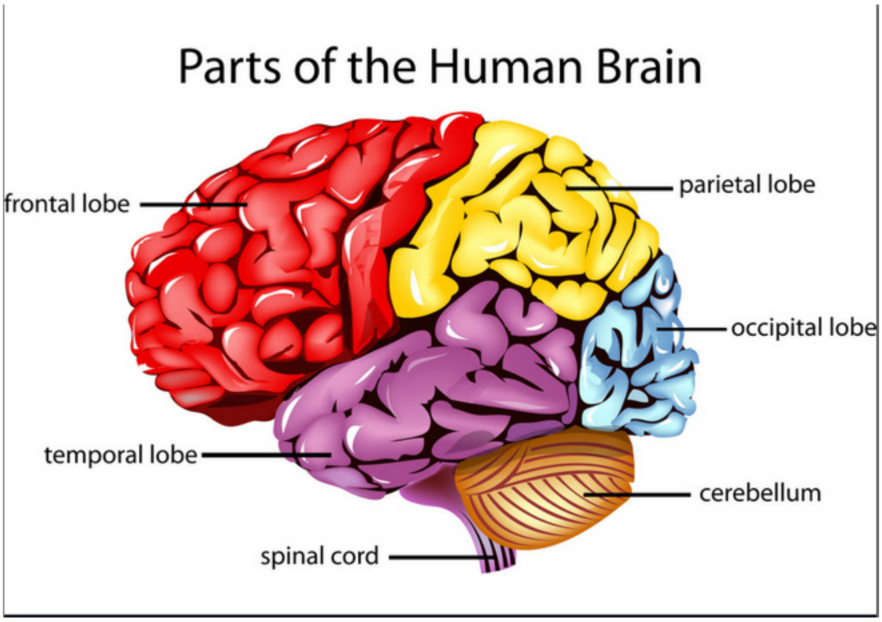

Figure 1.1: Human Brain

1.4.2 Cerebrum

The cerebrum is the biggest section of the brain and a deep groove and the longitudinal gap divides it into about symmetrical cleaned out and right half of the globe. As a petal, the asymmetry between the flaps is well-known. The corpus callosum is the largest of the five commissures that span the longitudinal gap on each side of the equator. The frontal projection, parietal flap, transitory flap, and occipital flap are the four principal projections on either half of the globe named after the skull bones that usually contains them. Despite the fact that there are a few utilitarian coverings between them each projection is linked to one or two specific abilities. Gyri and sulci are the divided parts of the surface of the brain. Many of which are called for their positions such as the frontal gyrus of the frontal projection or the central sulcus separating the central regions of the sides of the equator. Within the auxiliary and tertiary folds, there are several tiny variations. The cerebral cortex which is made up of dark matter structured in layers is the exterior section of the cerebrum. It is 2 to 4 millimeters thick (0.079 to 0.157 in) and deeply compressed, giving it a twisted look. The cerebral white matter lies under the cortex. The neocortex is the biggest section of the cerebral cortex with six neuronal layers. Three or four layers of allocortex make up the rest of the cortex. Brodmann's ranges are divisions in the cortex that map out roughly fifty different functional areas. When seen under a microscope these areas stand out especially. A motor cortex and a sensory cortex are the two basic utilitarian ranges of the cortex. The main motor cortex which transmits axons down to motor neurons in the brainstem and spinal cord is found in the frontal lobe's upper portion directly before the somatosensory zone. The thalamus's transfer cores carry messages from the sensory nerves and tracts to the major sensory centers. The visual cortex of the occipital projection, the sound-related cortex in sections of the worldly lobe and insular cortex and the somatosensory cortex inside the parietal flap are all primary sensory ranges. The association areas are the remaining sections of the cortex. These zones receive information from the sensory areas and lower sections of the brain and they are involved in sophisticated cognitive processes such as discernment and decision-making. Controlling contemplation, theoretical contemplating, conduct, issue confronting errands, bodily reactions, and identification are the most important skills of the frontal projection. The occipital projection is the smallest projection with visual collecting, visual-spatial handling, movement and color recognition as its primary functions. The cuneus is a smaller occipital lobule found inside the projection. Sound-related and visual memories, dialect and a few hearing and conversation are all controlled by the transitory projection. In a horizontal bisection of the head, cortical folds and white matter may be seen[7]. The cerebrum contains the ventricles which create and circulate cerebrospinal fluid. The septum pellucidum, a membrane that splits the lateral ventricles, is located below the corpus callosum. The hypothalamus is to the front and below, whereas the thalamus is below the lateral ventricle. The pituitary gland is related to the hypothalamus. The brainstem is placed between the thalamus and the midbrain. The basal ganglia is also known as the basal nuclei, which are a collection of neurons located deep within the brain that control movement and behavior. Despite the widespread problems seen in the striatum, the globus pallidus, substantia nigra, and the subthalamic nucleus are all toward the top of the list. Because function is fully dependent on which subdivision is being described, the striatum is divided into two subdivisions: ventral striatum and dorsal striatum. The ventral striatum is characterized by the nucleus accumbens and the olfactory tubercle, whereas the dorsal striatum has the caudate nucleus and the putamen. Internal pills divide the putamen and globus pallidus, located inside the lateral ventricles, from the lateral ventricles and thalamus, which are on the outer edges of the brain. The claustrum is a narrow neuronal sheet that runs between the insular cortex and the striatum in the deepest region of the lateral sulcus. A variety of basal forebrain regions are found underneath and in front of the striatum. Additionally, the medial septal nucleus, Broca's diagonal band, substantia innominata and the nucleus basalis are included. This neurotransmitter, acetylcholine, is produced by the facilities and is disseminated broadly throughout the brain. It is believed that the major cholinergic output of the nervous system to the striatum and neocortex is situated in the nucleus basalis, which is found in the basal forebrain.

1.4.3 Cerebellum

The cerebellum is often referred to as the "small brain". It is related to the midbrain and is placed behind the brain stem and behind the cerebrum. It has equator-like features as well as a grey matter cortex on the outside and a white matter core on the inside. The cerebellum is a prominent part of all vertebrates hindbrains and though it is normally smaller than the cerebrum, it may be as large as or even larger in some species. In humans, the cerebellum is responsible for motor control. Another way to think of it is that it's both engaged in tasks that range from the cognitive (like attention and language) to emotional (like managing fear and pleasure responses), but movement-related activities are most thoroughly studied. Although the human cerebellum influences coordination, precision, and exact timing, it does not generate movement. The brain as well as other components of the nervous system send information to the spinal cord, which incorporates it to fine-tune motor movement. In humans, cerebellar damage causes problems with pleasurable motion, balance, posture, and motor learning. The human cerebellum appears to be a distinct entity attached to the lowest part of the brain, nestled underneath the cerebral hemispheres from anatomical standpoint. In contrast to the cerebral cortex's vast irregular convolutions, its cortical floor has closely spaced parallel grooves. The fact that the cerebellar cortex is a non-stop thin layer of tissue firmly folded in accordion manner is hidden by those parallel grooves.Purkinje cells and granule cells are the most important neurons in this thin layer because they have a fairly regular organization. Because of its enormous signal-processing power, the cerebellar cortex is still responsible for very little of the output generated by the cerebellum. However, almost all of the cerebellar cortex's output is channeled to a cluster of deep nuclei situated within the white matter of the cerebellum. The cerebellum is required for several sorts of motor learning, such as mastering sensorimotor interactions. Many theoretical models have been presented to explain sensorimotor calibration in the cerebellum, each proposing a different explanation of synaptic plasticity. The concept proposes that every cerebellar Purkinje cell gets two kinds of input, with one type consisting of a high number of susceptible inputs and the other a relatively small number of resilient inputs. According to the Marr-Albus hypothesis, the trekking fiber works as a "teaching signal", which helps to train fibers to develop in parallel. The truth is still being debated about these theories because of the lack of long-term grief in parallel fiber inputs [5] .

1.4.4 Brain Stem

The brainstem is located next to the spinal cord, running parallel to it. The brainstem, which is found in the human brain, consists of the midbrain, pons, and medulla oblongata. The tentorial notch links the midbrain to the diencephalon's thalamus, and the diencephalon is contained inside the brainstem more often than not. A relatively small section of the brain (i.e., the brainstem) comprises just 2.6% of the whole brain's weight. This nutrient helps regulate the process of coronary heart rate and breathing charge. It also supplies the principal motor and sensory nerve supply to the face and neck by way of the cranial nerves. Cranial nerves ten in number arise from the brainstem. Sleep cycles, in addition to controlling the central nervous system, may also be used to regulate the body's circadian rhythms. It helps as well to send the paths consisting sensor and motor from the other parts of the brain to the rear and to the frame. The three pathways leading to the dorsal column medial lemniscus nucleus are the corticospinal tract, the dorsal column medial lemniscus route, and the spinothalamic tract [9] .

1.4.5 Mid Brain

The tectum, tegmentum and ventral tegmental area are all separated into three portions in the midbrain. The ceiling is formed by the tectum. The tectum is the dorsal covering of the cerebral aqueduct and is comprised of the paired structures of the advanced and inferior colliculi. The inferior colliculus is the principal midbrain nucleus of the auditory system, which gets input from several peripheral brainstem nuclei, including the auditory cortex. The inferior brachium of the brachial plexus reaches the medial geniculate nucleus of the diencephalon. Below the inferior colliculus lies the rostral midbrain, defined by the well advanced colliculus.

