# Brain Tumor Classification Using Deep Learning Through MRI Images

**Samaneh Sarfarazi**

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Engineering

Eastern Mediterranean University
July 2022
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

---

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Computer Engineering.

---

Prof. Dr. Hadi Işık Aybay
Chair, Department of Computer Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Computer Engineering.

---

Prof. Dr. Önsen Toygar
Supervisor

Examining Committee

---

1. Prof. Dr. Önsen Toygar ____________________

2. Assoc. Prof. Dr. Yıltan Bitirim ____________________

3. Asst. Prof. Dr. Mehtap Köse Ulukök ____________________

# ABSTRACT

A brain tumor is a dangerous neural illness produced by the strict growing of prison cell in the brain or head. The amount of persons suffering from brain tumor remains increasingly cumulative. Initial detection of wicked cancers is vital to provide cure to sickness, and early identification reduces the risk of death. If a brain cancer is not predicted in initial phase, it can assuredly cause to death. Hence, primary identification of brain tumors requires the usage of a mechanical means. The segmentation, analysis, and separation of unclean tumor parts from MRI images are the main source of anxiety. Nevertheless, the situation is a boring and slow procedure that radiologists or scientific professionals need to assume, and their act is only reliant on their knowledge. To report the segmented MRI images including tumor, the usage of computer-assisted methods come to be necessary.

In this thesis, a Convolutional Neural Network (CNN) approach is used to identify brain cancers in MRI images. The presented model focuses on improving accuracy because there has been a significant amount of research in this sector. This investigation is carried out using Python and Google Colab. Two datasets are used for this study, namely Kaggle Brain MRI dataset and Figshare Brain MRI dataset. Models of deep CNN, consisting of VGG16, AlexNet, and ResNet, are utilized to extract deep features. The classification accuracies of the aforementioned deep learning models are used to measure the efficiencies of the implemented systems.

For the Kaggle dataset, AlexNet achieves a 98% accuracy, VGG16 has 97% accuracy, and ResNet has 66% accuracy. Among these networks, AlexNet has provided the highest level of accuracy. In the Figshare dataset, AlexNet and VGG16 both achieve 99% accuracy, and ResNet has 96% accuracy. In terms of accuracy, AlexNet and VGG16 outperform ResNet. These performances aid in the early detection of cancers before they cause physical harm such as paralysis and other complications.

**Keywords:** CNN**,** deep learning, brain tumor detection.

# ÖZ

Beyin tümörü, beyindeki veya kafadaki hapishane hücresinin katı bir şekilde büyümesiyle meydana gelen tehlikeli bir sinir hastalığıdır. Beyin tümörü görülen insan sayısı kümülatif olarak artmaya devam etmektedir. Kötü huylu kanserlerin ilk tespiti, hastalığa çare bulunması için önemlidir ve erken teşhis ölüm riskini azaltır. Bir beyin kanseri ilk aşamada tahmin edilemezse, kesinlikle ölüme neden olabilir. Bu nedenle, beyin tümörlerinin birincil tanımlanması, mekanik bir yolun kullanılmasını gerektirir. Temiz olmayan tümör parçalarının MRI görüntülerinden bölütlenmesi, analizi ve ayrılması ana endişe kaynağıdır. Yine de durum, radyologların veya uzmanların üstlenmesi gereken sıkıcı ve yavaş bir prosedürdür ve eylemleri yalnızca bilgilerine bağlıdır. Tümör hücrelerini içeren bölütlenmiş MRI görüntülerini ortaya çıkarmak için bilgisayar destekli yöntemlerin kullanımı gerekmektedir.

Bu tezde, MRI görüntülerinde beyin kanserlerini tanımlamak için Evrişimsel Sinir Ağı (CNN) yöntemleri kullanılmıştır. Sunulan model, bu sektörde önemli miktarda araştırma yapıldığı için doğruluğu artırmaya odaklanmaktadır. Bu araştırma Python ve Google Colab kullanılarak gerçekleştirilmiştir. Bu çalışma için Kaggle Beyin MRI veri kümesi ve Figshare Beyin MRI veri kümesi olmak üzere iki veri kümesi kullanılmıştır. Derin öznitelikleri çıkarmak için VGG16, AlexNet ve ResNet'ten oluşan derin CNN modelleri kullanılmıştır. Herhangi bir derin öğrenme modelinin sınıflandırma doğruluğu, uygulanan sistemin verimliliğini ölçmek için kullanılmıştır.

Kaggle Beyin MRI veri kümesi için AlexNet %98, VGG16 %97 ve ResNet %66 doğruluğa sahiptir. Bu ağlar arasında en yüksek doğruluk seviyesini AlexNet sağlamıştır. Figshare Beyin MRI veri kümesinde AlexNet ve VGG16 %99, ResNet ise %96 doğruluğa sahiptir. Doğruluk açısından AlexNet ve VGG16, ResNet'ten daha iyi performans göstermiştir. Bu performanslar, kanserlerin felç ve diğer komplikasyonlar gibi fiziksel zararlara neden olmadan önce erken tespit edilmesine yardımcı olur.

**Anahtar Kelimeler:** CNN, derin öğrenme, beyin tümörü tespiti.

# ACKNOWLEDGMENT

Prof. Dr. Önsen Toygar, my supervisor, has my heartfelt gratitude for her guidance, wisdom, patience, and understanding throughout this time. My heartfelt gratitude also goes out to the Computer Engineering department's employees for their consideration and encouragement.

My thanks are extended to my parents and family for their unwavering support, prayers, and handcrafted task of developing me to be a whole human in terms of ethics, psychology, economics, and sociality. Your efforts will be recognized. I’d like to express my appreciation to my buddies in Cyprus and Turkey who have become my family.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AUC | Area Under ROC Curve |
| BTS | Brain Tumor Segmentation |
| CNN | Convolutional Neural Networks |
| CPU | Central Processor Unit |
| CAD | Computer-Assisted Diagnosis |
| CADe | Computer-Aided Detection |
| CRFs | Conditional Random Fields |
| DL | Deep Learning |
| DT | Decision Tree |
| DR | Digital Radiography |
| DNN | Deep Neural Network |
| DWT | Discrete Wavelet Transform |
| 3D CNN | 3 Dimensional Convolutional Neural Networks |
| FP | False Positive |
| FN | False Negative |
| FPR | False Positive Rate |
| GMM | Gaussian Mixture Model |
| GPU | Graphics Processing Units |
| GA | Genetic Algorithm |
| ICA | Independent Component Analysis |
| ISLES | Ischemic Stroke Lesion Segmentation |
| KNN | K-Nearest Neighbor |
| ML | Machine Learning |

| | |
|---|---|
| MRI | Magnetic Resonance Imaging |
| MTOP | Mild Traumatic Brain Injury Outcome Prediction |
| MCCNN | Multi-Cascaded Convolutional Neural Network |
| NN | Neural Networks |
| PCA | Principal Component Analysis |
| PSO | Particle Swarm Optimization |
| PNN | Probabilistic Neural Networks |
| RFC | Random Forest Classifier |
| RF | Random Forest |
| ROC | Receiver Operating Characteristics |
| SVM | Support Vector Machine |
| SGD | Stochastic Gradient Descent |
| TN | True Negative |
| TP | True Positive |
| WHO | World Health Organization |
| XG | Gradient Boosting |

# Chapter 1

# INTRODUCTION

A brain tumor is an imbalanced type of cell in the human brain. The brain of human is surrounded via a firm head. Slight development in such a minor part will cause intense problems. Tumors of brain may be malignant and nonmalignant. The gravity inside the head will increase such as nonthreatening or malicious cancers progress. This will consequence in enduring head damage or death of the person. Samples of brain MRI images are demonstrated in Figure 1.1 with a healthy MRI image and a brain MRI image with a tumor.

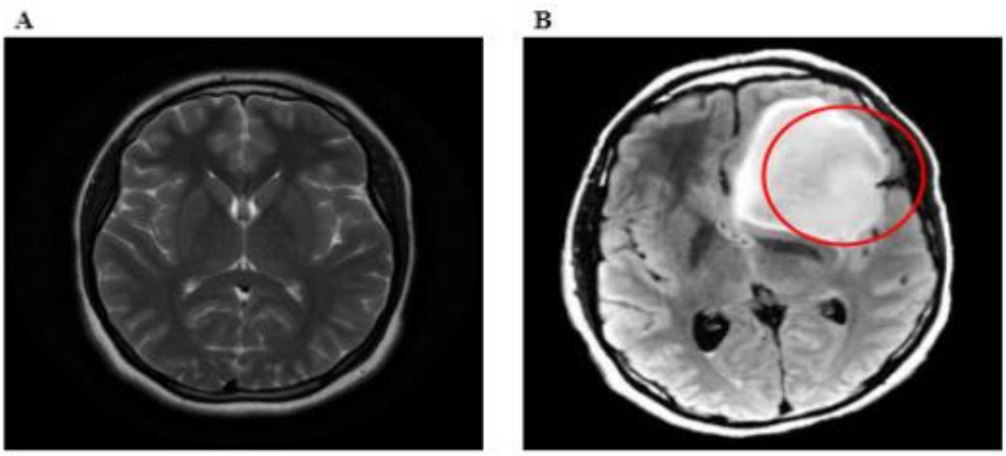


Figure 1.1:Samples of Brain MRI Images with (A) a Healthy Brain MRI (B) a Brain Scan with a Tumor

Experts and investigators have been studying to emerge complex methods and approaches aimed at diagnosing tumors of brain. While MRI depictions and Tomography of Computer (CT) are the two approaches with broadly usage which aimed at clarifying the anomalies in form, mass, or brain materials place that help doctors in identifying the cancers; Magnetic Resonance Imaging (MRI) picture is preferred more than the aforementioned methods by the specialists. Therefore, experts and scientists have absorbed Magnetic Resonance Imaging pictures. However, automatic methods, mostly applied by computer assisted – medicinal image processing methods, exist progressively helping surgeons for noticing tumors of brain.

In tumors of brain, Magnetic Resonance Imaging picture is generally made in three unlike forms that can be seen in Figure 1.2. The three dissimilar forms offer extra exact data about the form, material and capacity of the tumors of brain. MRI in different forms is specified in Figure 1.2.

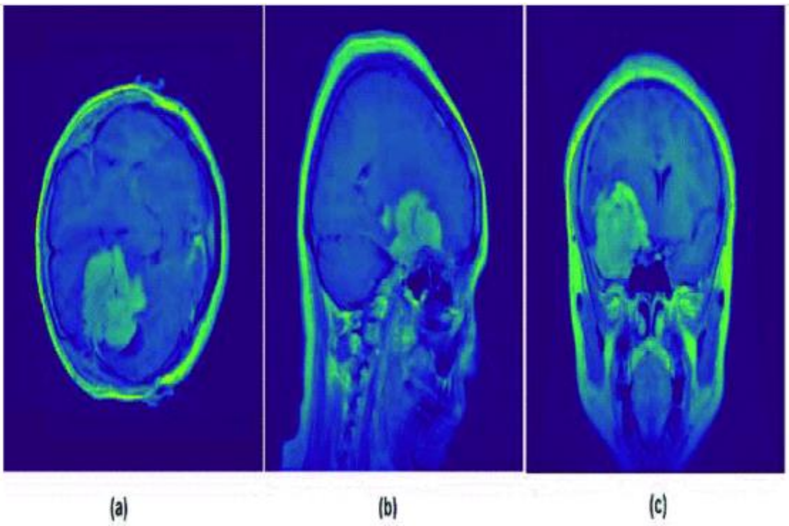

Figure 1.2: Brain MRI in Dissimilar Formats (a) Axial Form (b) Coronal Form (c) Sagittal Form

Handcrafted methods with Machine Learning (ML) classifiers are developed for training data examples. Handcrafted approaches are commonly being utilized in the area of health-informatics, estimating epidemic, assessing user experience in playing games, expecting shave asset. Also, numerous Handcrafted trainings are shown on health pictures to categorize tumors of brain. Medicinal image processing includes pre-processing and post-processing. These stages may be applied through the Handcrafted methods as perfect as the method of deep learning. In handcrafted methods, features are extracted to get consequences from pictures of test and the procedure is quick. In the Deep Learning (DL) methods, networks are adjusted through properly choosing the sum of layers, activation function, pooling. But, in both methods, new algorithms are possible to be employed to increase the system's accuracy in a wider viewpoint. Deep Learning techniques for identifying brain cancers in MRI scans are the topic of this thesis.

