EUROPEAN UNIVERSITY OF LEFKE
INSTITUTE OF GRADUATE STUDIES AND RESEARCH
DEPARTMENT OF ELECTRICAL AND ELECTRONICS
ENGINEERING

MASTER'S THESIS

BRAIN TUMOR CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

Bakary BADJIE

Assoc. Prof. Dr. Ezgi Deniz Ülker

LEFKE 2021

EUROPEAN UNIVERSITY OF LEFKE
INSTITUTE OF GRADUATE STUDIES AND RESEARCH
DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

MASTER'S of SCIENCE THESIS

BRAIN TUMOR CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

BY

Bakary BADJIE

SUPERVISOR: Assoc. Prof. Dr. Ezgi DENIZ ULKER

LEFKE 2021

EUROPEAN UNIVERSITY of LEFKE

INSTITUTE of GRADUATE STUDIES and RESEARCH

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

Bakary Badjie, a Master of Science student at the European University of Lefke, Institute of Graduate Studies and Research, defended his thesis entitled "**Brain Tumor Classification using Convolutional Neural Network"** on August 10, 2021 and has been found to be satisfactory by the jury members..

Jury Members

1. Prof. Dr. Sadık ÜLKER
   European University of Lefke
2. Assoc. Prof. Dr. İbrahim ERŞAN
   Girne American University
3. Assoc. Prof. Dr. Ezgi Deniz ÜLKER (Supervisor)
   European University of Lefke

Prof. Dr. Özgür Cemal ÖZERDEM
Director of the Institute of Graduate
Studies and Research

ABSTRACT

Deep Learning is the latest approach in the machine learning industry that has been used to carryout different operations; it has attracted a lot of interest from researchers worldwide. Deep learning has been adopted for addressing complex issues that necessitate a high level of human intelligence, particularly in the health industry. One of the most important applications in radiology is diagnosing brain tumors which are among the most severe and dangerous tumor disorders, with a short life expectancy if not treated early. Classifying tumors in radiographic images is a major concern in the health sector, but it is a complicated and time-consuming process that radiologists must undertake, while the accuracy of their analysis is solely dependent on their expertise. Today's radiology diagnoses, such as magnetic resonance (MR) examinations, are primarily based on personal judgment and may be insufficiently accurate and the risk that patients encounter may be very high. Therefore, leveraging Artificial Intelligence (AI) technology in order to reduce the inaccuracies in the diagnosing process is vital. This study, which focused on combining deep learning and radiomics, proposed an automated technique for classifying tumors in patients using MR images. The proposed technique includes the use of Convolutional Neural Network (CNN) with transfer learning as our deep learning model to perform a binary classification on our MR images. In other words, we took advantage of a pre-trained AlexNet model and transferred the knowledge to CNN architecture, and the model tested was evaluated for effectiveness and performance analysis using an image dataset that the model had never seen after its training and we obtained an accuracy of 99.62%, which is pretty much incredible.

Keywords: Brain tumor; MRI; Classification; Artificial Intelligence; Deep Learning; Convolutional Neural Network; Transfer Learning; Radiomics; AlexNet; Radiologist;

ACKNOWLEDGEMENTS

First and foremost, I express my gratitude and praise to Almighty Allah for allowing me to reach this point in my academic career. I owe a special debt of gratitude to my thesis supervisor, Assoc. Prof. Dr. Ezgi DENIZ ÜLKER, for her support and guidance throughout these challenging moments. Her sincerity, thoroughness, and perseverance have been a constant source of inspiration for me. My efforts have only come to light as a result of her conscientious efforts. I also do not want to miss this opportunity to acknowledge the contribution of all the staff of the department of electrical and electronics engineering for their endless support and excellent academic delivery from day one of my master's degree program up to date. Finally, my sincere gratitude goes to my family for their continuous prayers throughout the period of my study.

Contents

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	: 2-Dimensional
3D	: 3-Dimensional
AdaGrad	: Adaptive Gradient Algorithm
AI	: Artificial Intelligence
ADC	: Apparent Diffusion Coefficient
ANN	: Artificial Neural Network
CNN	: Convolutional Neural Network
CT	: Computerized Tomography
GPU	: Graphic Processing Unit
LDA	: Linear Discriminant Analysis
MR	: Magnetic Resonance
MRI	: Magnetic Resonance Imaging
NM	: Nuclear Medicine
PCA	: Principle Component Analysis
PET	: Positron Emission Tomography
PFS	: Progression-Free Survival
ReLU	: Rectified Linear Unit
RMSprop	: Root Mean Square prop
ResNet	: Residual Network
TTP	: Time-To-Peak
US	: Ultrasound
VGG	: Visual Geometry Group
NLP	: Natural Language Processing
EEG	: Electroencephalogram
WIR	: Wash-in-Radio

CHAPTER ONE

INTRODUCTION

1.1 Background

Artificial intelligence (AI) has revolutionized health care in a variety of applications, including disease diagnosis, surgical robots, distance surgical efficiency, etc. maximizing the efficiency of primarily health care services. AI in the health care market is simply expected to grow exponentially within the next five years; it's expected to reach around 45 to 50 billion U.S. dollars by 2026 from the current evaluation of 25 billion U.S dollars (Zhon et al., 2018). However, radiology is one of the common fields in the medical industry that is leveraging AI technology to carryout incredible performance in areas such as Magnetic resonance Imaging (MRI) scans, Computerized Tomography (CT) scans, Mammography scans, Nuclear Medicine (NM) scans, Positron Emission Tomography (PET), etc. (Afshar et al., 2019). These different types of medical imaging have played a very significant role in diagnosing diseases such as brain tumors, ovarian cancer, bone cancer, etc. However, brain tumor is one of the most disturbing tumor complications that the medical industry found very difficult with its diagnosis before the intervention of AI technology. A brain tumor is the abnormal growth of an extra cell in the brain. There are more than 145 types of brain tumors that have been documented and they are grouped into two; the first group is called the primary brain tumors which originate from the brain tissue itself, while the second group is called secondary or metastatic brain tumors which originate from other location of the body and later spread to the brain (Nagaraj et al., 2020). However, the rate of the development of tumor in the brain decides how it can impact the nervous system's functioning. The cause of the primary brain tumor is unknown, but specialists have identified two potential causes: ionizing radiation exposure and a family history of brain tumors, also known as genetic syndromes. Brain tumor is identified as a life threatening disease and has a worldwide record of approximately 350,000 cases each year. This disease is deathly and can restrict the normal growth or development of the brain, especially in children below the age of seven. According to the National Cancer Institute, 30,500 brain tumor patients are diagnosed in America out of which 17,125 are men and 13,375 are women in 2020, where 13,700 of these diagnoses result in death.

Recently, a subset of machine learning known as deep learning has been proven to be superior in diagnosing tumors in the human brain using a radiographic technology called magnetic resonance imaging that significantly improved the speed and the accuracy of the diagnosis. It has demonstrated an outstanding performance in the area of medical imaging since its introduction; it is a reliable and time-saving technique. However, before going further, let’s understand what artificial intelligence and machine learning is. In general, Artificial intelligence (AI) is a broad field of computer science dealing with creating autonomous systems that can execute functions that would otherwise involve high level human intelligence (Zhon et al., 2018). It aims to mimic or emulate intelligent human behavior in machines, while machine learning is a subset of artificial intelligence that enables computers to learn and develop on their own without having to be directly programmed. There are quite some number of A.I. medical imaging companies that are leveraging AI technology to diagnosis tumors in the brain and other diseases, companies such as IBM Watson, Brain Miner, and Visit A.I (Badza & Barjaktarovic, 2020). However, there is lots of research going on, involving the application of AI in health care; a lot of resources have been put into it because we can obviously save a lot of money and time using this technology in the health sector. In addition, there were numerous publications related to brain tumor classification in the last few years using different deep learning algorithms (See Figure 1.1).

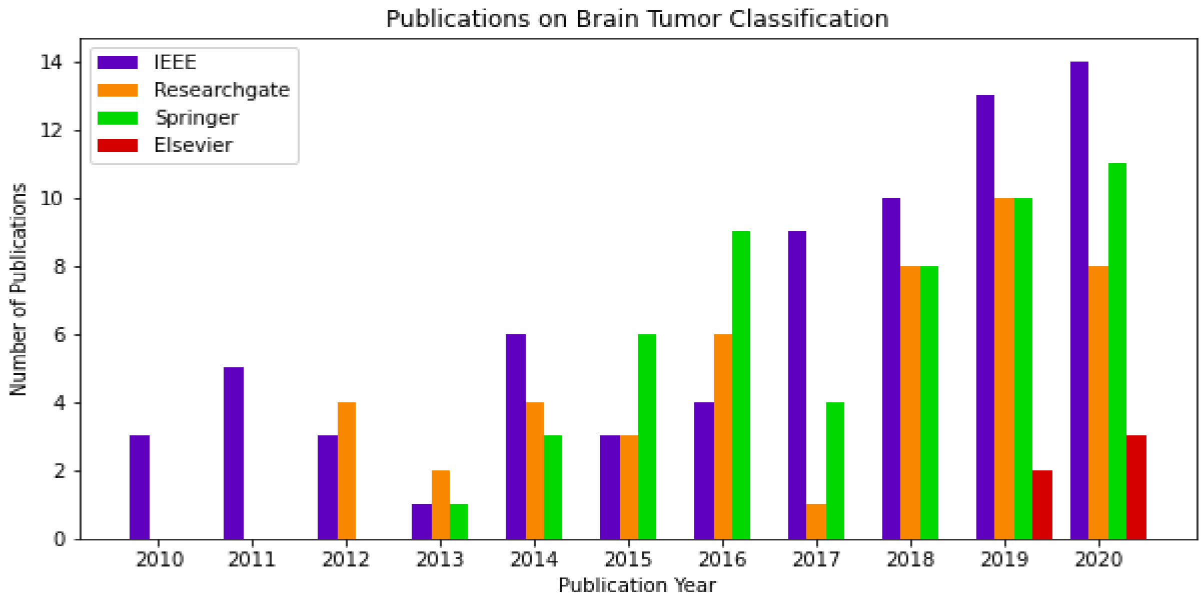


Figure 1.1: Number of publications on Brain Tumor Classification over the last 11 years in four major publication centers (Rehman et al., 2020).