It is interested in specific visual experience and deploys its advanced brachium to the lateral geniculate body of the diencephalon. The cerebral aqueduct is located ventral to the tegmentum which forms the midbrain's floor. This contains many nuclei, tracts and the reticular formation. A pair of cerebral peduncles make up the ventral tegmental vicinity (VTA). These are the axons of higher motor neurons that transfer information.

1.4.6 Pons

Between the medulla oblongata and the midbrain is the pons. The advanced pontine sulcus connects the pons to the midbrain whereas the inferior pontine sulcus connects it to the medulla. It consists of fibers that transport messages from the cerebrum to the medulla and cerebellum as well as sensory indications to the thalamus. The cerebellar peduncles are used to connect the pons to the cerebellum. The pontine breathing group which includes the breathing pneumotaxic centre and apneustic middle is housed in the pons. The cerebellar hemispheres actions are coordinated by the pons. The pons and medulla oblongata are hindbrain structures that form the brainstem.

1.4.7 Meddulla Oblongata

The medulla oblongata is sometimes known as the medulla for short. It goes parallel to the pons in the highest part of the brain stem. The medulla governs the heartbeat, respiration, and blood pressure. This medullary area after the salivary gland is the postrema, which carries out functions such as vomiting control.

1.4.8 The Lobs of The Brain

The brain's lobes were once thought to be only physical classifications but they have now been shown to be linked to distinct mental capacities. Both the cerebrum and the cerebellum, the biggest section of the human brain are divided into lobes. The cerebrum is referred to by the word "lobes of the mind", if it is not already clear. The cerebrum is divided into six lobes according to Terminologia Anatomica (1998) and Terminologia Neuroanatomica (2017). Every lobe of the brain is made up of different sub-areas that work together to form a comprehensive function for the entire brain.

1.4.9 Frontal Lobe

The frontal lobe is located on the front (anterior, APS) sides of the cerebral hemispheres above and in front of the parietal lobe, and is also located above and in front of the temporal lobe. It was Sir Niks Dhangar who found it initially. The lateral sulcus, known as the Sylvian fissural, separates the frontal lobe from the temporal lobe. Also, the central sulcus divides the temporal lobe from the frontal lobe. This region, located in the frontal lobe, is known as the precentral gyrus, and is where the main motor cortex (which directs voluntary actions of particular body parts) is found. The precentral area is home to the primary motor cortex (region 4) and the premotor cortex (region 6). In charge of all professional and postural activities are these sections. The frontal lobe is bordered by the frontal pole on the front (front-most) end and is composed of the prefrontal cortex, which is found at that location. It is very vital for your working memory and executive control, which is required for planning long-term goals and for making plans that include several subtasks.

1.4.10 Parietal lobe

One can trace the parietal lobe at the back of the frontal lobe and central sulcus and also over the occipital lobe. Spatial experiences and navigation (proprioception) are integrated in the somatosensory cortex immediately posterior to the central sulcus whereas the visual system's dorsal movement is handled in the dorsal occipital cortex. Thalamus is an area of the brain that carries messages about body sensation, such pain and temperature, from the skin to the parietal lobe. In various aspects, the parietal lobe is critical for language processing. Although this may seem as a distortion, the homunculus is just a depiction of how the bodily parts of a body map to the amount of somatosensory cortex dedicated to each body part. Spatial focus tends to be placed on more advanced regions of the parietal lobes, such as the anterior and inferior parietal lobules. A lesion in the right superior or inferior parietal lobule is the cause of hemineglect.

1.4.11 Occipital lobe

The occipital lobe is the bulk of the visual cortex's physical area, and contains the bulk of the mammalian brain's visual processing capabilities. is the major visual cortex, which is frequently referred to as V1 (a visual one). This refers to the calcarine sulcus on the medial side of the occipital lobe, and its complete volume is often found to cling to the occipital lobe's posterior pole. Additionally, the area around the stria of Gennari is also known as striate cortex, as it has a prominent stripe of myelin termed the Stria of Gennari. The extrastriate cortex is made up of areas outside of V1 that are being optically pushed. Extrastriate areas are located in each of the five main striate areas (vizuospatial processing, color discrimination, and so on) and each one is set aside for a particular visual function.

1.4.12 Temporal lobe

The temporal lobe of the mammalian brain is found under the lateral fissure of each cerebral hemisphere. to be very good in storing visual memories, excellent at comprehending language, and attuned to feeling The temporal lobe is engaged in making sense of sensory input and determining how the information is important. The hippocampus is a brain region that plays a role in the development of new memories and the research of new topics. A lot of study has been done on the hippocampus in regards to epilepsy, showing that it could cause injury. The link between temporal lobe configuration and rhythmic activity and cognition has only been tentatively established, but Chauvière (2020) demonstrates that there is a significant relationship between alterations in neural circuitry and temporal lobe configuration, which influences these types of activities and cognitive ability.

1.4.13 Brain Tumor

When abnormal cells grow in the brain, a brain tumor develops. Cancerous tumors and noncancerous tumors are the two main types of tumors. Primary tumors which is develop inside the brain. Secondary tumors which most usually have spread from tumors outside the brain are known as brain metastasis tumors. These are the two types of cancerous tumors. All types of brain tumors can cause different signs and symptoms depending on which section of the brain is affected. Headaches, seizures, eyesight problems, vomiting and mental abnormalities are some of the indications and symptoms [10]. The headache is usually worst in the morning and disappears after you vomit. Other symptoms might include issues with walking, speech or feelings. As the sickness worsens, it's possible that you'll become unconscious. Maximum brain tumors have no recognized etiology. The most frequent primary tumors in adults are Meningiomas and Astrocytomas, including Glioblastomas. Medulloblastoma is the most frequent kind of medulloblastoma in youngsters. Computed tomography (CT) or magnetic resonance imaging is often conducted at the same time as a professional evaluation (MRI). Biopsies are often utilized to find out the result. According to all the findings, the tumors are rated according to their severity. In addition to these three procedures, surgery, radiation therapy, and chemotherapy may all be included in the treatment strategy. Seizures may need anticonvulsant medications if they occur. The two medications that may be utilized to inhibit tumor development are dexamethasone and furosemide. Most tumors grow gradually, and in most situations no intervention is required. Currently, therapies that use the immune system to their advantage are being explored. Many factors, both tumor characteristics and tumor stage, influence the ultimate result. Even though benign tumors are the only kind to emerge, these tumors may be deadly, especially because of their location. It is known that glioblastomas have a severe impact whereas meningiomas have a positive prognosis. Overall, in the US, the five-year survival rate for all brain tumors is 33 percent. Brain tumors that have spread to other parts of the body (metastases) are roughly four times as prevalent as brain tumors that are the original disease (primary) and represent half of all metastases. About 250,000 individuals throughout the globe are affected with primary brain tumors each year, representing less than 2% of all malignancies. While childhood brain tumors are the second most common, acute lymphoblastic leukemia is the most common kind of cancer seen in children under the age of 15. Brain cancer has an estimated cost of \$1.9 million per case in Australia.

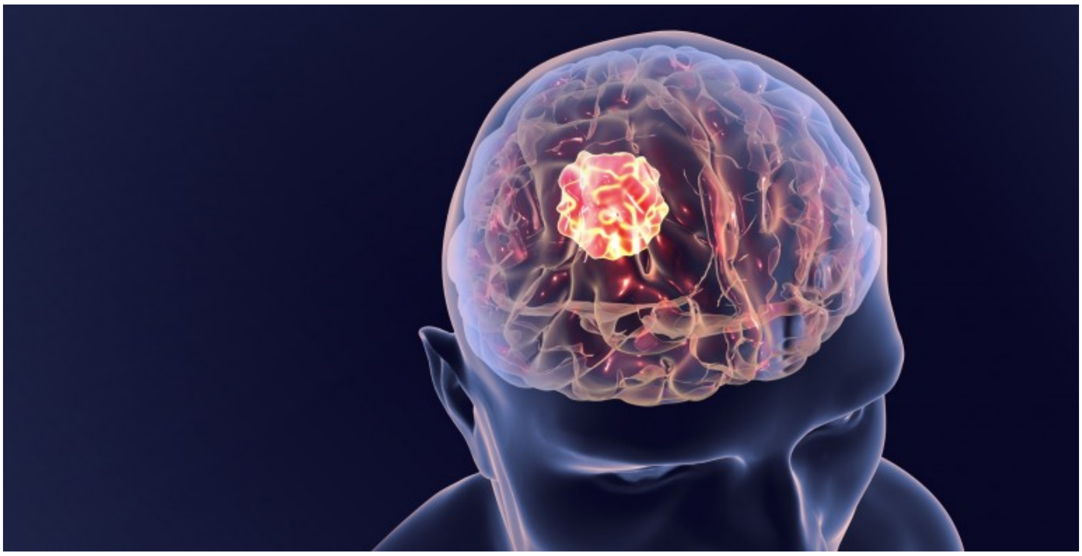

Figure 1.2: Brain Tumor

1.4.14 Primary Brain Tumor

Within the brain, a primary tumor began to form. The majority of initial tumors are benign. Our brain cells, meninges, nerve cells and glands all promoted this initial tumor. Gliomas and meningiomas are the most prevalent types of primary brain tumors.