The main part of this thesis focuses on finding tumors of brain through MRI images using methods of Deep Learning. Consequently, this thesis will offer the anticipated result i.e., effective deep learning method to distinguish tumors of brain by MRI images that will contribute medical experts to run appropriate cure.

The following is the structure of the thesis. The introduction of the thesis, background on brain tumor classification and the works done by other researchers using handcrafted and deep learning methods are presented in Chapter 1. Chapter 2 briefly explains deep learning applications. Later on, in Chapter 3, the thesis methodology and metrics for evaluations are presented. Chapter 4 presents the experimental results, discussion and the comparison with the state-of-the-art. Lastly, Chapter 5 gives the conclusions of this thesis and future directions.

## 1.1 Background on Brain Tumor Detection

This section in brief deliberates the researches that are shown to distinguish tumors of brain using the dissimilar advanced know-hows. Rehman et al. [1] recommended a different knowledge-founded technique for mini tumor of brain finding and type of tumor classification. The initial stage of their study focused on using a 3D CNN to abstract tumors in the brain, which were then transferred to a Convolutional Neural Network model having already received training to abstract model features. The features that have been extracted are carried out to a correlation-based election procedure, and the greatest characteristics are selected in place of the outcome. In the last classification, with the usage of feed forward neural network, the elected features are tested.

Amin et al. [2] used a Deep Neural Network (DNN)-based design for brain tumor segmentation. The proposed model has seven layers in the classification section, including three ReLU layers, three convolutional layers, and a SoftMax layer. Kebir et al. [3] suggested an approach that is supervised for identifying the anomalies of brain through the Magnetic Resonance Imaging images in several phases. The initial stage is to change a DL Convolutional Neural Network model, and after that a subclass of MRI brain images is completed via the k-mean process conformed by factor of brain grouping as standard or nonstandard groups in accordance with the advanced Convolutional Neural Network model. Alternatively, Vinoth et al. [4] introduced a Convolutional Neural Network-based automatic separation technique. At this point, classification was done with kernels, and Support Vector Machine classification was done with computed variables. Furthermore, MATLAB is used to extract and recognize malignancies from MRI images of the brain.

A Convolutional Neural Network founded on model of deep learning was effectively connected to the regarded issue of tumor of brain classification.

Classifier based on Convolutional Neural Networks constructions have the advantage of not requiring bodily sectioned tumor zones.

Sajjad et al. [5] proposed multi-grade tumors classification through exerting technique of data augmentation to images of MRI and at that time change it by means of a pre-trained model of VGG-19 Convolutional Neural Network architecture. They used k-Nearest Neighbor and multinomial Logistic Regression approaches for the classification of brain tumors.

Saed et al. [6] recommended a structure for classifying brain MRI images into natural and unnatural states, as well as a classifying structure for labeling unnatural brain images into tiny and large markings. The experimental results were presented with a 98.51 % training accuracy and an 84.19 % validation accuracy.

Talo et al. [7] classified normal and pathological Brain MRI photographs with 100% accuracy by means of the ResNet34 pre-trained Convolutional Neural Network model through a data augmentation technique of transfer learning. On the other hand, a model of the pre-trained ResNet50 Convolutional Neural Network was updated by Çınar et al. [8] by eliminating the preceding five levels and replacing them with eight new layers, matching the accuracy of prior pre-trained patterns as ResNet50, AlexNet, and GoogleNet. The reconstructed ResNet50 pattern achieved 97.2% accuracy indicating real repercussions.

## 1.2 Review of Handcrafted Methods Using Machine Learning Classifiers for Brain Tumor Detection

Handcrafted methods with Machine Learning classifiers have recently received many traction, and they have been used in numerous applications like spam detection, text mining, image categorization, video proposal, sight and sound thinking recovery. This section discusses Handcrafted methods using advanced Machine Learning classifiers for detecting and classifying brain tumors. A superlative hyper-plane considers Support Vector Machine (SVM) which is a widely used Machine Learning classifier. The SVM's algorithm try to separate data in the most discriminative way. As a result, SVM serves as the optimal border between the two classes. Another supervised machine learning approach, namely Logistic Regression, forecasts a binary result determined by a number of input features. The foremost goal of Logistic Regression is to find the best-fitting model to specify the relation among input and output features. On the other hand, k-Nearest Neighbor (kNN) is a classification technique used for supervised machine learning. kNN forecasts which known information sample relates to an explicit class or the other by computing the space among one sample and the others. The provided instance is assigned to the class with the most samples that are closest to it. k is a constant value and shows how many samples are enough to decide about for new sample. Naive Bayes (NB) is also a supervised method which is extensively employed for two class problems. Its concept comes from Bayes' theorem, which assumes individuality among forecasters. The NB classifier states that the attendance of one attribute in a class is totally different with the presence of another.

Decision Tree (DT) is also a supervised learning method. DT is employed for explaining binary classification problems. DTs acquire easy choice instructions derived from the feature values. Then they predict a target variable's value. Random Forest creates many DTs and finally combines the results of all trees together and makes a decision. Sometimes using DTs when the dataset is big implies over-fitting. So, RF is designed to evade over-fitting. Random Forest is suitable for both classification and regression problems. Stochastic Gradient Descent (SGD) is a method that decreases the cost function by using optimization algorithms. During the process, parameters and coefficients are changed. To support different cost functions, this algorithm uses a SGD learning routine. An advanced version of SGD is called Extreme Gradient Boosting (XGBoost). XGBoost shows how using GB trees can solve a supervised learning problem. It is a group of DTs which uses GB techniques to estimate result for an unseen data. This algorithm finds the features with low accuracy and modifies the problem somehow to predict with a high accuracy. The errors help the model to focus on a region that the previous model could not predict well.

There are several machine learning approaches for brain tumor classification and segmentation using MRI in the scientific literature. Hasan et al. [9] suggested an Image of an MRI brain scan categorization system based on deep and custom features. Pre-processed MRI is useful to an altered GLCM for extraction of statistical features. CNN extracts features automatically. SVM classification with Cross-validation 10-fold performed 99.30% accurately based on 600 sagittal MRI scans. While likened to new networks of transfer learning, such as GoogleNet, and AlexNet the recommended method performed fine by means of combining CNN and MGLCM features.

A Naive Bayes-based brain cancer identification approach uses maximum entropy segmentation [10]. The REMBRANDT dataset, which includes 114 MRI images, is used to exam the system. The recommended system has the advantage of detecting tumors anywhere in the brain, such as the temporal lobe. Sert et al. [11] proposes a different scheme aimed at diagnosing brain tumors based on CNN and maximum fuzzy entropy segmentation.

To improve resolution of MRI, the super resolution of a single image is employed. Pre-trained ResNet architecture is used to extract features. SVM Binary classification has a 95% accuracy rate. Edge Adaptive Total Variation [12] uses the mean shift clustering approach for brain tumor categorization segmentation. The proposed technique has 2 advantages: When utilizing the image, mean shift clustering, and EADTV preserves the edges mean shift clustering, unlike K-mean and fuzzy cmean, automatically updates cluster centers. In an integrated Approach of PSO with fusion features for tumor of brain diagnosis, a fine-tuned capsule network feature extraction and local binary pattern are applied.

SVM classification accuracy on the BRATS2018 and RIDER databases is 98.3 % and 97.9%, respectively. The new proposal has shown good results by combining hand-crafted and deep features. On the CEMRI dataset, SVM and kNN classifiers are used to assess pre-trained GoogleNet for deep feature extraction for 3 class classification into Glioma, Meningioma with accuracy of 97.8% and 98%, respectively. The BRATS 2017 dataset, which contains 48 images, is used to assess the accuracy of a multinomial logistic regression model for brain tumor classification. The system's performance, however, should be evaluated on larger datasets. The system works well with the genetic algorithm SVM classification method. In an effective brain tumor classification optimization technique, genetic algorithm (GA) is employed in order to segment tumors. With 91.23 %, SVM is provided GLCM texture characteristics [13].

For high-grade glioma (HGG) and low-grade gliomas (LGG) brain tumor categorization, Polly et al. [14] developed a k-means segmentation algorithm. From wavelet features, PCA is used to determine ten relevant features. On the way to discriminate between images that are normal and abnormal, the SVM algorithm is utilized. Once again, the SVM classification method is employed on the way to classify LGG and HGG tumors in aberrant pictures. On 440 photographs, the suggested technique achieves 99 %, but it needs to be evaluated on a larger database using added important data. A technique aimed at tumor of brain diagnosis by wavelet transform usages a morphological process by a method of threshold-based used for segmentation.

Amin et al. [15] suggested a new technique for identifying brain tumors using MRI. To reduce noise and smooth MRI, skull stripping and Gaussian filtering are used. Following K-means segmentation, GLCM texture characteristics are extracted. The system is tested on three datasets: local, AANLIB, and RIDER, using linear, RBF, and cubic SVM kernels. The linear kernel with 5-fold cross validation was found to have a 98% accuracy.

Minz et al. [16] provide a study that uses the Adaboost classifier to classify brain tumors. Following median filtering, threshold-based segmentation is used to reduce noise. Using GLCM characteristics, the system proposes texture-based classification.

PO outperforms the CSO technique in terms of accuracy, robustness, and execution time. Shankar et al. [17] proposes exploiting texture features to classify brain tumors using Gustafson-Kessel fuzzy clustering. A histogram-based approach is used to segment preprocessed images with the Wiener filter. G-K fuzzy approach is given GLCM texture features for binary classification with 95% accuracy.

The state-of-the-art techniques discussed so far are abridged in Table 1.1 for handcrafted method using machine learning classification. Table 1.1 shows the preprocessing, segmentation, feature extraction and classification techniques employed in each study. Additionally, the dataset details and the maximum accuracy reported are also shown.

Table 1.1: Summary of Handcrafted Methods Based Brain Tumor Detection

| Ref. | Publication year | Preprocessing and Segmentation | Features | Classification | Dataset | Accuracy |
|---|---|---|---|---|---|---|
| [9] | 2019 | Image enhancement and resizing | GLCM, CNN | CE-MRI | Iraqi-based research facility | 99.30% |
| [10] | 2019 | Maximum entropy threshold segmentation, morphological procedure, pixel subtraction | Intensity, Morphological | Naive Bayes | REMBRANDT | 94% |
| [11] | 2019 | Segmentation-Maximum fuzzy entropy for single picture super resolution for image enhancement (MFE) | ResNet's advanced features | SVM | TCIA | 95% |
| [12] | 2019 | Resize 224*224, min-max normalization | Deep aspects of GoogleNet | SVM, KNN | CE-MRI | 97.8% SVM, 98% KNN |
| [13] | 2018 | GA segmentation with the median filter | GLCM | SVM | Medical Dataset from Harvard | 91.23% |
| [14] | 2018 | Binarization K-means clustering, OTSU | DWT | SVM | BRATS 2013,BRATS 2017,Midas | 99% |
| [15] | 2017 | K-Means segmentation, skull stripping- BSE Gaussian filtering | Intensity, GLCM, and shape | SVM | AANLIB, RIDER, and Local | 98% |
| [16] | 2017 | Global adaptive segmentation, image enhancement-DSR-AD | RLCP | Naive Bayes | Local-JMCD,BRATS | 96% |
| [17] | 2016 | Histogram-based segmentation with Wiener filtering | GLCM | G-K Fuzzy system | - | 95% |

## 1.3 Review of Deep Learning Methods for Brain Tumor Detection

Over the most recent couple of years, deep learning model has been considered the best quality level in the ML association. Furthermore, it has become the most extensively utilized computational technique in the arena of Machine Learning over time. A deep learning system's ability to acquire enormous amounts of data is one of its advantages. Deep learning has progressed quickly in current times, and it is currently routinely employed to handle a variety of classic tasks. In a range of disciplines, such as natural language processing, cybersecurity, control and robotics, bioinformatics, and processing of pharmacological data, traditional machine learning methodologies have been overtaken by deep learning. Regardless of this, it has been donated numerous mechanisms studying the advances on deep learning. Every one of these mechanisms just embraced one component of the DL, which prompts an overall shortfall of data about it. Thus, in this impact, it is suggested to utilize a complete method so as to run an appropriate foundation from that to mature a complete sympathetic of deep learning.

