NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF ARTIFICIAL INTELLIGENCE ENGINEERING

MODIFIED DENSENET1.0 FOR BRAIN TUMOR CLASSIFICATION VIA MRI IMAGES IN THE IOT ERA

M.Sc. THESIS

Hadjer BENYAMINA

Nicosia
February, 2024

HADJER BENYAMINA

MODIFIED DENSENET1.0
FOR BRAIN TUMOR CLASSIFICATION
VIA MRI IMAGES IN THE IOT ERA

MASTER THESIS

FEBRUARY
2024

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF ARTIFICIAL INTELLIGENCE ENGINEERING

MODIFIED DENSENET1.0
FOR BRAIN TUMOR CLASSIFICATION
VIA MRI IMAGES IN THE IOT ERA

M.Sc. THESIS

Hadjer BENYAMINA

Supervisor

Prof. Dr. Fadi AL-TURJMAN

Nicosia

February, 2024

2

Approval

We certify that we have read the thesis submitted by Hadjer BENYAMINA titled **"Modified DenseNet1.0 for Brain Tumour Classification Via MRI Images in the IoT Era"** and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Applied Sciences.

Examining Committee	Name-Surname	Signature
Head of the Committee:	Prof. Dr. Fadi AL-TURJMAN	....
Committee Member:	Assist. Prof. Dr. Mubarak AUWALU SALEH	....
Committee Member:	Assist. Prof. Dr. Umar Ibrahim ABDULLAHI	....
Supervisor:	Prof. Dr. Fadi AL-TURJMAN	....

Approved by the Head of the Department

...../...../2024

Prof. Dr. Fadi AL-TURJMAN

Head of Department

Approved by the Institute of Graduate Studies

...../...../2024

Prof. Dr. Kemal Hüsnü Can Başer

Head of the Institute

Declaration

I hereby declare that all information, documents, analysis and results in this thesis have been collected and presented according to the academic rules and ethical guidelines of Institute of Graduate Studies, Near East University. I also declare that as required by these rules and conduct, I have fully cited and referenced information and data that are not original to this study.

Hadjer BENYAMINA

03/02/2024

Acknowledgments

In the hallowed halls of deep humility, I prostrate myself in profound gratitude unto the Almighty, ALLAH, whose divine guidance and presence have steered the course of this thesis' odyssey.

With an acute cognizance of bestowed blessings, I tender my heartfelt appreciation for the spiritual fortitude and enlightenment that have upheld me through the arduous pilgrimage of research and reflection. It is with a sense of profound reverence and acknowledgment that I discern the divine hand orchestrating this intellectual sojourn, rendering it possible through the grace of God. His wisdom has consistently illumined my path, and His benevolence has stood as a constant fount of inspiration.

May this opus stand as an enduring testament to the profound gratitude I hold for the divine providence that has graciously guided me through the labyrinthine tribulations and triumphant ascensions encountered in the noble pursuit of knowledge.

The deepest acknowledgments are extended to those luminaries, individuals, and institutions whose unfaltering support and sagacious counsel have been the compass guiding the fruition of this thesis.

Profound gratitude is extended to my supervisor and advisor, Professor Dr. Fadi Al-Turjman, whose mentorship, sage counsel, and unwavering encouragement have indelibly shaped the trajectory of this scholarly endeavour. His expertise and insights have not merely enriched the tapestry of this research but have also stood as an enduring source of academic inspiration. In contemplation of the human dimension of mentorship, I express sincere gratitude for the benevolence and patience extended during my academic pilgrimage.

In a spirit of sincere acknowledgment, I recognize Assist. Prof. Dr. Mubarak AUWALU SALEH for his sagacious advice and enlightened guidance, and for initiating me into the realm of Responsible AI. This serendipitous encounter has not merely altered my perspective towards AI but has strategically charted a significant deviation, a straight path, toward my future career. This unforeseen yet providential deviation is held in the highest regard.

Sincere acknowledgment is tendered to the members of my thesis committee, whose discerning feedback and scholarly contributions have substantially enriched the qualitative dimensions of this research.

Gratitude is particularly owed to Near East University for provisioning the indispensable resources and facilities that facilitated the execution of this project. The academic environment and research infrastructure at NEU have provided an environment conducive to intellectual growth and scholarly exploration. The first-hand knowledge acquired over more than two years will propel me profoundly in the research field and toward the realization of my career aspirations. A formal expression of gratitude is extended.

Special thanks are reserved for my dear friends and classmates whose collaborative spirit, insightful discussions, and shared dedication to the pursuit of knowledge have rendered this academic journey all the more rewarding especially “Ibrahim”, “Mahmoud”, and “Marcel”.

Gratitude is extended to my family: “Mother”, “Father”, “Sister”, and “Brother” for their continuous support, profound understanding, and unwavering encouragement. Their love and belief in my abilities have been the driving force behind this academic endeavour.

The gratitude of the highest order is extended to my beloved partner for his unlimited patience, infinite support, unconditional love, distinctive caring, and unending encouragement.

Lastly, I formally acknowledge the countless individuals, researchers, and authors whose seminal contributions laid the foundational stones for this work. Their luminary work has been a guiding beacon, shaping my understanding and providing a robust framework for the exploration undertaken in this thesis.

This opus would not have materialized without these individuals' collective support and contributions. Their enduring impact will resonate through the pages of this thesis and in my academic journey beyond.

I humbly offer my gratitude.

Hadjer BENYAMINA

Abstract

Modified DenseNet1.0 for Brain Tumour Classification Via MRI Images in the IoT Era

Benyamina, Hadjer

MSc, Department of Artificial Intelligence Engineering

February, 2024, 102 pages

Brain Tumour classification stands as a pivotal challenge within the realm of Computer-Aided Diagnosis. This thesis delves into the binary classification of brain images derived from varied angles in Magnetic Resonance Imaging scans of the human brain. The proposed classification model hinges on the paradigm of deep transfer learning, leveraging a pre-trained Modified DenseNet1.0 model and "Image Net" weights to autonomously extract features from the inputs of MRI brain images based on Densely Connected Convolutional Neural Networks. Our model was trained on a large dataset of two classes "Yes" and "No" taking into consideration the importance of hyper parameters tuning role and the usage of regularization techniques effectively. Therefore, our results suggest that our proposed model Modified DenseNet1.0 performs exceptionally well in accurately identifying instances of both classes, making it a robust solution for real-world tests and live usage through deploying the model in an accessible platform or IoT healthcare systems. Emphasizing the need for eXplainability, we build the decision-making processes of our AI system principle to display trust and transparency. To accomplish this objective, our extension Modified DenseNet2.0 will be associated with a comprehensive eXplainability of the anticipated outcomes through SHapely Additive explanations providing a visualized breakdown of the results.

Key Words: brain tumours, cnn, transfer learning, classification, explainability

Table of Contents

CHAPTER IV

List of Tables

List of Figures

List of Abbreviations

AI	Artificial Intelligence
MRI	Magnetic Resonance Imaging
NLP	Natural Language Processing
XAI	eXplainable Artificial Intelligence
CNN	Convolutional Neural Network
DenseNet	Densely Connected Convolutional Neural Network
SVM	Support Vector Machine
KNN	K-Nearest Neighbour
DL	Deep Learning
ML	Machine Learning
ELM	Extreme Learning Machine
SHAP	SHapely Additive eXplanations
DeepLIFT	Deep Learning Important FeaTures
LRP	Layer-wise Relevance Propagation
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
LIME	Local Interpretable Model-agnostic Explanations
PSI	Permutation Sample Importance
CAM	Class Activation Map
ECG	Electrocardiogram
ReLu	Rectified Linear Unit
AdaGrad	Adaptive Gradient Algorithm
RMSProp	Root Mean Square Propagation

CIFAR-10	Canadian Institute For Advanced Research – 10 Classes
CIFAR-100	Canadian Institute For Advanced Research – 100 Classes
SVHN	Street View House Numbers

CHAPTER I

Introduction

The meaning of the world and human existence was always a constant aim to look for, to feed our curiosity and broaden our horizons. This curiosity was and will always be the fuel of every mind looking for answers. Human curiosity led to finding out about fire passing by the invention of engines to make rockets reach what was unreachable one day.

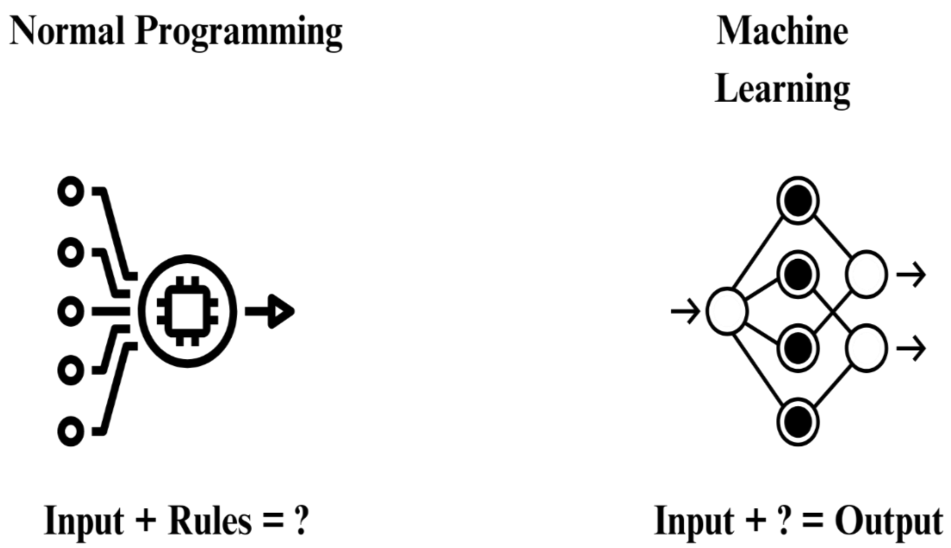
Computers played a big role in the process of every human history record, especially in the last decades. Since the invention of computers, they have been employed to achieve fast results and more accurate calculations in an exponential period compared to what humans could offer in the same given time. Throughout life, humans and scientists in particular were always in a race with time to uncover the mysteries of the universe and improve the quality of life for humanity's sake using knowledge, power, and computers.

The sky is not the limit. We have always believed that we can overcome all types of challenges if we employ science to our advantage. That is what computers and the Artificial Intelligence (AI) sector are found for.

One of the emerging areas of computer science is AI, which relies on processing a huge number of complex datasets using intelligent models. The pianist mathematician Marvin Minsky, one of the pioneers of AI, defines AI as the science of making machines do things that would require intelligence if done by men. On the other hand, Alan Turing proposed (Cooper & Leeuwen, 2012) that a computer can be considered intelligent if it can simulate human responses under specific conditions.

AI's main objective is to let machines handle tasks known as intelligence, in a more effective and faster way than humans through a learning process. The learning process reduces over time as more data is provided, and more systems are developed by adjusting old ones to get more accurate results. Based on biology, and how the brain learns through its environment “Surroundings”, scientists and especially neurologists thought that machines could be trained the same way to solve some kind of mathematical and statistical problems traditional methods could not solve yet. That marked the beginning of the new era of machine learning (ML). ML is a sub-field of AI that automatically detect patterns among data sets using detection and classification algorithms to extract recurring patterns to feed those algorithms through any type of data that could be stored digitally such as numbers, words, images, videos, statistics ... Etc.

Figure 1
Normal programming vs. machine learning.



In traditional programming, the output is determined by the input and predefined rules set by humans. In contrast, ML establishes rules autonomously based on input and desired output (See Figure 1). Supervised learning, unsupervised learning, and reinforcement learning are the three main machine learning paradigms, each is designed for a certain type of learning task. Supervised learning aims to discover relationships between input data, represented as dependent or independent parameters, by constructing a predictive model. The model is built from labelled training data, consisting of examples or instances. The learning process involves iteratively improving predictions through exposure to more labelled data, and enhancing accuracy with novel, unseen data (Maimon & Rokach, 2010). Noteworthy predictive supervised ML methods include Classification through Support Vector Machines, Decision Trees, Artificial Neural Networks, K-Nearest Neighbours, Prediction, and Estimation.

In unsupervised learning, classes are unlabelled, and the number of classes is unknown. This paradigm, often termed “Clustering”, involves identifying inherent structures within data (Han et al., 2006; Aggarwal, 2016). Real life applications of unsupervised learning are numerous and diverse. Companies specialized in transportation and logistics uses anomaly detection to identify obstacles or expose defective mechanical parts and condition monitoring (Del Campo Barraza, 2017). Anomaly detection can also be utilized in a medical setting to find abnormalities, lesions, or anomalies like brain tumours without having to specifically train the model for that particular disease. (Chatterjee et al., 2023), unsupervised anomaly detection in medical images that enables models to learn visual representations from unlabelled data (Iqbal et al., 2023). Unsupervised learning techniques are also applied in costumer and market segmentation to create costumer personas for better marketing and targeting campaigns (Paranavithana et al., 2021; Joppe, 2016), data preparations and visualization, recommender systems and engines such as Amazon's frequently bought together recommendations, Natural Language processing models for complex syntax learning, generating genuine grammatical sentences (Solan et al., 2005), and computer vision (Chen et al., 2022).

Reinforcement Learning (RL), a subtype of machine learning (ML), combines elements of supervised and unsupervised learning. In RL, an intelligent agent acts in the environment to maximize cumulative rewards. This process is essentially sequential decision-making (François-Lavet, 2018; Sutton & Barto, 2014). Various papers have proposed reinforcement learning-based trajectory optimization, motion planning, dynamic pathing, and scenario-based learning policies in highways for self-driving cars such as AWS DeepRacer. Google Data Centres also uses reinforcement learning agents to cool the system which helped reduce energy spending up to 40% (Gamble & Gao, 2018). NLP reinforcement learning is used for text summarization (Paulus et al., 2017), question answering (Choi et al., n.d.), and even machine translation (Grissom II et al., 2014). Reinforcement learning has a say in healthcare as well. RL can optimize treatment policies of chronicle disease and critical care based on previous experiences, it supports automated medical diagnosis, and other domains (Yu et al., 2020).

Self-learning category goes along with a neural network algorithm known as Crossbar Adaptive Array (CAA). It is learning with no supervisor or incentive mechanism. The CAA algorithm computes both actions and emotional decisions based on consequence situations (Al Sallab & Rashwan, 2012).

ML applications are various and diverse, they can be found almost everywhere and are a crucial part of our everyday life. Image recognition is the most common application of ML, it is applied in medical diagnosis, traffic predictions, and social media such as “Deep Face” by Facebook which is responsible for face recognition. Speech recognition is also one of the popular ML applications, it offers Google users the option of searching by voice and is the core of personal assistants such as Google Assistant, Alexa, Siri, and Cortana using speech recognition technology. Not to forget traffic prediction-assisted systems such as Google Maps. Moreover, ML techniques are used in recommendation systems in Amazon for example, self-driving cars also use unsupervised learning methods to detect objects and people while driving, Email spam and malware detection, online fraud detection, and automated translation such as Google neural machine translation that helps provide translations into our familiar languages. As we can see here, ML is everywhere and differs from one sector to another which means that there is a diversity in techniques as well.

Since our research is focused on supervised ML, we will only list supervised techniques in this section. Commonly, supervised ML algorithms are mainly classified into regression where algorithms provide predictions based on continuous data while classification predicts categorical values such as 1 or 0, cancer or no cancer. The outcome is always a class or category and not an actual value. Classification algorithms are also mainly classified into linear and non-linear models that are summarized in the following table (Table 1).

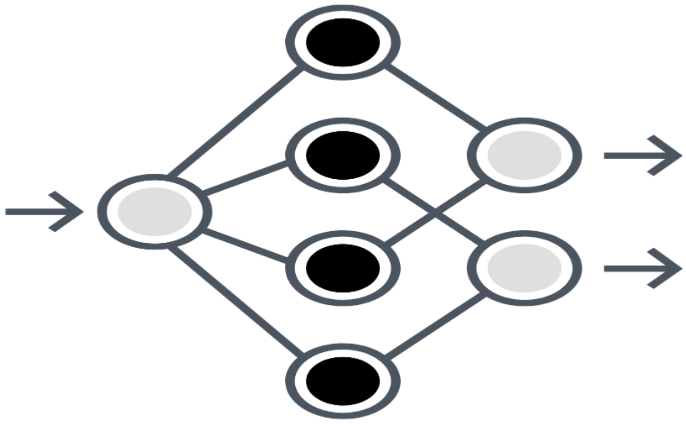
Table 1
Linear Models VS Non-Linear Models.