1.4.15 Glomas

The most frequent type of primary brain tumor is glioma. It starts in the glial cells which are involved in a variety of brain functions. Gliomas come in a variety of types, including low-grade glioma and glioblastoma.

1.4.16 Meningioma

Memorial Sloan Kettering's neurosurgeons, neuro-oncologists and radiation oncologists have extensive expertise treating meningioma at all stages. We help patients from all over the world who are suffering from the most difficult brain tumors which including meningiomas. The vast majority of meningiomas are benign. If they're huge or increasing, they're either detected through MRIs or eradicated. The primary goal of therapy is to eliminate the tumor. If this isn't often doable or if the tumor is very hostile, our doctors may employ a combination of radiation and other therapies to slow or stop the tumor from growing.

1.4.17 Secondary Brain Tumor

A secondary brain tumor (metastasis) is a tumor that develops inside the brain when a malignant tumor (cancer) spreads to another part of the body. These tumors developed from cells that had broken free from the main tumor and traveled to the brain through the circulation. The most common main cause is lung or breast cancer. However, they can also be caused by bowel, kidney, skin and other malignancies.

1.4.18 MRI

Magnetic resonance imaging (MRI) is a radiologic imaging technique that uses a strong magnetic field and radio waves to generate images of the anatomical and physiological processes of the body. MRI scanners utilize powerful magnetic fields, magnetic gradient areas, and radio waves to generate pictures of the organs inside the frame. Although MRI does not utilize X-rays or ionizing radiation, it is distinguishable from CT and PET scans because it uses magnetic resonance imaging (MRI). Nuclear magnetic resonance (NMR) imaging may also be done using the MRI clinical application MRI. While most scientific contexts appropriately regulate the hazards of ionizing radiation, an MRI might nevertheless be perceived as a stronger demand than a CT scan. In hospitals and clinics, MRI is widely utilized for medical diagnosis, staging and follow-up of illness without exposing the patient to radiation. When compared to CT, MRI may produce different data. MRI scans may be associated with risks and pain. In compared to CT scans, MRI scans take longer and are noisier and they require the patient to enter a narrow, restricting tube. Humans having surgical equipment or other non-detachable steel in their bodies may also be unable to appropriately undertake an MRI test. NMRI (nuclear magnetic resonance imaging) was the original name for MRI. But the "nuclear" was deleted to prevent bad connections[4]. While positioned in an outside magnetic discipline, some atomic nuclei may absorb radiofrequency power; the resulting developing spin polarization can cause an RF sign in a radiofrequency coil and so be noticed. Hydrogen atoms are most commonly utilized in clinical and research MRI to create a macroscopic polarization that is detected by antennas near the scenario being examined. In humans and other organic creatures, hydrogen atoms are abundant, particularly in water and fat. Maximum MRI scans effectively map the position of water and fats in the frame as a result of this. The nuclear spin power transition is excited by radio waves and magnetic area gradients concentrate the polarisation in space. To change the characteristics of the pulse sequence, different relaxation homes may be used for hydrogen atoms inside the tissues. MRI has demonstrated to be a versatile imaging method since its progress in the 1970s and 1980s. In biological research and clinical practice, MRI is utilized for the analysis of live tissue. However, it may also be utilized to get pictures of inanimate things. MRI scans, like accurate spatial photos may yield a ramification of chemical and physical data. The steady increase of MRI requests inside fitness systems has raised concerns regarding economic feasibility and misdiagnosis.

Figure 1.3: MRI Scanner

Chapter 2

The principles of Deep Learning, Artificial Neural Networks, CNN, Image Processing, Resnet-50 and VGG16

In this chapter, we discussed about, fundamentals of image Deep Learning, Artificial Neural Networks, CNN, Image Processing,Resnet-50,VGG-16.

2.1 Deep Learning

Machine learning is the underlying field for deep learning dealing with brain-inspired algorithms and a feature called artificial neural networks. Artificial intelligence is a broad word that relates to computer systems for imitating human behavior. Machine learning mirrors a collection of algorithms that are taught with data and is able to perform all of these functions. Machine Learning reflects an algorithm collection of data which makes all of this feasible. On the other hand, deep learning is merely a form of machine schooling that is influenced by the human brain's structure. By examining data with a logical structure continually, deep-learning computers attempt to draw similar findings to human beings. Deep learning uses a multi-layered algorithm system known as neural networks. At each level, students are trained to think in a deeper, more abstract manner while also learning to better understand and synthesize data. To convert the raw pixel matrix into a form that image recognition software can use, first expand the original image. First, the edge pixels can be encoded into the first representative layer, and then the pixel abstractions can be calculated from this representation. Pixels and edges can be taken from the first representative layer. The second layer can be arranged and the edge program. The nose and eyes may be encoded at the third layer. The fourth layer can detect faces in the frame. In particular, an in-depth learning process can learn which characteristics can be best suited at this point. Deep learning approaches allow supervised learning projects to focus on feature engineering by transforming data into compact, intermediate images that have the same types of critical components and enabling structures that remove duplicate data. Attended learning tasks such as automatic speech recognition and automatic text recognition can be performed with deep learning algorithms. This is a significant gain because the number of

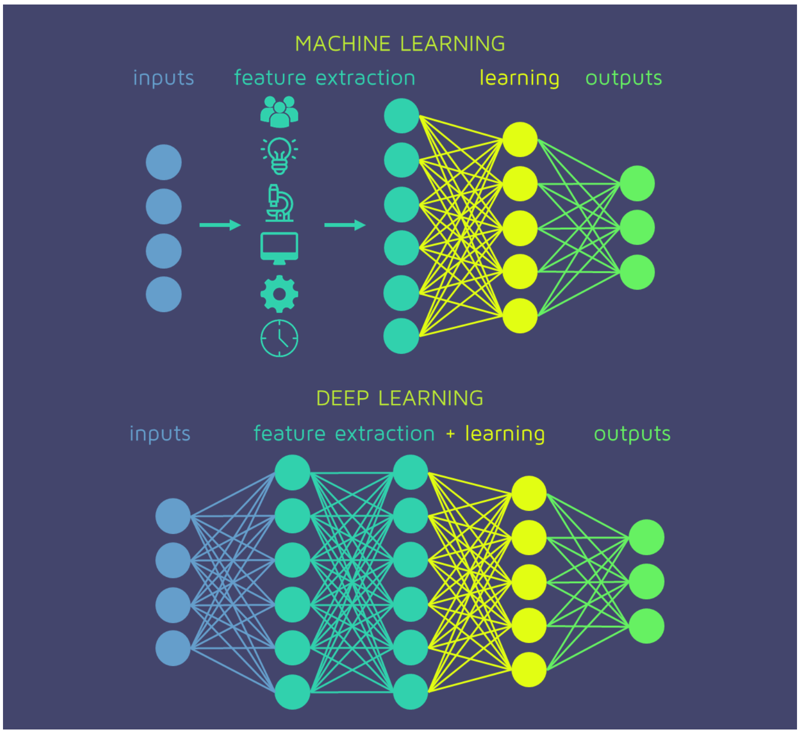

Figure 2.1: Difference between Deep Learning and Machine Learning

unlabeled datasets greatly outweighs the amount of identified ones. Neural context compressors and networks of deep beliefs are examples of deep systems which can be unattended. In terms of accuracy, deep learning models continue to get better and better with the amount of training data, however conventional models such as SVM and Naive Bayes reach a level of saturation after which they no longer improve. [22]

2.2 Artificial Neural Networks

The neural network's architecture is modeled after the structure of the human brain. By teaching neural networks to execute the same data operations as we humans use our brains to identify patterns and classify varied sorts of information, we can ensure our neural networks are never idle. In many circumstances, the separate layers of neural networks can be considered as a sort of filter that serves as a transition between gross and subtle data, which increases the ability to recognize it and ultimately leads to a desirable output. Just as the human brain does, the brain of a person functions. Whenever we gain new knowledge, the brain does its best to compare it to similar things we already know about. Another neural network-related method employed in deep learning is also employed. Neural networks allow a number of different tasks to be completed, such as clustering, classification, and regression. It's likely that the data used for this test has a similar structure to that of the unlabeled data used in a previous test, and so a neural network might be able to group or sort that data. When we work with the classification, we can first build a training dataset with several different sets of labeled data, then use the network to categorize all of the samples from this dataset. In neural networks, we are able to complete a variety of tasks, including classification, regression, and grouping. The unlabeled data that can be grouped or classified based on the commonalities in the samples across various data sets are referred to as "unsupervised" data. To handle the classified data, we first trained the neural network on a designated data set that consists of several categories of samples, after which we can use it to categorize the samples found in this particular data set. A neuron is just a graph representing a numerical value (e.g. 1.2, 5.0, 42.0, 0.25, etc.). Any link between two artificial neurons can be thought of as an axon in a biological brain. The neuronal relationships are determined by weights, which are only digital values. If an artificial neural network learns that the weights of its neurons fluctuate and that this fluctuation is proportional to the neural network's strength, we are also interested in finding certain fixed weights that enable the neural network to complete the classification task using training data and a specific task. Weights change according to task and dataset. Although the weight values cannot be predicted in advance, the neural network can learn. We can refer to the process of learning as a training procedure.