Using traditional Handcrafted methods with Machine Learning classifiers to do the classification task necessitates a number of steps. The first step is preparing data for classification which is called pre-processing. Feature extraction and feature selection are the next steps. Finally classification step will run. Also, feature selection has a significant influence on the presentation of machine learning approaches. The application of the biased feature can lead to erroneous classification of classes. In comparison to typical methods of machine learning, deep Learning allows for the automated learning of sets of features for a range of tasks. Deep learning causes classification and learning as shown in Figure 1.3. Deep learning has turned into an incredibly broad sort of ML algorithm lately because of the enormous advancement and development of the field of huge information. It is still under constant development in terms of new presentation for a wide range of Machine Learning errands, and it has worked on improving a few learning sectors, image super-resolution, object detection, and image detection, to name a few. Lately, performance of Deep Learning has come to outperform human follow up on errands like categorization of images as shown in Figure 1.4.

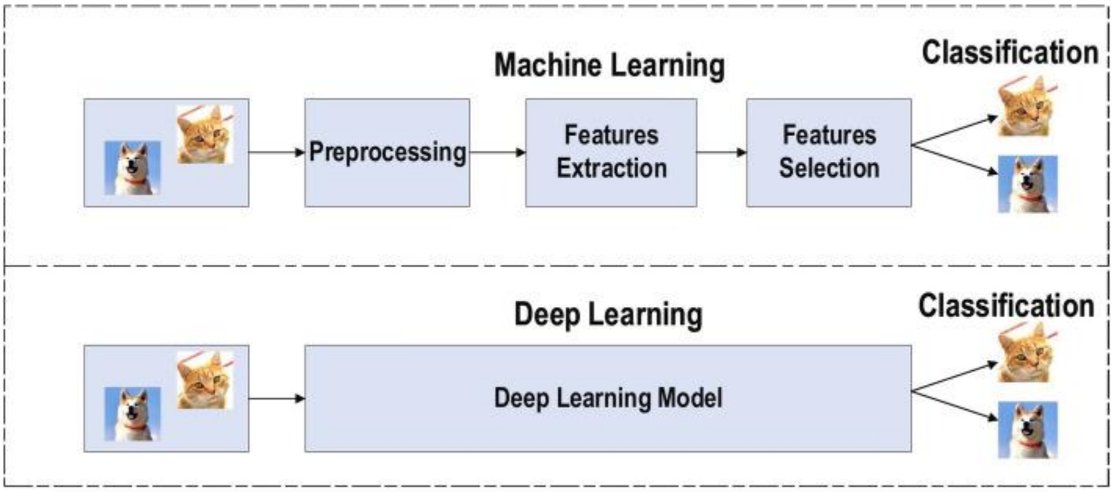


Figure 1.3: The Difference Among Old Machine Learning Methods and Deep Learning [18]

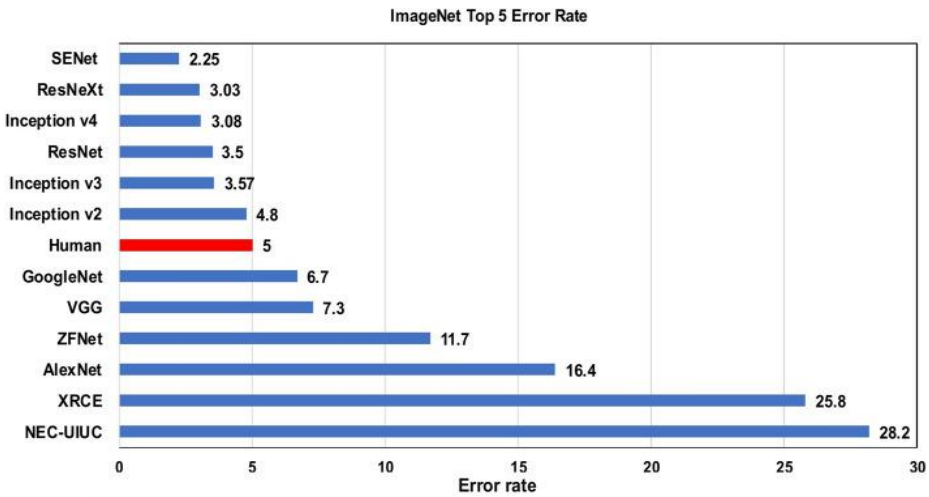


Figure 1.4: The Performance of Deep Learning Compared To that of Humans [18]

The detection of brain tumors using systems of deep learning is a cutting-edge subject of study. Researchers use a variety of deep learning architectures to automatically segment and classify brain tumors. For brain classification, The Regularized Extreme Learning Technique with Mix Features was recommended via Gumaei et al. [19].

A feature extractor of Hybrid PCA-NGIST can be utilized for 3-D feature extraction. The NGIST feature descriptor is a descriptor of standardized feature that is utilized to address image illumination and shadowing issues. RELM is a single hidden layer, input, and output feed forward neural network. The suggested technique is examined used for three kinds of tumors: neuroendocrine tumor, glioma, and pituitary tumor using CE-MRI database with 94.33 % after 5 fold cross-validation. Link Net is a small deep neural network design which is employed to classify brain cancers [20]. On a freely released UCI repository database, Binary classification achieved 91% accuracy.

The Multi-Layer Perceptron (MLP) classification system has a 96 % accuracy rate and a 0.65 Kappa Statistic. However, sparse auto encoder could be examined in the future when DNN is combined with other auto encoder versions such as denoising auto encoder. Latif et al. [21] Proposes a brain tumor classification method based on transfer learning. To suitthe VGG19 model, MRI images are scaled to 224*224 pixels. To update the weights, Fine tweaking of parameters such as learning rate, scheduling rate, and momentum is done block by block. The system has a 94.82% accuracy on the CE-MRI database. The disadvantage of this approach is that fine-tuning settings block by block takes 20-30 minutes to train the CNN classifier. MLP uses statistical and wavelet features to Classify brain tumors [22].

The scheme is assessed using both statistical and DWT characteristics, as well as a 2015 dataset, ANFIS is used to differentiate between normal and Glioma brain tumors. Methods of traditional classification similar to CNN and SVM produce errors of classification in Glioma images with low intensity, whereas ANFIS works perfectly in high and low levels of intensity Images of glioma. A scheme according to deep neural networks aimed at categorization of brain tumors is tested on the AANLIB dataset's 66 MRI images. For MRI segmentation, fuzzy C-mean clustering is employed. The relevant feature set from DWT features are selected using PCA. Normal, Glioblastoma, Sarcoma, and Carcinoma tumors are classified into four classes using a seven-layer DNN architecture.

Raju et al. suggested a unique technique for classification of brain tumor by means of Bayesian fuzzy clustering and HSC-founded multi SVN [23]. Information theoretic, scatter, and wavelet characteristics make up the feature vector. With Multi SVNN, the proposed approach provides four levels of classification. Normal and abnormal tumors are classified as level 1, edema tumors as level 2, core tumors as level 3, and progressed cancers as level 4. The HCS algorithm, which combines the Harmony and Crow search algorithms, is used to optimize the weights in SVNN. This strategy has the advantage of making Harmony search easier and Crow search faster to the global optimum.

Abdalla et al. [24] established a computer-aided design (CAD) structure aimed at ANN-based brain tumor detection. Using Haarlick texture features, a feed forward neural network is tested with 99 % accuracy on the AANLIB dataset, which has 239 images. Using a multi grade brain categorization method in [25], CNN is proposed. The Input Cascade CNN architecture is used to segment the tumor. Extensive data

Augmentation methods such as rotation, flipping, and embossing are used to expand data samples. The architecture of VGG19 is used to assess the system on a CE- MRI dataset with 94.58 percent accuracy. The proposed system uses substantial data augmentation approaches to overcome the issue of insufficient MRI picture availability. For increased performance, a CAD system based on lightweight CNN Building designed for brain tumor grading classification could be developed in the future.

The experiments with different preprocessing and deep learning approaches are conducted on different datasets and the accuracies are shown in Table 1.2. Additionally, Table 1.2 summarizes various Deep Learning approaches with the details related to preprocessing, classification, dataset and accuracy.

Table 1.2: Summary of Deep Learning Based Brain Tumor Detection

| Ref. | Publication year | Preprocessing | Classification | Dataset | Accuracy |
|---|---|---|---|---|---|
| [19] | 2019 | Contrast enhancement and normalization of intensity | RELM | CE-MRI | 94.33% |
| [20] | 2019 | Pixel subtraction, Average filter | CNN-Link Net | UCI | 91% |
| [21] | 2018 | DWT features,GLCM, Histogram | MLP | BRATS 2015 | 96.73% |
| [22] | 2018 | DWT features | DNN | AANLIB | 96.97% |
| [23] | 2018 | Bayesian fuzzy clustering segmentation, information theoretic, scatters and wavelet features | HSC based multi SVNN | BRATS | 93% |
| [24] | 2018 | Sharpening and smoothing filters, Threshold based segmentation, SGLD features | ANN | AANLIB | 99% |
| [25] | 2018 | Data augmentation, Input Cascade CNN segmentation | VGG19 | CE-MRI | 94.58% |

# Chapter 2

# DEEP LEARNING APPLICATIONS

A lot of DL applications are now unavoidable wherever you go in the world. These are the programs comprising analysis of social network, healthcare, visual data analyzing methods (for example computer vision and analysis of multimedia data), speech and audio processing (similar identification and improvement), and sentence classification and translation, and so on as shown in Figure 2.1. These applications can be categorized into five groups:

1. Classification
2. Localization
3. Detection
4. Segmentation
5. Registration.

Despite the fact that each of these vocations has its own aim, as seen in Figure 2.2, they all have a common goal. The pipeline implementation of these programs has a lot of commonality.

A theory that divides a bunch of data into categories is known as classification. Recognition is a technique for locating noteworthy things in an image while taking into account the background. During detection, bounding boxes encircle many objects, maybe from different classes. The concept of localization is used to find an object that has a single bounding box. Outlines surround the target object bounds in segmentation (semantic segmentation). Registration is the process of recognizing and appropriate a solitary image (It could be two-dimensional or three-dimensional) onto another. One of the best significant and diverse applications of DL is healthcare. Because of its connection to human existence, this field of research is extremely important. Furthermore, DL has excelled in the field of healthcare. As a result, we use DL applications to define DL applications within the domain of medical image analysis.