Linear Models	**Non-Linear Models**
Support Vector Machines (SVMs)	Decision Trees Classification Kernel SVM
Logistic Regression	K-Nearest Neighbours
	Naïve Bayes

ML models generally learn properly through processing structured data, small amounts of entries, and require low-end hardware and low cost whereas unstructured data and the huge amounts out there that need classification and more complex processing procedures require more advanced and high-performance algorithms than what we were introduced to so far. In other words, Deep Learning (DL) jumps to the scene with all its capabilities including the ability to learn from unstructured data, large datasets, and large-scale problem-solving. However, the most important characteristic of DL that makes it stand out compared to ML is that feature engineering is done implicitly and automatically without the need of a data scientist interference thanks to the core of DL algorithm that learns patterns on their own.

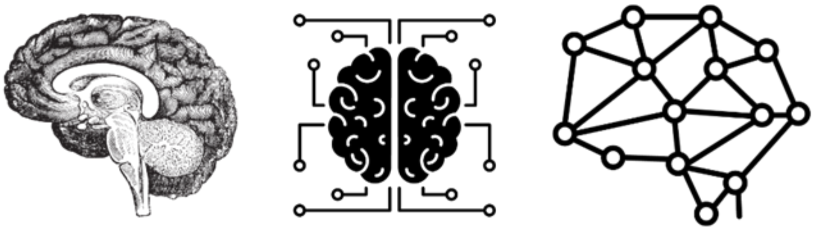
Based on scientific research concerning neurology, an approach in the field of computer science was built to simulate the bio neurons in a mathematical-based model called the Artificial Neural Networks (ANNs). ANNs are also known as Neural Networks, connectionist models or parallel distributed processing. The first ever interest was raised in 1943 after the neuroscientist Warren S. McCulloch and the logician Walter Pitts published their paper entitled: “A logical calculus of the ideas immanent in nervous activity” in the Bulletin of Mathematical Biophysics. In that paper, the two scientists tried to simplify the human brain analogy through a highly simplified mathematical model of neurons called “MCP Neuron” or “Threshold Logic Units” which makes the unit of an artificial neural network connected. That is exactly how the brain is made (Kröse & Van, 1996, McCulloch-Pitts Neurons, 2019).

Figure 2
Artificial Neural Network.



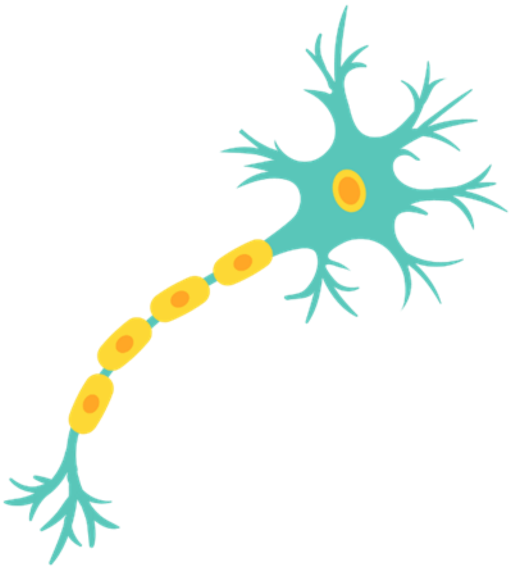
We can define a neural network as a computer program operating in an inspired manner of the natural human brain’s neural network (Haykin, 1999). The term “Neural Network” is more than a computational model. It is a paradigm of the human brain and mind.

Figure 3
Human Brain Analogy.



The computation of the human brain differs from that of a traditional digital computer due to its tremendous complexity, non-linearity, and parallel architecture. It is a huge neural network with billions of neurons, each one connected to thousands of other neurons (Nwadiugwu, 2020).

Figure 4
Natural Human Neuron.



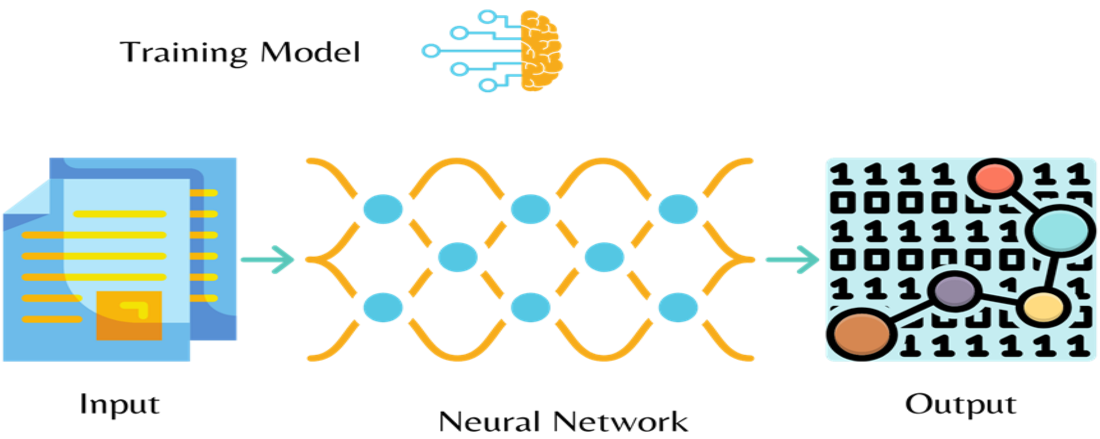
The human brain can organize its neurons and build up its own rules from birth. It can perform certain computations, such as speech recognition and image processing, and much more, many times faster than any powerful computer in existence for the time being (Haykin, 1999).

Scientists consider the human brain as the most powerful computer-like machine in existence. So theoretically, the McCulloch approach concludes that any operation that is made by the computer can also be made through a network of MCP (ANN). Simulating the brain functioning in artificial neural networks could encode many complicated computer programs and mathematical issues in many fields like biology and astrophysics. Even though the MCP model and ANN have not yet reached the real human brain functioning level, they contributed hugely to the development of computers throughout history from classical digital computers to non-classical computers that are playing a big role in the scientific research industry.

The objective of neural networks (Mehlig, 2021) is to perform cognitive functions such as problem-solving and ML. Research and development R&D data is growing exponentially, and in order to find hidden data or information that could be properties or entries for learning algorithms, advanced analytic techniques are needed. One of the many different tools available to meet the demands of R&D needs is the ANN. In contrast to the conventional regression method, the ANN can be used in modelling complex non-linear relationships proving excellent fault tolerance, fast and high scalability with parallel processing.

The architecture of the human brain serves as an inspiration for ANN learning algorithms. Just like the brain, neural networks are trained by learning patterns through trial and error. They are organized into layers which are made of several interconnected nodes containing an activation function. (Omondi & Rajapakse, 2006).

Figure 5
Simplified Neural Network Learning Algorithm Approach.



What features the neural network algorithms is that the knowledge (Output) is distributed throughout the network itself instead of explicit programming. The output then is modelled as connections between the artificial neurons and the adaptive weights of each of these connections. The network starts learning through various situations. The majority of Artificial Neural Networks (ANNs) techniques are supervised in nature. Nevertheless, depending on the desired result and the outputs, both paradigms can be used on a neural network. There are three major learning algorithms categories for ANN Learning. Gradient descent-based neural networks are used to find the local minimum of a function such as Back Propagation Neural Network and Radial Basis Function Neural Network. (Dkhichi & Oukarfi, 2014)

Consider Back Propagation Neural Networks, which begin with computing the partial derivative of the loss function while accounting for the last layer's parameters, which are ineffective with respect to any other network parameter. Due to the chain rule, this is not a complicated computation. (Dimililer, 2013, Qi et al., 2019)

Genetic algorithm-based neural networks are the most common evolutionary algorithm. To mimic the processes of natural selection, neural networks based on genetic algorithms select the most efficient rules for problem-solving (prediction) until a predictive model is obtained. Using genetic algorithms with ANN helps optimize the learning process and enhance the results of the process. It is applied to evolutionary neural networks, Bayesian neural networks, and others (Vié, 2020). Natural selection in biology serves as an inspiration for evolutionary neural networks. They are population-based and, like convolutional neural networks (CNNs), generate new candidates by recombining candidates from a population (Zhang et al., 2011).

There are two distinct aspects to the CNN architecture: the first is made up of several convolutional layers that do automatic extraction. “Feature Learning” is the term for this phase, which turns unstructured data into a collection of useful functions. Then the “Dense Layer” phase starts which is responsible for classification. The classifier modules contain fully connected feed-forward (recursive) neural networks. (Baldominos et al., 2019)

Teaching a machine to recognize images was a huge challenge for decades until research was able to finally simulate the human brain's visual recognition system. In 1963, Lawrence Roberts (Roberts, 1963) proposed the idea of using a 2D perspective to extract 3D geometrical information. This marked the first breakthrough of the field of computer vision. Several algorithms and models were suggested along this journey such as the generalized cylinder model, pictorial structure model, and CNNs or ConvNets.

ConvNets are not much different from classical neural networks. They are both structures on neurons based on weights learned from data. A CNN is composed of three main layers: the convolution layer, the pooling layer, and the fully connected layer which is the output layer.

Neurons in a CNN can be arranged in three dimensions: width, height, and depth. Using an activation function, each layer converts the three-dimensional input into a three-dimensional neuronal output. To summarize, CNNs are deep neural networks that share their parameters across space. Moreover, CNNs accept matrices as an input differently from multi-layer perceptron that only accept vectors which guarantee the preservation of the spatial structure of the image during the process. These characteristics and more are what allowed CNNs to be top candidates when addressing computer vision problems. In this chapter, we have discussed AI history along with applying ML techniques to AI models to achieve outstanding results. As you will see in the state of the art, we have mentioned several research papers where CNNs performance achieved considerable results in various fields and sectors mainly medical imaging where our research in particular has its full focus.

Background of Study

In a medical context, a tumour is caused by uncontrolled growth and abnormal spread of body cells that can result in death if not treated. Cells typically divide and die in an orderly manner so that new cells can take their place. A tumour, which is a mass or lump, might, nevertheless, continue to grow abnormally. The second most common cause of death in America, behind heart disease, is cancer, according to the Centres for Disease Control and Prevention. The American Cancer Society reported on 2022 an estimated number of 18.280 deaths versus 25.050 new cases of brain and other nervous system cancer.

In both adults and children, one of the most dangerous and deadly types of cancers are brain tumours. Moreover, they have the lowest survival rate. Consequently, early diagnosis' role in preventing more spread of the tumour and identification of the right treatment plan is highly important. Analysing Magnetic Resonance Imaging (MRI) scans of the brain with tumours is a useful method for differentiating between brain tumours, which can be challenging even for skilled neurosurgeons or oncologists. This means that even a small mistake in the diagnosis could have unfavourable effects. Notably, the number of brain tumour cases is increasing gradually which makes manual techniques tedious, time-consuming, and erroneous.

MRI is well-known for effectively finding some cancers like breast, brain, and spinal cord tumours. Using MRI, medical professionals can even tell sometimes if a tumour is or is not cancerous. Furthermore, it helps plan treatment such as surgeries or radiotherapy (American Cancer Society, 2021).

Brain tumour diagnosis begins usually with MRI scans. When the MRI shows that the brain has a tumour then professionals head towards further investigations for example performing a biopsy procedure or surgery depending on the tumour type and severity of the patient's condition (Cancer.Net, 2019). And here comes computer-aided systems that are always a handful for medical professionals to avoid errors, and time-consuming procedures and reduce the need for manual-based investigations. In other words, developing systems that can classify brain tumours is an important key for more accurate results, early diagnosis and even making treatment plans which is possible through CNNs that have proven interesting results in treating classification problems and the computer vision field globally.

Research Problem and focus

The rapid growth in the AI discipline and medical imaging processing has led to the need for accurate and reliable healthcare AI-based systems. In the field of ML, medical image classification tasks have proven a high level of importance in the healthcare sector where accuracy and precision of diagnosis is crucial but also the explainability and interpretability of the outcome. In many applications, consistency, credibility, and determining why ML model generates a particular prediction can be as crucial as the prediction's accuracy. Therefore, it should be one of the main keys along high accuracies when dealing with people's lives. Suppose the medical practitioner is unable to trust or reasonably explain the decision-making process, would it be possible for the patient to fully trust the physician and the AI model that is deciding their health state or determining their medical condition at the moment? The answer is an absolute NO. Therefore, improving the transparency of the model's effectiveness and reliability is highly recommended.

To introduce explainability in a computational context, we have to refer to linguistics and psychology first. According to the Oxford English Dictionary, the word “Explain” means describe, clarify, define, disclose, analyse, demonstrate, and any other word that leads to providing more information about a specific topic. Based on Lombrozo (Lombrozo, 2006) explanations are the means we have to exchange beliefs. Nevertheless, linguists claim that the explanation depends on the question (Bromberger et al., 1992)

However, it is difficult to provide a precise definition for explainability as it is interchangeably used in the meaning of interpretability. On the other hand, we can define the properties that explainability should cover. Explainability can be evaluated mainly based on two concepts: Interpretability and completeness. Interpretability explains the midst of the trained model in a way that humans can comprehend. Whereas completeness aims to capture in detail the whole process accurately (Gilpin et al., 2018). In summary, a well-interpreted model is easy to understand by humans but if the model is too complex, it requires a complete explanation that might be too hard for humans to understand. So far, gathering interpretability and completeness at the same time is very difficult because interpretable explanations don't reach a good level of completeness and complete explanations are most of the time not very understandable (Kanerva, 2019). Human understanding of the results produced by ML models becomes increasingly challenging as it becomes more complex. The history of philosophy of science indicates that debates have arisen about what constitutes an explanation. Causation is linked to explanation. As Aristotle's theory of causation contends, determining the cause is necessary to explain an occurrence or a phenomenon. Nevertheless, the theory of explanation treated explanation in either an epistemic or a realist sense following the development of the philosophy of science in the 20th century. The entities or processes that an explanation postulates are actual, according to a realist interpretation of explanation. The explanation is an accurate portrayal of the outside world. According to an epistemic interpretation, these entities or processes are merely helpful for arranging the data from scientific tests and human experience, rather than necessarily existing in a literal sense. Explanation, here, is only to facilitate the construction of a consistent empirical model, not to furnish a literal description of reality (Randolph Mayes, 2023). That is, in general, the outcome of scientific philosophers trying to comprehend the nature of modern theoretical science. According to a philosophical approach, the reasoning process is classified based on two opposite approaches: Inductive and deductive reasoning.

Deductive reasoning based on Aristotle's syllogism (Kulicki, 2020) is a method where you start with a series of premises that would lead to a certain conclusion. While the father of empiricism, Francis Bacon who established and popularized the scientific method of research into natural phenomenon, introduced inductive reasoning approach by evaluating and improving theories through measurement, experimentation, and observation (Bacon, 1620). Scientific reasoning theory was generally defined as problem-solving process that involves a critical thinking process when conducting scientific research about procedural, content, and epistemic knowledge (Simon & Newell, 1971, Kind, 2013). It is studied through different sectors such as cognitive sciences, education, and to identify a suitable reasoning mechanism in a scientific context such as deductive reasoning, inductive reasoning, problem-solving, and causal reasoning (Dunbar and Fugelsang, 2005).

Moreover, as mentioned in the article published by (Miller, 2019), the process of explanation is composed of two parts: Cognitive process where the explanation of a given event is determined and that is what we call the explanandum according (Hempel, C. G., & Oppenheim, P., 1948). Group interaction is represented by the second part, which is the social process of knowledge transfer between the explainer and the explainee (Kim et al., 2021). In other words, Causality is known as the highest degree of an interpretation process of a natural phenomenon where we try to understand logically and scientifically how and why it happened. In other words, cause and effect science. Back to Aristotle again, the easiest way to represent causality is that more than one cause may be a non-coincidental cause of something. Based on this approach, the eXplainability of Artificial Intelligence (XAI) models has seen the light and witnessed the present attempts to improve previous methodologies and a new set of algorithms that serve to explain and interpret a ML model's outcome in a way that is valid, faithful, trustworthy, and transparent.