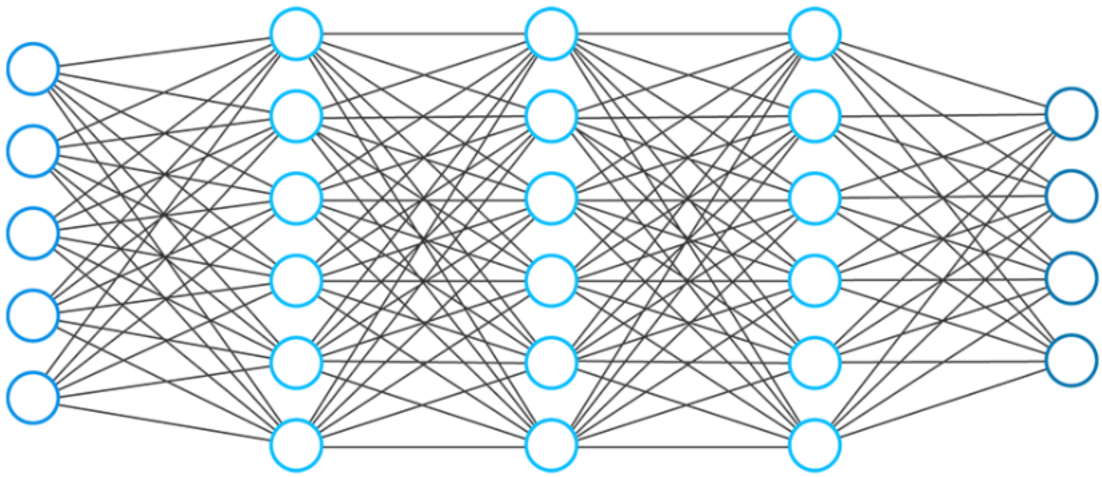

Figure 2.2: An Artificial Feedforward Neural Network

2.3 Convolution Neural Network

In various fields linked to model recognition, from image processing to voice recognition, the Convolutional Neural Network had groundbreaking breakthroughs in the last decade. The most helpful characteristic of CNN would be a drop in ANN parameters. This achievement led to broader models for researchers and developers to solve complex tasks that standard ANNs could not accomplish. There are no spatially dependent characteristics in the key supposition of problems handled by CNN. In other words, we must not, for example, focus on where the faces are in the photos in a face detection application. The main concern is to detect them in the photos, irrespective of their position. The abstract qualities when an input spreads towards deeper levels are another crucial part of CNN. For example, in image classification, the edge might be detected in the first layers, and then the simpler shapes in the second layers, and then the higher-level features such as faces in the next layers. [12]

2.4 Convolutional

The input for programming a CNN is a tensor with form (size of image number) x (size image) x (size of image) x (input channels). Then the painting is abstracted to a function map, with form (number of images) x (height of the map) x (width of the map) x (feature map channels). The following features should be used for a convolutionary layer within a neural network:
Convolutionary kernels with width and height definition (hyper-parameters). The number of channels of input and output (hyper-parameter). The Convolution filter depth (entry channel) must be equivalent to the input function map channels (depth).
In general, input is merged using convolutional layers and then passed to the next layer. This is analogous to a neuron's response to a specific stimulus in the eyepiece. Each neuron processes information exclusively for its own receptive zone. While fully connected neural networks can be used to learn characteristics as well as classify data, this design cannot be used for photographs. Due to the vast amount of pictures in the input, many neurons are needed, even in a shallow architecture (in contrast to deep). A significant variable is each pixel. For example, a completely linked layer of 100 by 100 pictures for each neuron in the second layer has 10,000 weights. This problem is resolved by the convolution procedure, which lowers the number of free parameters and makes the resonance of the network less effective. Tiling 5×5 regions, each with the same weight shared, only requires 25 learning parameters regardless of the image size. During backpropagation in conventional neural networks, regularized weights over less than regularized parameters are prevented from disappearing gradient and expanding gradient concerns. Completely connected layers connect each neuron to each neuron in another layer in one layer. The classic multilayer perception network is basically the same (MLP). In order to identify the photos, the flattened array passes through a completely connected layer.

2.5 Receptive field

Each neuron receives input from certain places in neural networks in the previous layer. Each neuron receives information in a fully connected layer from every portion of the last layer. Only a small sub-area of the previous layer provides the input for neurons in a convolutionary layer. The sub-area is usually of a square form (e.g., size 5 by 5). It is referred to as a neuron input field. Thus, the reception field in a fully connected layer is the whole prior layer. In a convolutionary layer, the sensory region is smaller than the entire final layer. The original input picture subarea in the receptive field increases with increasingly deeper network design. This is because the convolution is repeatedly applied, taking into account the relevance of a single-pixel and other environmental pixels.

2.6 Weights

Each neuron in a neural network adds a certain feature to the values of the preceding layer's receptive field, and this serves as the final output value. The input values are determined by weight and a bias vector (typically real numbers). In a neural net, recurrent modifications in these weights and distortions advance learning. Filters and specific characteristics are called the weight vector and the inclination (e.g., a specific shape). A characteristic aspect of CNNs is that numerous neurons are able to share the same filter. Reduced memory footprint is achieved by the use of a single bias and a single vector of weights, which are applied to all receptive fields that use that filter.

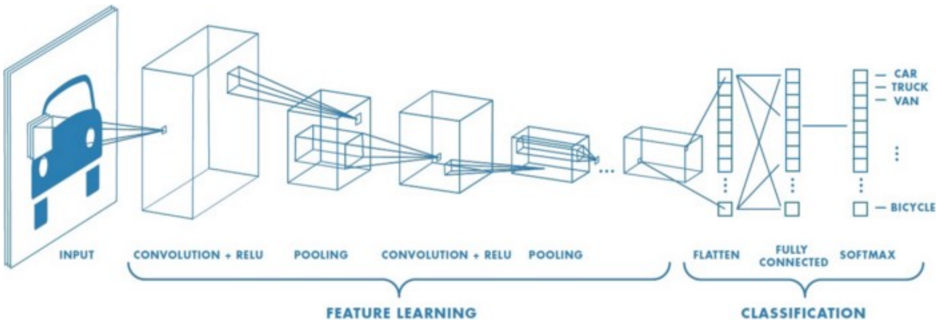

Figure 2.3: Neural network with many convolutional layers

2.7 The Operation of Convolutional Neural Networks

Similar to neural networks, CNNs have learning weights and bias built-in. When a neuron takes in multiple inputs, adds them up, and applies the result to a calculation, that cell will activate and then respond with an output.

2.7.1 Input image

Red, Green, and Blue, colored images are shown in the figure. Several color spaces, such as Grayscale, RGB, HSV, CMYK, etc., exist in which images exist. To illustrate how complex things will be once photos reach the size of 8K, just think about it (76804320). The ConvNet's job is to take the images and to change them in such a way that processing is easy, but still retains important characteristics for prediction. While building a highly scalable architecture is very important, designing a system that is capable of learning feature information and able to scale up to very large datasets is essential. [23]

2.7.2 Convolution Layer

When it comes to removing features from an input image, the first layer to do so is called Convolution. Photorealistic images can be learned by using small input data squares that keep the pixels from moving out of place.In order to complete this task, we'll want two inputs: a filter and kernel and the picture matrix. Length in pixels of the image is five (height) by five (large) by one (length in pixels) (Number of channels, e.g. RGB). That green display appears to be our image 5x5x1. The kernel/filter, which is illustrated in the color yellow, is the initial step in the convolution process. The three-by-three-by-one matrix has been selected. The matrix multiplication is done between K and the section P of the image over which the kernel is travelling, results in a movement of the kernel 9 times because of Stride Length = 1. (Non-Stride) [6] .

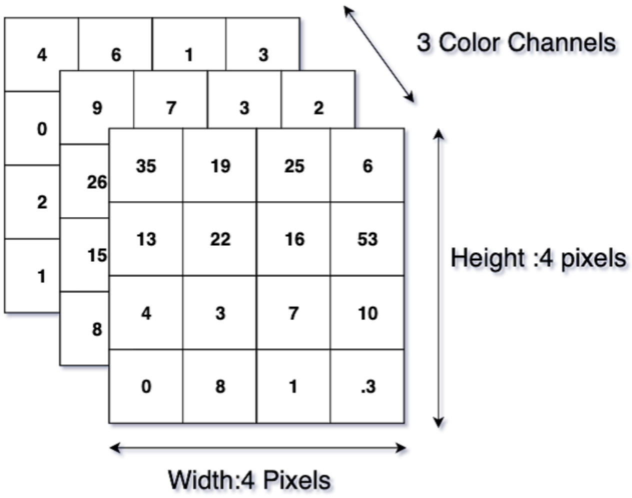

Figure 2.4: RGB 3 Channels.

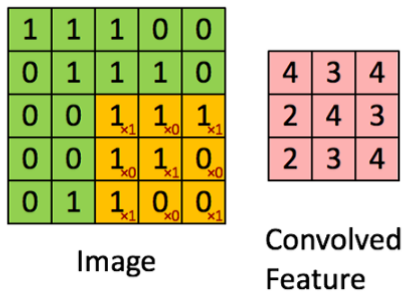

Figure 2.5: Convoluting a 5x5x1 image with a 3x3x1 kernel to get a 3x3x1 convolved feature.