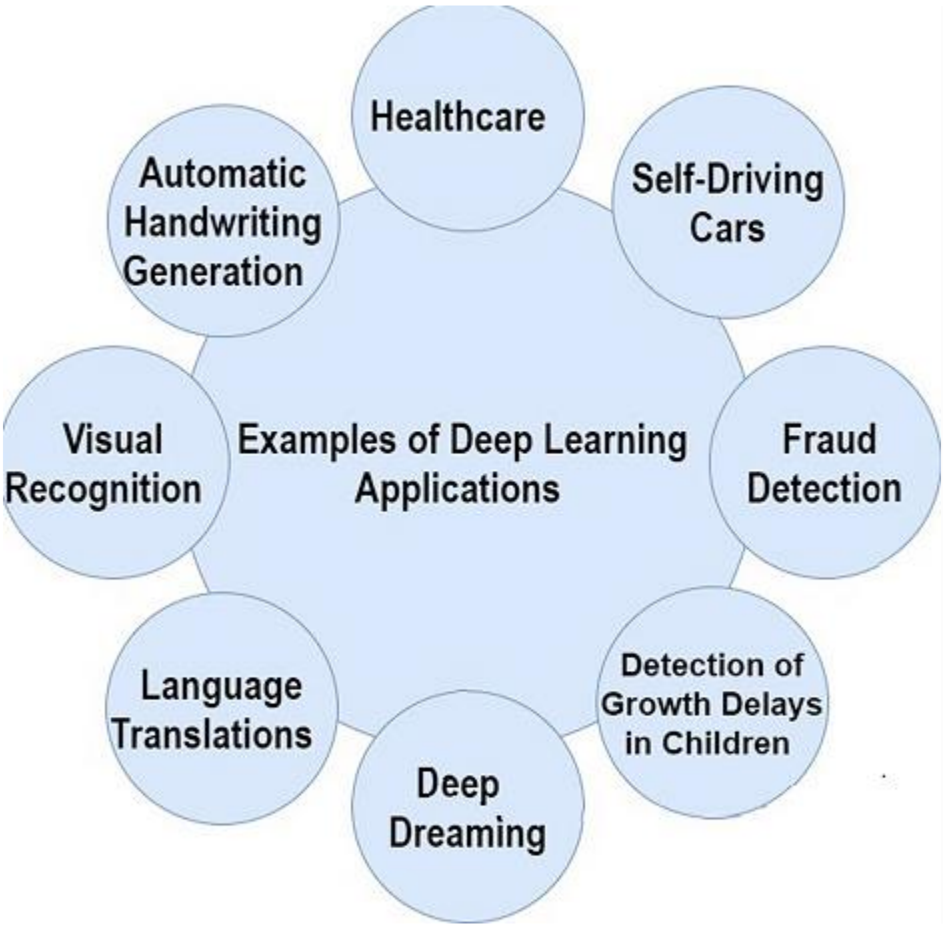


Figure 2.1: Examples of Deep Learning Applications [26]

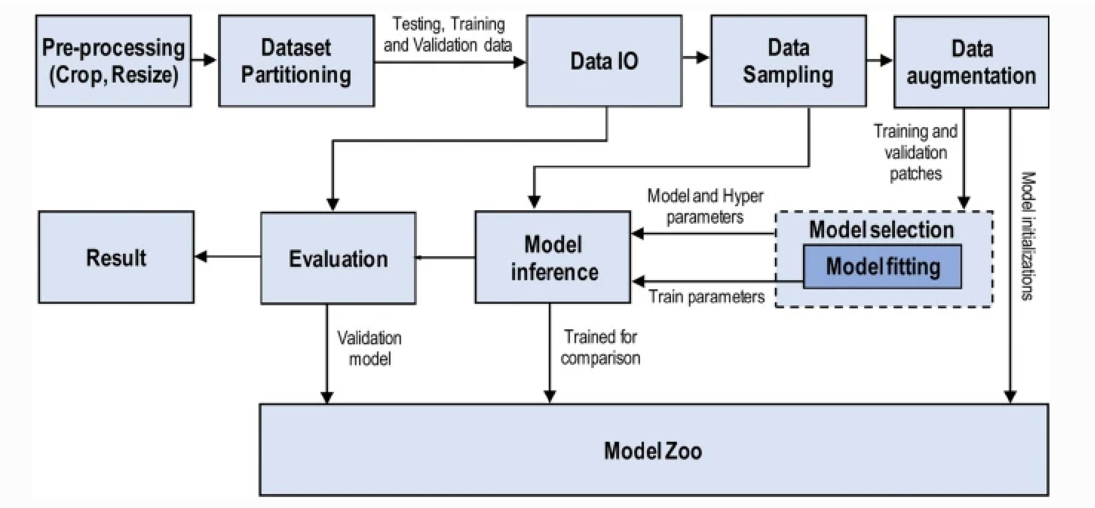


Figure 2.2: Task Flow for Deep Learning Applications [26]

## 2.1 Classification

Computer-Aided Analysis is another term for classification that is sometimes used. Using a chest X-ray dataset, Bharati et al. [27], used a CNN to diagnose lung diseases. Another study used CNN to attempt to read X-ray pictures. The ease with which these images may be accessed has most certainly aided the growth of DL in this media. For the training and testing procedures, the researchers utilized an upgraded pre-trained GoogLeNet CNN with over 150,000 images. From 1850 chest X-rays, this dataset was expanded. The developers restructured the location of image into side and front views, achieving about 100 precent accuracy. The therapeutic utility of this research on orientation classification is limited. Data augmentation combined with pre-trained models can help to analyzing the structure of photos in a totally automated framework.

Chest infection, sometimes known as pneumonia, is a common yet treatable illness that affects individuals all over the world. CheXNet, an enhanced variant of DenseNet with 121 convolution layers, was advanced by Rajpurkar et al., was used., to classify 14 different diseases. The CheXNet14 database, which contains 112,000 photos, was employed by these researchers. This network performed exceptionally well in detecting fourteen distinct diseases. By means of receiver operating characteristics (ROC) analysis, pneumonia classification achieved an AUC of 0.7632. Also, a trio of radiologists and four different radiologists fared as well as or better than the network. CNN has been accepted by Zuo et al. [28]. for candidate classification in lung nodules. To categorize lung nodules, CNNs were used to train SVM and Random Forest (RF) classifiers by Shen et al. In each of the three simultaneous CNNs, they used two convolutional layers. The LIDC-IDRI (Lung Image Database Consortium) database, has Lung images labeled with 1010, was accustomed to categorize the two kinds of lung nodes (benignm and alignant).Every CNN extracted features at different scales from the picture patches, the learnt features were utilised to build the output feature vector. Some of these vectors were categorized as malignant and some other as benign. These categorization was obtained by using RF classifier. In some cases SVM was eployed. As a kernel for SVM, radial basis function (RBF) is used. Despite a variety of noisy input levels, The model was successful in classifying nodules with an accuracy of 86%. On the other hand, 3D CNNs were utilized in the model to fill in the image data gaps between PET and MRI pictures.

In the “Alzheimer Disease Neuroimaging Initiative” (ADNI) dataset, they examined 830 individuals and MRI images of Alzheimer's patients. The 3D CNNs are trained using PET and MRI images, first as input and subsequently as output. Furthermore, the 3D CNNs used the training images to reconstruct PET scans for patients who lacked them. These reconstructed photos came close to matching the originals.

## 2.2 Localization

Although structural education applications may rise in popularity, professional clinicians are more possible to be attracted in the localization of usual anatomy.

Outside of human influence, radiological pictures are independently analyzed and characterized, while localisation could be valuable in fully automated programs that run from beginning to end. In [29], Roth et al. designed a CNN which had 5 convolutional layers. The authors classified the pelvis, legs, lung, liver, and neck into five categories. They were able to get an AUC of 0.998 after using data augmentation approaches, and the model's classification error rate was 5.9%. The spleen, kidney, heart, and liver must all be located.

## 2.3 Detection

A strategy aimed at locating information is Computer-Aided Detection (CADe). Observing a lesion on a scan might have disastrous effects for both the practitioner and the patient. As a result, the science of detection demands both precision and sensitivity. Chouhan et al. [30] recommended a different framework of deep learning for pneumonia detection according to the sense of transfer learning. Their strategy had a 96.4% accuracy and a 99.62% recall on unseen data. Numerous convolutional neural network algorithms for automatic discovery from X-ray pictures have been proposed in the field of COVID-19 and pulmonary ailment, with outstanding results.

## 2.4 Segmentation

Organs like knee cartilage, prostate, and liver have been explored in MRI and CT image segmentation investigations, Despite the fact that the majority of research has focused on brain segmentation, specifically tumors. This is a critical issue in surgical planning since it is hard to determine the particular tumor limits for the most efficient surgical excision. Neurological abnormalities such as cognitive damage, emotionlessness, and limb issues may emerge from the extreme sacrifice of critical brain areas during surgery. Anatomical segmentation used to be done by hand in medicine, with the doctor drawing lines slice by slice throughout the entire stack of the CT or MRI volume. As a result, it is ideal for putting in place a system that automates this time-consuming process. Wadhwa et al. gave a concise summary of brain tumor MRI picture segmentation.

Convolutional neural networks were utilized by Chen et al. [31] to precisely segment brain malignancies. As part of their goal for greater feature learning, they used the Deep Medic algorithm. It is an advance dual-force training structure. The loss function in this method is a label spreading-founded. Also a post-processing which is based on Multi-Layer structure and Perceptron-based conept is used. The BRATS 2017 and BRATS 2015 databases, which are the two most recent brain tumor segmentation datasets, were used to evaluate their technique.

Moeskops et al. [32] used three similar processes CNNs, for each with a different size 2D input patch, to segment and classify MRI brain images. The scans were separated into many matter sorts, containing cerebrospinal liquid, grey matter, and white matter, and comprised 35 individuals and 22 pre-term children. The use of three distinct sizes of input patches concentrates each patch on gathering different visual characteristics; the bigger sizes gave spatial information, however the smaller patch dimensions centered on resident textures. The method's Dice coefficients are usually between 0.82 and 0.87, and it is extremely accurate.

## 2.5 Registration

The process of translating multiple sets of data into the same coordinate system is known as image registration. These images can have rigid (translations and rotations), affine (shears, for example), homographs, or complicated deformable models as spatial relationships. The steps of registration are listed below.

Given two input images, the 4 key steps in the image registration job's canonical approach are:

1) Goal Selection: shows the determined input image that must be superimposed correctly on the second counterpart input image.
2) Feature Extraction: decides how many features based on every input image can be extracted.
3) Feature Matching: lets for the detection of similarities between previously collected attributes.
4) Pose Optimization: shortens the distance between two images. The registering process generates the appropriate geometric alteration (e.g., scaling, rotation, translation, etcetera.) that aligns both of the input images in the same system of coordinates with the smallest space among them, resulting in the best amount of superimposition/overlapping.

In general, there exist several studies that apply image registration with different deep learning architectures in the literature.

# Chapter 3

# METHODOLOGY FOR BRAIN TUMOR CLASSIFICATION

This chapter reviews several Convolutional Neural Networks based Deep Learning architectures in which three of them are implemented for brain tumor classification in this study. Different deep learning architectures are explained in the next section. The methodology used in this thesis for the implementation of brain tumor classification employs AlexNet, VGG16 and ResNet architectures. The evaluation metrics for the presentation of the experimental results are then discussed in the following section.

## 3.1 Convolutional Neural Networks

The CNN is the most well-known and extensively utilized approach in the field of Deep Learning. CNN's key benefit over its predecessors is that it accurately characterizes relevant characteristics with almost no human intervention. Face recognition, computer vision, audio processing, and other applications have all benefited from convolutional neural networks. The development of Convolutional Neural Networks was invigorated by neurons in human and creature minds, like a CNN.

A typical type of Convolutional Neural Networks has many convolution pooling layers, similar to a multi-layer perceptron, except the end layers are Fully Connected layers. Figure 3.1 depicts how to construct image classification using Convolutional Neural Networks [26].

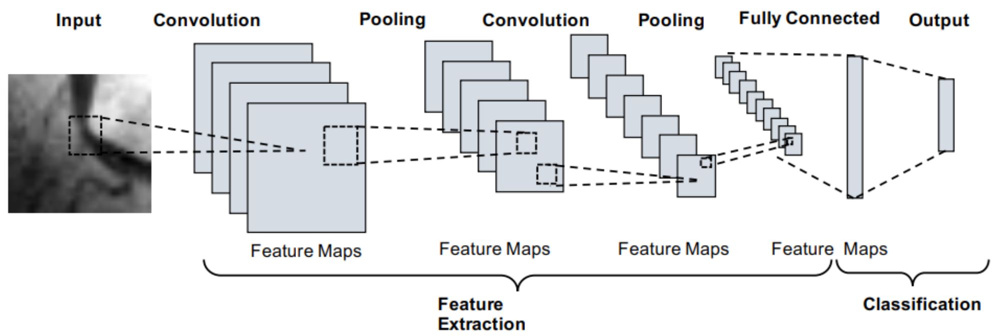

Figure 3.1: Convolutional Neural Network Design [26]

Throughout the most recent 10 years, various Convolutional Neural Network designs have been introduced. Model engineering is a basic issue in working on the exhibition of various applications. Numerous adjustments have been acknowledged in architecture of CNN from 1989 until now. It should be noted that the critical improvement in CNN implementation happened primarily as a result of reorganization of the handling unit and the construction of new blocks. The use of network depth was particularly important in making substantial breakthroughs in the architectures of Convolutional Neural Networks. In this sector, we glance the generally famous architecture of CNN.