To propose an explainer algorithm, we need to consider three aspects: the presentation of the explanation, the type of the targeting model that the algorithm can be used for, and the relations between the explainer and the DL model.

Research Aim and Objectives

The task of imparting image recognition capabilities to machines has historically been a substantial challenge, persisting for several decades. Recent strides in research have, however, achieved a commendable simulation of the complex visual recognition system inspired by the human brain. Presently, within the realm of AI, and particularly in the context of Responsible AI, a strong focus is on clarifying the models' decision-making procedures. This highlights the importance of explainability, making it a necessary component for the smooth integration of DL models in the medical field. Given that these models play essential roles in decision-aid systems, this integration is extremely important. This research aims to efficiently classify and detect brain Tumours by utilizing a supervised DenseNet-based modified DL model "Modified DenseNet1.0" with automatic features extraction to detect the brain Tumour provided with an explanation integrated in our future extension "Modified DenseNet2.0" based on the usage of SHapely Additive eXplanations (SHAP) and visualized graphs supporting the predictions of the model. The following research objectives would facilitate the accomplishment of this aim:

- Proposed a binary classification deep transfer learning model using DenseNets named: "Modified DenseNet1.0".
- Maintaining high accuracy of the results to promote safety, reliability, ensure robustness and the success of decision-making processes.
- Propose an extension "Modified DenseNet2.0" that will efficiently provide explained prediction accompanied by visualized graphs using SHAP values for explainability.

Research Framework

The core of this research is to efficiently classify two datasets of two different sizes carrying brain tumour MRI images, the datasets are open source and available on the Kaggle platform for research purposes. The research attempts to classify the MRI images into two classes through a binary classification model "Modified DenseNet1.0" to decide whether the MRI image bears a tumorous brain or a non-tumorous one. The research will be based on using transfer learning techniques, normalization and regularization techniques to classify and predict the state of a brain MRI image and then provide a visualized explanation for the prediction in a future extension "Modified DenseNet2.0" that complies with international standards.

Thesis Organization

The thesis report is divided into five chapters, and the research objectives are thoroughly elaborated in the following chapters which can be read independently from each other. The chapters are organized as follows:

Chapter 1: In this chapter, we are presenting the background of the study including Artificial Intelligence (AI), Machine Learning (ML) paradigms, neural networks, classification techniques, explainability philosophically and scientifically, eXplainable Artificial Intelligence (XAI), and Responsible AI. The chapter also discusses the research problem and focus, aims and objectives, and scope of the study.

Chapter 2: In this chapter, we provide a review of Deep Learning (DL), Convolutional Neural Networks (CNNs), transfer learning, medical imaging, brain tumours, and MRI imaging. It also provides a thorough and concise literature review related to convolutional networks, other classifiers, and explainability.

Chapter 3: Here we describe the methodology and approaches employed to build our proposed model, portray the datasets utilized, introduce the explainability techniques, and discuss the evaluation metrics and study tools.

Chapter 4: This chapter demonstrates the results and findings derived from our research, encompassing quantitative analysis and visualization of the outcome, presenting the explainability of the model's predictions, and related discussion. Additionally, it includes a comparative study with previous works and similar approaches.

Chapter 5: The thesis conclusion and suggestions for further research are provided in this chapter.

In the coming section, we have selected some interesting works that have processed related topics focused on supervised learning using CNNs supported by other techniques to classify different datasets of various sizes and nature including brain tumours. And explain predictions using different approaches such as LIME, Grad-CAM, and SHAP Values.

CHAPTER II

Literature Review

Several machine-learning techniques are used to classify brain tumour datasets using Support Vector Machines (SVMs) (Ozsoz et al., 2021), K-Nearest Neighbour Algorithms (KNNs) (Altman, 1992), and Convolutional Neural Networks (CNNs) (Akella et al., 2023). Our research focuses on using CNNs as the state of art proving the high efficacy of using CNNs in computer vision generally (Mubarak et al., 2022) and medical imaging specifically (Salehi et al., 2023).

Theoretical Framework

The winners of ImageNet Challenge 2014 (Simonyan & Zisserman, 2014) submitted a CNN model known as VGGNet16 and VGGNet19 that secured them first place in the localization track and second place in the classification one. Moreover, the model generalizes well on other datasets and even a wide range of tasks. Such achievement marked a significant milestone in computer vision with the proposed VGGNet16 and VGGNet19 CNN models. The VGGNet architecture showcased its effectiveness in handling diverse challenges within the competition. This versatility highlighted the robust nature of the VGGNet models, making them valuable tools not only for image classification and localization but also for various applications beyond the initial challenge.

A year later, (He et al., 2015) introduced an innovative approach to DL with their CNN architecture known as Deep Residual Learning. The proposed model, particularly the ResNet50 and ResNet101 variants, achieved remarkable success by securing first place in both the image classification and localization tracks. Notably, the introduction of residual learning, which involves the use of residual blocks, allowed for the creation of significantly deeper networks without suffering from vanishing gradient problems. This breakthrough design facilitated the training of extremely deep neural networks, leading to improved performance and generalization across various datasets and tasks.

In 2016, a ground-breaking contribution to Deep Learning (DL) was presented by (Huang et al., 2016) that addresses the challenges and limitations of conventional convolutional networks by introducing the Dense Convolutional Network (DenseNet). DenseNet's fundamental innovation is its dense connectivity pattern, which effectively addresses problems associated with the vanishing-gradient problem, improves feature flow through the network, promotes feature reuse, and drastically lowers the number of parameters when compared with conventional architectures. This resulted in DenseNet demonstrating superior performance across a range of tasks and establishing its significance in advancing the state-of-the-art in DL.

The focus of this thesis goes to the depth of the CNN architecture that leads to better accuracy compared to previously published research in the same field.

Brain Tumours Detection Using AI

A CNN classifier for brain MRI images was proposed by (Pathak et al., 2019) to indicate any presence of tumour in addition to Tumour area calculation using the watershed algorithm for segmentation operation. The proposed classifier achieved a very good accuracy with low complexity. Another paper using DL to segment Tumour areas from brain MRI images was proposed by (Sajjad et al., 2019) after applying data augmentation to treat the lack of data issue while training the model based on a novel CNN system for multi-grad brain Tumour classification task. The proposed system was evaluated on both raw data and augmented data achieving convincing performance compared to existing methods. (Ayadi et al., 2021) suggested a multi-layer CNN to perform brain tumour MRI image classification. The suggested model evaluated on three datasets attained satisfying performance. The proposed method in (Pashaei et al., 2018) consists of extracting hidden features from brain Tumour MRI images using CNNs then a Kernel ELM algorithm performs a feature-based classification task which has shown promising results compared to other classifiers such as support vector machines and radial base function. (Abiwinanda et al., 2018) implementation of a simple CNN architecture that recognizes the three most common brain tumour types trained on a publicly available dataset. The architecture has no prior region-based segmentation, yet it achieved higher accuracies compared to the one that applied segmentation algorithms. (Naseer et al., 2021) focuses on early diagnosis possibility through a convolutional network trained on a benchmark dataset containing brain tumours MRI images. The proposed system outperforms other suggested systems by achieving a very satisfying average accuracy and almost perfect specificity while hitting 100% correct diagnosis for two datasets that were evaluated. (Alanazi et al., 2022) built a novel transfer deep-learning model to classify various types of brain tumours by re-using a 22-layer binary classifier to adapt the neuron weights to their model through a transfer learning technique.

To approach the computational time issue lying within CNNs approaches, (Bacanin et al., 2021) suggested a metaheuristic method based on an improved firefly algorithm (Jati et al., 2013) developing an automatic system for brain tumour multi-class classification that has outperformed similar approaches with considerable classification accuracy.

For a more structured study, we have classified the works mentioned below according to the techniques used, the learning paradigm which is exclusively supervised in this study since we are treating classification problems using CNNs. Table 2 represents the classification of the related works.

Table 2
Related works classification.

ML Types	Supervised Machine Learning						
Approach	Convolutional Neural Networks						
Techni-que	Optimization	Segmentation	Data Augmentation	Kernel methods	Transfer Learning	Traditional	XAI
Simonyan & Zisserman, 2014		•					
Pashaei et al., 2018				•			
Abiwinanda et al., 2018						•	
Pathak et al., 2019		•					

Sajjad et al., 2019		•					
Bacanin et al., 2021	•						
Ayadi et al., 2021						•	
Naseer et al., 2021						•	
Alanazi et al., 2022					•		
Our Proposed Model	•				•		•

Employing XAI in ML for predictions

We believe that the explainability step, where the trustworthiness of the final prediction must be considered seriously, is absent from the state of the art when it comes to employing CNNs to classify MRI brain pictures. Rather than that, most research focuses on achieving high training and testing accuracy while neglecting the opinion and judgment of the concerned individuals on this accuracy itself and if it is trustworthy to decide a patient's diagnosis or even to be part of medical-aided systems and decision-making process. Therefore, we expanded our research and studied several practical XAI methods that achieved significant progress in the medical field. In particular, classification tasks run through DL algorithms such as CNNs as mentioned in the research paper of (Tang et al., 2019) where a DL pipeline concept is proven to be a good identification of specific neuro-pathologies using automated segmentation and CNNs for training and evaluation of the classification task conducted in the research achieving good correlation scores based on prediction confidence maps that provide high-resolution morphology distributions. Another study of brain aging-related impairment was done by (McKenzie et al., 2022) using a DL-trained model that predicts cognitive impairment existence with modest accuracy. Furthermore, attention-based interpretation studies of the features were conducted to find out how the model identified the myelin pallor which demonstrates a scalable platform that helps in locating unanticipated pathological elements in neurobiological disorders and cognitive impairment in general. Research by (Zhang et al., 2018) proposes a modified conventional ConvNet into an interpretable CNN to explain knowledge representations in high convolutional layers without annotations. The latter is thus assigned during the training process. This method applies to various CNNs regardless of their structures. (Tomita et al., 2019) designed a novel approach to train a dataset of images taken from the endoscopic oesophagus and gastro-oesophageal junction mucosal biopsies. The model was trained based on a CNN approach and grid-based attention network and then evaluated on a different testing set of four classes achieving a better performance compared to the existing state-of-the-art taking into consideration that the detection part was based only on tissue-level annotations.

An alternative framework proposed by (Apicella et al., 2020) where researchers applied sparse dictionary methods to classify inputs using middle-level properties to build blocks for the image classification explanation without relying on the input’s low-level features to explain the ML model. Another low-level feature and annotation's role in prediction tasks, (Hu et al., 2021) still believe that the latter can be improved to enhance CNN performance by proposing a simplistic and wide attention convolutional approach to discover each label's local characteristics and low-level properties. A clinical open-access dataset of intensive care unit (ICU) medical records was used to evaluate the approach, and it performed noticeably better than the state-of-the-art. (Sattarzadeh et al., 2021) propose an algorithm to visualize various layers. In order to achieve total explainability, the CNN uses an attribution-based input sampling approach and aggregation along with an empirical analysis of low-level features efficacy to enhance the mentioned attributions.

In order to simultaneously achieve high accuracy values and explainability criteria, (Duffy et al., 2022) propose an interpretable approach that respects the standard clinical flow in a DL prediction based on a cardiac junction assessment and improved segmentation method using the frame-by-frame 3D depth-map technique. According to the researchers, this method is easily interpreted and the diagnosis can even be corrected by clinicians at any point.

Another design interpretable CNN by (Ismail Fawaz et al., 2019) to classify surgical skills using pattern extraction technique to classify surgical skills resulting from trainees' motion while performing robotic surgery. Using the Class Activation Map (CAM) technique, the model minimized the black box effect and was evaluated on the JIGSAWS dataset. This allows surgeons to receive very accurate and objective skill assessments and helps them identify the surgical procedures that had an impact on the skill assessment. Moreover, a novel personalized feedback technique.

Another ML approach is proposed by (Oviedo et al., 2019) to conduct a prediction task on crystallographic space group and dimensionality from a given set of thin-film X-ray diffraction data patterns. This research study uses simulated data from the Inorganic Crystal Structure Database (ICSD) and experimental data to combine a model-agnostic and physics-informed data augmentation strategy with a supervised machine learning approach. The testing phase included 115 samples with 3 dimensionalities and 7 classes. The algorithm achieved excellent accuracy. Furthermore, a high level of model interpretability is ensured by the CAM approach on an average pooling layer, which addresses the misclassification problem logically. To address the understanding part of model performance, (Ancona et al., 2018) propose a novel evaluation metric and perturbation-based attribution method to test the gradient attribution methods on different datasets of texts and images through classification methods.

Back to medical fields and computational pathologies, this paper (Tosun et al., 2020) presents enabled applications in anatomic pathology workflow based on XAI mechanisms by recognizing the regions of interest providing an efficient and accurate diagnostic that promotes safety and reliability.

(Spinner et al., 2020) researchers are presenting a new test and evaluation methodology of XAI methods using a time series that demonstrates the efficacy of SHAP techniques and robustness for all models where DeepLIFT, LRP, and saliency Maps are selective. In the field of cyber-physical systems, XAI is used to explain the performance of the system and the decision-making process. However, those methods lack applicability according to (Jha, 2022). As a result, the author proposes semantic technologies using contextual information, user feedback, and visualization techniques based on knowledge graphs, for further understandability.

In the energy utility field, ML technologies are used highly to improve the sales support systems' performance. Yet, customer relationship and cross-selling are key points that were not augmented enough compared to recommender systems and targeting for example.

In that context, (Haag et al., 2022) consider the SHAP values the most relevant feature attribution method to generate more meaningful explanations for a cross-selling task. (Iliadou et al., 2022) propose a framework based on RNN and attention-based LSTM that provides comprehensive profiles for patients with hearing issues using various factors such as personal behaviour, environmental, and other factors that significantly improve patients' satisfaction and their quality of life as well as decrease the number of counselling sessions with the audiologist.

In this master's dissertation on clinical decision support systems predictions, the author (Lourenço das Neves, 2020) proposes an explicative pipeline based on two architectures: interpretable KNN and black box classifier CNN explained by PSI, LIME, and SHAP techniques on a sample-based approach. Then the obtained model using performance decrease and Jaccard index resulted in promising findings.

Another CAM-based approach was presented by (Zhang et al., 2022) by training a model on multi-input and multi-task to retrieve explanation faithfulness which has improved the user task performance and reached helpfulness as well, which can support a robust platform for applications with data biases.

In (Spinner et al., 2020), The interactive framework incorporates steering and monitoring features together with an interactive XAI pipeline under a visualizer that uses explainable ML instantiated within the TensorBoard (TensorBoard | TensorFlow, 2019) toolkit environment. The proposed framework provides a well-informed machine-learning process that could open room for future integrated systems and extensions.

DL models are not the only available options that can be used to train data and provide explainability, but tree-based ML techniques also provide interpretability (Pedretti et al., 2021), the proposed analog CAM aims to accelerate the model's inference.

To improve the medical decision-making process which is already difficult to explain or interpret, (Eder et al., 2022) designed an algorithm that uses

MRI scans to predict the survival rate of individuals with brain tumours and implemented it on CNN architectures and SHAP technique to improve the explainability regardless of the accuracy.

(Wäldchen et al., 2022) proposed an XAI method based on ANN architecture and SHAP technique as attribution methods help evaluate characteristic functions and measure each component's importance according to the function's output. The ECG waveforms have been successfully diagnosed and interpreted by the designed CNN-based XAI model, which uses gradient weighted class activation mapping (Grad-CAM) approach and electrocardiogram monitoring data (Taniguchi et al., 2021). Another interactive web application for facial analysis, based on the Grad-CAM approach and XAI, is designed to facilitate interaction with application users and provide clarification on the ML process and findings.