2.7.3 Padding

Convolution is designed to extract high-level characteristics like edges from the image as an output. ConvNets layers do not simply have to be limited to one. The first ConvLayer, which is known as the low-level features (for example, color, gradient direction, etc.) are often captured in the first ConvLayer. With more layers, the architecture also adjusts to increase High-Level Tasks' efficiency, giving us a network that is totally familiar with all of the photos in the dataset. If the curved function is lowered in dimension relative to the input, then two types of outputs are available. Instead, the second happens to either increase or retain the same dimensions.

Same Padding

To find out if our 5x5x1 image now has dimensions of 6x6x1, we first convolve the 5x5x1 picture with the 3x3x1 kernel. This gives us a 5x5x1 picture with dimensions of 6x6x1.

Valid Padding

To obtain the matrix with dimensions of the kernel, the same operation is performed, and the result is used (3x3x1).

2.7.4 Non Linearity (ReLU)

ReLU is a nonlinear operation. The maximum rate is (x) = (0, x). In our ConvNet, ReLU seeks to introduce non-linearity. The linear values are not negative since, in the real world, our ConvNet wants to understand. In addition to the RELU nonlinearities, additional nonlinear functions such as sigmoid may be employed. In most cases, data scientists use the ReLU activation function, as it's more efficient than the other two.

2.7.5 Pooling Layer

The scaled-down characteristic of the pooling layer is adjusted. Reducing the number of dimensions minimizes the computing power required for data processing. Also, prominent elements should be extracted that is constant regardless of rotation or position so that the model is more successfully retained. Pooling can be accomplished in two ways: by having a large swimming pool or by doing simple pooling. The max value is returned from the Kernel component of the image in Max Pooling. However, on the other side, averaging over all image section data yields the average noise level of the entire kernel. It is as good as or better than Peak Pooling. It has the capacity to reject disruptive behaviors and de-noise with the lowering of dimensionality. Average Pooling cuts the number of dimensions as a technique of reducing noise. Max Pooling is substantially superior to Normal Pooling, as Max Pooling results in a considerably more complete picture of the dataset than Normal Pooling. Along with the CNN layer, the Convolutional Layer and Pooling Layer comprise the CNN layer. With respect to the complexity of the images, the number of such layers can be increased, but this comes at the expense of greater processing power. After following the above process, the model can correctly process the features. Additionally, we will also be a part of the final output and input it to a standard neural network for categorization.

2.7.6 Fully Connected Layer

Now let us level the perceptron and prepare to flip the image so that it will be represented as a column vector. Once training is complete, the network receives the incoming output and feeds it into a feed-forward neural network, which is then propagated back propagation every iteration. The model is capable of distinguishing between dominant and certain low-level features in images and classifying them using the Softmax Classification algorithm across a series of epochs. [19] There

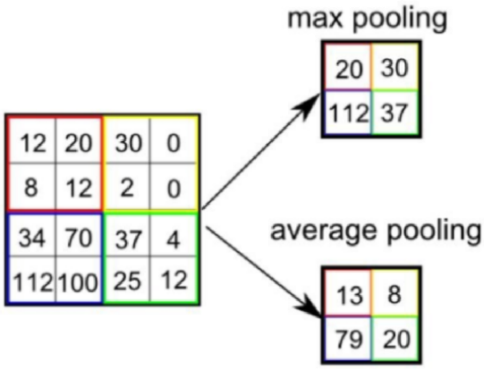

Figure 2.6: Types of Pooling.

are different architectures of CNNs out there, and they've proven crucial in the development of algorithms that support and power artificial intelligence in the near future.
Some of them have been listed below:
*LeNet
*AlexNet
*VGGNet
*GoogLeNet
*ResNet
*ZFNet

2.8 Image Processing

Digital image processing is the process of processing digital images with the aid of a digital computer. Concerning how image enhancement is done using computers, we can also state that it is the utilization of computer methods, in order to extract additional information from the image. Digital image processing is the process of processing digital images with the aid of a digital computer. Concerning how image enhancement is done using computers, we can also state that it is the utilization of computer methods, in order to extract additional information from the image. These three tasks entail image processing:
• Scanning or photographing the image using an optical scanner or digital camera.
• Looking for patterns that are not visible to the naked eye like satellite images.
• Output is the last stage in which the outcome of an image analysis process can be changed, for example, to a final image or report.
For image processing, there are two basic approaches: analog and digital. In hard copy photography, for example, analog image processing can be utilized. [16] Digital image processing techniques enable the computer-based alteration of digital images. Three stages are involved in the processing of digital images:
* Low-level image processing
* Mid-level image processing

* High-level image processing

Low-Level Image Processing

The primary goal of this step is to reduce noise, sharpen images, enhance color, and so on.

Mid-Level Image Processing

A segmentation and classification process occurs during the image analysis phase..

High-Level Image Processing

This is the final stage we may utilize in order to visualize the output image, as merging the successfully discovered items has been completed.

2.9 Stages of Digital Image Processing

Processing involves numerous steps, and the end result is achieved by doing multiple operations in a systematic way. Image processing phases that are nearly universal are described below:

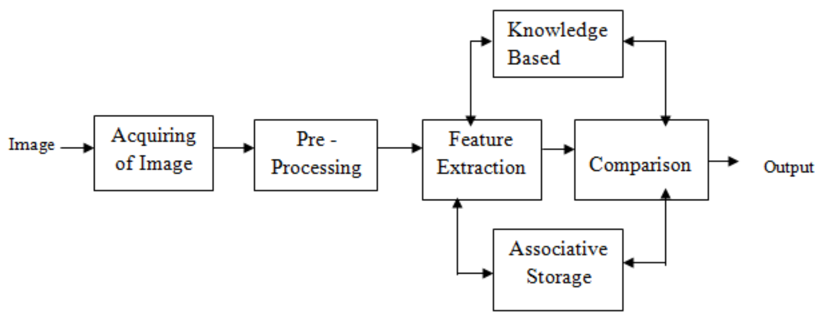

Figure 2.7: Steps in digital image processing.

Image Acquisition

This is the initial stage in digital image processing's four-step process. An illumination source is used to light the scene and a chemical reaction or absorption is applied to the various aspects of the scene to yield the final images. To perceive the image, we use the sensor for lighting conditions. The image sensor is called the image acquisition process. The pre-processing stage (e.g., scaling, translation, and/or rotating) takes place before the picture acquisition process begins.

Image Enhancement

Enhancing photographs is the most simple and preferred application for digital image processing. The practice of modifying an image to make it appropriate for various uses is called image enhancement. Enhancement shows which parts of the image you find appealing. When it comes to image processing, there is no such thing as an objective rule of thumb. You'll be surprised by the variety of enhancement techniques that are available and how many distinct image processing methodologies are used.

Image Restoration

It is about an image's effectiveness in evoking an emotional response (Restauration is based on an image or a mathematical/probabilistic model.).

Color Image Processing

As the number of digital photographs on the internet has expanded, color picture transformation has been an increasing subject of interest. Colors are employed to reveal the qualities of pictures.

Wavelets and Multi resolution Processing

A vital role in portraying images of varied resolutions is played by wavelets, which allow you to accurately partition photos of any size, ranging from web-sized photos to retail shop shelf-sized ones. [1] Multi-resolution analysis is a technique used in signal processing to represent a signal in more than one resolution/scale.

Compression

Two primary approaches used in image compression 1. Saving an image can be made easier by reducing the storage needed. 2. Using the bandwidth required for transmission while reducing the image size (JPEG Standard) is recommended when transmitting an image.

Morphological Processing

Shape representations and descriptions are made easier using the components extracted from image data. Morphology extracts images that are useful in describing and representing shapes. Shape or morphology features are processed by non-linear morphology-related processes. In morphology, just the pixel ordering relative to one another is used, and that only while working with binary images. Grayscale photographs can also be treated with morphological techniques to remove the knowledge of their light transfer functions, leaving only the relative pixel values.

Segmentation

Segmentation software divides a picture into its constituent components or pieces. Segmentation is one of the most challenging tasks in image processing. The segmentation approach gets the process moving in the right direction with regards to resolving imaging problems that necessitate object identification for each individual object.[11] Segmentation can be classified into the following types:
*Region based.
*Edge based.

*Threshold.
*Feature based clustering
*Model based.

Representation and Description

The output data of the segmentation phase, raw pixel information, are always followed by representation and design. The representation of the data will determine whether it is considered to be a border or an entire country. Boundary representation is more concerned with the properties of the structure of the environment, such as corners and actions. Internal features such as texture or skeleton shape are important when creating region representation. Another name for feature selection is descriptive feature extraction, which addresses how to locate information that results in some quantifiable piece of information or is important for differentiating distinct classes.