This paper, proposed a very special and unique strategy of classifying brain tumors through MRI scan images by taking advantage of the AlexNet transfer learning model in which we frozen its weights and add our own weights and since this model has been previously trained on huge number of image samples, we transferred its parameters to a new CNN model which we used to perform binary classification on our brain MIR scan image dataset to predict whether each of those images has tumor or not. However, because we are dealing with binary classification, our predicted results are presented as zeros and ones, which is also a categorical output, where the zeros represent healthy images and the ones represent unhealthy images with high performance accuracy (See Figure 3.3 a & b).

This operation is possible by developing a very good deep learning model, and if one may ask, how to acquire a good deep learning model? Due to the enormous difficulties in image classification, a model with a high learning potential is required to learn about thousands of artifacts from dozens of images. Even a dataset as large as ImageNet

(Krizhevsky & Hinton, 2021) is insufficient to address this issue, so a model should acquire a lot of previous information to substitute for the data we don't have, in order to secure a very good deep learning model. However, CNNs may vary their size to regulate their ability, but they often create clear and accurate conclusions about their predictions. They have far less correlations and parameters than normal artificial neural networks with a similar number of layers, making them easier to train while their outputs are expected to be marginally high in performance (Deepak & Ameer, 2019) and CNNs have remained incredibly costly to apply to high-resolution images on a massive scale. This research is fortunate in that it makes use of Graphic Processing Units (GPUs) in conjunction with a well-optimized architecture of 2D convolution, which is powerful enough to allow the training of massive datasets with CNN models (Deepak & Ameer, 2019), and employs Rectifier Linear Unit as an activation function, back-propagation techniques, lower learning rate, proper percentage drop out, batch normalization, and good optimization which improved the overall accuracy of our model.

Deep-CNN models, on the other hand, can perform very poorly if they are not trained with a large dataset and also if the data is not well structured, because the model will not perfectly learn the correlation between the pixels in the images during training, causing the model to memorize the data rather than understanding it, and as a result, the model will perform very poorly when tested with another datasets it has never seen before (Bodapati et al., 2021). In addition, lack of rescaling or translating the input data can badly change the model's prediction to a very poor prediction because this decreases the generalization capability of the network exponentially. CNN models can also perform very poorly due to over-fitting and under-fitting. Over-fitting occurs when the model fits too much to the noise from the training data, which, as a result, has a negative impact on the model during evaluation. Under-fitting occurs when a model does not capture the underlying trend of the data and does not fit the data well enough, which also has a negative impact on its performance. However, there are various ways of countering or preventing over-fitting and under-fitting which are explicitly explained in this paper. The flowchart of our proposed technique is demonstrated in the figure below.

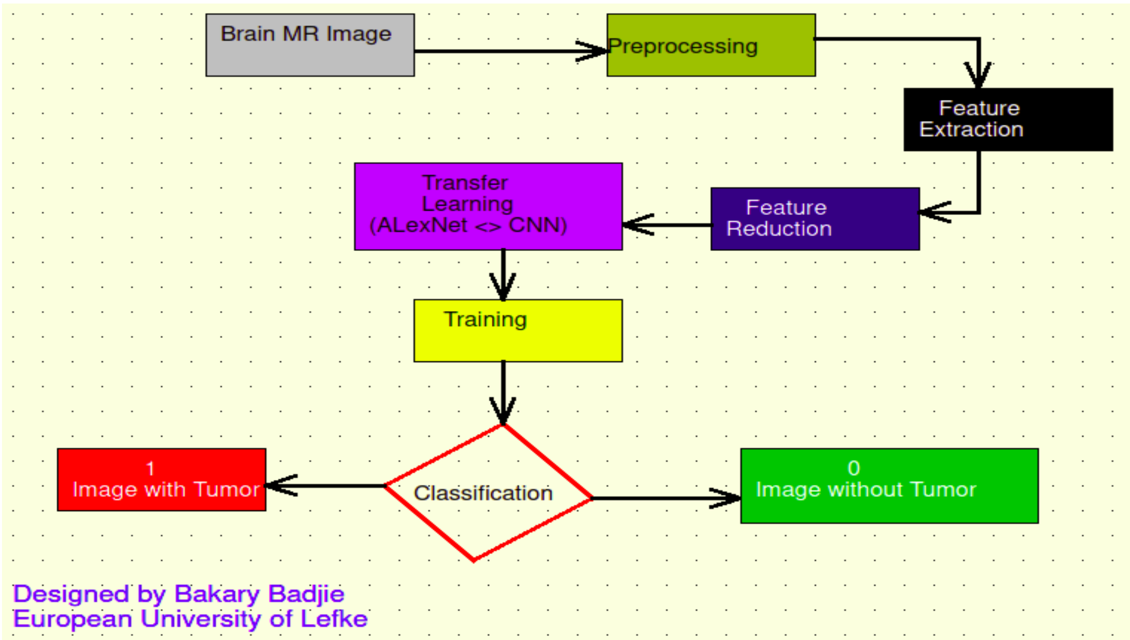


Figure 1.2 Flowchart of the proposed technique

The techniques for image processing and analysis have dramatically improved and computer aided algorithms are used to reveal hidden information in the image that is unobservable to the naked eye (Anaraki et al., 2019). Therefore, the use of the recent deep learning techniques in collaborations with radiography gives way to better diagnosis and prediction of different diseases which are threatening to human wellbeing.

1.2 Objectives of the Proposed Technique

Brain tumor is one of the most tumor complications that have claims many lives over the past decades and its diagnosis requires high level human intelligence which is a time consuming approach. Therefore, the proposed study is aimed at attaining the subsequent objectives;

1. To increase robustness, accuracy, and efficiencies in the health care sectors by developing a deep learning model capable of performing brain tumor classification using MRI scan images in order to assist radiologists in diagnosing brain tumors.

2. To produce high performance accuracy that will allow us to deploy its functionalities in the real world.

3. To understand the theory and intuition behind CNN and transfer learning in deep neural network.

1.3 Significant of the Proposed Technique

The proposed technique is quite an important technique in the field of radiology because it has the power to leveraged artificial intelligence technology to combine a modern computerized technology with radiography to diagnose brain tumor patients, which improved efficiency exponentially in the medical industries. Instead of radiologists spending tens of hours diagnosing brain tumor patients, which may even be inaccurate due to the fact that we humans are prone to errors and failures, which can claim lives, our proposed technology produces a highly accurate diagnosis within the blink of an eye, which is also less time consuming and has a very low risk of endangering patients' lives. It is almost close to impossible for radiologists to analyze thousands of MRI scan images at the same time and know whether they are affected by tumor or not. Therefore, leveraging the proposed technique to automate this entire process is very useful in today's medical environment to solve time-consuming and inaccuracy issues.

1.4 Scope of the Thesis

The scope of the thesis includes the brief summary of the proposed technique, the introductory section to assist readers in gaining a broad knowledge of the topic, to understand our aims and objectives of carrying out the research and also to provide the history of all the previous works that had been carried out on the subject matter. Within the scope is also to understand our adopted methodology to perform the brain tumor classification. A systematic view point is offered in conjunction with the topic matter that needed specifics such as the intuition behind deep learning, transfer learning, convolutional neural network, and computer based image processing. In addition, discussions on the outcome and achievements of the research are also within the scope of the thesis and finally, the conclusion of the research, which provides general discussions on the entire thesis as well as our recommendations for future studies, are also within the scope of the study.

1.5 Study Problem

This study's problem is focused on early diagnosis of brain tumors for effective patient treatment. Many people die as a result of late identification of tumors, particularly brain tumors. Individuals will have a better opportunity of being treated and cured if brain tumors are detected and classified at its premature stage.

1.6 Technology behind Radiography (Medical Imaging)

Radiography is a medical imaging technology that uses radiant energy to provide images of human tissues, organs, bones and vessels using different scanning and x-ray techniques. Technology in radiography offers non-invasive detail on brain cancer disease. The three most often employed diagnostic imaging techniques are computerized tomography (CT), ultrasound (US) and magnetic resonance imaging (MRI) (Afshar et al., 2019). CT photographs reflect the characteristics of the tissue through various X-ray reflection coefficients in the human body. CT is more useful in medical imaging because it has a brief scanning period, fast tumor localization and convenient imagery. However, stronger radiation may inflict damage on humans during radiographic process but CT images are less than MRI in resolution (Nagaraj et al., 2020).

Another medical imaging technique for the detection of ovarian tumor is ultrasound. Different tissues reflect, diffract, and interact with ultrasonic waves during this process. This highlights the conditions of the pelvic bodies which can then be used for further processing, including photography, by the effect of obtained ultrasound signal. The ovarian injuries are identified in the ultrasound picture by detecting variations and properties in the tissue. The ultrasound imaging technique is non-radioactive, affordable and recommended. However, the image quality generated is far lower than the images generated by MRI. Consequently, the detection of ovarian cancer by ultrasound imaging is restricted to certain simple roles, such as classifying benign and malignant tumors (Nagaraj et al., 2020)

The images from MRI scans are generated from the presence of the hydrogen nucleus and the powerful magnetic field. MRI gives outstanding images of pelvic structures without any radiation, which makes it special for imaging. By using numerous emission strategies, such as the pulse periodicity, the recurrent time, and the reversal time, MRI data sets are obtained. Watch the labels common fluid recovery sequences include T1-weighted imaging (T1W1) (Ghassemi et al., 2020), T2-weighted imaging, diffusion weighted imaging (T2W1), and apparent diffusion coefficient (ADC). There are differences in the MRI sequences for different issues.

The brain tumor disease is better seen with an MRI image than with an ultrasound or CT. Also, the T2WI series of MRI photos is made to assist the radiologists in locating tumor in the brain due to the fact that MRI has better criteria and a higher resolution

(Afshar et al., 2019). However, MR images of brain tumor are studied in this thesis. MRI has so many benefits over the other imaging methods, which makes radiologists to prefer it over all the other techniques.