(Lu et al., 2021) presented an innovative crowdsourcing technique for human-based evaluation inspired by the "peek-a-boom" game which provides quantitative performance measures for automated evaluation schemes. (Klein et al., 2022) proposed a formalization of interpretable ML along a sophisticated statistical process that renders an explained correlation as a medium towards interpreted models or post-training explanatory methods which reduce the gap between the correlation and the causation. XAI was used to mark, along with a visual interpretation, the association between each class of biomarkers and the likelihood of developing breast cancer. The findings helped in the development of precise treatment plans based on the Body Mass Index (BMI) and biomarker levels of the patient. (Idrees & Sohail, 2022)

XAI was used to mark, along with a visual interpretation, the association between each class of biomarkers and the likelihood of developing breast cancer. The findings aid in the development of precise treatment plans based on the Body Mass Index (BMI) and biomarker levels of the patient.

The extent to which the suggested qualities are, in fact, the most important factors that have contributed to a particular decision and how this may be measured with reliability are the two main areas of research that need more attention.

To address the matter effectively, we've concluded that the explainability and interpretability of DL models are still in their early steps and need to have the light spot on to satisfy the need for answering the questions: "Why and How" through a CNN-based explainable model using shapely values focused on features importance aspect which will be our target in our extension “Modified DenseNet2.0”.

CHAPTER III
Methodology

This section explains the practical aspects of the research including dataset sources, image pre-processing techniques, and the Convolutional Neural Network (CNN) architecture used to train the most stable and reliable generated model “Modified DenseNet1.0” based on pre-trained DenseNet and evaluated by comparison to other baselines models such as VGGNet16, VGGNet19, and ResNet50. Also, it includes insights about our research approach and an evaluation of the model's performance for more credibility of the research. Finally, a sneak-peek introduction of the explainability extension of the output prediction generated by our upcoming “Modified DenseNet2.0”.

Data Selection

Both datasets used in the experiments below are downloaded from the Kaggle website. We have a small size dataset that is imbalanced as shown in the table below (Table 3), which will be named Dataset I (Figure 6), and a large size dataset that is balanced with an equal number of samples for both classes and will be named as Dataset II (Figure 7). Both datasets are composed of two classes “Yes” and “No” where class “Yes” represents the class with Tumorous brain MRI images and class “No” represents non-tumorous brain MRI images.

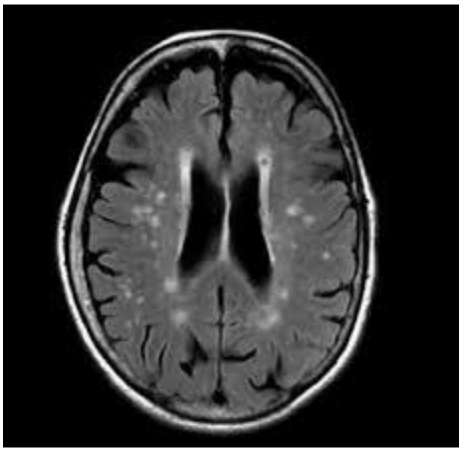
Table 3
Datasets content count.

Class	No	Yes
Dataset I	98	155
Dataset II	1500	1500

In the first part of our methodology, we will use both datasets to compare our model's performance across the two datasets during the training, validation, and testing phases. Then the most suitable dataset will be selected for the model evaluation compared to other baselines architectures.

Figure 6

A sample of an MRI image from class no and class yes respectively, from the imbalanced dataset I.



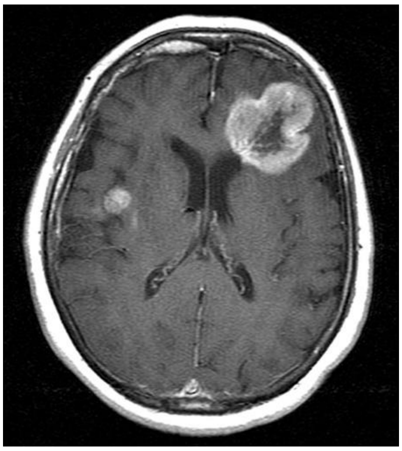
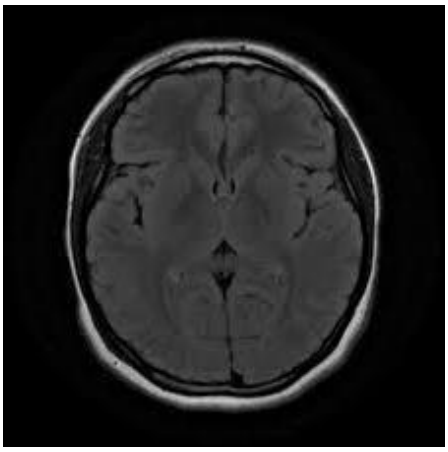
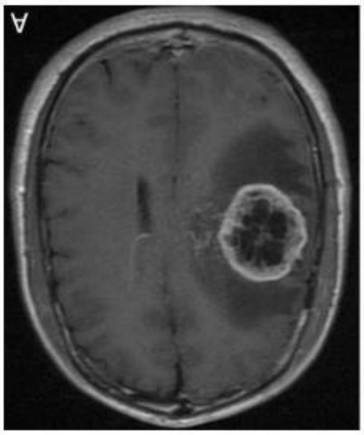


Figure 7

A sample of an MRI image from class no and class yes respectively, from the balanced dataset II.





Research Design

The meticulous planning and preparation of the data and model architecture is critical to the performance of Deep Learning (DL) models. In this research, our research design passes by critical steps to optimize the model's performance, address data limitations, and ensure robustness. The key components of our research design include image pre-processing, data augmentation, choice of activation functions, regularization techniques, loss function selection, optimizer usage, and the adoption of transfer learning with DenseNets.

Image Pre-processing

We conducted feature scaling on both dataset I and dataset II, where we used image normalization and standardization, in order to enhance the performance of the training phase thereafter and satisfy our model's design requirements for input data.

For image standardization, we rescaled the brain images to have the same heights and widths as required by our model's architecture which is 224x224.

For image normalization, most image pixel values range between 0 and 255. However, inputs with huge values could disrupt or slow down the learning process which prompts us to divide all image values by the largest pixel value which is 255.

Data Augmentation

We used the data augmentation technique only with dataset I because of a lack of data. Data augmentation is a transformation method used to create several copies of the same image by applying changes to different parameters such as random shifting, flipping, rotating, and brightness changes.

Activation Functions

An activation function in a neural network defines the sum of the input weights and converts it into an output from one or more nodes in an output layer, also known as a hidden layer. ReLu is the same activation function that is typically used by hidden layers. Additionally, the output layer has a distinct activation function, in our case, it's the Softmax function.

ReLu (Rectified Linear Unit) (1) is a linear function that, in the event that the input is positive, produces the value directly. Otherwise, it outputs 0. It falls in the range of 0 to +∞.

$$\sigma\ (x) = \begin{cases} 0 & ,\ x < 0 \\ \max\ (0, x), & x \geq 0 \end{cases} \tag{1}$$

Softmax (Softargmax) (2) is a multidimensional generalization of the logistic function that takes the form of a normalized exponential function. It has a range of 0 to 1.

$$\sigma\ (\vec{z})\ \mathrm{i} = \frac{e^{z_j}}{\sum_{j=1}^{k} e^{z_i}} \tag{2}$$

Regularization

In neural networks, regularization techniques are used to prevent overfitting. As a result, it contributes to raising the trained model's accuracy so that it can make more consistent, trustworthy, and objective predictions when processing unseen data.

During all our experiments on both datasets I and II, we used the regularization technique: Dropout, which is a technique that drops out or deactivates a precise probability of random neurons during the training phase which helps simplify the network and reduce its complexity which works on reducing the overfitting. We stress the importance of regularization as one of the strength points of our proposed model.

Loss Function

In neural networks, the loss function is a measurement that evaluates how good the model at predicting an expected output is. The loss function chosen for all our experiments is the binary cross entropy (3) as we are treating a binary classification matter. Binary Cross Entropy finds a difference of either 0 or 1 between each anticipated probability and the actual class output. Next, depending on the difference between the actual and expected values, it computes the score that penalizes the probability.

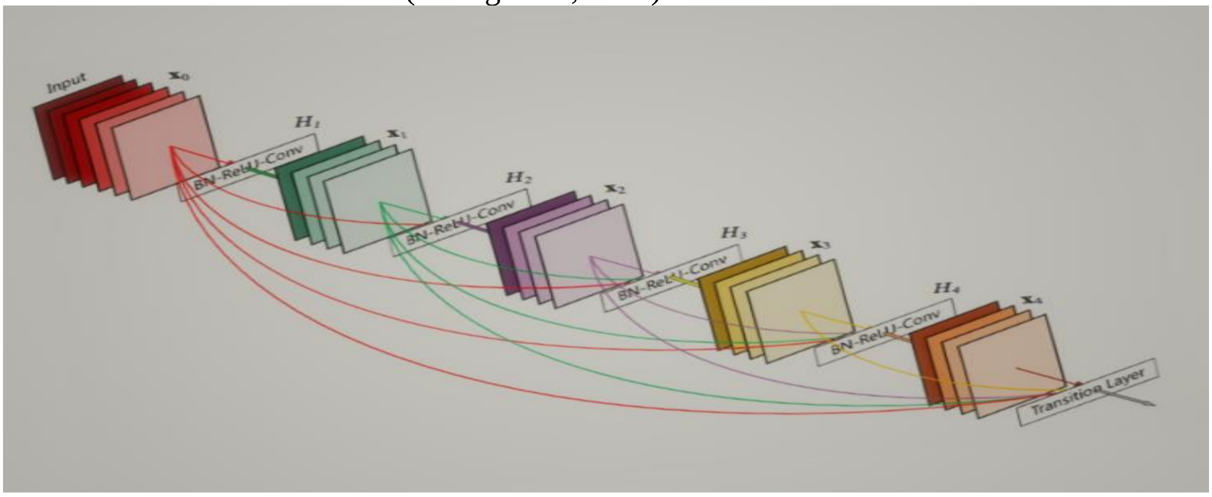
$$Loss(y,z) = \begin{cases} z - zy + \log(1 + e^{-z}), & z \geq 0 \\ -zy + \log(e^{z} + 1) & , z < 0 \end{cases} \tag{3}$$

During the model training phase, we employ optimizers to minimize the loss function and update the weights of the neurons for each iteration. The neural network's weights and learning rate are two examples of the parameters that these algorithms or functions change in order to lower loss and increase accuracy. In our case, Adam Optimizer is utilized. It is a hybrid of the Root Mean Square Propagation (RMSProp) and Adaptive Gradient Algorithm (AdaGrad) gradient descent techniques. To specifically enhance the performance of computer vision issues, both approaches retain a per-parameter learning rate. Adam optimizers are largely famous as the best handlers of medical images (Hassan et al., 2022).

Transfer Learning

In order to avoid having to train the model from start, we will now employ the transfer learning technique. Within the ML field, transfer learning constitutes a whole research problem. First, a source model that is comparable to the target domain must be chosen, and it must then be modified to fit the target model. To put it another way, it's leveraging a previously trained model as a springboard to help our model complete a new task. We can achieve faster results and better performance with this method than if we trained our model from start. In order to attain good performance, raw modelling is computationally expensive and needs a large amount of data. As a result, it is preferable to use pre-trained models such as ImageNet, AlexNet, VGGNet, ResNet, DenseNet, and Inception that have the basis of transfer learning and know how to perform classification tasks. Moreover, they're computationally efficient and provide better results even when using small datasets which will be proven in the following experiments. In our research, we adopted Densely Connected Convolutional Networks (DenseNets) for their connectivity feature as mentioned in the state of the art (See Figure 8).

Figure 8
DenseNets architecture (Huang et al., 2016).

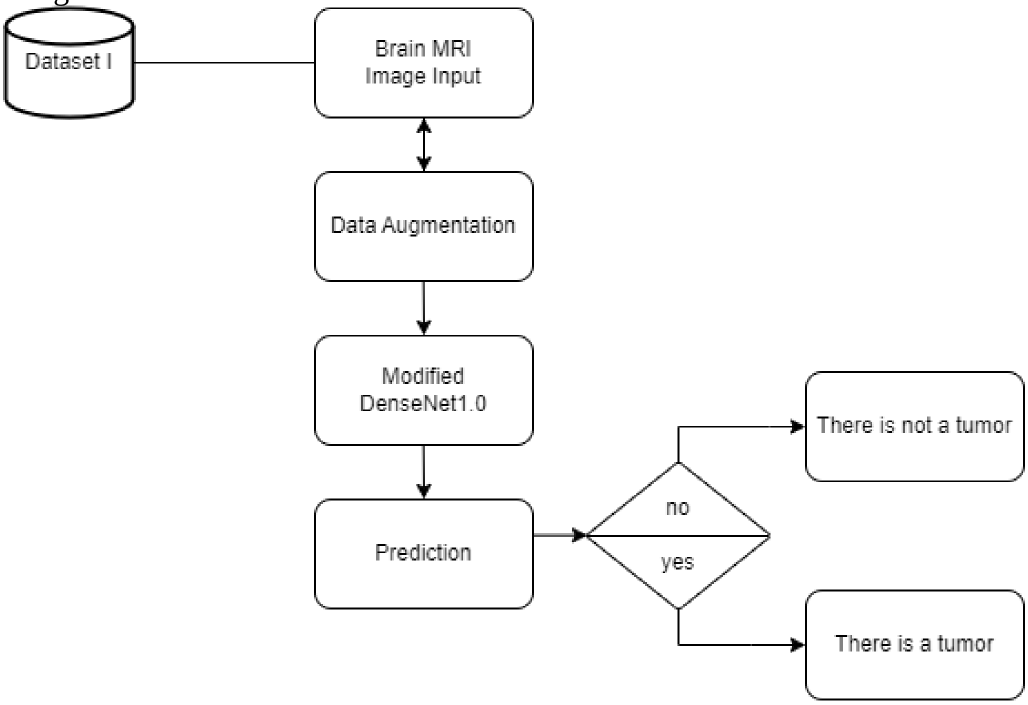


DenseNets are regarded as one of the best vision model architectures available today. It is frequently used to carry out tasks involving detection and classification using transfer learning techniques. But when it comes to applying these kinds of models to determine the state of human health, accuracy and credibility is crucial especially when processing large datasets that can't be checked manually. Moreover, we recommend an additional extension to be integrated in our model "Modified DenseNet1.0" through shapely values to provide the outcome prediction a plausible explanation "Modified DenseNet2.0".

During our experiments, we have used as mentioned previously two different datasets I and II to train our proposed model using transfer learning technique based on the DenseNet pre-trained model.

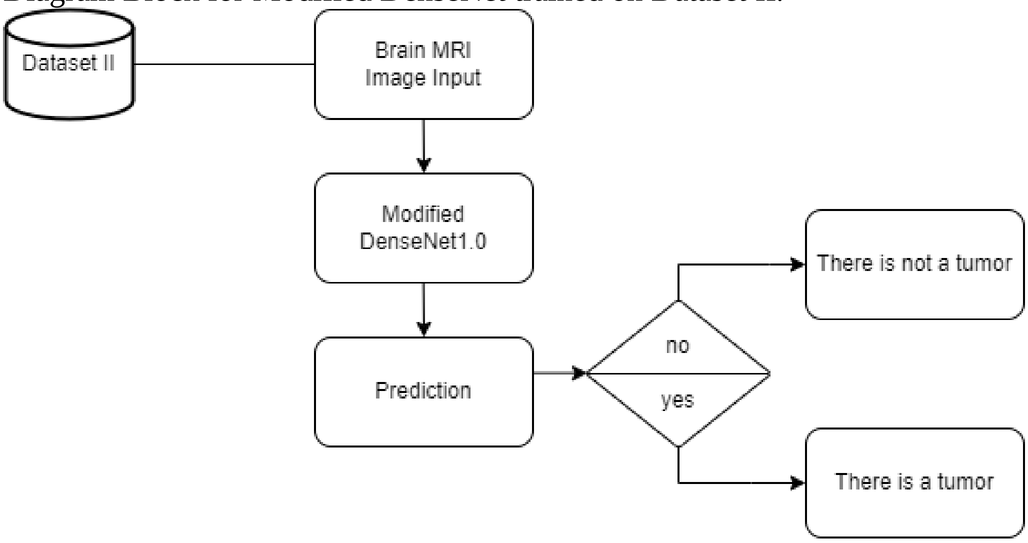
The following flowchart (Figure 9) represents the diagram block of our DenseNet-based model trained on dataset I which is a small dataset of 2 classes (yes and no). The data augmentation strategy is essential since it improves our model's performance despite the scarcity of data.