Object Recognition

Recognition is the procedure by which an object is assigned the label on the basis of its descriptors' information. It covers, as does human understanding:
*Detection
*Description
*Classification
*Identification
*Understanding

2.9.1 Resnet-50

In terms of analyzing or processing visual images, we use deep machine learning. Convolutional Neural Network or CNN is a notable category of deep neural network. Here in this project, we will be detecting brain tumor with the help of two convolutional neural networks; one of which is Resnet-50. Resnet-50 falls under deep neural network category. It is a convolutional neural network that consists of 50 layers. In this case, I can run a previously trained network by the help of an image database of more than a million images. Then it will categorize the image according to 1000 contents and identify the test image. This results in proficiency in identifying test images from an extensive range of options. Primarily, Resnets were used to identify a particular image. But because of its accuracy, it is later widely used for image recognition frameworks in other extracted image processing.

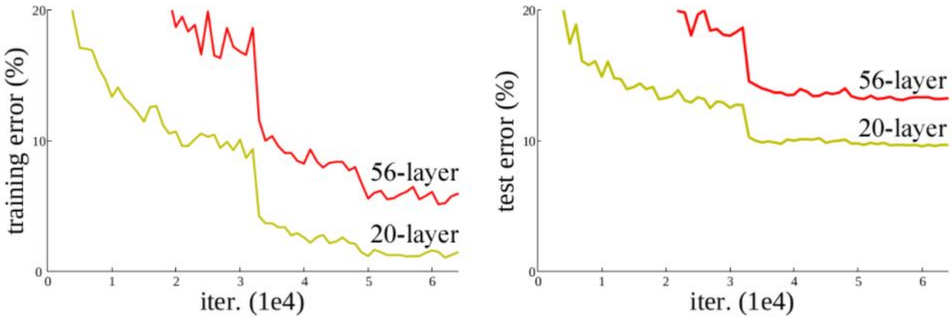

Figure 2.8: Steps in digital image processing.

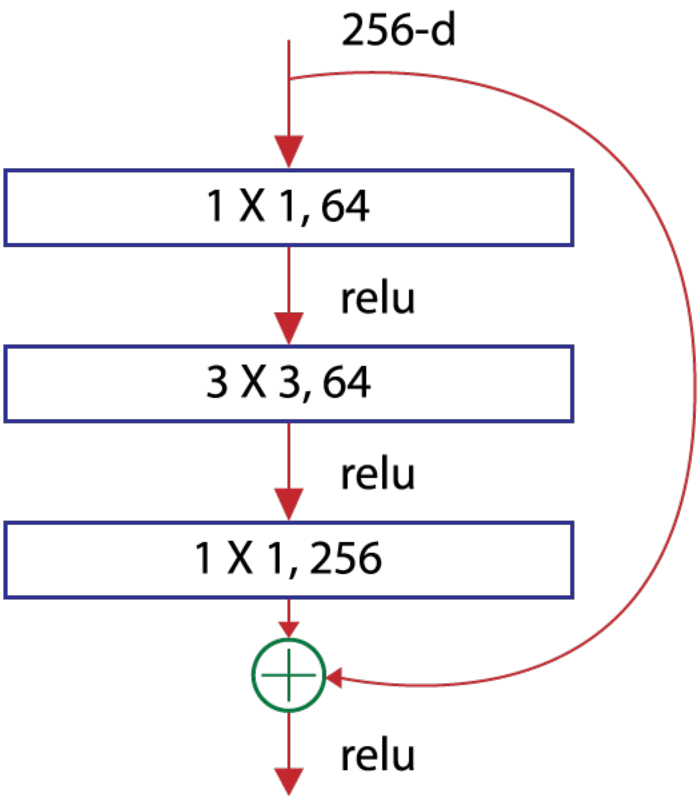

Figure 2.9: Steps in digital image processing.

Previously, Resnets with lesser layers were being used thoroughly. But later it was suggested that by adding more layers, the network provides more accurate results while identifying an image. But as you can see in the graph below, Resnet with more layers tends to generate more errors than the one with lesser layers. In order to resolve this issue, the experts came up with an easier process to complete identify mappings. Moreover, they sticked to the network with only 50 layers rather than the network with more layers as Resnet-50 balances the accuracy and errors perfectly. In Resnet-50, instead of skipping two layers, three layers were skipped and 1*1 convolution layers were added for more accuracy and less errors. Resnet-50 contains 1 layer of a convolution with a 7 * 7 sized kernel, 9 layers of three times repetitive two 64 and one 256 kernel, 12 layers of four times repetitive two 128 and one 512 kernel,

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Figure 2.10: Steps in digital image processing.

18 layers of 6 times repetitive two 256 and one 1024 kernel, 9 layers of two 512 and one 2048 kernel and finally 1 fully connected layer. The total layer by layer process of Resnet-50 can easily be observed from the following Resnet-50 Architecture table.

2.9.2 VGG-16

A proficient convolutional neural network model is VGG-16 which consists of 16 layers. This is said to be the most effective neural network architecture. VGG-16 only focuses on a few padding and pooling layers instead of huge deep layers. Unlike Resnet-50, the execution process of this network ends with two fully connected layers. Approximately, VGG-16 has around 138 million parameters.

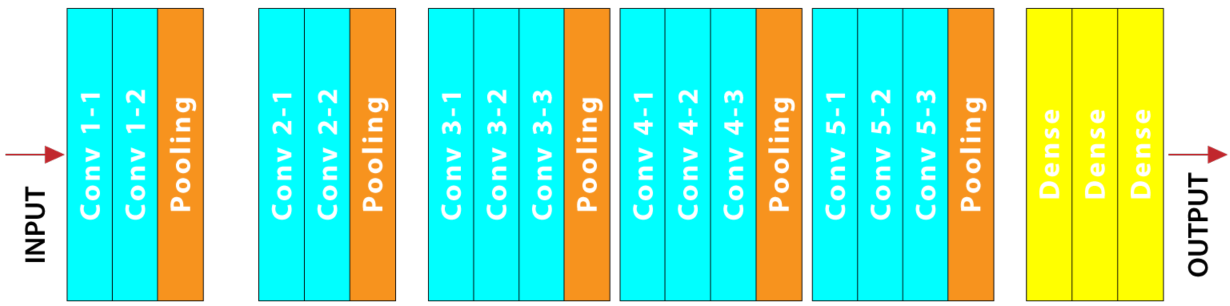

Figure 2.11: Steps in Digital Image Processing.

VGG-16 was first brought to light by Karen Simonyan and Andrew Zisserman of Oxford University in 2014. The accuracy level of this model is significantly high as its pertained image database is very rich. In VGG-16, the network takes a 224 * 224 sized RGB channeled image as an input. Then it generates the output as a vector of 1000 grades. The vector classifies all the values pixel by pixel. After error elimination, the vector calculates the best five values and generates the output with zero error. As you can see, there are five Pooling layers and three additional Dense layers. The architectural structure of VGG-16 will convey a better work process experience. It contains two 3*3 sized 64 channeled layers with similar padding, two 256 channeled layers, then again two 256 channeled layers, two pairs of three padding layers. All these layers are of similar padding and in between these sets of padding layers there is a max-pooling layer. This image processing ends with a set of three fully connected layers. Here, the first layer takes an input from the generated vector and the first two layers give an output vector. But the third layer, which is a fully connected layer, gives a vector output from its pre-stored network. From all these deduction and eliminations, the output vector is extracted after the final soft-max pooling. The total layer by layer process of VGG-16 can easily be observed from the following VGG-16 Architecture table.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 2.12: Steps in digital image processing.

Chapter 3

Proposed Method

We used VGG16 and ResNet50 model architecture of Convolution Neural Network to detect brain tumor. There are two parts of our approach. One is classification and segmentation using VGG16 and another is classification and segmentation using ResNet50. We run these two approaches separately on same dataset in our model. We first train and validate our model from dataset which contains MRI images of brain and then applied VGG16 to detect brain tumor and same for ResNet50. Finally, we also plotted train loss, validation loss, train accuracy and validation accuracy with accuracy rate of both architecture after detecting brain tumor.

3.1 Preparation of the Dataset

This paper users MRI images to detect brain tumor. We have collected the dataset from Kaggle [1]. We tried to use such a dataset which has best and worst-case data to increase accuracy. Many contributors also used the data from Kaggle to implement their model in this dataset.[2] The dataset contains MRI images in the format PNG, JPG and JPEG. There are two flies in the main dataset folder. One of them named "no" contains MRI images of brain with no brain tumor and another one named "yes" contains MRI images of brain with tumor. The no folder contains 89 files, and the yes folder contains 135 files. There is total 224 files.