1.7 Related Research

Numerous scholars from all over the globe have been concerned with the study of malignant tumors. Dozens of studies on brain tumors and various techniques for early diagnosis are released each year. In their suggested efforts, most of these studies depend on image processing methods such as segmentation. Many do similar jobs with the help of artificial intelligence structures. In other kinds of studies, identification was carried out using a mix of various detection techniques. A classification method built on rippled connection artificial neural networks was described in (Bodapati et al., 2021). MRI brain tumors are classified using the technique provided. Gabor features were extracted from the area of interest in the image. In comparison to previous research classifiers, the classifier was shown to be quite accurate. Badza & Barjaktarovic (2020) demonstrated their work on using ANN to identify brain tumors using electroencephalograms (EEG). The EEG is cited as a useful tool for determining brain function. The classifier was tested and outcomes were generated using the feed forward back propagation method.

Deepa Arunadevi and three other PhD students in the University of Science and Technology of China (USTC) suggested an intense machine learning-deep neural technique for brain tumor classification from 3D MR images in the month of December, 2020. The accuracy of his approach was 92.94 percent, the sensitivity was 92.07 percent, and the unique rate was 98.02 percent [8]. Arunadevi et al. used a multiclass classification technique to detect tumor in patients, using a dataset of 428 brain MR images. Initially, the author used ANN and later incorporated Principle Component

Analysis (PCA) algorithms with ANN in his process and saw an increase in classification accuracy from 77% to 93.2%.

In addition, Kumar and Vijayakumar proposed a PCA techniques in conjunction with a kernel based Support Vector Machine (SVM) algorithm for classifying and segmenting tumors in human brain, claiming an accuracy of 95.98 percent, a 95 percent overlap fraction, and a 0.025 percent extra fraction. This approach has a classification accuracy of 94% for identifying tumor type, with a total error rate of 7.5 percent. Abodapati et al. have introduced a credible method for classifying brain tumors from MR images that claims to be 92% accurate. He combined ANN with radiography to perform this operation.

Approaches in natural language processing (NLP), computer vision, and speech analysis as well as other sub-fields, have emerged over the past decade (Aliyu et al., 2020). With some new activation functions including support for graphics processing units (GPUs), produced dramatic improvements in ImageNet, an image dataset with a million levels. However, the applications of deep learning in numerous contexts have been improved because of its recently impressive performance in medical image processing. Qiu et al developed a multi-resolution automatic classification approach to classify using MRI based CNN model. Pathak et al proposed a deep learning model with convolutional layers that speed up the classification process. The cascading architecture connected the last network's output with the subsequent networks. They were using small-sized kernels for classifying each and every pixel in MR image. Using the small-sized kernel greatly lessened the risk of over-fitting in creating a larger network. Data enhancement and normalization also occur before training during pre-processing the image dataset (pathak et al., 2020). Dropout regularization of over-fitting usually the deep learning has been used. The algorithm has worked several times in magnetic resonance imaging (MRI).

Recent brain tumor research has largely relied on machine learning or medical image processing, a technique known as radiomics. The term radiomics refers to a new paradigm established by these two disciplines that combines radiology and bioinformatics. Radiomics also includes quantitative extraction method of maximum health from medical images as well as extracting the features of the image to gather the necessary data (Afshar et al., 2019). Composite neural network may be used to achieve certain targets, such as clinical material, disease prognosis, etc. In 2012, Lambin became the first to propose the theory of radiomics. He first proposed the radiomics theory and research method, then used image analytics to forecast the correlation between expression of genes and the tumor attributes, and also the correlation that ally the features of the image with the cancer prognosis.

Majority of current findings into computerized diagnostic tasks for brain tumor uses a radiomics approach. Kazerooni et al used linear discriminant analysis (LDA) classifiers to classify tumors in 65 patients. The time-to-peak (TTP) and wash-in-rate (WIR) features were found to have important sensitive point in distinguishing between malignancies and benign brain tumors in their analysis (Kazerooni et al, 2018). In a study of 101 patients with brain cancer, Rizzo et al assessed the correlation between radiomics attributes on CT images and prognostic factors. Their findings revealed that CT radiomics features could accurately predict the relapse rate in brain cancer cohorts 24 months after care. Two classes of CT brain images were compared by Qiu et al (pre-treatment and post-treatment). Three primary characteristics, tumor length, density, and density variance, were proven to be helpful in predicting 6-month progression-free survival (PFS). When these three features were combined, their model was able to predict PFS with an accuracy of 83%. Despite the fact that recent studies have aided, there are also a few limitations in these findings when it comes to diagnosing brain cancers by radiologists. To continue, we are uncertain of any past efforts on MRI brain tumor classification using medical images. The Type one and Type two classifications, as mentioned earlier, have significant clinical significance and act as a valuable framework for the laying a comprehensive treatment strategies.

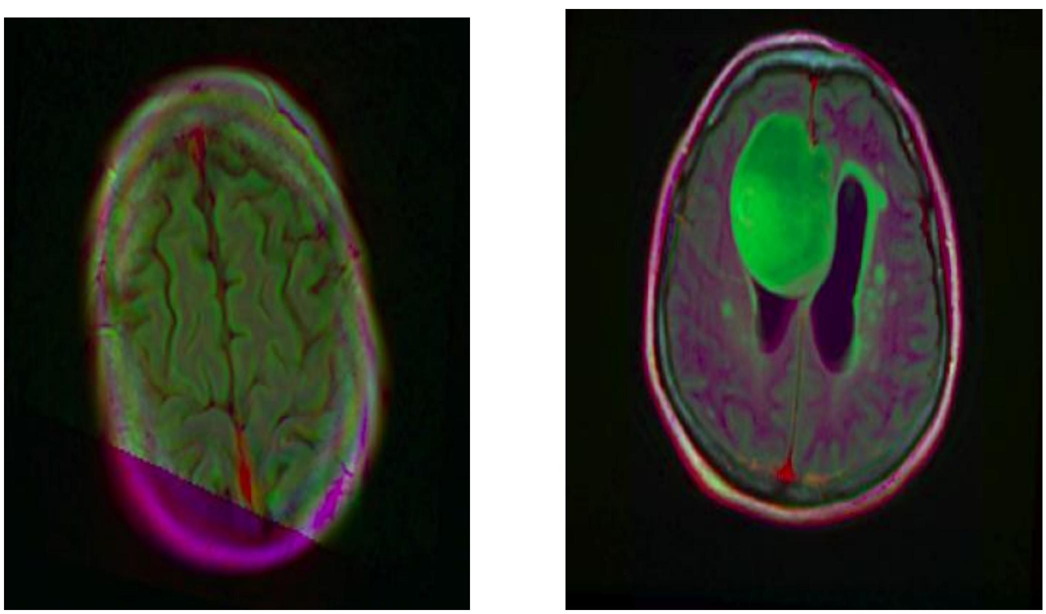
In addition, whereas the used of deep learning to evaluate clinical data of many diseases has increased exponentially, it has not yet been used to examine medical images of other diseases, we are unaware of any deep learning applications for brain cancer. When compared to radiomics, deep learning has a few more benefits. Deep learning is particularly useful for end-to-end model building because it has a high capacity for extracting features from images. Previous bran cancer research did not use deep learning, which limited the model’s accuracy.

CHAPTER TWO

RESEARCH METHODOLOGY

2.1 Data Pre-processing

The dataset used in this study was obtained from the kaggle database. The dataset contains 3929 brain MR images, 2756 of which were found to be affected by tumors, while the remaining images were found to be healthy. There are 1290 malignant tumors in the 2756 unhealthy images, and 1466 benign tumors. This indicates that this study used an unbalanced dataset to perform the classification task.



(a) Healthy brain MR image (b) Unhealthy brain MR image

Figure 2.1: Sample brain MR images used in this study

2.2 Introduction to Convolutional Neural Network

A convolutional neural network is a type of deep learning technique that is widely used in image processing and classification. It is often recognized as a shift invariant or space invariant neural network due to the fact that it is built on the shared-weight design of convolutional layers and max pooling layers that change the input features and give translation split answers (Deepak & Ameer, 2019). Many convolutional neural network architectures, contrary to popular belief, are only direct and straightforward for image processing. Picture and video identification, recommendation system services, image detection, segmentation techniques, medical image interpretation, natural language processing, brain-computer interfaces, and financial time series are only a few that the convolutional neural network is used for as a deep learning technique (Rehman et al., 2021). Figure 2.2 demonstrates the overall CNN architecture.

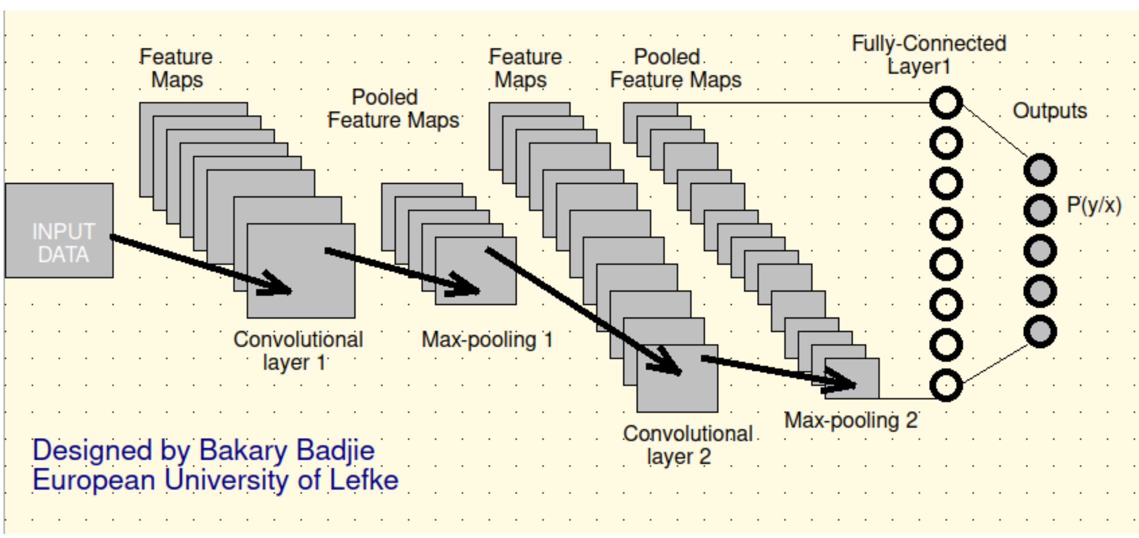


Figure 2.2: Convolutional Neural Network Architecture

The Convolutional neural network is made up of five convolutional layers and three fully connected layers. The power of CNNs lies in their ability to learn rich representations from raw input images. Each convolutional layer integrates its input feature maps with those derived features from the previous layers to provide higher-order feature maps that are further processed into a fully-connected layer. CNNs synthesize their feature extractors and train them according to their essence (Pashaei et al., 2018). The potential of CNN to acquire complicated non-linear feature map pings from a vast number of training samples justifies their widespread use in computer vision applications (Anaraki et al., 2019).