CNN architectures are the supreme extensively used framework of deep learning. CNNs are used in a varied sort of applications, containing natural language processing and computer vision. We will go over each type of well-known CNN architectures in considerable detail. In the next subsections, several CNN architectures, namely AlexNet, Network-in-Network, ZefNet, Visual Geometry Group (VGG), GoogLeNet, Highway Network and ResNet, are reviewed. In this thesis, three of the aforementioned architectures, namely AlexNet, VGG16 and ResNet are implemented for brain tumor classification.

### 3.1.1 AlexNet

AlexNet was first proposed by Krizhevesky et al. [33], who improved Convolutional Neural Network skill learning by raising its depth and employing multiple augmentation approaches. The historical backdrop of deep Convolutional Neural Networks started with the presence of LeNet [34] as shown in Figure. 3.2. Around then, the Convolutional Neural Networks were restricted to written by hand digit acknowledgment assignments, it is not scalable to all picture classes. AlexNet is a well-known name in the design of deep Convolutional Neural Networks, having achieved groundbreaking achievements in image identification and classification. Figure 3.3 shows the fundamental plan of the AlexNet design [35].

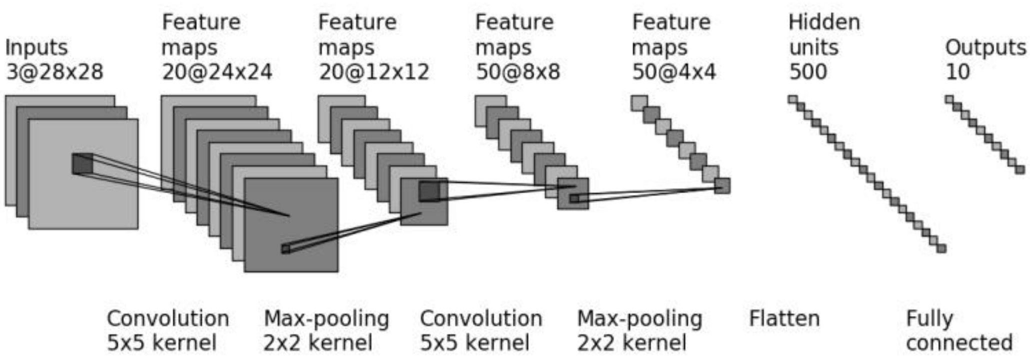


Figure 3.2: The LeNet Architecture [34]

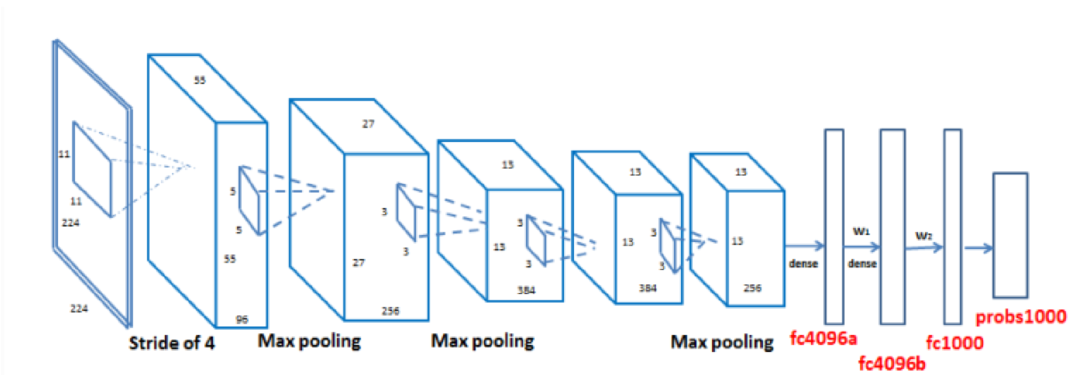


Figure 3.3: The Architecture of AlexNet [35]

Because of hardware limitations, the deep Convolutional Neural Network's learning ability is currently limited. To overcome these restrictions, AlexNet was trained in parallel on two GPUs (NVIDIA GTX 580). The quantity of feature extraction layers was also improved starting 5 in LeNet toward 7 in AlexNet on the way to improve the Convolutional Neural Network's applicability to different sorts of images. Despite the detail that complexity improves generalization for numerous resolutions of image, it was as a matter of fact overfitting that addressed the primary weakness connected with the depth. To discuss this issue, Krizhevesky et al. [33] used Hinton's impression. To ensure that the traits created by the algorithm were exceptionally strong, algorithm of Krizhevesky et al. [33] arbitrarily ignores numerous groundbreaking elements through the training phase. Furthermore, to increase the performance of model ReLU which is a non-soaking activation function is used. It reduces the problem of vanishing gradient. Indigenous response overlapping subsampling and normalization were also achieved to promote generalization by reducing over fitting to improve prior network performance, different modifications were finished by utilizing huge size channels (5×5and11×11) (5×5and11×11) in the past layers. AlexNet has broad importance in the momentum Convolutional Neural Network ages, as well as start a best in class research period in applications of Convolutional Neural Network.

### 3.1.2 Network-in-network

The wonderful and easy idea of employing 1x1 convolutions to add more combinational power to the features of a convolutional layer came from Network-in-Network (NiN). After each convolution, the NiN design used spatial MLP layers to better blend features before moving on to the next layer. One could believe that the 1x1 convolutions go against LeNet's initial concepts, however they actually help to aggregate convolutional features in a more efficient method, which is impossible to achieve by simply stacking more convolutional layers. This is not the same as using raw pixels as the following layer's input. After convolution, 11 convolutions are used to spatially integrate features across features maps, resulting in extremely few parameters that are shared across all pixels of these features.

### 3.1.3 ZefNet

Prior to 2013, the CNN learning system was mostly built on an experimentation basis, which made it difficult to understand the specific determination made after the development. This subject limited the deep CNN follow up on convoluted pictures. Accordingly, DeconvNet (a multilayer de-convolutional neural structure) was introduced by Zeiler and Fergus in 2013 [36].

### 3.1.4 Visual Geometry Group (VGG)

Convolutional Neural Networks were resolute to be active in the ground of image detection when they were resolute to be active in the ground of image detection, a specific and efficient strategy rule for Convolutional Neural Network was offered via Zisserman and Simonyan [37]. This state-of-the-art strategy was named VGG. It was a multilayer prototype with nineteen additional layers than ZefNet and AlexNet to mimic network representational capacity cousins in complexity.

ZefNet, on the other hand, was the outskirts organization in the 2013-ILSVRC competition, ensuring that channels of small sizes could enable the execution of Convolutional Neural Networks. Concerning these outcomes, VGG embedded a layer of the pile of 3×33×3 channels instead of the 5×55×5 and 11 × 11 channels in ZefNet. This suggested that the equal effort of these tiny dimensions' channels may have a comparable effect as the massive size filters. As such, these little size channels made the responsive field comparatively productive to the huge size channels (7×7and5×5)

(7×7and5×5). By diminishing the quantity of variables, an additional and advantage of lessening computational intricacy was accomplished by utilizing little size filters.

These findings established a new review pattern for using small size filters in Convolutional Neural Networks. VGG also solves network difficulties by introducing 1×11×1 convolutions in the convolutional layers' foci. It learns a direct collection of the element mappings that go with it. A “maximum pooling layer” is added after the convolution layer to tune the network followed by cushioning to keep the spatial objective. Overall, VGG got remarkable consequences for picture classification and limiting concerns. While it did not accomplish ahead of all comers in the 2014-ILSVRC contest, it gained a standing because of its expanded profundity, same topology, and simplicity. In any case, due to its utilization of over 140 million borders, VGG's computing cost was expensive, revealing its most critical fault. The network's design is depicted in Figure 3.4 [38].

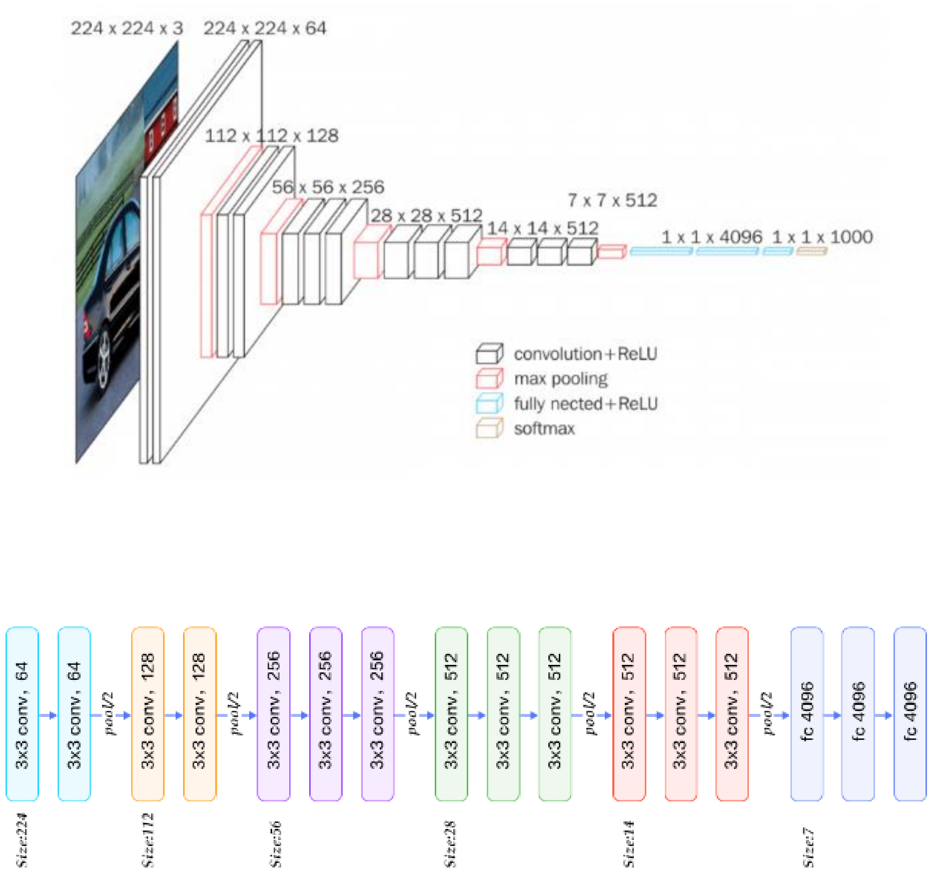


Figure 3.4: The Architecture of VGG16 [38]

VGG (Group of Visual Geometry) is a standard CNN with multiple layers. The word "deep" mentions to the quantity of layers in VGG-19 or VGG-16, which have 19 or 16 convolutional layers, correspondingly. The design of VGG is used to create cutting-edge object recognition models. Outside of ImageNet, the VGGNet outperforms baselines on several datasets and tasks as a DNN. Furthermore, it is arguably the most well-known image recognition design today.

Zisserman and Simonyan [37] suggested the model of VGG, or VGGNet that upholds sixteen layers is likewise denoted to as VGG16, which is CNN model. The VGG16 obtains roughly 92.7% accuracy in ImageNet. ImageNet is a collection of above fourtheen million photographs organized into over 1000 categories. Furthermore, one of the most famous prototypes sent to ILSVRC-2014 was VGG16. It substitutes the filters of large kernel-sized by numerous 3×3 filters of kernel-sized consistently, subsequently making huge advancements over AlexNet. The VGG16 model was prepared involving Nvidia Titan Black GPUs for a long time.

As referenced over, the VGGNet-16 backings sixteen layers and can order pictures into 1000 article gatherings, counting console, creatures, pencil, mouse, and so on Also, the model has a picture input size of 224-by-224.

On the other hand, the idea of the VGG19 model (likewise VGGNet-19) is equivalent to the VGG16 With the exception of that it upholds 19 layers. The numbers "16" and "19" indicate how several layers of weight in the model (convolutional layers). This implies that VGG19 has three a bigger number of convolutional layers than VGG16.