3.2 Implementation of VGG16 and ResNet50 architecture of CNN

We first prepare our dataset then we make the classification using VGG16 and ResNet50 separately. After our model takes tumor detected images, it removes the artifact and filter images. Then we applied multilevel thresholding to extract the feature to detect the brain tumor. We done model with the help of TensorFlow and Keras libraries. Steps are shown below-

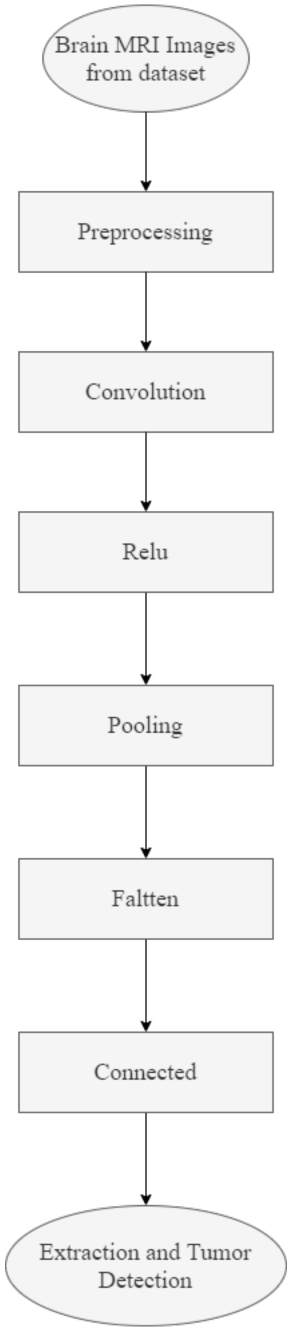

Figure 3.1: Workflow of the Proposed Model

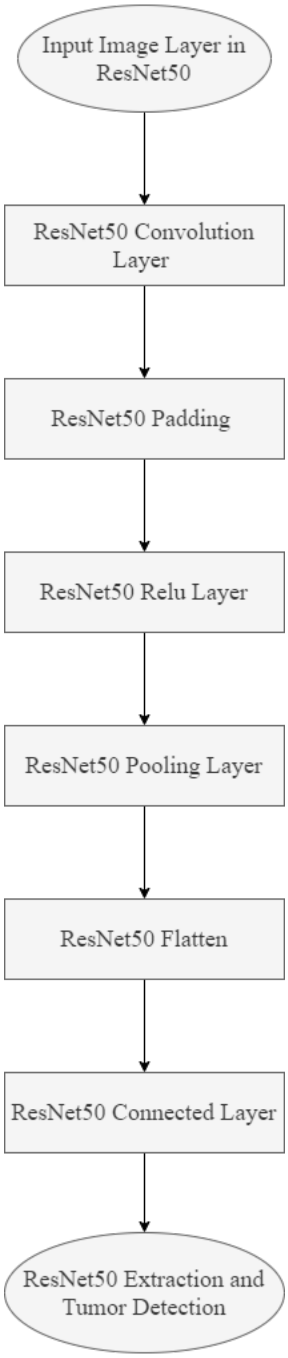

Figure 3.2: Implementation of VGG16

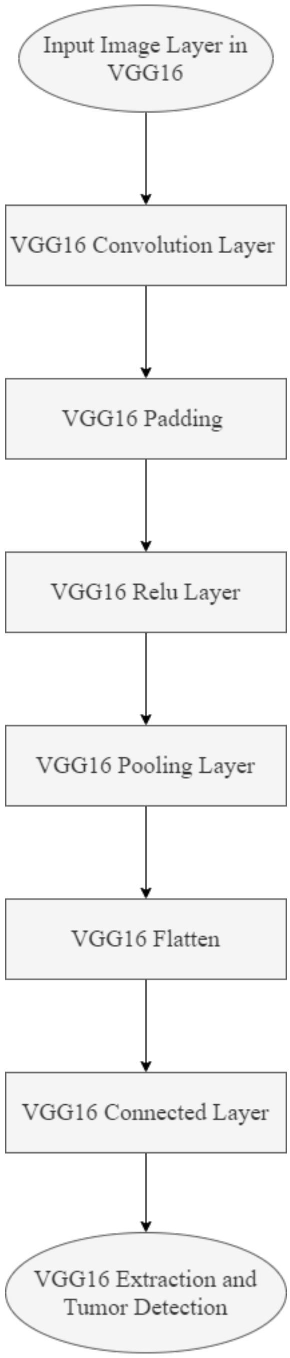

Figure 3.3: Implementation of ResNet50

3.2.1 Input Image Layer in both VGG16 and ResNet50

As our dataset contains JPG, PNG, JPEG format data so we don't need to convert the files. VGG16 takes input and crop it to 224x224 RGB image files at the first layer. But for the ResNet50 we don't need to crop to any size. This architecture takes full image input at the first layer.

3.2.2 VGG16 and ResNet50 Convolution Later

The convolution layer of VGG16 is 16 layers deep. After taking input it convoluted the images to a filter matrix. I (x, y) is the format that express images and F (x, y) is the format that express the filter. On the other hand, Resnet 50 is 48 layers deep. The layers contain 1 max pool and 1 average pool layer.

3.2.3 VGG16 and ResNet50 Padding

In the first layer we took input 255x244 images for VGG16 and it passes through the convolution layer. So, in this stage for VGG16 the model selects 3x3 small field. Where the padding happens to every smaller size like up down, left right and center. ResNet50 does the same thing, but it does on every field with a ratio of 16x16.

3.2.4 VGG16 and ResNet50 Relu Layer

The conviction of relu layer is same for both the architecture. An elementwise non-linearity is applied to the outcome of the kernel convolution to obtain features that are non-linear modifications of the input. When a non-linearity is stated, such as the sigmoid, hyperbolic tangent, and rectified linear functions, there are various alternatives. Convolution with a negative filter output may result in a negative convolution result. Unvalued images have no relevance unless they represent the lowest level of activation. To make things right, we need to change the values of those things. This layer outputs non-linear results to the next layer. More computationally efficient is the rectification method, as it uses the sigmoid and tanh functions. Conv2D allows you to supply a string after doing the convolution that specifies the name of the activation function you want to use. We've previously employed the activation parameter of Conv2D, known as sigmoid, to serve as the activation value of Conv2D.

3.2.5 VGG16 and ResNet50

Convolution layers in a CNN architecture can be considered as a pooling layer as the pooling happens between convolution layer. By doing this, the number of parameters and computations in the network is limited, and overfitting is avoided. This layer consists of two processes: pooling and unpooling. Maxpooling, as the name implies, is a technique that picks out only the largest piece from a pool. At every stride, the maximum parameter is removed, and the rest is dropped. For VGG16 pooling happens in 3x3 and for ResNet50 pooling happens 16x16 matrices.

3.2.6 VGG16 and ResNet50 Flatten

In the last stage of a CNN, it is a classifier. It is simply an ANN classifier known as a dense layer. As with any other classifier, an ANN classifier requires unique features. It needs a feature vector, hence it needs a feature. In a VGG16 and Resnet, each feature map channel that emerges from a CNN layer is a "flattened" 2D array created by the results of multiple 2D kernels.

3.2.7 VGG16 and ResNet50 Classification and Training

Here at classification stage, data comes from connected layer. This takes 2D output and convert to 1D. We run 50 epochs for both VGG16 and ResNet50 with default train and test generator length. So, the model runs 50 times to train and runs same time for test data.

Chapter 4

Result Analysis

We have experimented with brain tumor MRI images of 224 real brain pictures developed by radiographers using data from real-life patients publicly available. Our information is divided into training, validation and testing. 165 training pictures, 41 validation images, 18 test pictures to assess our model accuracy. To begin, we do data augmentation to improve our dataset by making tiny changes to our MRI pictures and extracting these enhanced images from our suggested model. The experiment is done using TensorFlow and Keras libraries in python on a CPU having a 2.40 GHz core i3 7th gen processor with 8 Gb of ram. On our training data, our suggested model was 92 percent accurate, and on our validation dataset, it was 90 percent accurate. On the same dataset, we trained pre-trained ResNet-50 to compare the accuracy of our model. ResNet-50 achieved 84 percent accuracy rate on training data and an 88 percent accuracy rate on validation data.