2.3 Introduction to AlexNet Transfer Learning Model

As mentioned earlier, AlexNet is a deep learning model that is a modified version of the convolutional neural network. Alex Krizhevsky proposed this model as part of his research. Geoffery E. Hinton, another well expert in the profession of deep learning research, oversaw his work. Their main finding was that, the model's depth was absolutely necessary for it to perform well. This was computationally expensive, but using multiple Graphical Processing Units (GPUs) made it possible during training (Alom et al., 2018). AlexNet was the very first convolutional network to use the graphics processing unit (GPU) to improve its performance during training. Its architecture is quite an amazing deep neural architecture and, because of this architecture, research has started thinking differently, which has led to the invention of other deep neural architectures such as VGG16, VGG19, ResNet50, GoogLeNet and much other architecture (Deeper & Ameer, 2020).

AlexNet was initially designed for the Kaggle ImageNet competition. The competition was organized to perform classification on the ImageNet dataset with 1000 categories and, therefore, this model was designed with 1000 outputs (Krizhevsky & Hinton, 2012). However, this study solved a binary classification problem using the brain MR image dataset with only two categories and, therefore, in order to use this architecture, we twisted the output layer to a binary output layer for the benefit of this research. Using the transfer learning technique, we froze all its layers and added new layers on top of them, with the help of Keras and Tensorflow2.0, which enabled us to achieve about 91,00,000 parameters, which is indeed quite a massive network to train our data with.

2.4 AlexNet Architecture

The AlexNet convolutional neural network is a transfer learning model with eight layers, including five convolutional layers and three fully connected layers, as proposed in this paper. Max-pooling layers are placed after the first, second and the fifth convolutional layers, while a flattened layer is placed between the final max-pooling layer and the fully connected layer (See Figure 2.3). The AlexNet network employs the Rectified linear Unit (ReLU) mechanism as an activation function, which outperforms the sigmoid and Tanh activation functions more than five times (Zhon et al., 2018). Below is the overall architecture of the AlexNet convolutional neural network.

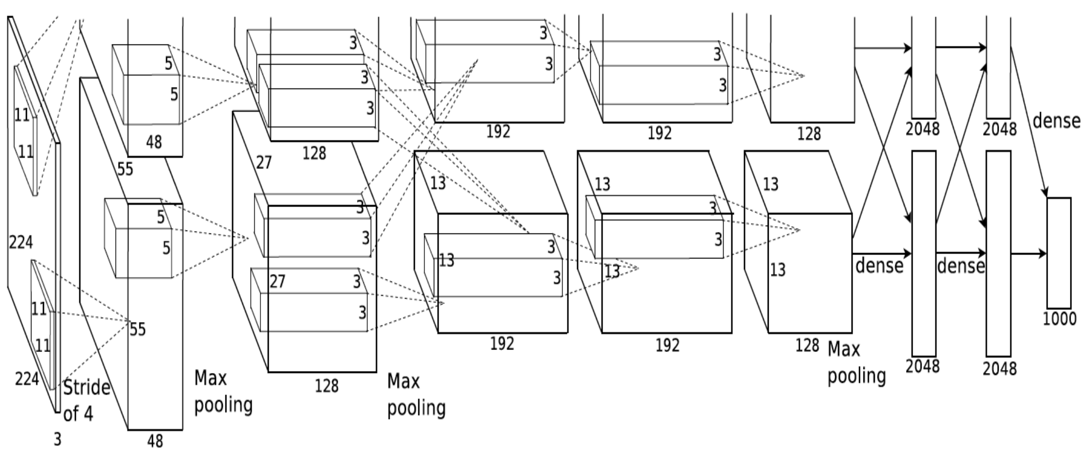


Figure 2.3: AlexNet-CNN architecture designed on ImageNet dataset. (Krizhervsky et al., 2012)

The reason for adopting AlexNet as our transfer learning model is that it empowers us to train our data on multiple GPUs. That is to say, it placed some of the neurons of the model on different GPUs, which accelerates the training process of our model by reducing the training time. Another advantage of AlexNet is that it leverages a rectified linear unit (ReLU) activation function in each convolutional layer to solve the vanishing gradient issue.

2.5 Data Preparation

After obtaining the brain MR image dataset from the Kaggle database, we prepared the data for the training in such a way that, we split the data into train, test and validation sets where 80% is allocated for training and 10% is allocated for both testing and validation respectively. We used the train-test-split function to organize the data into these sets.

2.6 Binary Classification on Brain MR Image Dataset

The binary classification on brain MR image dataset comprises of the following techniques as shown in the subsections below.

2.6.1 Leveraging Transfer Learning Technique

This research proposed a transfer learning technique to perform a binary classification on our brain MR image dataset to classify tumors in the images. It is a very unique technique in the field of deep learning because leveraging this technique, allowed us to train our model in a massive network. Another advantage is that, all the feature extraction layers are already pre-trained, and therefore, this dramatically reduces the computational time required in training of our model.

We transferred the knowledge of this pre-trained network to a new convolution neural network from one convolution layer another then we added a new dense fully connected artificial neural network (See Figure 2.4). The dense layers and their features were then frozen, implying that we did not necessarily train these features. The new network used the transferred knowledge to perform a binary classification task on our dataset, allowing us to predict whether or not the images show the presence of a tumor.

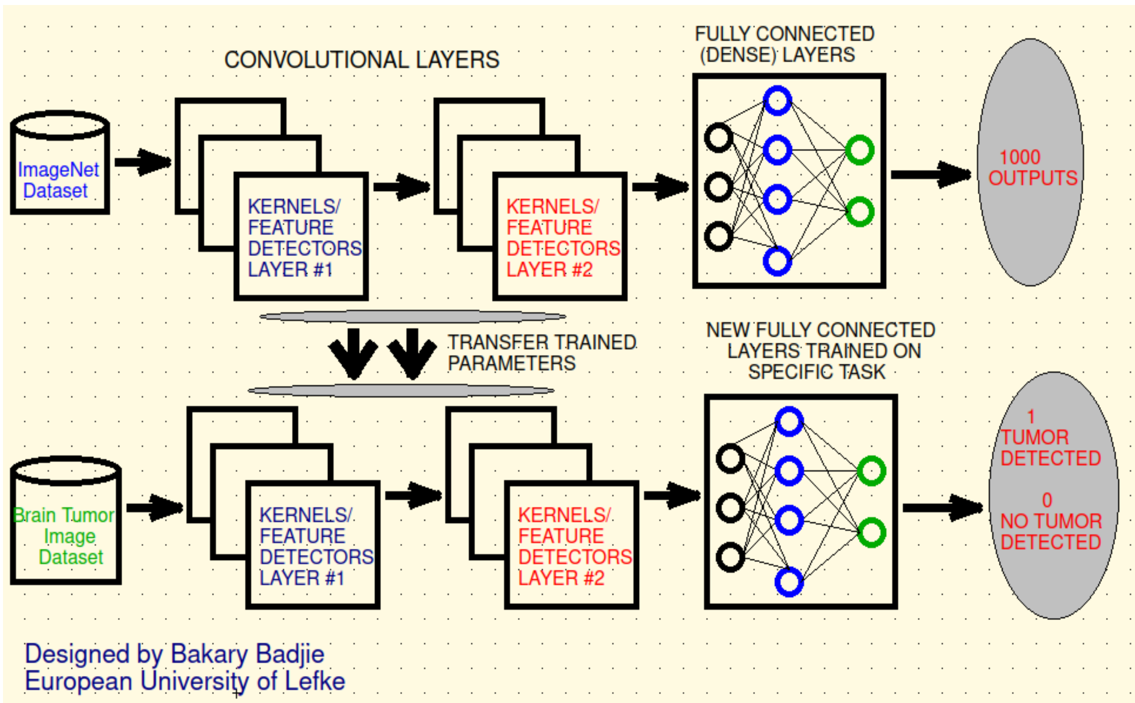


Figure 2.4: Transfer Learning Architecture

The transfer learning is actually extremely powerful because it extremely accelerates the training process where we achieved an incredible performance in just couple of epochs, which is pretty much fascinating.

2.6.2 Feature Extraction Technique used in this Study

After transferring the pre-trained parameters of the AlexNet model to our new CNN network, we fed the training and validation datasets into this new network. This training data consist of 2319 brain MR images with a pixel size $256 \times 256 \times 3$. Our CNN model received this training set as its input features through the first convolutional layer, and a convolutional operation was applied on these images, using a kernel filter size

11×11 pixels, striding of 4 steps and 96 kernel filters with zero padding. A max-pooling operation is applied on the output feature maps of the convolutional layer with a pooling size of 4×4 and a stride of 2 which down-sampled the image to $27 \times 27 \times 96$ pixels.

The down-sampled feature maps are further fed into the second convolution layer, where the same convolutional operation was performed, using a kernel filter sized 5×5 pixels, striding of 1 step and padding of 2 and with 256 kernel filters, and this operation outputted feature maps of size $27 \times 27 \times 256$ pixels. Here, the shape of the image remained the same because we used padding in this layer. However, on these convolved feature maps, a second max-pooling operation is performed with a pooling size of 3×3 pixels, zero padding, and a stride of 2 steps. This further down-sampled our image to $13 \times 13 \times 256$ pixels.

In addition, a third and a fourth convolutional operations were performed on the output feature maps from the second layer, using a kernel filter sized 3×3 pixels, padding of 1 and striding of 1. Each of these convolutional operations extracted the image features to feature maps of size $13 \times 13 \times 384$ pixels using a kernel filter size of 384 and 256 filters respectively. At the final convolutional layer, we obtained output feature maps of size $13 \times 13 \times 256$ pixels, and a final max-pooling operation was performed on these feature maps, which finally down-sampled the images to $6 \times 6 \times 256$ pixels. However, a ReLU activation function is applied after every convolutional operation to help update the kernel filters before performing a max-pooling operation.