Figure 9
Diagram Block for Modified DenseNet1.0 trained on Dataset I.



Our second flowchart (Figure 10) represents the diagram block of a DenseNet-based model trained on dataset II which is a larger dataset of 2 classes (yes and no) as well.

Figure 10
Diagram Block for Modified DenseNet trained on Dataset II.



Proposed Modified DenseNet1.0 Classifier Architecture

To build our proposed model, we utilized the DenseNet architecture without training the network from the beginning. We also excluded the output layers which allow us to add our layers. Then we added two hidden layers and one output layer. The first added hidden layer contains an average pooling function of 4 x 4 size, a flattening function, a dense layer with 64 neurons, and a dropout function to avoid overfitting. The second hidden layer contains another dense layer with 32 neurons and a dropout again but with a smaller parameter. An activation function of a rectified linear unit is used by both hidden layers. The last layer is dense layer with a layer of 2 neurons since we have a binary prediction of yes or no output using a softmax activation function.

The model architecture was preserved while training Dataset II as both classes bear the same number of samples. However, Dataset I, as it bears imbalanced data, we needed to use an extra step which is data augmentation with one parameter from Keras image data generator which is rotation degree by 30 and filling mode to replace the void created after applying the rotation process. Our third flowchart (Figure 11) represents the final diagram block of the Modified DenseNet1.0 DenseNet-based model trained on dataset II and the upcoming explainable Modified DenseNet2.0.

As our research consists of explaining the prediction that our previous model is outputting and by taking into consideration that the focus lays on feature importance to help estimate the data contribution value to the model's prediction, we suggest applying the shapely values technique to the model with the best evaluation result and overall best-obtained accuracy to hit two achievements at the same time, high testing accuracy and well-justified prediction.

Explainable Modified DenseNet2.0 extension

It is the model's explanations for the predictions or decisions. These explanations may take various forms, including textual description, visualization, or highlighting important features in the input data that impacted the outcome. In our study, we have selected the SHAP (SHapely Additive eXplanations) technique. Cooperative game theory inspiring the idea of the shapely values (Von Neumann, J., & Morgenstern, O.,1944) have been applied in XAI to highlight each feature's contribution to the final prediction providing a fair way of distributing the credit for a prediction among the overall input features.

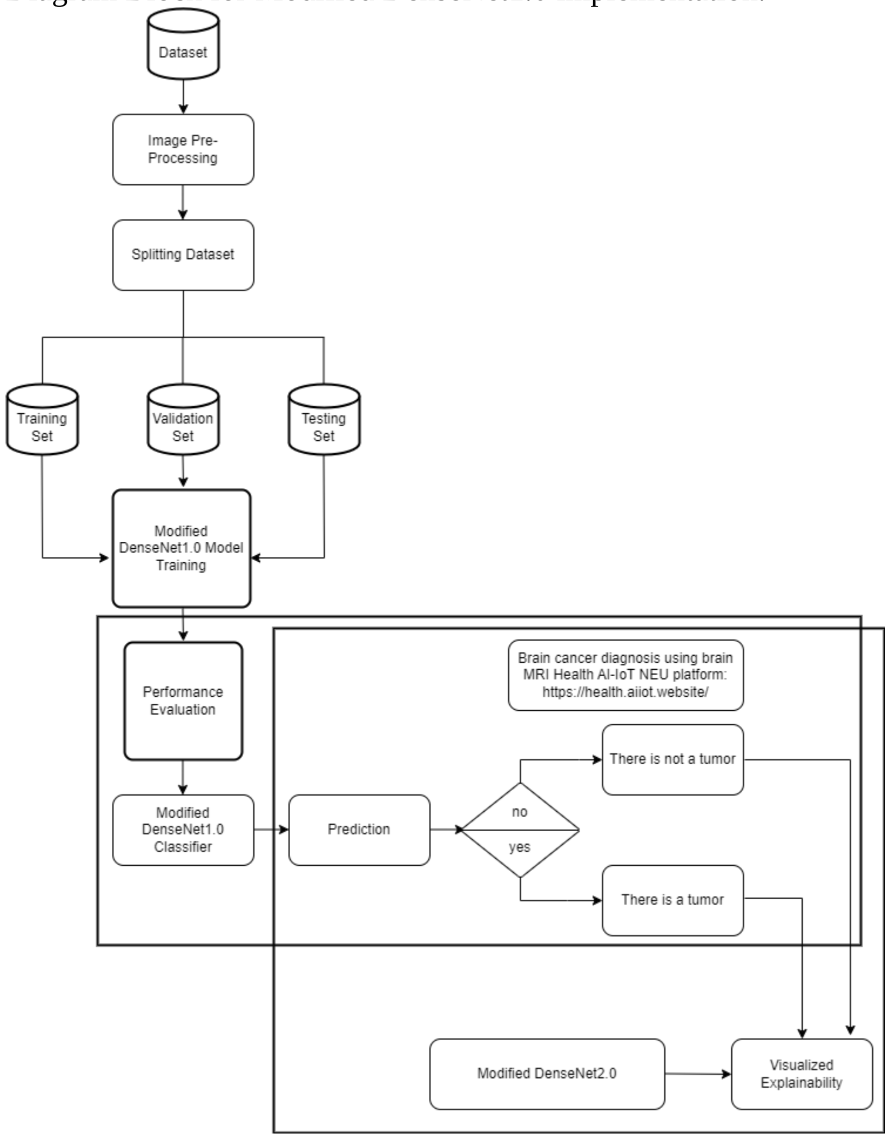
Shapely values

The optimal Shapley values in a coalitional game are the source of SHAP. Predictions are explained by seeing each aspect of an instance as a distinct "player" in a game, in which the prediction represents the overall prize. A technique to fairly divide this "payout" across the different attributes is provided by Shapley values (Molnar, 2023). Every feature in machine learning is given a significance value that indicates how much it contributes to the model's output. SHAP values illustrate the relative importance of each feature in relation to other features, the impact of each feature on each final prediction, and the dependence of the model on feature interaction (Datacamp, 2011).

A value function Val of players in S (12) defines the Shapley value. The payment contribution of a feature value, weighted and totalled over all potential feature value combinations, is its Shapley value:

$$\phi_j (val) = \sum_{S \subseteq \{1,\ldots,p\} \setminus \{j\}} \frac{|S|! \cdot (p-|S|-1)!}{p!} \left(val (S \cup \{j\}) - val (S)\right) \quad (12)$$

Figure 11
Diagram Block for Modified DenseNet1.0 implementation.



Evaluation Metrics

Evaluation metrics and criteria are tied parameters and crucial parts of any learning process regardless of the method or the algorithm used during the process. It is a required phase to evaluate our models and give credibility to the outcome. There are different metrics according to the task or the classification we want to perform. We list the most important ones to our research as follows:

Accuracy

The accuracy (4) measures the probability of correct predictions among the overall number of predictions. It is calculated as follows:

$$\frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

Confusion Matrix

A more thorough explanation of the correct incorrect classifications for each class can be found in a confusion matrix or confusion table. The ground truth labels are represented by the rows of the matrix, while the prediction is represented by the columns.

Precision

The precision metric (5) indicates the quality of truly positive predictions among the positive predictions. This is how it's calculated:

$$\frac{TP}{TP+FP} \quad (5)$$

Recall

Commonly referred to as the true positive rate (6). It is widely used as a classification model performance metric that outcome the correctly classified positives. It is computed as follows:

$$\frac{Correctly\ classified\ positives}{TP+FP} \quad (6)$$

F1-Score

Known also as the harmonic mean (7), it consists of recall and precision together. It is calculated as follows:

$$\frac{2}{Precision^{-1}+Recall^{-1}} \quad (7)$$

Sensitivity

It is the model evaluation ability to predict truly positive cases of each class in the dataset (8). The following formula is used to compute it:

$$\frac{TP}{TP+FN} \quad (8)$$

Specificity

It is the evaluation ability of the model to predict truly negative cases of each class in the dataset (9). It is calculated as follows:

$$\frac{TN}{FP+TN} \quad (9)$$

ROC AUC

The ROC curve and AUC (10) are well-known concepts. An AUC (Area Under the Curve) or ROC (Receiver Operating Characteristic) curve is frequently used to illustrate this concept. A binary classification model's trade-off between true positive rate *(sensitivity)* and false positive rate *(1 - specificity)* at different thresholds is represented graphically by the ROC curve (Bishop, 2006). Regarding the threshold t, the inverse function of the false positive rate is denoted by *FPR−1(t)*.

$$\text{AUC} = \int_0^1 TPR\big(FPR^{-1}(t)\big)\, dt \tag{10}$$

The AUC for the ROC is usually calculated using trapezoidal rule (11) which sums up the areas of trapezoids formed by adjacent points on the ROC curve. It is calculated as follows:

$$\text{AUC} = \sum_{i=1}^{N-1} \frac{1}{2}\,(x_{i+1} - x_i).\,(y_{i+1} + y_i) \tag{11}$$

Where (x_i, y_i) are the coordinates of the ROC curve at point *i* and *N* is the total number of points on the curve.

Comparative Study

Using a CNN-based classification model and the transfer learning technique, we demonstrate that our suggested model has the optimum architecture, we have selected 3 classification approaches: VGGNets, ResNets, and DenseNets with different number of layers, epochs and kept the same batch size that has proven efficacy for our selected Dataset during the dataset selection process.

VGGNets

VGG stands for Visual Geometry Group. It is a multi-layered, conventional deep CNN architecture. The VGGNet, a deep neural network, outperforms baselines on a wide range of tasks and datasets, going beyond ImageNet. Furthermore, it remains one of the most often used architectures for image recognition. In our study, we have included VGGNet16 and VGGNet19. VGGNet uses a very small 3 x 3 receptive filter layer through the entire network which supports the receptive area when it is large, along a stride of 1 pixel. In addition to these 3 filter layers, there are also 3 non-linear activation layers which make the decision function more discriminative and increase the network ability to converge faster. The ReLu activation function is used to add non-linearity, which lessens overfitting during the training phase. VGGNets are considered one of the excellent vision model architectures in the present time. It is widely used to perform classification and detection tasks through transfer learning techniques.

ResNets

ResNet stands for Residual Networks. They are a type of deep CNN architecture known for handling deeper architectures while addressing the vanishing gradient problem associated with training very deep networks (He, et al., 2015). Unlike traditional architectures, ResNets introduce the concept of residual learning, allowing for the training of extremely deep networks without facing issues such as vanishing gradients. This approach is beneficial in the context of deep networks. ResNets use residual blocks, which add, rather than transform, the input from the preceding layer to the output. This residual connection enables the network to learn the residual mapping and capture complex features. These residual connections enable the training of deep networks. ResNets architecture also includes the use of 3 x 3 convolutional filters throughout the network. Similar to VGGNets, this helps maintain a small receptive field and supports the learning of complex features. Additionally, ResNets introduces the concept of bottleneck layers, which involve the use of 1 x 1 convolutions to reduce the dimensionality before applying 3 x 3 filters. To introduce non-linearity, the activation function ReLU is used by ResNets to help solve the vanishing gradient issue and hasten the network's convergence during training. The combination of residual connections and ReLU activation contributes to ResNet's ability to build highly accurate and easily trainable deep networks. ResNet's effectiveness extends to tasks beyond image recognition, and it's frequently utilized as a base model for transfer learning methods along VGGNets, showcasing its versatility and robustness in different applications.

DenseNets

Densely Connected Convolutional Networks is most commonly referred to as DenseNet. Feature reuse problems, vanishing gradients, and parameter efficiency are the three main goals of this deep network design (Huang, et al., 2016). In contrast to conventional convolutional network designs, DenseNets architecture presents a novel connectivity pattern with the aim of resolving issues with feature reuse, vanishing gradients, and network depth. DenseNets are characterized by their dense connectivity topology, in which all layers are feed-forward connected to all other layers. This connectivity results in $L(L+1)/2$ direct links in a network with L layers. DenseNet allows each layer to transfer its feature maps to all following layers and receive input from all previous layers, in contrast to typical architectures where information flows through layers sequentially.

DenseNet achieves several benefits because of this dense connectivity architecture. Since information can go directly across layers, it reduces the issue of vanishing gradients. The model's representational efficiency is improved by this extensive connection, which also promotes feature reuse. In addition, DenseNets uses fewer parameters than conventional architectures.

DenseNet, like VGGNets and ResNets, uses 3 x 3 convolutional filters across the network. DenseNets also uses Batch Normalization (Ioffe & Szegedy, 2015) and ReLU activation functions to handle non-linearities. The combination of these aspects enhances the model's capacity to learn complicated patterns and features efficiently.

DenseNets has outperformed other object recognition benchmark datasets, including CIFAR-10, CIFAR-100, SVHN, and ImageNet. The architecture's ability to handle diverse tasks extends its applicability beyond image classification to tasks like object detection and segmentation.

The success of DenseNet has established it as a powerful and versatile architecture in the realm of DL. Its impact is not limited to image recognition, as DenseNets serve as a strong foundation for transfer learning techniques, showcasing its adaptability and effectiveness in different applications, we have adopted DenseNets to build our proposed model Modified DenseNet1.0.

CHAPTER IV

Findings and Discussion

The experiments and analysis along with model generation conducted during this study were run with Python programming language 3.11.7 using Visual Studio code editor 1.85.2 on a lab desktop at the R&D classroom provided by the Innovation Centre at Near East University with specifications of Core i9-9900K CPU, RAM capacity of 32 GB and a 64-bit Windows operating system.

Datasets Splitting Process

Both datasets utilized in the experiments were randomly split with a ratio of 90% and10% for training and testing respectively, along with an independent random split of 90% for validation to ensure the model's effectiveness.

Training Process

In this study, our Deep Learning (DL) model was trained over different metrics and evaluated on two distinct datasets to assess its performance. The evaluation process involved the utilization of three different architectures, namely VGGNet16, VGGNet19, and ResNet50. To comprehensively assess the model's learning capabilities, evaluation sessions were conducted for 50 epochs and subsequently extended to 100 epochs. This approach seeks to provide an understanding of the model's performance across various datasets and training durations, allowing for an in-depth evaluation of its robustness and generalization capabilities.

An epoch in ML refers to a single training phase sweep across the whole training dataset. Training and Validation Loss refers to the evaluation of the model's performance throughout several epochs on the training and validation sets.

The loss is the metric used to indicate how well the model is doing its job. This value is usually minimized. The validation loss shows how well the model generalizes to new, unseen data, whereas the training loss shows how well the model fits the training data.

Dataset Evaluation

Dataset I was trained over 50 then 100 epochs (See Figure 12 and Figure 13) with batch size equals to 8 as the size of the total number of samples was critical and the data augmentation technique is one of the solutions we suggest to maximize the model’s performance by adding a rotation parameter by 30 degrees which helped suppress the overfitting issue under the 100 epochs case.

Consideration should be given to the balance between improved performance on training and validation sets and the overfitting issue when deciding on the value of training epochs. Therefore, we have decided to train our proposed model on 50 and 100 epochs on a larger dataset using regularization technique to prevent overfitting in such a longer training process expecting a better performance and that will be proven through the following experimentations.

Dataset II was trained over 50 and 100 epochs and batch size equals to 27 as the size of the total number of samples was big enough to expect a considerably high possible accuracy without any signs of overfitting following 100 training steps as the following curves indicate so clearly in Figure 14 and Figure 15.

Figure 12
Dataset I evaluation on 50 epochs and batch size = 8.