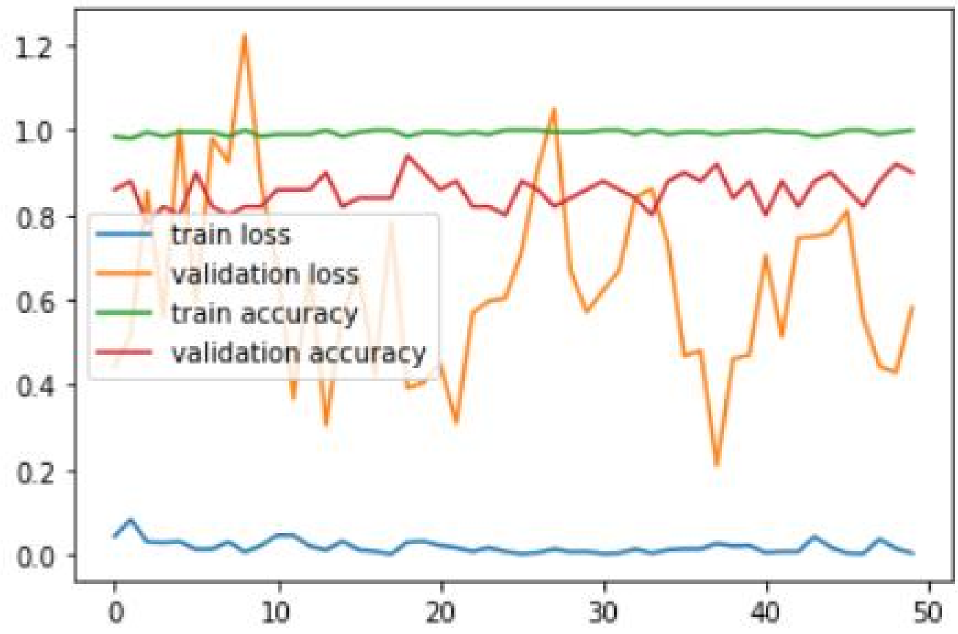

Figure 4.1: VGG-16 Model

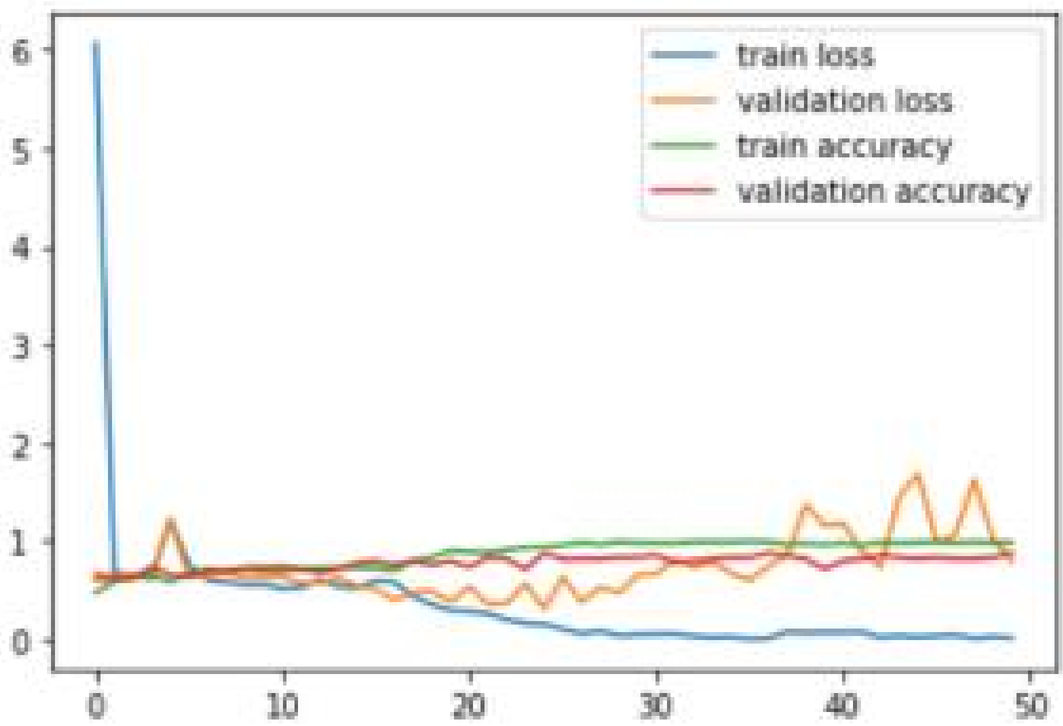

Figure 4.2: ResNet-50 Model

We performed a blind test on unseen testing data on our model. In the Table, True positive and true negative labels organize correct classification, with TP denoting abnormal brain images and TN denoting normal brain images. On the other hand, false positive and false negative are responsible for classifying incorrect results, which are displayed when FP classifies normal brain images as positive tumors and FN classifies abnormal brain images as negative tumors.
Here we have calculated the Accuracy, Precision, Recall, and F1-Score of proposed models. Accuracy is the measurement of actual true classifications. We plotted all our results of Accuracy, Precision, Recall and F1-score in the table in an evaluative way. Our proposed system may have prognostic value in the detection of tumors in patients with brain tumors. To further improve the model's efficiency, extensive hyper parameter tuning and a more sophisticated preprocessing technique could be considered. The method can be further developed for categorical classification problems such as classifying different types of brain tumors, such as Glioma, Meningioma, and Pituitary tumors, or can be used to find other types of brain abnormalities.

Model	True Positive	True Negative	False Positive	False Negative
VGG-16	28	18	3	1
ResNet-50	28	14	5	3

Table 4.1: Model performance

Model	Accuracy	Recall	Precision	F1-score	Time(sec)
VGG-16	.92%	1.47	1.33	1.4	960
ResNet-50	.84%	0.90	0.85	0.87	1004

Table 4.2: Results from VGG-16 and Resnet-50

Chapter 5

Conclusion and Future Work

5.1 Concluding Remark

Finding the dataset of MRI image which will work in VGG 16 and ResNet-50 architecture pattern, was a bit challenging. Data needs to crop and resize before train, test, and validate. Our proposed model has prepared raw data properly. We used VGG16 which is a CNN architecture are using nowadays for small dataset. In our model, MRI brain images are trained using different libraries of VGG16 and we got 92% accuracy in the model. Later, we put the same data set on ResNet-50 where we got an accuracy of 84% with much longer time to process the data taking 1004 seconds where VGG-16 model takes 960 seconds. Our proposed system has the potential to be effective in the early detection of dangerous disease in other clinical domains involving medical imaging, particularly lung and breast cancer, which have extremely high global mortality rates. We can extend this approach to other scientific fields where large data sets are scarce, or we can use different transfer learning methods with the same proposed technique. The proposed method achieved optimistic result which can be further improved and worked on in the near future.

5.2 Future Works

We will work on extending this proposed model in future by using more architecture of CNN like VGG19 and so on. Increasing the accuracy of the model will be target. We will also work on detecting different types of tumor in human body and as well as to find the actual size. In this model we used a small dataset but our future will include to work on more to test, and train so the performance become more accurate.

Bibliography

[1] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[2] V. Chandnani, T. Yeager, T. DeBerardino, K. Christensen, J. Gagliardi, D. Heitz, D. Baird, and M. Hansen, "Glenoid labral tears: Prospective evaluation with mri imaging, mr arthrography, and ct arthrography.," *AJR. American journal of roentgenology*, vol. 161, no. 6, pp. 1229–1235, 1993.

[3] R. Beck, "Image processing in the context of imaging science," in *Proceedings., International Conference on Image Processing*, IEEE, vol. 2, 1995, pp. 304–307.

[4] D. I. Hoult and B. Bhakar, "Nmr signal reception: Virtual photons and coherent spontaneous emission," *Concepts in Magnetic Resonance: An Educational Journal*, vol. 9, no. 5, pp. 277–297, 1997.

[5] D. M. Wolpert, R. C. Miall, and M. Kawato, "Internal models in the cerebellum," *Trends in cognitive sciences*, vol. 2, no. 9, pp. 338–347, 1998.

[6] P.-S. Liao, T.-S. Chen, P.-C. Chung, *et al.*, "A fast algorithm for multilevel thresholding," *J. Inf. Sci. Eng.*, vol. 17, no. 5, pp. 713–727, 2001.

[7] A. L. Rhoton Jr, "The cerebrum," *Neurosurgery*, vol. 51, no. suppl_4, S1–1, 2002.

[8] J.-L. Danger, S. Guilley, S. Bhasin, and M. Nassar, "Overview of dual rail with precharge logic styles to thwart implementation-level attacks on hardware cryptoprocessors," in *2009 3rd International Conference on Signals, Circuits and Systems (SCS)*, IEEE, 2009, pp. 1–8.

[9] B. Kolb and I. Q. Whishaw, *Fundamentals of human neuropsychology.* Macmillan, 2009.

[10] D. Longo, *Seizures and epilepsy harrison's principles of internal medicine*, 2012.

[11] R. Yogamangalam and B. Karthikeyan, "Segmentation techniques comparison in image processing," 2013.

[12] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, IEEE, 2017, pp. 1–6.

[13] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using u-net based fully convolutional networks," in *annual conference on medical image understanding and analysis*, Springer, 2017, pp. 506–517.

[14] S. Hussain, S. M. Anwar, and M. Majid, "Brain tumor segmentation using cascaded deep convolutional neural network," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2017, pp. 1998–2001.

[15] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation," *Medical image analysis*, vol. 36, pp. 61–78, 2017.

[16] M. B. Shakir, M. A. Hossain, K. M. A. Shams, and F. R. Akib, "Early detection, segmentation and quantification of coronary artery blockage using efficient image processing technique," Ph.D. dissertation, BRAC University, 2017.

[17] H. Shen, R. Wang, J. Zhang, and S. McKenna, "Multi-task fully convolutional network for brain tumour segmentation," in *Annual Conference on Medical Image Understanding and Analysis*, Springer, 2017, pp. 239–248.

[18] S. Cui, L. Mao, J. Jiang, C. Liu, and S. Xiong, "Automatic semantic segmentation of brain gliomas from mri images using a deep cascaded neural network," *Journal of healthcare engineering*, vol. 2018, 2018.

[19] R. Ezhilarasi and P. Varalakshmi, "Tumor detection in the brain using faster r-cnn," in *2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2018 2nd International Conference on*, IEEE, 2018, pp. 388–392.

[20] S. Khawaldeh, U. Pervaiz, A. Rafiq, and R. S. Alkhawaldeh, "Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks," *Applied Sciences*, vol. 8, no. 1, p. 27, 2018.

[21] X. Fan and H. Markram, "A brief history of simulation neuroscience," *Frontiers in neuroinformatics*, vol. 13, p. 32, 2019.

[22] W. Saib, D. Sengeh, G. Dlamini, and E. Singh, "Hierarchical deep learning ensemble to automate the classification of breast cancer pathology reports by icd-o topography," *arXiv preprint arXiv:2008.12571*, 2020.

[23] K. Sood and J. Fiaidhi, "Capsule networks: An alternative approach to image classification using convolutional neural networks," 2020.