The subsequent sections gave detailed working principles of each layer in our CNN architecture, which enabled us to successfully extract the important features of our images before feeding them to the fully-connected artificial neural network for prediction.

2.6.2.1 Convolutional Layer

In this study, we performed a 2D convolution on every convolutional layer. At its core, 2D convolution is a rather straightforward process: we begin with a kernel filter, which is essentially a tiny matrix of weights (Alom et al., 2018). This kernel "slides" through our 2D input feature map, doing pixel-wise multiplication with the portion of the input feature maps it is presently on, and then combining the outcome into a single output pixel (See figure 2.5). In other words, this layer detects local concordances between previous layers' features and maps their appearance to feature maps.

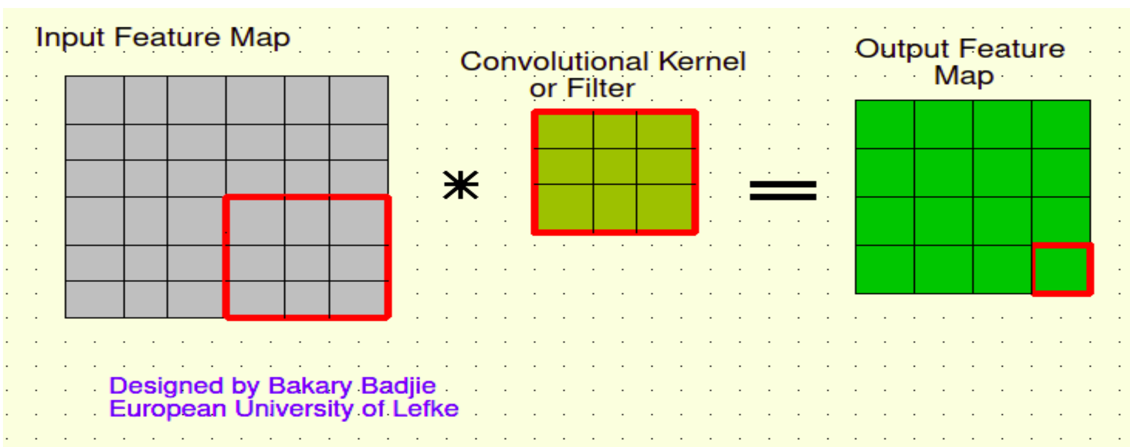


Figure 2.5: Feature Extraction Technique

With a lot of neurons involved, the image is transformed into a perceptron which produces local receptive fields, and then is compressed into feature maps of a particular size. The convolutional operation is mathematically represented as:

$B(i,j) \sum_{m=0} \sum_{n=0} K(m,n) \ * A(i-m,\ j-n)$, (Alom et al., 2018) [1]

Where;

K= Kernel, A= Input features and B= Output features

The primary goal of the convolution operation is to decrease the number of model parameters while focusing on the local connection. The properties of this layer are required to meet two standards which are, weight sharing and local perception.

The subsequent subsections explain various steps that are carried out or that are used to extract the features during convolutional operation in each convolutional layer in our CNN architecture.

2.6.2.1.1 Striding

Striding is an important building block of the convolutional neural network. It is the amount of steps we move the kernel filter both horizontally and vertically on top of the input feature maps in order to perform a pixel-wise multiplication in the convolutional and the max-pooling layer (Aliyu et al., 2020). The proposed architecture used a stride of metrics sizes in the five convolution layers respectively and a stride of metrics size at every max-pooling layer.

2.6.2.1.2 Kernel Filters

These are the convolutional filters that are convolved with the input feature maps during a convolutional operation in our proposed architecture. They are represented as a matrix format with a desired size. In this study, kernel filters were used in each of the five convolutional operations, they are feature extraction filters that extract important features from our images (Abiwinanda et al., 2019).

2.6.2.1.3 Padding

Padding is an image processing technique proposed in this research to allow the output feature maps to have the same pixel size as the input features after a convolutional operation on an image. We used padding in this study to add extra vertical and horizontal pixels to the input feature maps depending on the number of paddings. It helps us not to lose any information about the input image after applying the kernel filters to the images (Pashaei et al., 2018).

2.6.2.1.4 Batch Normalization

This research proposed a technique of introducing additional layers to a deep neural network to make it quicker and more reliable during training. The standardizing and normalizing procedures are performed by the newly added layers on the output of an activation function which puts all the data on the same scale (Belaid & Loudini, 2020). This operation reduces the generation error, and also eliminates internal covariate shift, by enabling a normal distribution of the input data for every layer around the same mean and standard deviation, which prevents the weights from being imbalanced. However, this whole process occurs in each batch we set up before the training.

2.6.2.1.5 Applying Rectified Linear Unit (ReLU)

Rectified linear units are a unique design that incorporates non-linearity and rectification layers into a deep neural network. It is applied to every convolution layer in our architecture to avoid the vanishing gradient problem (Ismael & Abdel-Qader, 2018). That is to say, it prevents the weights of our network from not being updated during back-propagation. As a result, the vanishing gradient prevents the weights from updating after each epoch or iteration during training. During forward propagation, “ReLU” helps in the activation of the weights for the features to be transferred from the input layer to the hidden layer or from one hidden layer to another (Alom et al., 2018). “ReLU” is described as:

$Y_i = max(0, x)$ and $x = w_i\, x_i + b_i$, (Ismael & Abdel-Qader, 2018) [2]

Where;

w_i = Weights of the network, x_i = input image and b_i =bias term

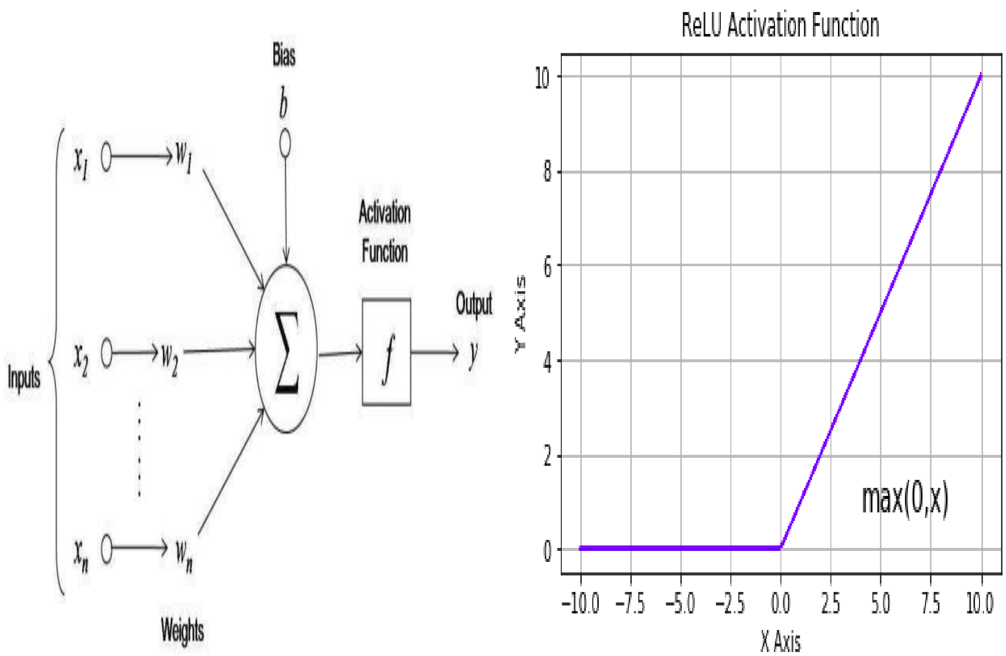


Figure 2.6: (a) Mathematical structure of ReLU. (b) Graphical representation of ReLU.

2.6.2.2 Max-Pooling Layer or Feature Pooling Layer

On a typical CNN, max-pooling pooling is the most reliable pooling technique. A feature pooling layer sub-samples the feature maps it receives, resulting in the same collection of feature maps with smaller dimensions (Cheng et al., 2015). It uses a down-sampling technique called the "location Invariant", which is, it places a pooling filter on top of the convolved feature from the convolutional layer and takes the highest pixel value at every stride and placed it in the corresponding output matrix (See Figure 2.7). During back-propagation, the pooling filters get updated with the help of the ReLU activation function (Rehman et al., 2021).

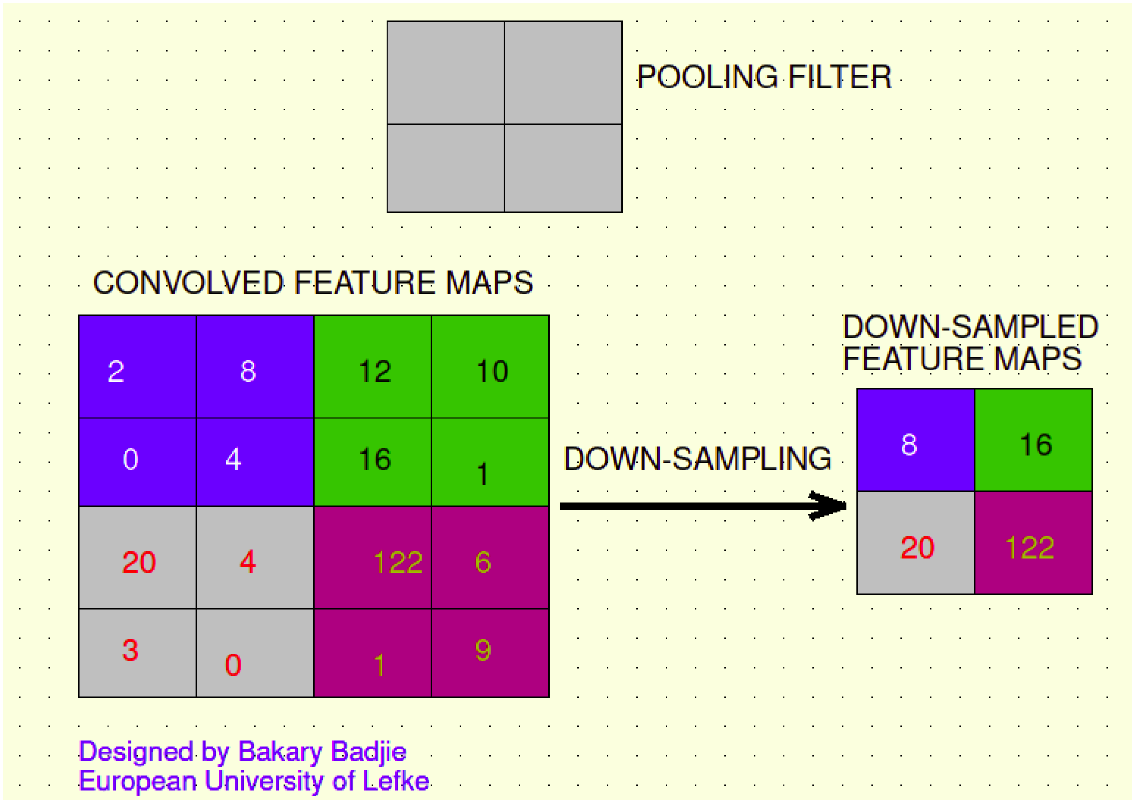


Figure 2.7: Max-pooling operations on the convolved features maps

After this operation the relative locations of the features are maintained and the image main properties are not lost. This operation is necessary because it reduces the parameters network which gave us chance to avoid over-fitting.