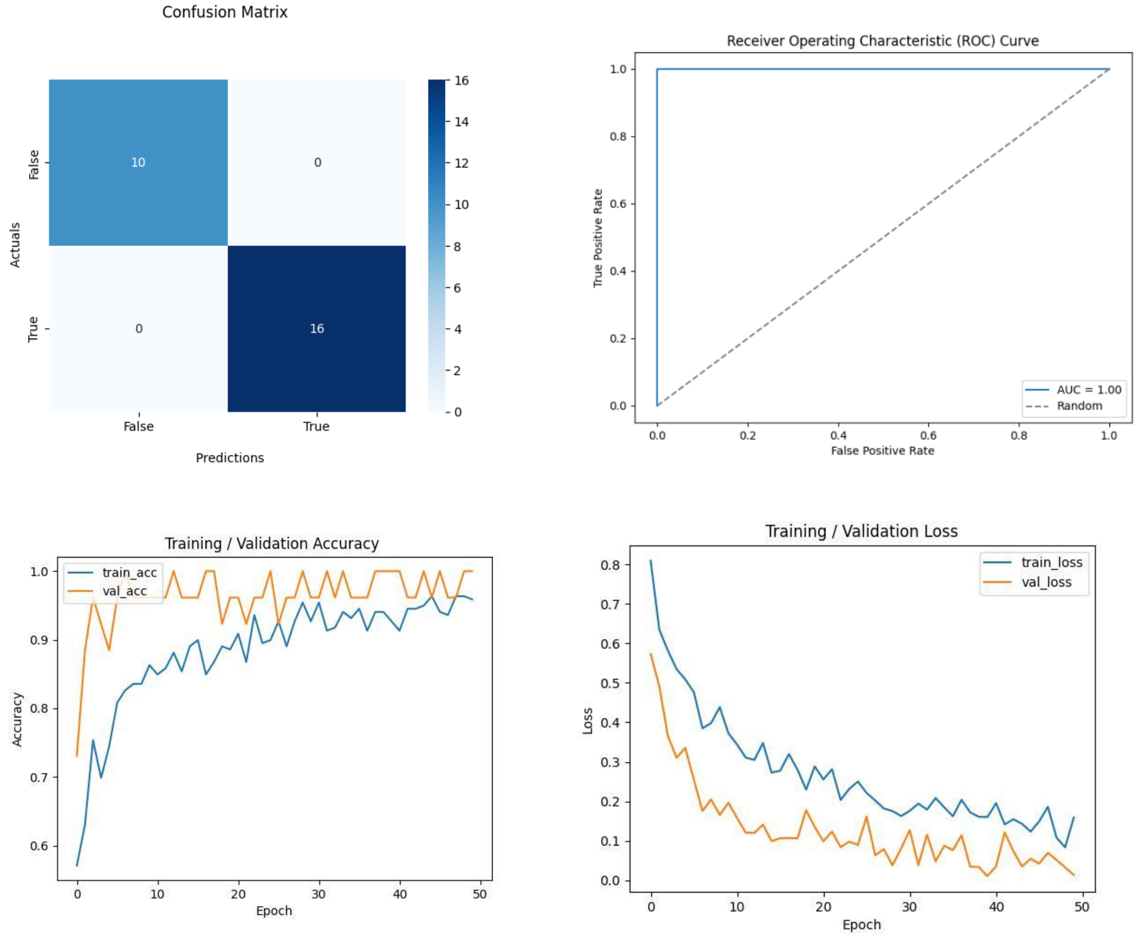
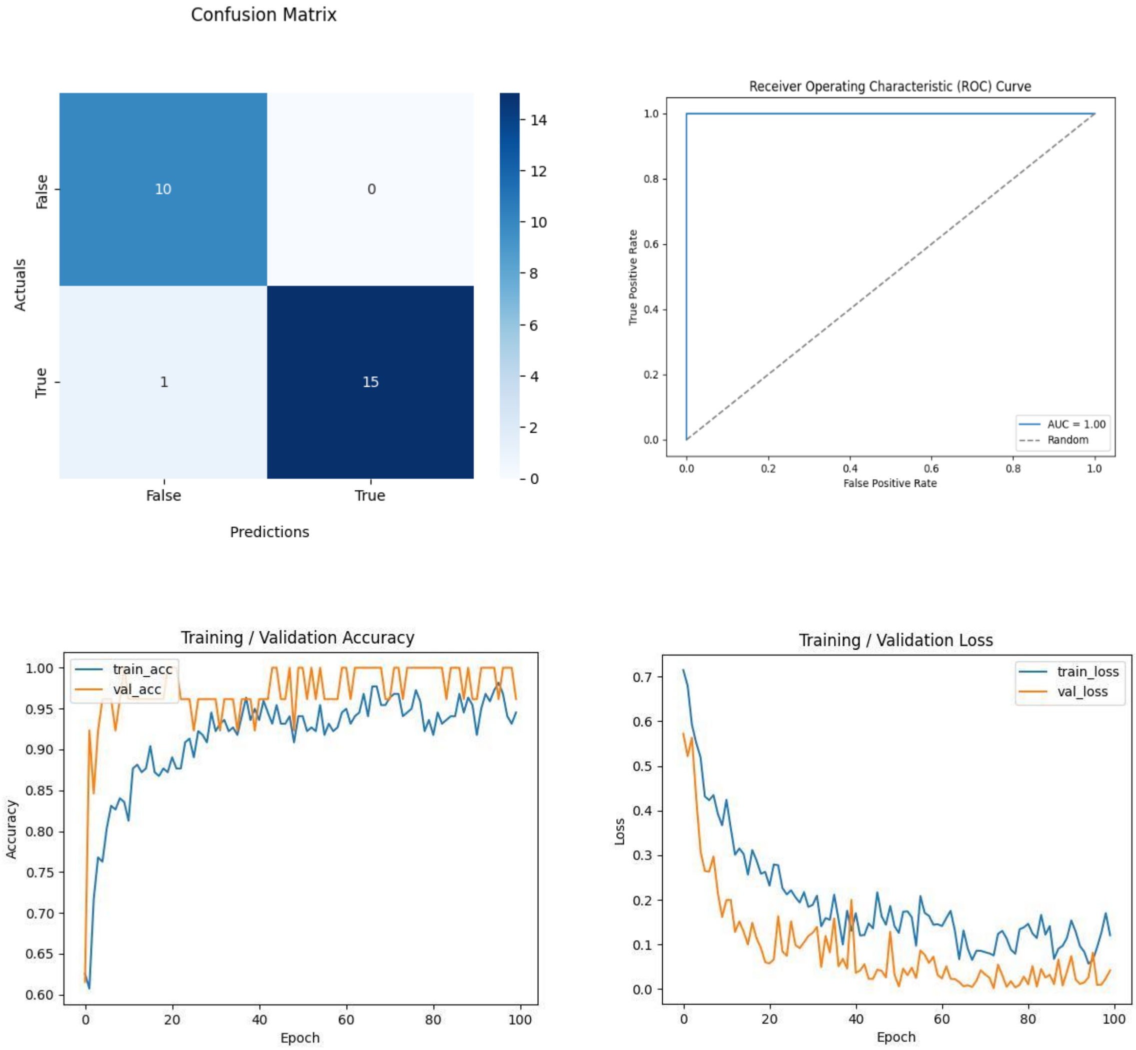


Figure 13
Dataset I evaluation on 100 epochs and batch size = 8.



From an accuracy perspective, we read that our Modified DenseNet1.0, trained for 100 epochs, achieved a higher training accuracy (99.93%). This indicates that extending the training phase improved the model's performance on the training data. Our proposed model also achieves a higher validation accuracy (99.12% and 99.78) when evaluated on dataset I and dataset II respectively reinforcing the idea that additional training epochs have contributed to better generalization. While our model has achieved a high testing accuracy when evaluated on both datasets, it reached a perfect accuracy of 100% when evaluated on dataset I using 50 epochs indicating crystal clear overfitting. The loss metric evaluation shows that model has the lowest training value when evaluated on dataset II over 100 epochs and the lowest validation loss evaluated on dataset I over 100 epochs, indicating improved convergence during training. In fact, our model reaches the lowest testing loss when evaluated on dataset I over 100 epochs, and the highest training accuracy when evaluated on dataset II over 100 epochs.

In conclusion, the model trained over 100 epochs on a larger dataset (Dataset II) generally shows improved performance in terms of training accuracy and validation accuracy.

Figure 14
Dataset II evaluation on 50 epochs and batch size = 27.

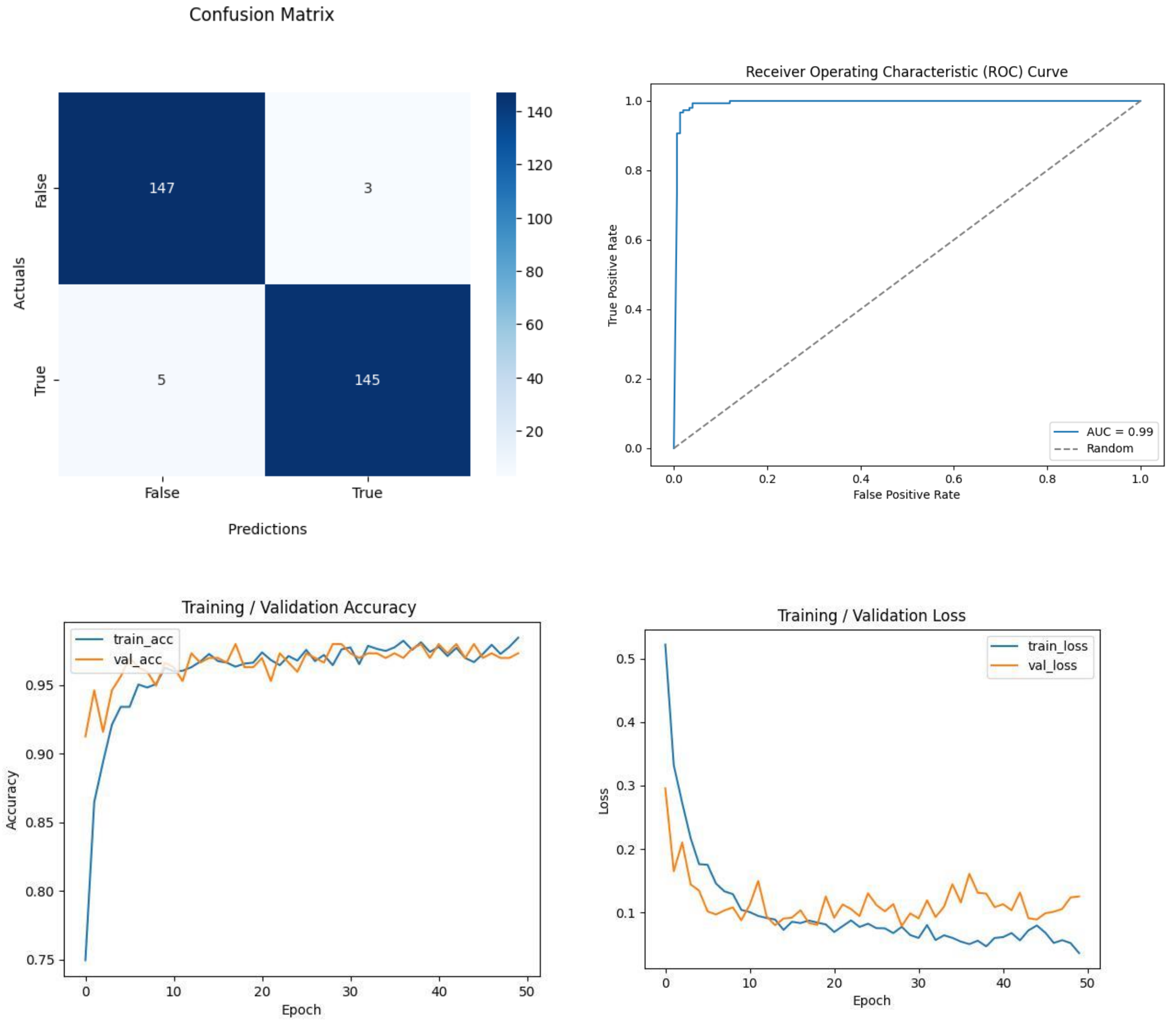
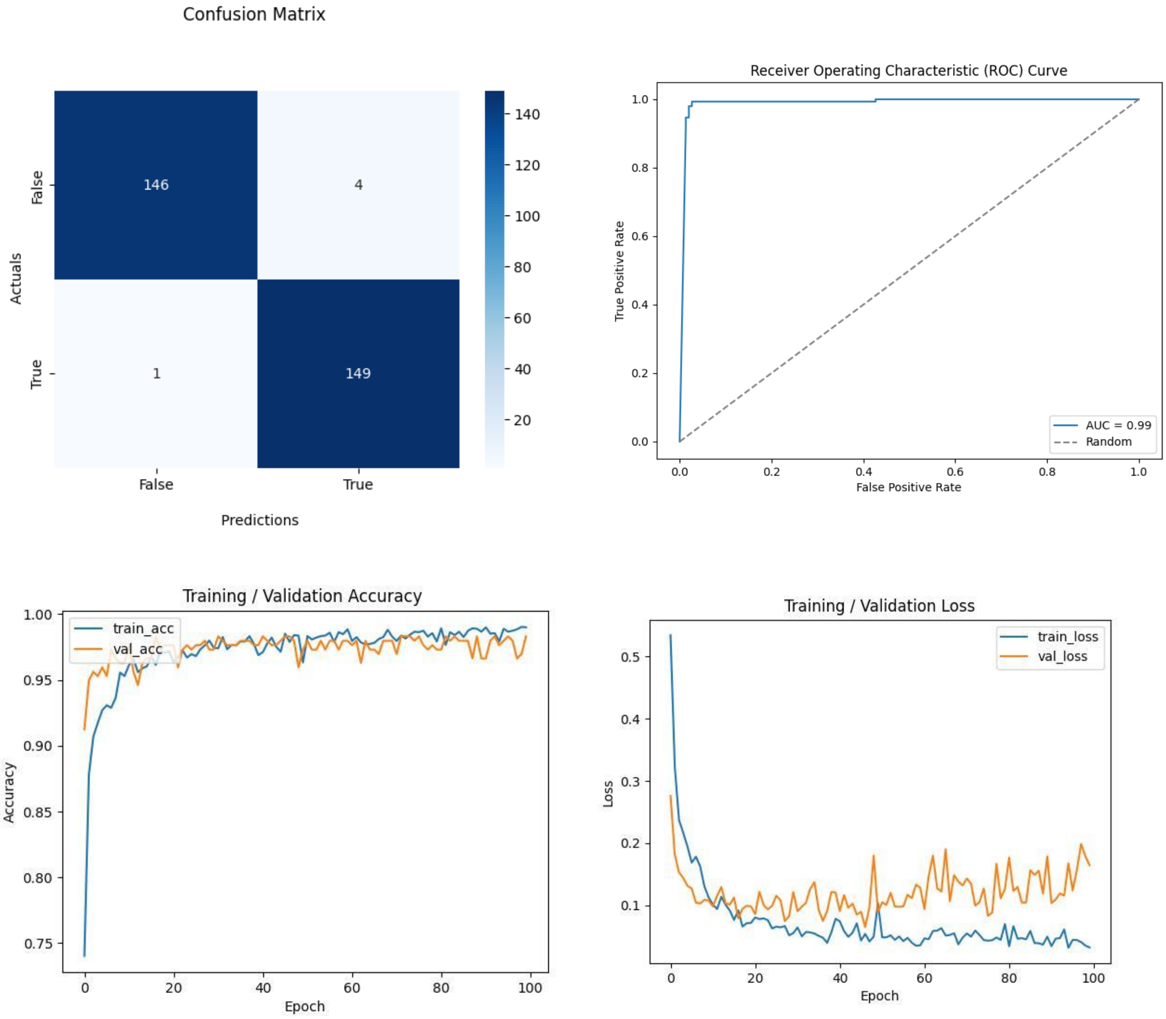


Figure 15

Dataset II evaluation on 100 epochs and batch size = 27.



While paying attention to any signs of overfitting, especially as the model complexity increases with more epochs, we could state that our proposed model performed well on both datasets. Table 4 summarizes the results from the two experiments.