### 3.1.5 GoogLeNet

GoogleNet (sometimes referred to as Inception-V1) triumphed in the 2014-ILSVRC rivalry. The primary goal of the GoogleNet architecture is to achieve irrefutable level exactness while reducing computing cost. It suggested another commencement block (module) thought in the Convolutional Neural Network setting. Since it consolidates different scale convolutional changes by adding union capacity to it. For feature extraction, change, and split capacities are also used. Figure 3.5 depicts the GoogleNet architecture [39].

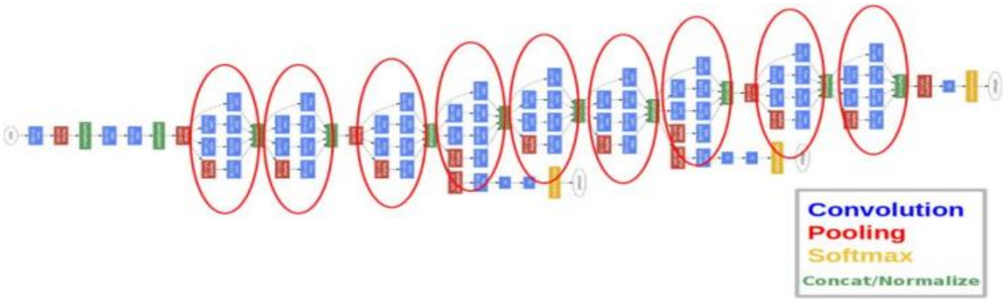

Figure 3.5: GoogleNet Architecture [39]

### 3.1.6 Highway Network

Developing the organization profundity works on its demonstration, generally for complex undertakings. Conversely, the organization preparing turns out to be hard. The inclusion of numerous layers in increasingly complicated networks can consequence in a small gradient values for down layers in back-propagation steps. In [39], Srivastava proposed another Convolutional Neural Network design, the Highway Network, in 2015 to defeat this problem. The principle of cross-availability underpins this strategy. Highway Network enables an uninterrupted data stream by training two gating units within the layer. Figure 3.6 describes the Highway network architecture [40].

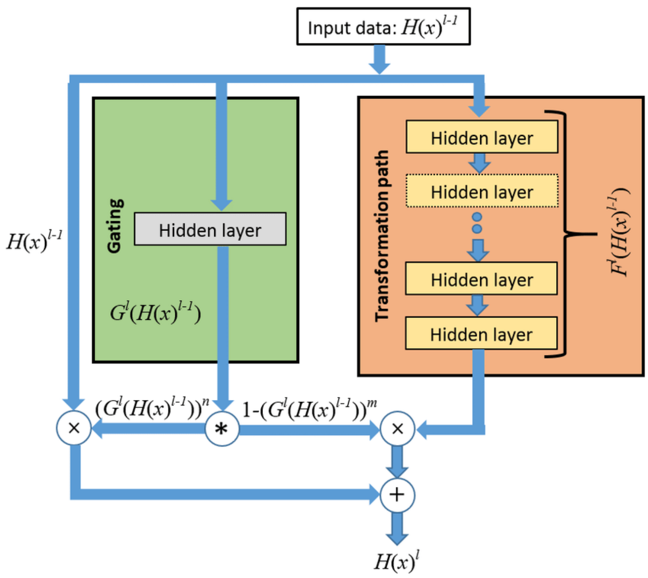

Figure 3.6: Highway Network Architecture [40]

### 3.1.7 ResNet

He et al. [40] created Network of Residuals, which took initial place in the ILSVRC 2015 rivalry. Their goal, in comparison to previous networks, is developing an extraordinarily network that there is no need to gradiant. A few different varieties of ResNet were established because to the vast sum of layers (beginning with 34 layers and increasing to 1202 layers). ResNet50 was the majority of well-known version, with 49 layers plus a solitary FC layer. Despite the fact that there were 3.9 million MACs, there were 25.5 million network weights. ResNet's fundamental notion is to apply the stay away from pathway principle. To solve the difficulty of training a high level network, a Highway Nets in 2015 [41] was proposed as shown in Figure. 3.7. This is a network with both traditional feed forward and a relict link. The (l-1) th(l-1) th productions, that are carried as of the previous layer (xl-1) (xl-1), can be recognized as the remaining layer productions. The residual network contains a large number of essential surviving squares. The activities in the remaining block are also adjusted depending on the type of residual network architecture.

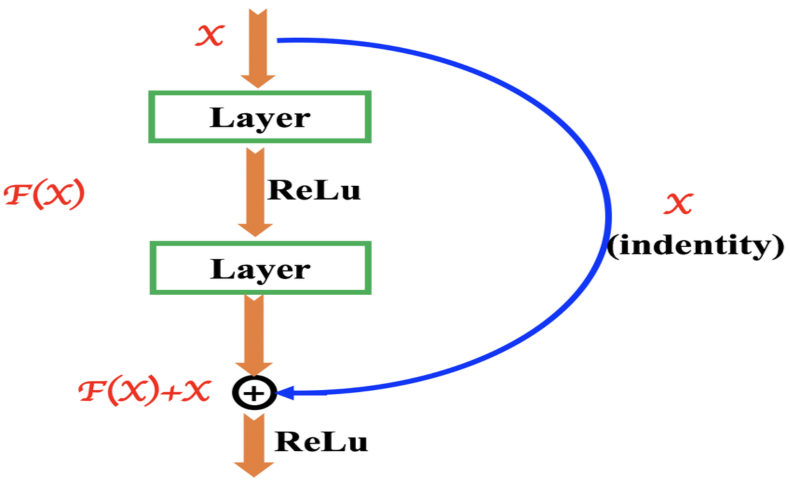

Figure 3.7: ResNet Block Diagram [41]

In contrast to the highway network, ResNet featured alternate way connections inside layers to enable non-parametric, data-independent multi-layer networks. When a gated alternate route in the highway network is closed, the layers clearly indicate non-residual approaches. Surprisingly, the easy avenues to freedom are rarely closed, but residual data is often accepted in ResNet. Also, because the related links speed up convergence, ResNet can prevent gradient diminishing difficulties. ResNet won the 2015-ILSVRC event Using 152 depth layers, which is eight times and 20 times more than VGG and AlexNet. Although ResNet with a large depth has a lower computation cost than VGG, ResNet with a large depth has a lower computation cost.

## 3.2 Evaluation Metrics Used

This section discusses metrics for evaluation, which are accustomed to assess the value of a model statistically. Handcrafted methods with Machine Learning classifiers and Deep Learning approaches must be evaluated in any study. A number of evaluation measures can be employed to show the value of a model.

The evaluation metrics used in DL tasks are critical in determining the best classifier. They are employed in the testing and training stages of a typical data classification process. During the phase of training, it is employed to improve the algorithm of classification. This indicates that the assessment measure is utilized to distinguish between options and select the best one, such as a discriminator, which can yield a more precise estimate of future evaluations when used in conjunction with an exact classifier. In the meantime, the assessment metric is accustomed to analyze the developed classifier's effectiveness, such as a hidden data evaluator during the model test phase. The number of effectively classified negative and positive instances is denoted by the letters TN and TP, correspondingly. Furthermore, the amounts of misclassified positive and negative cases are defined as FN and FP, respectively. The following are around of the supreme famous evaluation metrics.

**1) Accuracy:** Computes the percentage of correctly forecast classes in relation to the overall number of samples that were tested. Accuracy can be calculated as follows:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{4.1}$$

wherever TP or True Positive is the number of Digital Radiography (DR) pictures that are perfectly recognized. The total of perfectly identified non-DR pictures is recognized as True Negatives (TN). False Positive (FP) denotes the quantity of DR Images that are wrongly recognized in place of positive but are really non-DR. False Negative (FN) is the sum of falsely detected non-DR that are truly DR.

Accuracy values are in the range [0,100] percent. If we divide that range evenly, 100-87.5% equals very good, 87.5-75% equals good, 75-62.5% equals satisfactory, and 62.5-50% equals bad. In reality, we regard numbers between 100 and 95% to be excellent, 95 to 85% to be good, 85 to 70% to be satisfactory, and 70 to 50% to be "needs to be improved" for brain tumor recognition.

**2) Recall or Sensitivity:** The percentage of successfully classified positive patterns is calculated using sensitivity or recall as shown below:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} \tag{4.2}$$

where TP or True Positive is the total of Digital radiography (DR) pictures that are perfectly recognized. False Negative (FN) is the total of falsely detected non-DR that are truly DR. The recall is calculated as TP/FN, in which TP represents true positives and FN represents false negatives. The recall of a classifier refers to its ability to locate all samples that are positive. 1 is the best value while 0 is the worst.

3) **Specificity:** is used to calculate the percentage of incorrectly classified Negative Patterns. The formula of specificity is as follows:

$$\text{Specificity} = \frac{\text{TN}}{\text{FP+TN}} \quad (4.3)$$

where the quantity of perfectly identified non-DR pictures is equal to True Negatives (TN). False Positive (FP) denotes to the quantity of DR images that are wrongly known as positive but are truly non-DR.

In an ideal world, the model would have a high specificity or true negative rate. A greater specificity score would imply a higher real negative rate and a decreased rate of false-positives. A reduced specificity score indicates a lower genuineness score.

**4) Precision:** is used to figure out which positive patterns in a positive class are the most common. Precision is calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (4.4)$$

where TP, or True Positive, is the total of fully recognized Digital radiography (DR) images. The sum of DR images that are wrongly recognized as positive but are essentially non-DR is referred to as False Positive (FP). Precision can be used as a measure of quality. When an algorithm's precision is higher, it produces more relevant outcomes rather than irrelevant ones.

**5) F1-Score:** also known as F-score and F-measure, is a model's accuracy on a dataset. It is used to assess binary classification systems that categorize examples as positive or negative.

F1-Score is calculated as follows

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.5}$$

where recall (also called sensitivity) is the percent of related examples identified, and precision for positive predictive value is the proportion of applicable examples found among the improved instances. The greatest rate of an F-score is 1.0, which implies faultless accuracy and recall, while the minimum value is 0 if neither precision nor recall are 0.

## 3.3 Comparison of the Evaluation Metrics

The most fundamental metric is accuracy, which is distinct as the quantity of accurate forecasts separated by the whole quantity of estimates, multiplied by 100. In many cases, classification accuracy is ineffective in predicting model performance. When the class distribution is unbalanced, this is one of the circumstances. In this circumstance, even if the most common class is forecasted in all samples, it will have a high level of accuracy, which makes no logic, because the model does not learn anything and simply predicts everything as the best class. As a result, it must also consider performance measures particular to each class. Precision is one of these criteria. Recall is another significant parameter; this is the percent of data from a category that the model predicts correctly.

Depending on the situation, it might need to favor recall or precision. Here are numerous instances wherein both precision and recall are essential. As a result, it only makes sense to think about how to combine the two into a single statistic. The F1-score is a famous metric for combining precision and recall. If we wish to increase precision too much, we will observe a decrease in recall rate, and vice versa.

Two more popular measures in the medical and biological sectors are sensitivity and specificity. A laboratory test's sensitivity indicates how frequently it is positive in patients with a specific condition. A laboratory test's specificity indicates how frequently the test is negative in patients who do not have the condition in question. A test with 100% sensitivity accurately identifies everyone who has the condition, while a test with 100% specificity correctly identifies everyone who does not.

# Chapter 4

# RESULTS AND DISCUSSION

This chapter describes how the algorithms used in this thesis were implemented. It covers everything from how to prepare datasets for system training to how each algorithm generates evaluation measures.

## 4.1 Setup for Experiments

This part goes over the setup used to obtain the results. Python 3.7 is used to implement the algorithms. The pathological brain images were taken from the Kaggle [42] dataset where its link is https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri, and there are two folders called test and train to analyze the performance of each prediction model. In the train folder, there are 1200 images labeled as “yes” and 1200 images labeled as “no”, for a total of 2400 images in the train folder. In the test folder, there are 300 images labeled as “yes” and 300 images named as “no” and there is a total of 600 images in the test folder. Additionally, there is another dataset named Figshare [43], with two files titled “test” and “train”. In the train folder, there are 4117 images considered as “yes” and 1588 images termed as “no” and there is a total of 5705 images in the train folder. In the test folder, there are 906 images considered as “yes” and 405 images called “no” and there is an overall of 1311 images in the test folder. Table 4.1 shows the amount of images used in the training and testing in Kaggle dataset and Table 4.2 presents the number of images used to train and test in Figshare dataset. However, because of the quality of some images in Figshare dataset which is low, some of the images are not used in the experiments. Therefore, in this study, 5600 train images and 1400 test images are used in the experiments.