2.6.3 Fully Connected Layer

The fully-connected layer is the last and final layer in our proposed architecture. It is also known as a dense layer in a deep neural network. Our proposed architecture has three fully-connected which takes in the end results of the last pooling layer via a flattening layer (See Figure 2.8). That is, we receive the flattened feature vectors after the flattening operation and combine the features into two attributes or neurons at the output layer to make the actual binary classification decisions on our brain MR image dataset.

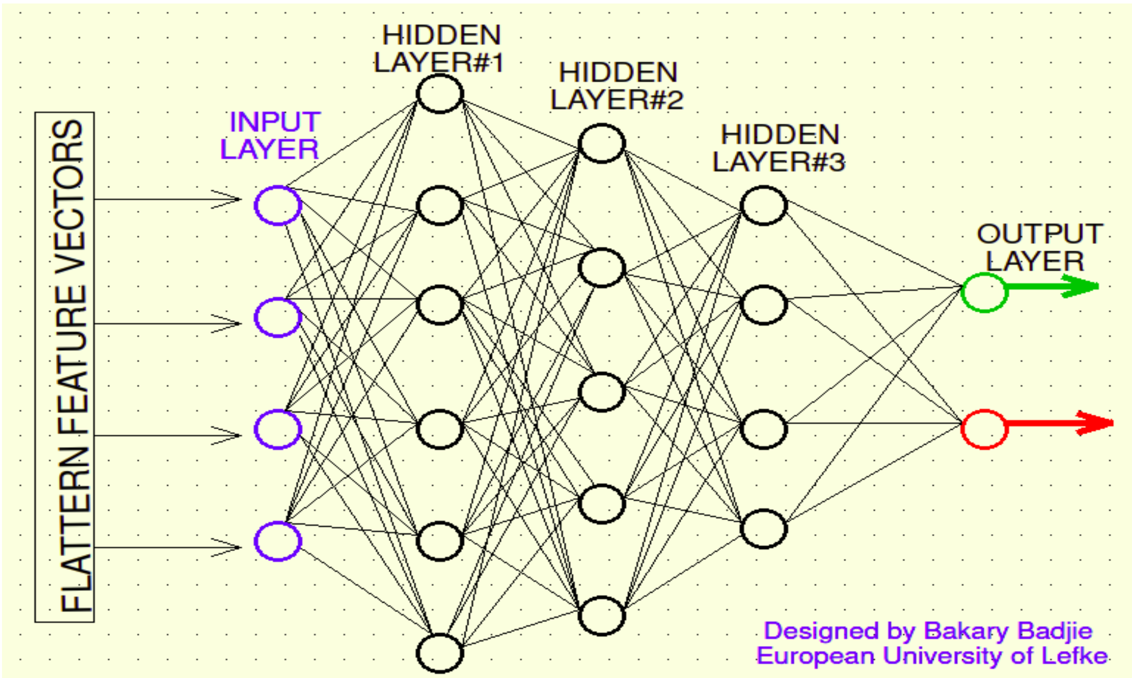


Figure 2.8: Fully-connected layer of CNN classification model.

Without this layer, our CNN model would be unable to predict the actual classes of our image dataset. In this layer, every input is connected to each and every other output by weight. The fully connected layer is mathematically expressed as;

$y_i^{(l)} = f\left(z_i^{(l)}\right)$ With $z_i^{(l)} = \sum_{j=1}^{m_1^{(l-1)}} w_{ij}^{(l)}\, y_i^{(l-1)}$, (Rehman et al., 2021) [3]

Where;
$m_1^{(l-1)} \times m_2^{(l-1)}$ = Flatten feature vectors
$y_i^{(l)} = f\left(z_i^{(l)}\right)$ = Fully connected layer
$w_{ij}^{(l)}$ = Weights
$y_i^{(l-1)}$ = Output layer

This layer is intended for fine-tuning the weight parameters $w_{ij}^{(l)}$ in order to generate a stochastic probability representation of each class centered on the activation maps produced by the sum of convolutional, non-linearity, rectification, and pooling layers. However, in the fully-connected layer there are various activities that occur during the training process. These activities are detailed in the subsequent subsections.

2.6.3.1 Forward Propagation

Feed-forward is a process in a deep neural network where all the information from the input layer is passed on to one hidden layer to another hidden layer to the output layer of the network by multiplying input features by the weights plus the bias. This process is fully connected, which means every neuron in one layer is connected to every other neuron in the next layer and, therefore, the outputs of one perception are directly fed as inputs to another perception. Basically, this allows the network as a whole to learn about interactions and relationships between features in the different layers, where the first layer is the input layer, which directly receives the data. It is expressed as;

$fw = w_i \times m_i + b$, (Rehman et al., 2021) [4]

Where; w_i is the weights, m_i is the input features and b is the bias.

2.6.3.2 Dropout

Dropout can essentially be thought of as a form of regularization technique used in the proposed architecture to help prevent the over-fitting of the model to the training data during training. It helped us randomly drop 40 percent of our input features and the neurons, along with their weights.

2.6.3.3 Loss Function

This study used cross entropy as a loss function to evaluate the performance of the neurons on our brain MR images in our network. It adds up the average differences between the original and predicted values and squares them up. In other words, it measures how our model is diverging from the gradient curve, which means, it tells us how wrong our model is in finding the relationship between the original and the predicted features. This loss function is mathematically expressed as;

$C = \left(\frac{-1}{n}\right) \sum_{i=1}^{n} (y \times \ln(a) + (1 - y) \times \ln(1 - a))$, (Belaid & Loudini, 2020) [5]

Where y represents the True value and a represents the model predictions

2.6.3.4 Back Propagation

Back propagation is a technique proposed in this research to help the optimizer to update the weights of our network in each hidden layer in response to the output of the loss function at every epoch during the training process. It allows the training process to go backward and update the weights via the activation function by taking the difference between the weights and the product of the learning rate and the derivative of the loss function with respect to the weight. Weight updating is calculated mathematically as follows:

$w_{new} = w_{old} - \zeta \frac{dC}{dw_{old}}$, (Alom et al., 2018) [6]

Where;
w = Weights, C = Cross entropy loss function, and ζ = Learning rate

In other words, back propagation calculates the negative gradient of a loss function and then distributes it back through the network layers for the weights to get updated.

2.6.3.5 Adam Optimizer (Adaptive Moment Estimation)

This study used an Adam optimization algorithm to minimize the loss of the model during training by updating the model in response to the output of our loss function. It is a combination of Adagrad and RMSprop optimizers. The reason for choosing this optimization algorithm is because its hyper-parameters have an intuitive interpretation and typically require little or no tuning. Secondly, it is computationally efficient and requires little memory usage. It is mathematically expressed as;

$\theta_t = \theta_{t-1} - \propto \frac{\widehat{M_t}}{\sqrt{\widehat{V_t}} + \varepsilon}$, (Bodapati et al., 2021) [7]

Where $\widehat{M_t}$ and , $\widehat{V_t}$ are the bias adjustment term at time t , ε (10^{-8}) is a small positive number to ensure that there is no zero division when the RMSprop becomes zero, and $\propto$ is the learning rate.

2.6.3.6 Learning Rate

This study uses a learning rate of 0.001 which is used by the Adam optimizer at every epoch to find the global minima of our loss function during training. However, a higher learning rate causes the model to learn poorly, while a lower learning rate helps the model to learn correctly, though the training process will be slow.

CHAPTER THREE

RESULTS and DISCUSSIONS

3.1 Results

The results of the work in this study are presented in both a graphical and a tabula format, showing the performance of our model during and after training. In other to judge the performance of the proposed model, we need to analyze the model's performance in different ways using different approaches. As mentioned earlier, the model is trained on a massive network with millions of parameters and 3929 brain MR image samples were used to train the model. The classification results are based on our validation accuracy report, which tells us how perfectly our model behaved during training. When we fed the model a dataset that it had never seen before, the true performance of the model capability was evaluated. After being tested, the model performed admirably (see Appendix 2 in page 41).

After completing the training process, the model was able to correctly classify almost all the images in our training set. Out of the 2319 images fed into the model, only seven images were wrongly classified. Appendix 3 in page 42 demonstrates the number of images that are correctly and wrongly classified by the model. The confusion matrix shows the true and false positives which tell us the right predictions of the model. It also shows us the true and false negatives which tell us the wrong predictions the model has made on our dataset.