Table 4

Model Learning Evaluation I.

Learning phase		Training		Validation		Testing	
Epochs				50			
Evaluation Metrics %		Accuracy	Loss	Accuracy	Loss	Accuracy	Loss
Modified	Dataset I	97.80	04.70	97.81	04.70	100	01.35
DenseNet1.0	Dataset II	**99.70**	**01.04**	**99.63**	**01.61**	**98.33**	**06.98**

Table 5

Model Learning Evaluation II.

Learning phase		Training		Validation		Testing	
Epochs				100			
Evaluation Metrics %		Accuracy	Loss	Accuracy	Loss	Accuracy	Loss
Modified	Dataset I	99.56	0.98	99.12	01.42	96.15	04.16
DenseNet1.0	Dataset II	**99.93**	**0.39**	**99.78**	**02.08**	**98.33**	**16.28**

As a consequence, balanced data is an important key to more reliable and stable models in ML experiments which is proven by the different evaluation metrics (see Table 5 to 6).

Table 6
Evaluation metrics on 50 epochs.

	Evaluation Metrics %	Precision		Recall		F1-score		Sensitivity		Specificity		AUC	
	Dataset	I	II	I	II	I	II	I	II	I	II	I	II
Class	No	100	99	100	98	100	98	100	99	100	98	100	100
	Yes	100	98	100	99	100	98						

Table 7
Evaluation metrics on 100 epochs.

	Evaluation Metrics %	Precision		Recall		F1-score		Sensitivity		Specificity		AUC	
	Dataset	I	II	I	II	I	II	I	II	I	II	I	II
Class	No	91	99	100	97	95	98	94	98	100	97	100	99
	Yes	100	97	94	99	97	98						

Another key to improving the model's learning performance is the number of epochs or iterations and batch size that have proven effective with both datasets in reducing the loss ratio and suppressing the overfitting. The chosen epochs effectively address overfitting, guaranteeing that the model generalizes significantly well beyond the training data. Figure 15 implies that visual representations accompany these assertions, providing a graphical understanding of how the training and validation accuracy evolve over epochs. This insight not only enhances the interpretability of the model's learning trajectory but also aids in informed decision-making regarding computational resources and training time.

After conducting experiments with different epoch values, it has been observed that using 50 to100 epochs is effective in achieving balance. We conducted extensive experiments to strengthen the foundation of our research and evaluate the performance of our Dense Network-based model. The model may become under-fitted if there are too few epochs, failing to capture the complexity of the data, or over-fitted if there are too many epochs, performing poorly on new data. The effectiveness of our model is evaluated against a variety of well-known classification architectures, such as ResNet50, VGGNet16, and VGGNet19, taking into account the number of epochs, using Dataset II, and a batch size of 27.

Experimentation Phase I

Our primary objective is to compare and analyse the model's performance across these architectures, focusing on key metrics such as accuracy, loss, and other relevant evaluation metrics. The models underwent an initial training for a total of 50 epochs to capture complex patterns within the dataset.

ResNet50

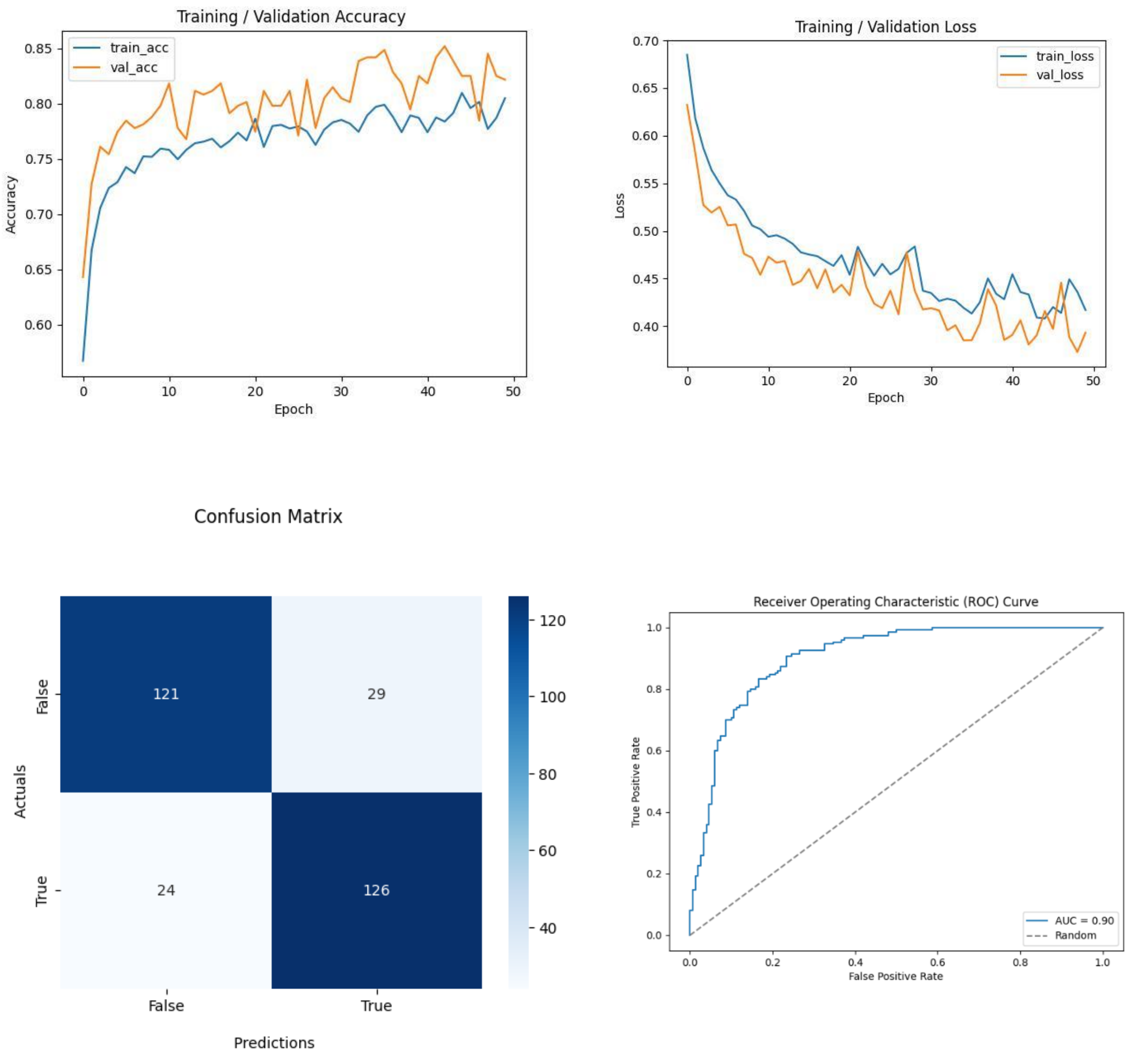
During the learning phase, the effectiveness of the ResNet50 architecture was evaluated on dataset II, with an emphasis on training, validation, and testing sets. The results that were collected are included in the table that follows (Table 8)

Table 8
ResNet50 Performance Evaluation Summary I.

Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	81.74
	Loss	37.53
Validation	Accuracy	82.07
	Loss	37.15
Testing	Accuracy	82.33
	Loss	39.06

The results indicate a fine accuracy across all phases, showcasing a consistent trend between the training, validation, and testing sets. The training accuracy of 81.74% suggests that the model learned from the dataset. During the validation phase, the accuracy slightly improved to 82.07%, indicating good generalization to previously unseen data after 50 epochs.

Figure 16
ResNet50 Performance Evaluation I.



However, it is essential to note a marginal increase in loss during validation and testing, reaching 37.15% and 39.06%, respectively. This increase in loss suggests a certain level of overfitting, where the model performs well on the training set but finds it difficult to generalize to new data. A confusion matrix was used to undertake an in-depth evaluation of the ResNet50 model's performance, in addition to evaluating accuracy and loss metrics.

For both classes (yes and no), the model achieves a balanced performance with similar precision, recall, and F1-score. The confusion matrix shows how the distribution of True Positives (126), True Negatives (121), False Positives (29), and False Negatives (24) reflects the trade-off between recall and precision values in accurately predicting images with and without tumours. The model's 90% AUC score indicates that it can effectively discriminate between positive and negative classes and has a comparatively excellent discriminatory capacity.

Even though the model's overall accuracy is 82%, its ability to accurately recognize “yes” positive image samples stands out in particular, as seen by its greater recall of 84%. However, there is still an opportunity for improvement in terms of lowering False Positives (29), which would help to increase the accuracy of the model.

In conclusion, the ResNet50 architecture demonstrated promising accuracy after 50 epochs, but the observed increase in loss prompts a careful consideration of potential overfitting issues.

VGGNet16

During the learning phase, the VGGNet16 architecture's performance on the dataset was evaluated, paying particular attention to the training, validation, and testing sets. Table 9 provides a summary of the outcomes that were obtained.

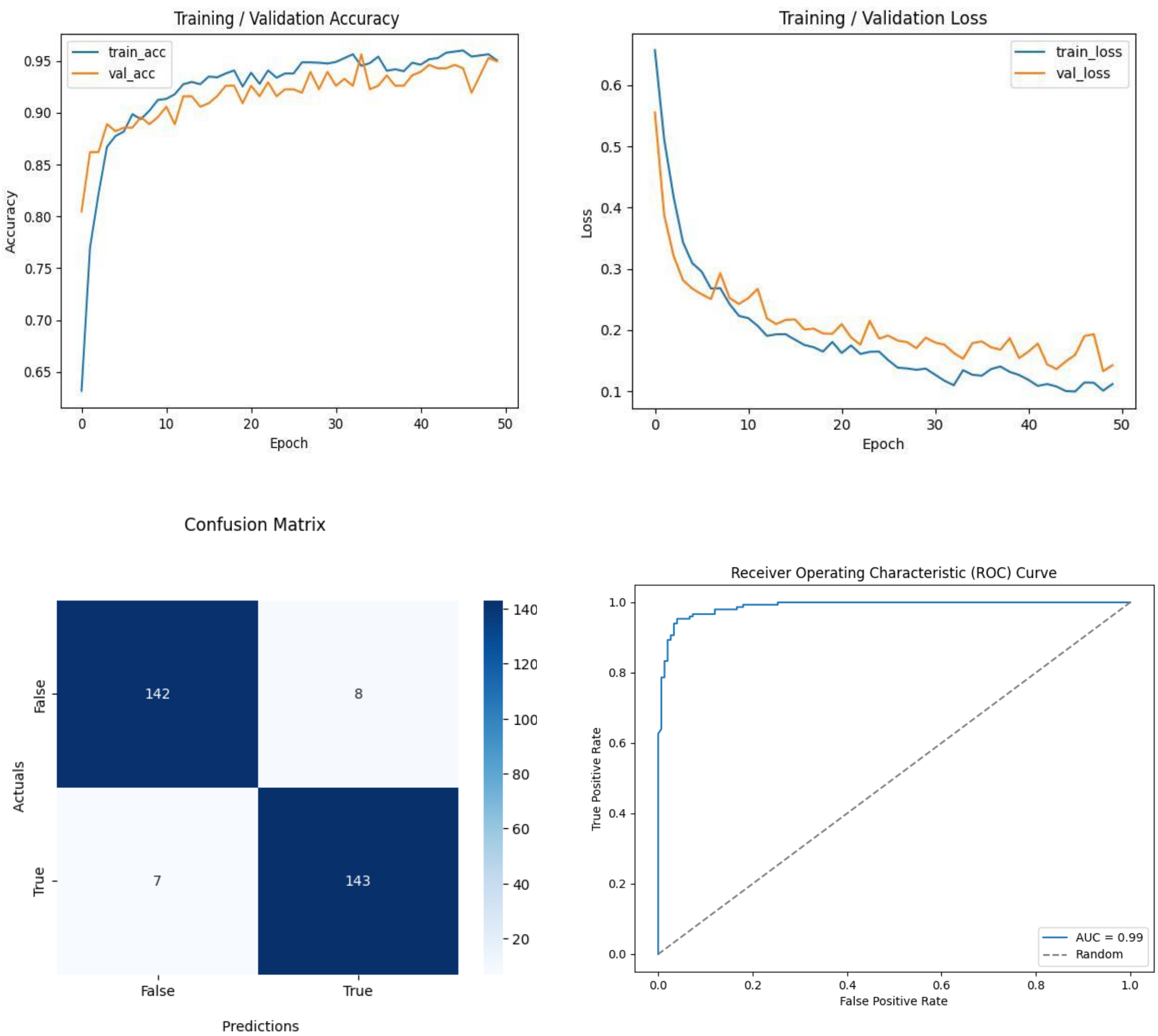
Table 9
VGGNet16 Performance Evaluation Summary I

Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	99.04
	Loss	04.24
Validation	Accuracy	98.63
	Loss	05.31
Testing	Accuracy	95.00
	Loss	14.15

The VGGNet16 model shows competitive performance as expected across various evaluation metrics, highlighting its capability to learn and generalize from the provided dataset. The model reaches a high accuracy of 99.04% in the training phase, indicating that it is adept at identifying intricate patterns in the training data. The low training loss of 04.24% further emphasizes the model's effectiveness in minimizing errors during the learning process.

Figure 17

VGGNet16 Performance Evaluation I.



With an accuracy of 98.63%, the model retains its resilience during the validation stage. Even though it was somewhat more than the training loss, the matching loss of 05.31% nonetheless points to a successful generalization to never-before-seen validation data.

Testing further demonstrated the model's strong generalization, as it obtained a 95% accuracy rate with a 14.15% loss. A confusion matrix was used to undertake a thorough performance evaluation of the VGGNet16 model, in addition to assessing accuracy and loss metrics.

The VGGNet16 model showcase a great performance with high precision, recall, and F1-score values for the two classes “no” and “yes”. The distribution of Truly Positives (143), Truly Negatives (142), False Positives (8), and False Negatives (7) is demonstrated in the confusion matrix. The model's ability to correctly identify MRI images with and without tumours is demonstrated by its balanced precision and recall values, which are anticipated of a first- and second-place award winner. The overall accuracy of 95% underscores the model's proficiency in making correct predictions. These metrics analyses affirm the robustness of the VGGNet16 architecture in binary classification tasks. The high precision and recall values, along with a balanced F1-score, indicate how well the model can distinguish between positive and negative images. The AUC value of 99% indicates excellent discriminatory power, underlining the model's capacity to distinguish between classes. The evaluation metrics for VGGNet16 after 50 epochs reveal a well-performing model with high accuracy and discriminatory ability. These results contribute to the model's reliability and suitability for binary classification tasks. Moreover, it makes it the top competitor of our Modified DenseNet1.0.