The first dataset used is called Kaggle Brain MRI dataset. Kaggle is an available data science and machine learning community. Kaggle lets operators to find and post databases, study and form models in a web-based data knowledge situation, connect with new data experts and specialists of machine learning, and participate in data knowledge competitions. Kaggle began with competitions of machine learning in 2010 and has nowadays grown to encompass a platform for open data, a data science desk in the cloud, and AI courses. Anthony Goldbloom and Jeremy Howard were significant members of the team. The company was valued at $25 million when equity was raised in 2011. Google announced the acquisition of Kaggle on March 8, 2017. Figure 4.1 presents the samples of healthy brain MRI images of Kaggle and Figure 4.2 shows the samples of unhealthy brain MRI images of Kaggle Database.

Table 4.1: Numbers of Train and Test Images Used in Kaggle Dataset

| Database name | Kaggle https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri |
|---|---|
| Image dimension | 128*128 |
| Number of train images | 2400 |
| Number of test images | 600 |

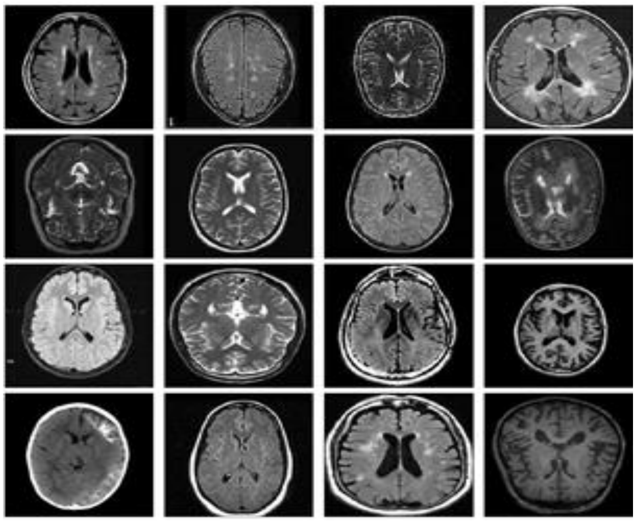

Figure 4.1: Samples of Healthy Brain MRI Images of Kaggle Database

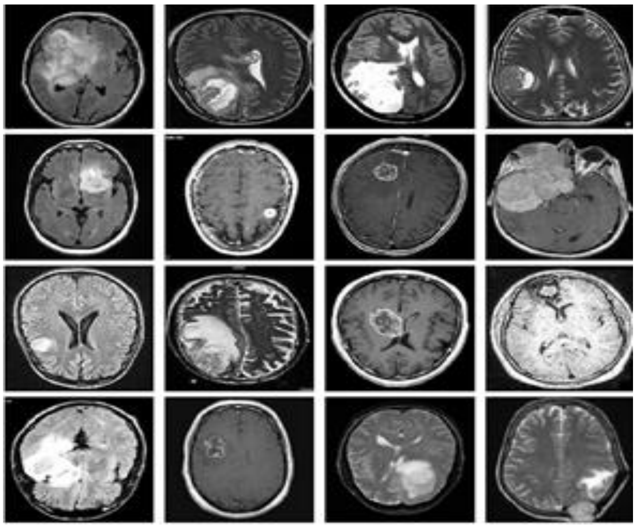

Figure 4.2: Samples of Unhealthy Brain MRI Images of Kaggle Database

The second dataset used is called Figshare Brain MRI dataset. Figshare has been aiding researchers in making their data publically accessible for more than 10 years, during which time they have watched data sharing development across disciplines as funders and publications demand it and investigators desire credit for all of their research's findings. This dataset has some benefits, including the option to publish data and materials related to a specific publication or research project, as well as the flexibility to structure and descriptions, and the ability to create a collection with a single DOI that points to all of the items. Figure 4.3 shows the samples of healthy brain MRI images of Figshare database and Figure 4.4 demonstrates the samples of unhealthy brain MRI images of Figshare database. To evaluate the performance of three deep learning models for diagnosing brain tumors, 6 metrics are used in this thesis to show the effectiveness of the three deep learning models. These metrics are accuracy, precision, recall, sensitivity, specificity, F-Measure. Moreover, two databases are used in the experiment and also the number of Epochs is 100 and all the images used are the same size. The size of all images used is 128 x 128.

## 4.2 Deep Learning Models Used

AlexNet, VGG16 and ResNet deep learning models are used in the experiments. VGG16 is a 16-layer architecture that includes two convolution layers, a pooling layer and a fully connected layer at the end. The VGG network is based on the concept of considerably deeper networks with smaller filters. The number of layers in VGGNet has grown from eight in AlexNet. ResNet, a new architecture presented by Microsoft Research in 2015, established a new architecture called Residual Network. This architecture introduces the concept of the Residual Network to overcome the problem of the vanishing gradient.

Table 4.2: Numbers of Train and Test images Used in Figshare Dataset

| | |
|---|---|
| Database name | FigShare<br>https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/5<br>Cheng, Jun (2017): brain tumor dataset. figshare. Dataset.<br>https://doi.org/10.6084/m9.figshare.1512427.v5 |
| Image dimension | 128*128 |
| Number of train images | 5600 |
| Number of test images | 1400 |

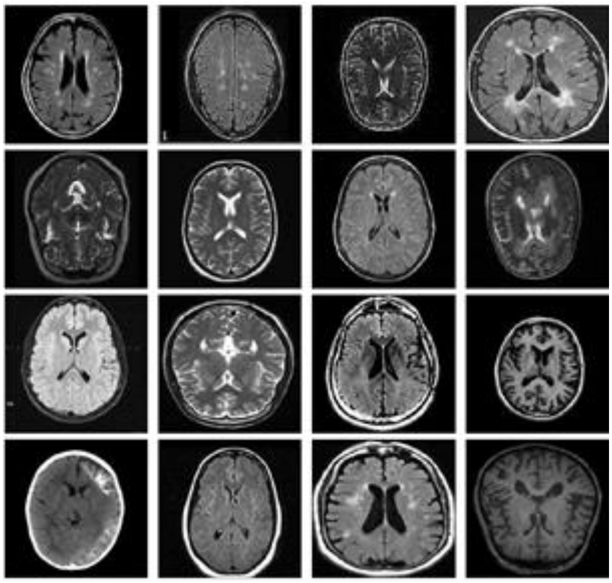
Figure 4.3: Samples of Healthy Brain MRI Images of Figshare Database

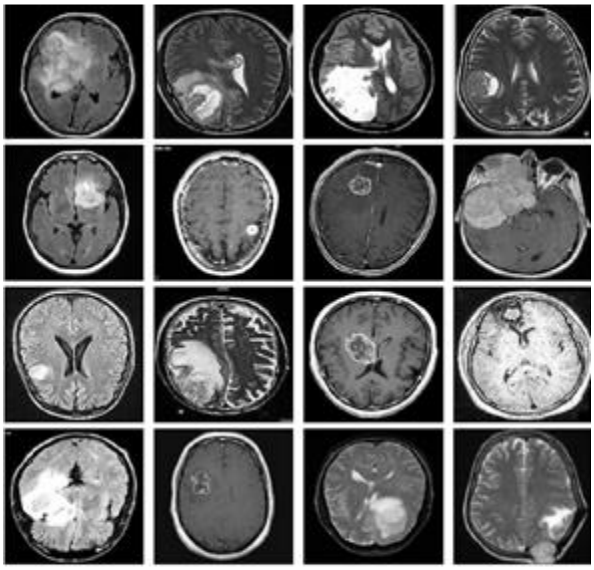
Figure 4.4: Samples of Unhealthy Brain MRI Images of Figshare Database

## 4.3 Results and Metrics

All the results and metrics as well as the details used are shown in Table 4.3 and Table 4.4.

Table 4.3: Results and Metrics for Kaggle Dataset

| Method Used | | | |
|---|---|---|---|
| Measure | AlexNet | VGG16 | ResNet |
| ACCURACY | 0.9883 | 0.9733 | 0.6667 |
| PRECISION | 0.9895 | 0.9767 | 0.6185 |
| RECALL(SENSITIVITY) | 0.9861 | 0.9703 | 0.9258 |
| SPECIFICITY | 0.9904 | 0.9764 | 0.3897 |
| F-MEASURE | 0.9878 | 0.9735 | 0.7416 |

Table 4.3 shows the metrics for AlexNet and VGG16 and ResNet.the results show in all metrics AlexNet has a good performance in comparision with VGG16 and ResNet ResNet has the worst performance compared to VGG16 and AlexNet.

Table 4.4: Results and Metrics for Figshare Dataset

| Method Used | | | |
|---|---|---|---|
| Measure | AlexNet | VGG16 | ResNet |
| ACCURACY | 0.9943 | 0.9915 | 0.9658 |
| PRECISION | 0.9960 | 0.9928 | 0.9818 |
| RECALL(SENSITIVITY) | 0.9960 | 0.9948 | 0.9700 |
| SPECIFICITY | 0.9900 | 0.9840 | 0.9553 |
| F-MEASURE | 0.9960 | 0.9938 | 0.9759 |

When Figshare dataset is used, the results for three different models (AlexNet, VGG16 and ResNet) are better than when Kaggle dataset is used. The reason for this is that the number of images in Figshare dataset are more than the amount of images in Kaggle dataset and also the quality of images in Figshare dataset is better than the quality of images in Kaggle dataset and finally these models are more compatible with Figshare Dataset compared to Kaggle dataset.

## 4.4 Comparison with the State-of-the-Art

Expert radiologists perform the crucial task of brain tumor segmentation and classification. As decision-making aids, radiologists can use machine learning and deep learning approaches. This publication outlines a number of cutting-edge methodologies for classifying brain tumors automatically. Brain tumor classification results are compared on Kaggle Brain MRI and Figshare Brain MRI datasets in Table 4.5.

Table 4.5: Comparison with the Stae-of-the-Art on Kaggle and Figshare Datasets

| Ref | Publication year | Preprocessing and Segmentation | Features | Classification | Dataset | Accuracy |
|---|---|---|---|---|---|---|
| [44] | 2022 | segmentation, image enhancement | Pixel-based feature extraction | CNN | Kaggle | 97.79% |
| [45] | 2020 | deep fusion | PCA | fused deep features and SVM | Kaggle | 97.89% |
| [46] | 2020 | - | VGG16 | | Figshare | 98.69% |
| [47] | 2021 | - | DenseNet121, ResNet50 | | Figshare | 98.91%, 99.02% |
| This Study | 2022 | - | AlexNet | | Kaggle | 98.83% |
| | | - | VGG16 | | Kaggle | 97.33% |
| | | - | ResNet | | Kaggle | 66.67% |
| | | - | AlexNet | | Figshare | 99.43% |
| | | - | VGG16 | | Figshare | 99.15% |
| | | - | ResNet | | Figshare | 96.58% |

In recent years, there are several state-of-the-art studies for the classification of brain MRI images using Kaggle Brain MRI dataset and Figshare Brain MRI dataset. Comparison with the state-of-the-art methods in Table 4.5 indicates that the results on Kaggle Brain MRI dataset show that most of the deep learning architectures, such as

AlexNet and VGG16, achieve better results compared to handcrafted methods. Similarly, the results on Figshare Brain MRI dataset show that AlexNet achieves the best accuracies for the classification of brain tumors. The best accuracies obtained using AlexNet on Kaggle Brain MRI dataset and Figshare Brain MRI dataset are 98.83% and 99.43%, respectively.

# Chapter 5

# CONCLUSION

MRI-based medicinal image study for brain tumor investigations has recently attracted a lot of researchers due to a growing demand used for effective and impartial valuation of huge sums of medicinal data. Since brain tumors have such a high fatality rate, it is serious to discover them first and treat them. Because of the intricacy of brain tissue, physical identification of brain and tumor components takes time and is operator-dependent.