The classification report presents the precision, recall, and F1-score of our model after the training. According to figure 3.3, our model shows a precision, recall and F1-score very closed to one which shows that the model's prediction is very accurate and very closed to perfection. One of the standard ways of judging the capability of a deep learning model is by analyzing the accuracy and loss plot of the model. The accuracies are the training and validation accuracies and losses are also presented as training and validation losses. The

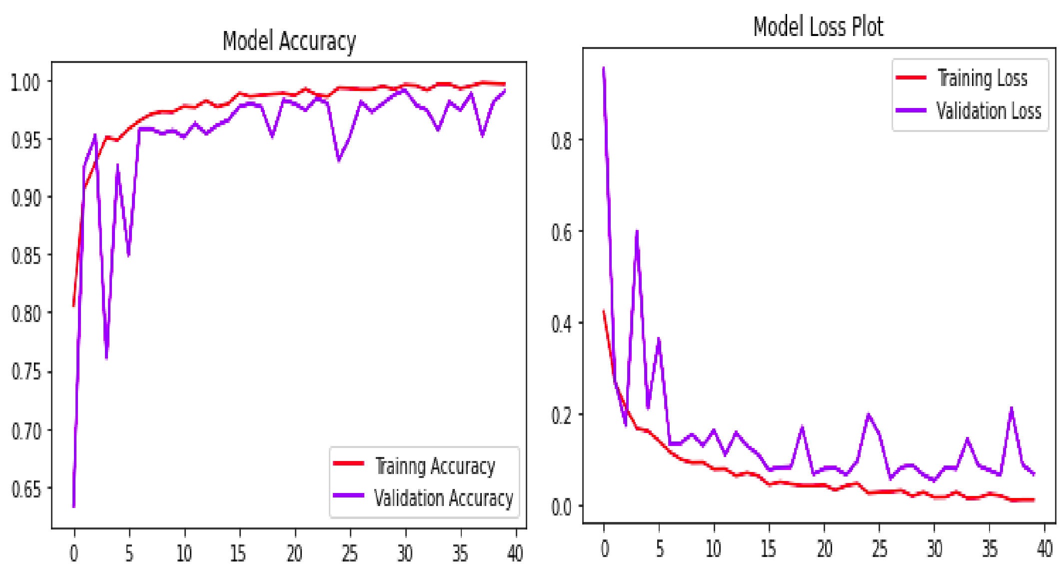


Figure 3.1: (a) Model's Accuracy plot (b) Model's Loss Plot

According to the above plots, the model obtained a training accuracy of 80.09% at the first epoch and, as the training was going on, this accuracy kept on increasing and, at the final epoch, the model obtained an overall training accuracy of 99.70%. Looking into the validation accuracy, the model attained an accuracy of 60% at the first epoch and as we trained over more epochs, this accuracy kept on increasing while overlapping at some epochs and, at the final epoch, the model obtained a validation accuracy of 99.62%. However, considering the loss plot, the training loss at the first epoch was 4.2% and as the model was learning, this loss kept on decreasing and at the final epoch, the training loss was almost at the minima on the gradient curve with a loss of 0.13%. Looking into the validation loss, the model performed badly in the first ten epochs and as we moved into the training of more epochs, the model had a validation loss of 0.19%. However, a comparison analysis is being performed between our proposed techniques with some recently proposed techniques performed by different researchers (See Appendix 1).

3.2 Discussions

This study demonstrates an outstanding performance in classification brains with great confidence. The performance was analyzed and evaluated using different standard required to determine a good prediction. We were motivated to carry out this study because of a growing concern in the alarming rate of inefficient diagnosis in the health sectors using human intelligence which causes which causes many losses of lives. The subsequent subsections are the standard tools for evaluating any machine learning/deep learning models.

3.2.1 Classification Report

The classification report is used in this research to evaluate how well our model actually performs on our brain MR image dataset. This evaluation is performed by understanding the key classification metrics, which are accuracy, recall, precision and F1 score. Typically, in any classification test, a model can only achieve two results; either the model is correct in its prediction or it is incorrect in its prediction. However, these metrics can be expanded in situations where we have multiple classes, such as in multi-classification tasks. Nonetheless, this research is performed on binary classification and, therefore, we are dealing with only two classes, which show whether or not the image has a tumor. The subsequent subsections highlight the concept of these key classification metrics.

3.2.1.1 Accuracy

Accuracy is one of the most common classification metrics and it is obtained in classification problems by the number of correct predictions made by the model divided by the total number of predictions made by the model (See Appendix 2). Accuracy essentially answers the question of how many predictions the model gets right as a percentage. It is expressed as:

$$Accuracy = \frac{Correct\ Predictions}{Total\ Predictions} \times 100\%$$

$$= \frac{True\ Positives + False\ Positives}{True\ Positives + False\ Negatives + True\ Negatives + \ False\ Negatives} \times 100$$

However, before using AlexNet model, we used the dataset on two other state-of-art deep learning algorithms which are VGG19 and InceptionV3 and we were not able to achieve accuracy greater than the accuracy of the previous works that were performed on this topic. Blow is the table that shows the accuracy of the all the models that we used performed for this classification problem while looking for the best model.

Table 1.

Deep Learning Model	**Dataset**	**Performance Accuracy**
VGG19	3929 Brain MR images	99.20%
InceptionV3	3929 Brain MR images	91.59%
AlexNet	3929 Brain MR images	99.62%

This is the main reason we choose AlexNet as our proposed transfer learning model to perform this particular task.

3.2.1.2 Recall

Talking about recall, it simply means the ability of a model to find all the relevant instances within a dataset (See Appendix 4) and, precisely, recall is obtained by the number of true positives divided by the sum of the true positives and false negatives. It is expressed as;

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

3.2.1.3 Precision

In terms of performance metrics, precision is the ability of a classification model to identify only the relevant data points; in other words, it expresses the proportion of the data points that the model has identified as relevant and is calculated by dividing the number of true positives by the sum of the number of true positives and the number of false positives (See Appendix 4). It is expressed as:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

3.2.1.3 F1 Score

F1 score is essentially the combination of recall and precision. F1 score is the harmonic mean of precision and recall taking both metrics into account (See Appendix 4). It is express as;

$$F1\ Score = \frac{2 \times (Precision\ \times Recall)}{Precision + Recall}$$

The reason the harmonic mean is used instead of just a simple mean or simple average is because the harmonic mean, lets us punish or eliminate extreme differences between precision and recall and gives you a fair assessment of that tradeoff between precision and recall. A model with a precision of 1 means essentially perfect precision and a recall of zero, essentially the worst record possible.

3.2.2 Confusion Matrix

In this research, the classification metrics are evaluated using a confusion matrix. A confusion matrix provides us with a two by two matrix which contains information about both true and false positives, as well as both the true and false negatives of the model's performance (See Appendix 3). True positive and negative show us the number of correct or right predictions the model has made on our test MR image dataset, which is fed into the model after completing the training, while false positive and negative tell us the number of wrong or incorrect predictions that the model has made on the data. The confusion matrix helps us understand the tradeoff between recall and precision because accuracy on its own is very straightforward.

The question many students ask is, at what percentage of accuracy a model becomes a good model? The answer is, it really depends on the context of the situation we are building our model. However, from our recommendation, we would like to say that 90 percent accuracy of a model is good enough for any situation.

CHAPTER FOUR

CONCLUSION

4.1 Summary

In conclusion, we proposed a technique for classifying brain tumors using a convolutional neural network. Our CNN model utilized the pre-trained features of AlexNet deep-CNN to extract deep features from brain MR images and the features were down-sampled and were passed into a fully-connected ANN to predict the final output of our network. From the final down-sampling stage, down-sampled feature maps of our images are fed into a flattening layer, which flattens them to feature vectors and passes them into the fully connected layer and a 40% dropout is applied to each fully connected layer to avoid over-fitting during the training process and a sigmoid activation function is employed at the last layer which activates the output to a binary output, through which, the network is able to perform the prediction on our dataset and tell us if an image has a tumor or not. The practical phase of this study was carried out with the help of the Python programming language and the Google-colab integrated development environment (IDE).We pre-processed the data, prepared the data into three different sets, we constructed the model, started the training process and finally, we evaluated the performance of the model using different evaluation techniques. After testing the model on a new brain MR image dataset, the experimental results show that our model's predictions are 99.10% correct. Transfer learning is one of the techniques that keeps me awake at night when it comes to AI because, unfortunately for humans, we die with all of our memories, which is one of the major problems or issues for us. For example, Einstein's brain, we were unable to preserve his intelligence, so he essentially took all of his knowledge with him when he died. It is close to impossible for other humans to try to capture the knowledge that he died with. Therefore, leveraging AI technology is essential in today’s world.

4.2 Recommendations

As the research mentioned earlier, we often have a precision-recall tradeoff, which means we essentially need to decide if the model should focus on fixing false positives versus false negatives. So, at the cost of decreasing false negatives, we are likely to increase false positives and vice versa. Therefore, the following recommendations must be put into consideration.

- We recommend that, it is probably better to try to minimize the number of false negatives at the cost of increasing the number of false positives, so that it goes in the direction of these false positives. The reason for this is that we want to ensure that we correctly classify as many cases of the disease as possible, and we don't want to turn away someone who has brain cancer by telling them they don't have it.

- In addition, we recommend that, all the images that have the presence of a tumor in them according to our model's predictions should actually go through to the next step of an invasive test in a real-life scenario. The confusion matrix lacks a universal truth, which is the truth that applies to all problems. Therefore, it is really a collaborative process where we should always be consulting experts in the domain we are making predictions before taking a final decision.

- For future work on this topic, it is recommended that, a brain MR image dataset should be collected in such a way that it contains an extract column which must contain the type of tumor, so that if we classify the presence of a tumor in any images, we can build another model to classify the type of brain tumor in a particular image.
- Because a tumor goes through different stages as it grows in our brain tissues, we recommend using a dataset with multiple categories in the future so that we can perform a multi-classification task to detect and classify the tumor at each stage.

REFERENCES

Gumaei, A., Hassan, M. M., Hassan, M. R., Alelaiwi, A., & Fortino, G. (2019). A Hybrid Feature Extraction Method With Regularized Extreme Learning Machine for Brain Tumor Classification. *IEEE Access*, *7*, 36266–36273. https://doi.org/10.1109/access.2019.2904145

https://www.kaggle.com/datasets?search=MRI+brain+tumor+dataset&datasetsOnly=tru

Afshar, P., Plataniotis, K. N., & Mohammadi, A. (2019). Capsule Networks' Interpretability for Brain Tumor Classification Via Radiomics Analyses. *2019 IEEE International Conference on Image Processing (ICIP)*. Published. https://doi.org/10.1109/icip.2019.8803615

Rehman, A., Naz, S., Razzak, M. I., Akram, F., & Imran, M. (2019). A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning. *Circuits, Systems, and Signal Processing*, *39*(2), 757–775. https://doi.org/10.1007/s00034-019-01246-3

Ghassemi, N., Shoeibi, A., & Rouhani, M. (2020a). Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. *Biomedical Signal Processing and Control*, *57*, 101678. https://doi.org/10.1016/j.bspc.2019.101678

Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., Wang, Z., & Feng, Q. (2015). Correction: Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. *PLOS ONE*, *10*(12), e0144479. https://doi.org/10.1371/journal.pone.0144479

Soumik, M. F. I., & Hossain, M. A. (2020). Brain Tumor Classification With Inception Network Based Deep Learning Model Using Transfer Learning. *2020 IEEE Region 10 Symposium (TENSYMP)*. Published. https://doi.org/10.1109/tensymp50017.2020.9230618

Deepak, S., & Ameer, P. (2019). Brain tumor classification using deep CNN features via transfer learning. *Computers in Biology and Medicine*, *111*, 103345. https://doi.org/10.1016/j.compbiomed.2019.103345

REFERENCE (Cont.)