Deep learning algorithms that are used in context to promote health diagnosis have shown to be effective. Correct brain tumor diagnosis is defined by the World Health Organization (WHO) as "the identification, diagnosis, and tumor classification based on malignancy, grade, and kind criteria." Detecting the tumor, identifying the tumor regarding to position and type, and finding the tumor site are totally portion of this investigational endeavour in the detection of brain diseases applying Magnetic Resonance Imaging (MRI). This methodology has been investigated by means of using one model for categorizing brain MRI on several classification tasks instead of using a different version according to each classification test. Multi-task categorization using Convolutional Neural Networks (CNN) is able to classify and identify tumors. So, by using deep learning with brain MRI scans, we can manage the diagnosis of the tumor automatically when it is a vital issue in tumor detection.

In this thesis, three Deep Learning models are used, namely Alexnet, VGG16 and ResNet, to classify brain tumors employing MRI images. The performances of these models have been investigated using two datasets, namely Kaggle and Figshare, and also five metrics are used to calculate their performances. AlexNet achieved 98% accuracy on the Kaggle dataset, VGG16 had 97% accuracy, and ResNet got 66% accuracy. AlexNet has offered the highest level of accuracy among these networks.

On the other hand, AlexNet got 99% accuracy in the Figshare dataset, VGG16 got 99% accuracy, and ResNet has 96% accuracy. AlexNet and VGG16 outperformed ResNet in terms of accuracy. These accuracies allow for the early detection of abnormalities before they create physical harm such as disability or other complications.

The experimental results reveal that deep learning models perform well on Figshare Brain MRI dataset and Kaggle Brain MRI dataset, however better accuracy is obtained on the Figshare dataset. The reason for this is that we have more images in the train and test sections when Figshare dataset is used. Therefore, the efficiency is increased and better results are obtained on Figshare dataset.

As a future work, since identifying the exact location of a brain tumor is very important and the location of the tumor determines the need of surgery to remove malignant tumors, other segmentation methods can be investigated. Additionally, more powerful and efficient deep learning architectures, such as ResNet50, can be used to increase the accuracy of brain tumor classification.

# REFERENCES

[1] A. Rehman, M.A. Khan, T. Saba, Z. Mehmood, U. Tariq, and N. Ayesha, ‘‘Microscopic brain tumor detection and classification using 3D cnn and feature selection architecture,’’ *Microsc. Res. Techn.*, vol. 84, no. 1, pp. 133–149, 2021.

[2] J. Amin, M. Sharif, M. Yasmin, and S.L. Fernandes, ‘‘Big data analysisfor brain tumor detection: deep convolutional neural networks,’’ *FutureGener. Comput. Syst.*, vol. 87, pp. 290–297, Oct. 2018.

[3] S.T. Kebir and S. Mekaoui, ‘‘An efficient methodology of brain abnormalities detection using cnn deep learning network,’’ *in Proc. Int. Conf.Appl. Smart Syst*. (ICASS), Nov. 2018, pp. 1–5.

[4] R. Vinoth and C. Venkatesh, ‘‘Segmentation and detection of tumor in mri images using cnn and svm classification,’’ *in Proc. Conf. Emerg. Devices Smart Syst.* (ICEDSS), Mar. 2018, pp. 21–25.

[5] M. Sajjad, S. Khan, M. Khan, W. Wu, A. Ullah, S.W. Baik, “Multi-grade brain tumor classification using deep cnn with extensive data augmentation,” *J. Comput. Sci*., 30 (2019), 174–182.

[6] S. Khawaldeh, U. Pervaiz, A. Rafiq, R. Alkhawaldeh, “Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks”, “*J. Appl. Sci*., 8 (2018), 27.

[7] M. Talo, U.B. Baloglu, O. Yıldırım, U.R. Acharya, “Application of deep transfer learning for automated brain abnormality classification using mri images,” *Cognitive Systems Research*, 54(2019), 176–188.

[8] A. Çınar, M. Yıldırım, “Detection of tumors on brain mri images using the hybrid convolutional neural network architecture,” *Med. Hypotheses*, 139 (2020), 109684.

[9] A.M. Hasan, H.A. Jalab, F. Meziane, H. Kahtan, A.S. Ahmad, “ Combining deep and handcrafted image features for mri brain scan classification,” *IEEE Access*, pp. 79959–79967,2019.

[10] H.T. Zaw, N. Maneerat, K.Y. Win, “Brain tumor detection based on naïve bayes classification,” *International Conference on Engineering, Applied Sciences and Technology*, pp.1-4,2019.

[11] E. Sert, F. Ozyurt, A. Doğantekin, “A new approach for brain tumor diagnosis system: single image superresolution based maximum fuzzy entropy segmentation and convolutional neural network,” *Medical Hypothesis*, pp.1-9,2019.

[12] S. Deepak, P.M. Ameer, “Brain tumor classification using deep cnn features via transfer learning,” *Computers in Biology and Medicine*, pp.1-7,2019.

[13] T.L. Narayana, TS. Reddy, “An efficient optimization technique to detect brain tumor from mri images,” *International Conference on Smart Systems and Inventive Technology*, pp.1-4, 2018.

[14] F.P. Polly, S.K Shil, M.A. Hossain, A. Ayman, Y.M. Jang, “Detection and classification of hgg and lgg brain tumor using machine learning,”*International conference on Information Networking*, pp.813-817,2018.

[15] J. Amin, M. Sharif, M. Yasmin, S. Fernandis, “A distinctive approach in brain tumor detection and classification using mri,” *Pattern Recognition Letters*, pp.1-10, 2017.

[16] A. Minz , C. Mahobiya , “Mri image classification using adaboost for brain tumor type,” IEEE 7th International Advance Computing Conference, pp.1-5,2017.

[17] A.S. Shankar, A. Asokan, D. Sivakumar,“Brain tumor classification using gustafson–kessel (G-k) fuzzy clustering algorithm,” *International Journal of Latest Engineering Research and Applications*, pp.68-72,2016.

[18] L. Alzubaidi, J. Zhang, A.J. Humaidi et al. “Review of deep learning: concepts, cnn architectures, challenges, applications, future directions.” *J Big Data 8*, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8.

[19] A. Gumaei, M.M. Hassan, M.R. Hassan, A.N. Alelaiwi, G. Fortino, “A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification,” *IEEE Access*, pp. 36266-36273, 2019.

[20] G. Hemanth, M. Janardhan, L. Sujihelen, “Design and implementing brain tumor detection using machine learning approach,” *Third International Conference on Trends in Electronics and Informatics*, pp.1-6,2019.

[21] G. Latif, D.N.F. Iskandar, J.M. Alghazo, N. Mohammad, “Enhanced mri image classification using hybrid statistical and wavelets features,” *IEEE Access*, pp.9634-9644, 2018.

[22] H. Mohsen, E. Sayed, E. Dahshan , A. Badeeh, M. Salem, “ Classification using deep learning neural networks for brain tumors future”. *Computing and Informatics Journal*, pp.68-73, 2018.

[23] A.R. Raju, P. Suresh, R.R. Rao, “Bayesian hcs-based multi-svnn: a classification approach for brain tumor segmentation and classification using bayesian fuzzy clustering,” *Biocybernetics and Biomedical Engineering*, pp.646-660, 2018.

[24] H.E.M. Abdalla, M.Y. Esmail, “Brain tumor detection by using artificial neural network,” *International Conference on Computer, Control, Electrical, and ectronics Engineering*, pp.1-6, 2018.

[25] M. Sajjad, S. Khan , K. Muhammad ,W. Wu , A. Ullah , S.W Baik, “Multi-grade brain tumor classification using deep cnn with extensive data augmentation,” *Journal Of Computational Science*, pp.174-182, 2018.

[26] O.M. Emmanuel, G.A.C. Juan, C.A. Ivan, R.P. Jose. “Transfer learning for stenosis detection in x-ray coronary angiography,” *Mathematics 2020*, 8, 1510; doi:10.3390/ math 8091510,2020.p.1-20.

[27] S. Bharati, P. Podder, M.R.H. Mondal, “Hybrid deep learning for detecting lung diseases from x-ray images.” *Inform Med Unlocked*. 2020; 20:100391.

[28] W. Zuo, F. Zhou, Z. Li, L. Wang. “Multi-resolution cnn and knowledge transfer for candidate classification in lung nodule detection.” *IEEE Access*. 2019; 7:32510–21.

[29] H.R. Roth, C.T. Lee, H.C. Shin, A. Seff, L. Kim, J. Yao, L. Lu, “Summers rm. anatomy-specific classification of medical images using deep convolutional nets.” *In: 2015 IEEE 12th international symposium on biomedical imaging* (ISBI). IEEE; 2015. p. 101–4.

[30] V. Chouhan, S.K. Singh, A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damaševičius, V.H.C. De Albuquerque. “A novel transfer learning based approach for pneumonia detection in chest x-ray images”. *Appl Sci*. 2020; 10(2):559. doi: 10.3390/app10020559.

[31] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille. “Deep lab: semantic image segmentation with deep convolutional nets, aatrous convolution, and fully connected crfs.” *IEEE Trans Pattern Anal Mach Intell*. 2017;40(4):834–48.

[32] P. Moeskops, M.A. viergever, AM. Mendrik, L.S.De. Vries, MJ. Benders, I. Išgum. "Automatic segmentation of mri brain images with a convolutional neural network." *IEEE Trans Med Imaging*. 2016;35(5):1252–61.

[33] A. Krizhevsky, I. Sutskever, G.E. Hinton. "Imagenet classification with deep convolutional neural networks." *Commun ACM*. 2017; 60(6):84–90.

[34] Z. Yiren, S. Sibo, C. Ngai-Man, "On classification of distorted images with deep convolutional neural networks," *Singapore University of Technology and Design*, 2017.p.1-5.

[35] W. Huafeng, P. Haixia, W. Huafeng, Z. Yanxiang , C. Yehe, "Deep learning for image retrieval: what works and what doesn't," *Conference Paper* · November 2015, p.1-22.

[36] M.D. Zeiler, R. Fergus, "Visualizing and understanding convolutional networks." *In: European conference on computer vision*. Springer; 2014. p. 818–33.

[37] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition"; 2014. *arXiv preprint arXiv:1409.1556*.

[38] https://www.kaggle.com/code/blurredmachine/vggnet-16-architecture complete-guide/notebook.

[39] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov. "Dropout: a simple way to prevent neural networks from overfitting." *J Mach Learn Res*. 2014; 15(1):1929–58.

[40] H. Kaiming, Z. Xiangyu, R. Shaoqing and S. Jian. "Deep residual learning for image recognition." *In: Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016. p. 770–8.

[41] W. Fei, G. Mali, "Single image super-resolution by residual recovery based on an independent deep convolutional network," *IEEE Access*, VOLUME 4, pp. 1–10, 2020.

[42] N. Chakrabarty, "Brain mri images dataset for brain tumor detection," *Kaggle, 2019. Available from*: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection.

[43] Figshare brain tumor dataset, https://doi.org/10.6084/m9.figshare.1512427.v5 , Accessed date: December 2018.

[44] S. Manav, S. Pramanshu, M. Ritik, G. Kamakshi ,"Brain tumour detection using machine learning ," *Journal of Electronics and Informatics,* December 2021, Volume 3, Issue 4, Pages 298-308,2022.

[45] K.S. Prabira, K.B. Santi, "A data constrained approach for brain tumour detection using fused deep features and svm," *Multimedia Tools and Applications* (2021) 80:28745–28760, 2021.

[46] R. Arshia, N. Saeeda, I.R. Muhammad, A. Faiza, I. Muhammad, “A deep learning-based framework for automatic brain tumors classification using transfer learning,” *Circuits, Systems, and Signal Processing* (2020) 39:757–775,2020.

[47] P. Özlem, G. Cafer, “Classification of brain tumors from mr images using deep transfer learning,” *The Journal of Supercomputing* ,volume 77, pages7236–7252 (2021).