Badža, M. M., & Barjaktarović, M. (2020). Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. *Applied Sciences*, *10*(6), 1999. https://doi.org/10.3390/app10061999

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A. S., & Asari, V. K. (2019). A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics, 8*(3), 292. https://doi.org/10.3390/electronics8030292

Bodapati, J. D., Shaik, N. S., Naralasetti, V., & Mundukur, N. B. (2020). Joint training of two-channel deep neural network for brain tumor classification. *Signal, Image and Video Processing*, *15*(4), 753–760. https://doi.org/10.1007/s11760-020-01793-2

R, T., & S, D. G. (2020). Accuracy based Correlation of ANN, CNN and PNN in Analysing Brain Tumor Classification. *Journal of Advanced Research in Dynamical and Control Systems*, *12*(05-SPECIAL ISSUE), 1047–1053. https://doi.org/10.5373/jardcs/v12sp5/20201855

Kumar, R. L., Kakarla, J., Isunuri, B. V., & Singh, M. (2021). Multi-class brain tumor classification using residual network and global average pooling. *Multimedia Tools and Applications*, *80*(9), 13429–13438. https://doi.org/10.1007/s11042-020-10335-4

Deepak, S., & Ameer, P. (2020). Retrieval of brain MRI with tumor using contrastive loss based similarity on GoogLeNet encodings. *Computers in Biology and Medicine*, *125*, 103993. https://doi.org/10.1016/j.compbiomed.2020.103993

Sultan, H. H., Salem, N. M., & Al-Atabany, W. (2019). Multi-Classification of Brain Tumor Images Using Deep Neural Network. *IEEE Access*, *7*, 69215–69225. https://doi.org/10.1109/access.2019.2919122

Chethan, K. (2020). An Efficient Medical Image Retrieval and Classification using Deep Neural Network. *Indian Journal of Science and Technology*, *13*(39), 4127–4141. https://doi.org/10.17485/ijst/v13i39.1621

Rehman, A., Khan, M. A., Saba, T., Mehmood, Z., Tariq, U., & Ayesha, N. (2020). Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture. *Microscopy Research and Technique*, *84*(1), 133–149. https://doi.org/10.1002/jemt.23597

REFERENCE (Cont.)

Badža, M. M., & Barjaktarović, M. (2020b). Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. *Applied Sciences, 10*(6), 1999. https://doi.org/10.3390/app10061999

Kabir Anaraki, A., Ayati, M., & Kazemi, F. (2019). Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *Biocybernetics and Biomedical Engineering*, *39*(1), 63–74. https://doi.org/10.1016/j.bbe.2018.10.004

Abdulkarim, H. A., Abdul Razak, M. A., Sudirman, R., & Ramli, N. (2020). A deep learning AlexNet model for classification of red blood cells in sickle cell anemia. *IAES International Journal of Artificial Intelligence (IJ-AI), 9*(2), 221. https://doi.org/10.11591/ijai.v9.i2.pp221-228

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Pashaei, A., Sajedi, H., & Jazayeri, N. (2018). Brain Tumor Classification via Convolutional Neural Network and Extreme Learning Machines. *2018 8th International Conference on Computer and Knowledge Engineering (ICCKE)*. Published. https://doi.org/10.1109/iccke.2018.8566571

Ismael, M. R., & Abdel-Qader, I. (2018). Brain Tumor Classification via Statistical Features and Back-Propagation Neural Network. *2018 IEEE International Conference on Electro/Information Technology (EIT)*. Published. https://doi.org/10.1109/eit.2018.8500308

Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A., & Mengko, T. R. (2018). Brain Tumor Classification Using Convolutional Neural Network. *IFMBE Proceedings*, 183–189. https://doi.org/10.1007/978-981-10-9035-6_33

Zhou, Y., Li, Z., Zhu, H., Chen, C., Gao, M., Xu, K., & Xu, J. (2019). Holistic Brain Tumor Screening and Classification Based on DenseNet and Recurrent Neural Network. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 208–217. https://doi.org/10.1007/978-3-030-11723-8_21

Afshar, P., Plataniotis, K. N., & Mohammadi, A. (2019a). Capsule Networks for Brain Tumor Classification Based on MRI Images and Coarse Tumor Boundaries. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and*

APPENDIX

Appendix 1

Author	Classification Approach	Accuracy
Rajan & Sundar, 2019	Support Vector Machine (SVM)	98%
Shree & Kumar, 2018	Probabilistic Neural Network (PNN)	95%
Arumachalam & Royappan, 2017	Feed-Forward propagation Neural Network	99.59%
Ullah et al., 2020	Feed-forward Neural Network	95.8%
B. ural 2018	PNN	90%
Preethi & Ashwarya, 2019	Deep Neural Network	99.3%
Francisco et al., 2021	Multi-pathway CNN	97.3% 97.3%
Deepak & Ameer, 2019	Deep Transfer Learning (GoogleNet)	98%
Ahmet &Mohammad, 2020	CNN	97.20%
Hemanth et al., 2019	Deep Transfer learning (ResNet50)	94.50%
Saxena et al., 2019	Deep Transfer learning (InceptionV3)	95%
Das et al., 2019	Deep Transfer learning (VGG19)	94.39%
Paul et al., 2017	Fully-Connected CNN	91.43%
Proposed	Deep Transfer learning (AlexNet)	99.62%

Classification performance on Brain MR Image dataset proposed recently

Appendix 2

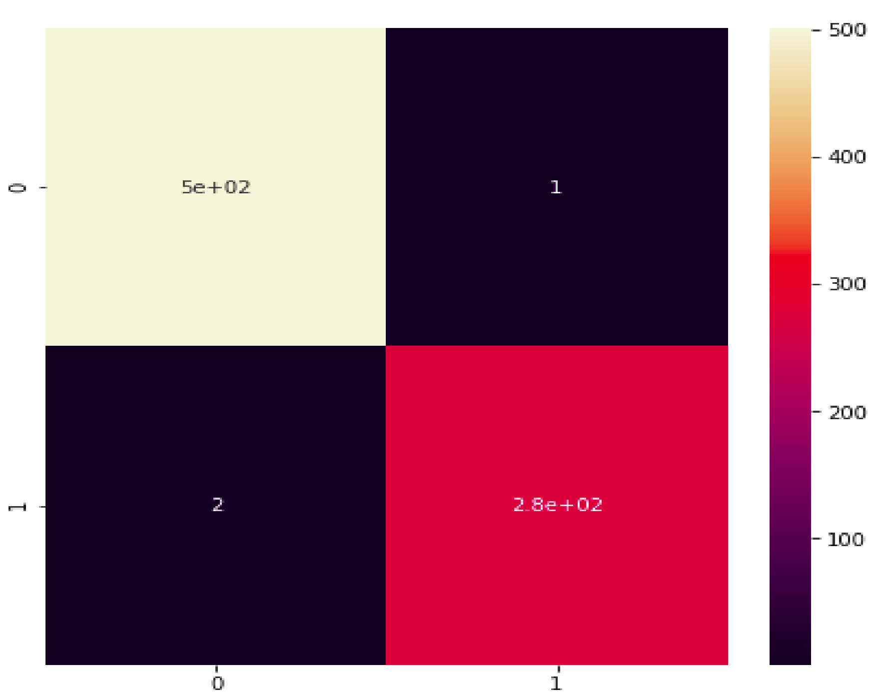
```
[30] # Obtain the accuracy of the model
     from sklearn.metrics import accuracy_score

     accuracy = accuracy_score(original, predict)
     print(f'The Accuracy score is: {accuracy*100:.2f}%')

     The Accuracy score is: 99.62%
```

Accuracy of the proposed model.

The accuracy of the model is mathematically calculated as;

$$Accuracy = \frac{Correct\ Predictions}{Total\ Predictions} \times 100\%$$

Appendix 3



Appendix 4

```
[32] from sklearn.metrics import classification_report

     report = classification_report(original, predict, labels = [0,1])
     print(report)
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       502
           1       1.00      0.99      0.99       278

   micro avg       1.00      1.00      1.00       780
   macro avg       1.00      1.00      1.00       780
weighted avg       1.00      1.00      1.00       780
```

Classification Report

Appendix 5

```
dense_6 (Dense)              (None, 1000)              4097000
_________________________________________________________________
batch_normalization_16 (Batc (None, 1000)              4000
_________________________________________________________________
activation_16 (Activation)   (None, 1000)              0
_________________________________________________________________
dropout_5 (Dropout)          (None, 1000)              0
_________________________________________________________________
dense_7 (Dense)              (None, 2)                 2002
_________________________________________________________________
batch_normalization_17 (Batc (None, 2)                 8
_________________________________________________________________
activation_17 (Activation)   (None, 2)                 0
=================================================================
Total params: 91,782,754
Trainable params: 91,761,614
Non-trainable params: 21,140
```

Model's Parameters

Signed Plagiarism Form

Student’s Name & surname: Bakary Badjie

Student’s Number: 20340009

Program: Masters in Electrical and Electronics Engineering

□ Master’s without Thesis X Master’s with Thesis □ Ph.D.

I hereby declare that I have fully cited and referenced all material that are not original to this work as required by these rules and conduct. I also declare that any violation of the academic rules and the ethical conduct concerned will be regarded as plagiarism and will lead to a disciplinary investigation which may result in expulsion from the university and which will also require other legal proceedings.

…………………………..

Signature