VGGNet19

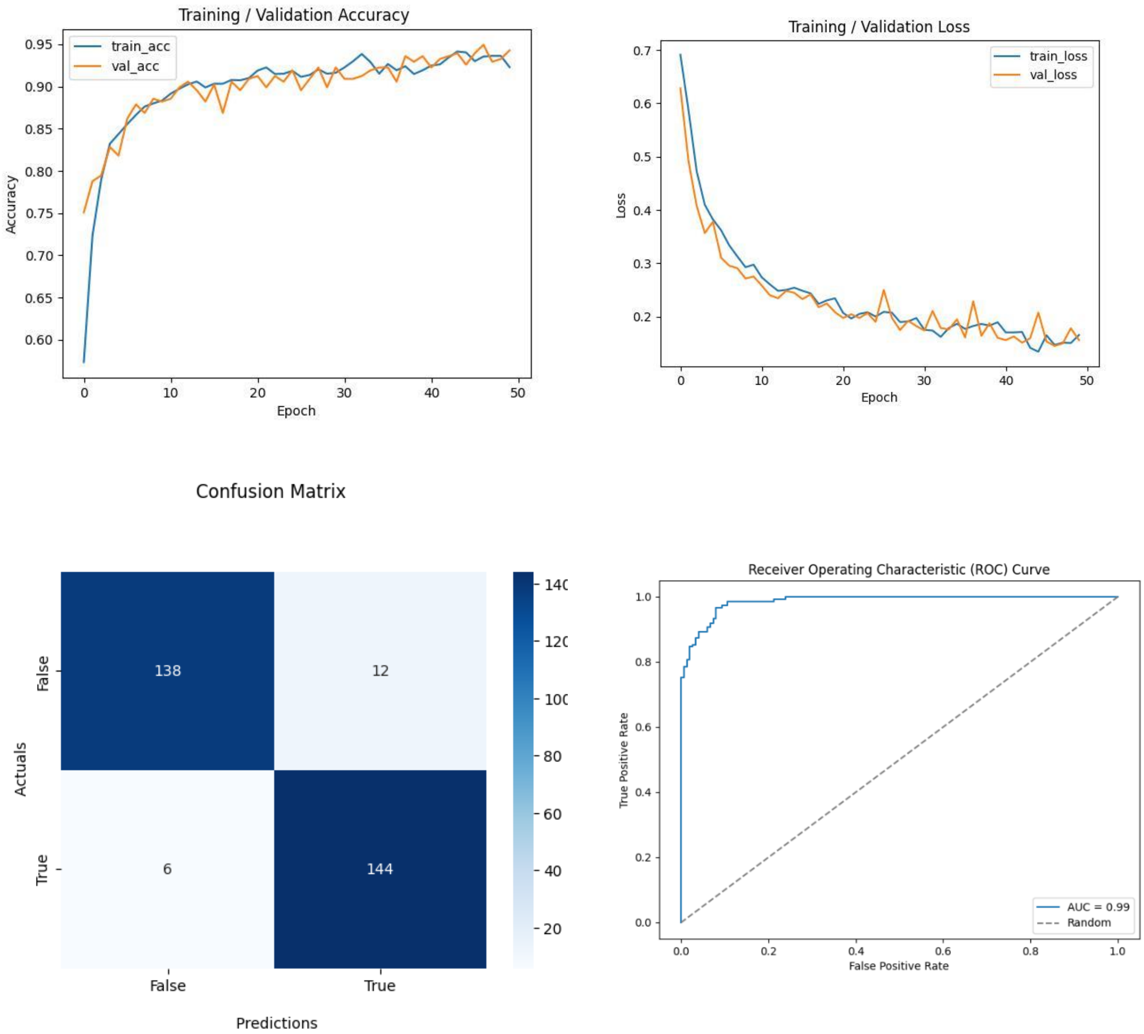
During the learning phase, a detailed evaluation of the VGGNet19 architecture's performance on the dataset was conducted, with an emphasis on training, validation, and testing sets. Table 10 provides a summary of the outcomes that were acquired.

Table 10
VGGNet19 Performance Evaluation Summary I.

Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	97.37
	Loss	07.29
Validation	Accuracy	97.15
	Loss	07.95
Testing	Accuracy	94.00
	Loss	16.35

The VGGNet19 model also shows a great performance across various evaluation metrics, indicating its proficiency in capturing complex patterns within the dataset. The model achieves, during the training phase, a high accuracy level of 97.37%, suggesting successful learning from the training data. The corresponding low training loss of 07.29 reflects the model's capability to minimize errors during the learning process.

Figure 18
VGGNet19 Performance Evaluation I.



The model maintains its robustness during the validation phase, with an accuracy of 97.15%. Despite a slight increase in loss to 07.95, the validation findings show how effectively the model generalizes to new, unknown data. However, during the testing phase, the accuracy slightly drops to 94%, and the loss increases to 16.35. This divergence between training and testing metrics indicates a potential issue, such as overfitting.

In addition to evaluating accuracy and loss metrics, we conducted a scan for the VGGNet19 model's performance confusion matrix. The VGGNet19 model has good performance considering then achieved precision, recall, and F1-score values for both classes “yes” and “no”. The balanced precision and recall scores show how the model can accurately identify MRI scans with and without tumours. An overall accuracy of 94% shows that the model is capable of producing accurate predictions. The observed confusion matrix highlights the model's ability to strike a balance between sensitivity and specificity, contributing to its effectiveness in binary classification. The AUC value of

99% also indicates excellent distinction ability, highlighting the model's ability to differentiate between the two classes.

Modified DenseNet1.0

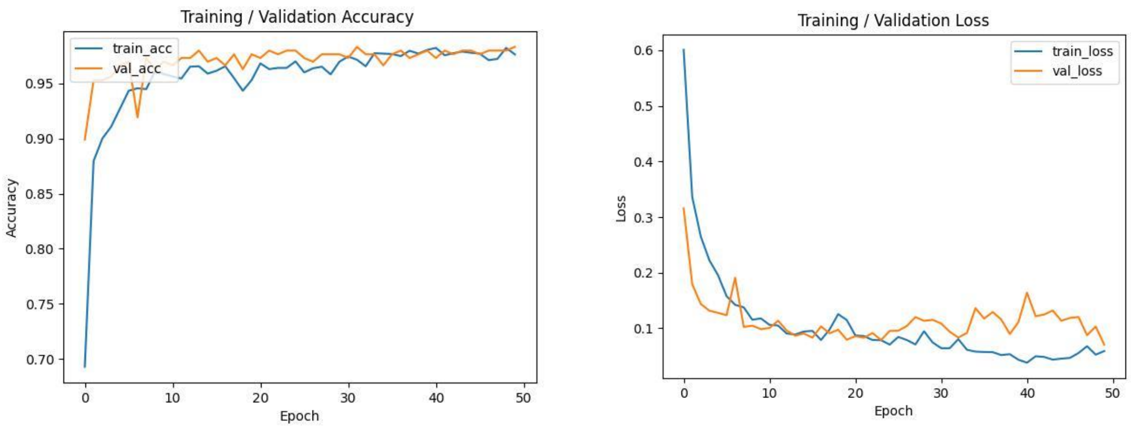
During the learning phase, the Modified DenseNet1.0 architecture's performance on the dataset was carefully evaluated, with an emphasis on training, validation, and testing sets. Table 11 provides a summary of the outcomes that were collected.

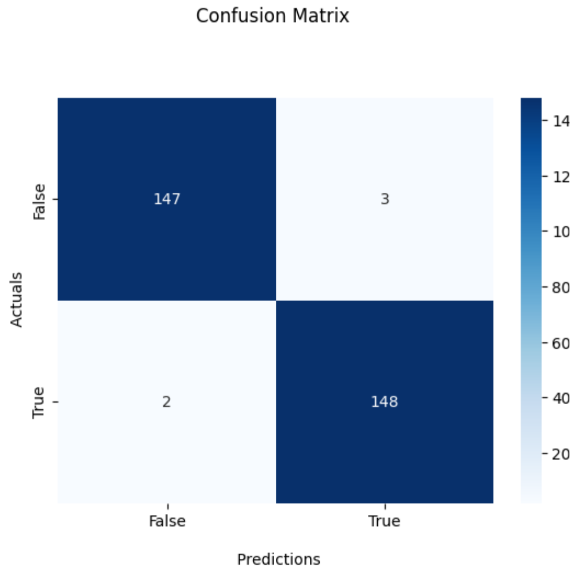
Table 11 Modified DenseNet1.0 Performance Evaluation Summary I.

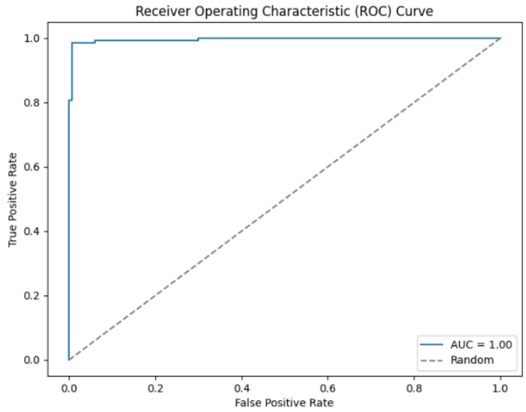
Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	99.70
	Loss	01.04
Validation	Accuracy	99.63
	Loss	01.61
Testing	Accuracy	98.33
	Loss	06.98

Our Modified DenseNet1.0 model exhibits remarkable performance across all evaluation metrics, underscoring its proficiency in learning complex patterns within the dataset. The achieved accuracy during training is an impressive 99.70%, suggesting that the model has successfully captured intricate features and nuances present in the training data without any overfitting signs and reducing testing loss compared to other baseline models' performance.

Figure 19
Modified DenseNet1.0 Performance Evaluation I.







Despite a comparatively lower testing loss of 06.98, it is essential to contextualize these values within the specific requirements of the application. In certain real-world scenarios, the achieved accuracy may be considered outstanding, and the observed loss might be acceptable depending on the consequences of misclassifications.

Our proposed Modified DenseNet1.0 model demonstrates exceptional performance, as indicated by the high precision value, recall, and F1 score for the two classes “no” and “yes”. The confusion matrix also emphasizes the model's capacity to correctly classify MRI images with and without tumours. The precision values for both classes are remarkable. With a precision of 99% for “no” and 97% for “yes” the model exhibits a high accuracy in correctly identifying MRI images belonging to each class. The recall values are also impressive, reflecting the model's ability to capture a significant proportion of true positive images. With recall values of 97% for class “no” and 99% for class “yes” the model effectively minimizes false negatives.

The balanced F1-scores of 98% for both classes emphasize the model's equilibrium between precision and recall, showcasing its robustness in making accurate predictions. The 98% overall accuracy shows how accurate the model is in classifying the entire dataset. The model's performance is comprehensively outlined in the confusion matrix. The program correctly classified 146 photos as “non-Tumorous” and 149 images as “Tumorous” demonstrating its remarkable prediction ability. The correctness of the model is further demonstrated by the low values of (1) and (4) False Negatives.

In summary, the Modified DenseNet1.0 model excels in its ability to distinguish between MRI images with and without tumours. The high precision, recall, and F1-score, coupled with the balanced confusion matrix, exhibit the model as a highly reliable and effective tool for sensitive binary classification tasks. These results suggest that the Modified DenseNet1.0 model is well-suited for practical applications in medical image analysis.

Tables 12 and 13 provide a complete summary of our experiment findings and a detailed comparison of the CNN model's performance on various classification approaches.

Table 12
Selected Baseline Models Learning Evaluation.

Learning Phase Epochs = 50		Training		Validation		Testing	
Evaluation Metrics %		Accuracy	Loss	Accuracy	Loss	Accuracy	Loss
Classification Architectures	ResNet50	81.74	37.53	82.07	37.15	82.33	39.06
	VGGNet19	97.37	07.29	97.15	07.95	94	16.35
	VGGNet16	99.04	04.24	98.63	05.31	95	14.15
	Modified DenseNet1.0	**99.70**	**01.04**	**99.63**	**01.61**	**98.33**	**06.98**

Table 13
Selected Baseline Models Evaluation metrics on 50 epochs.

Evaluation Metrics % Epochs = 50		Class	Precision	Recall	F1-score	Sensitivity	Specificity	AUC
Classification Architectures	VGGNet16	Yes	95	95	95	95.33	94.67	99
		No	95	95	95			
	VGGNet19	Yes	92	96	94	96	92	99
		No	96	92	94			
	ResNet50	Yes	81	84	83	84	80.67	90
		No	83	81	82			
	Modified DenseNet1.0	**Yes**	**98**	**99**	**98**	**98.67**	**98**	**100**
		No	**99**	**98**	**98**			

In the initial phase of experimentation, we conducted training sessions with 50 epochs, followed by a subsequent set of experiments extending the training duration to 100 epochs. This two-step approach allows for a comprehensive exploration of model performance across varying epoch intervals, providing insights into the impact of extended training on classification accuracy and loss and other evaluation metrics.

Experimentation Phase II

To serve the same objective as mentioned in the experimentation Phase I, the models underwent training for a total of 100 epochs to capture complex patterns within the dataset and improve the performance of our proposed model compared to other models' architectures in the state-of-art over a prolonged training session approach.

ResNet50

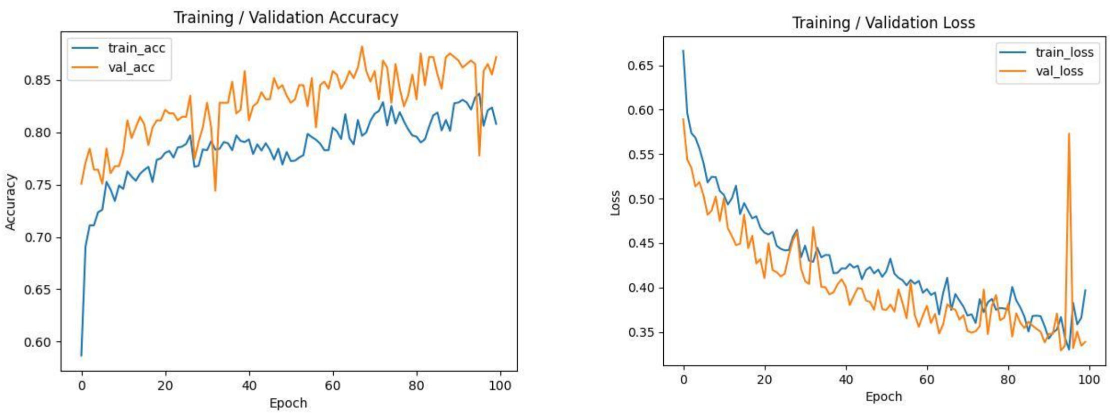
In the extended learning phase with 100 epochs, the ResNet50 architecture showcased consistent performance. Table 14 provides a summary of the outcomes that were collected.

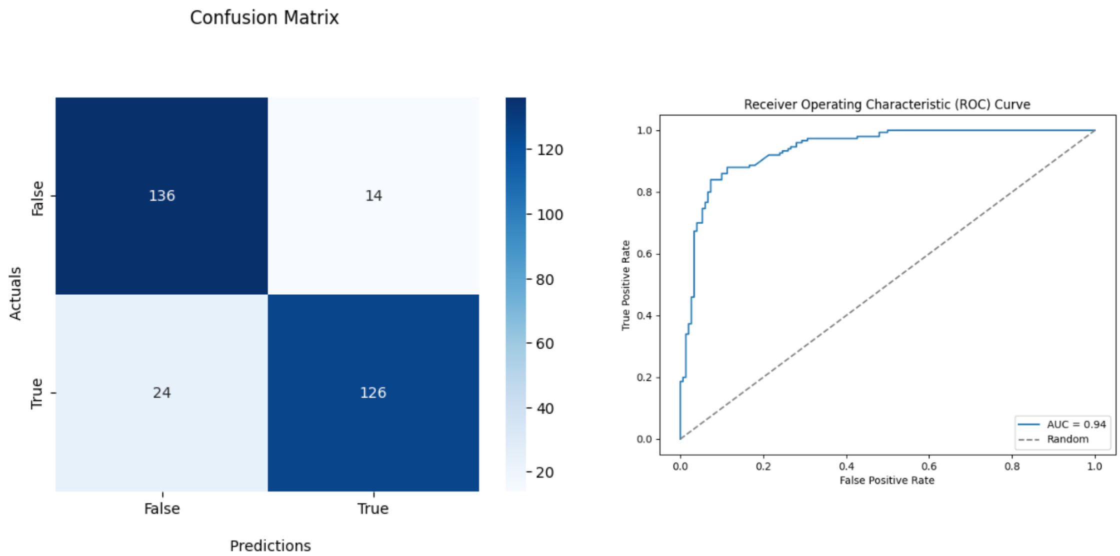
Table 14
ResNet50 Performance Evaluation Summary II

Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	88.15
	Loss	30.85
Validation	Accuracy	88.41
	Loss	30.78
Testing	Accuracy	87.33
	Loss	33.66

The training accuracy reached 88.15%, with a corresponding loss of 30.85. During validation, the accuracy further improved to 88.41%, accompanied by a slightly reduced loss of 30.78.

Figure 20
ResNet50 Performance Evaluation II.





The model demonstrated a fine robustness on the testing set, achieving an accuracy of 87.33%, albeit with a slightly increased loss of 33.66. In addition to assessing accuracy and loss metrics, a confusion matrix was used to provide an in-depth evaluation of the ResNet50 model's performance (See Figure 20).

VGGNet16

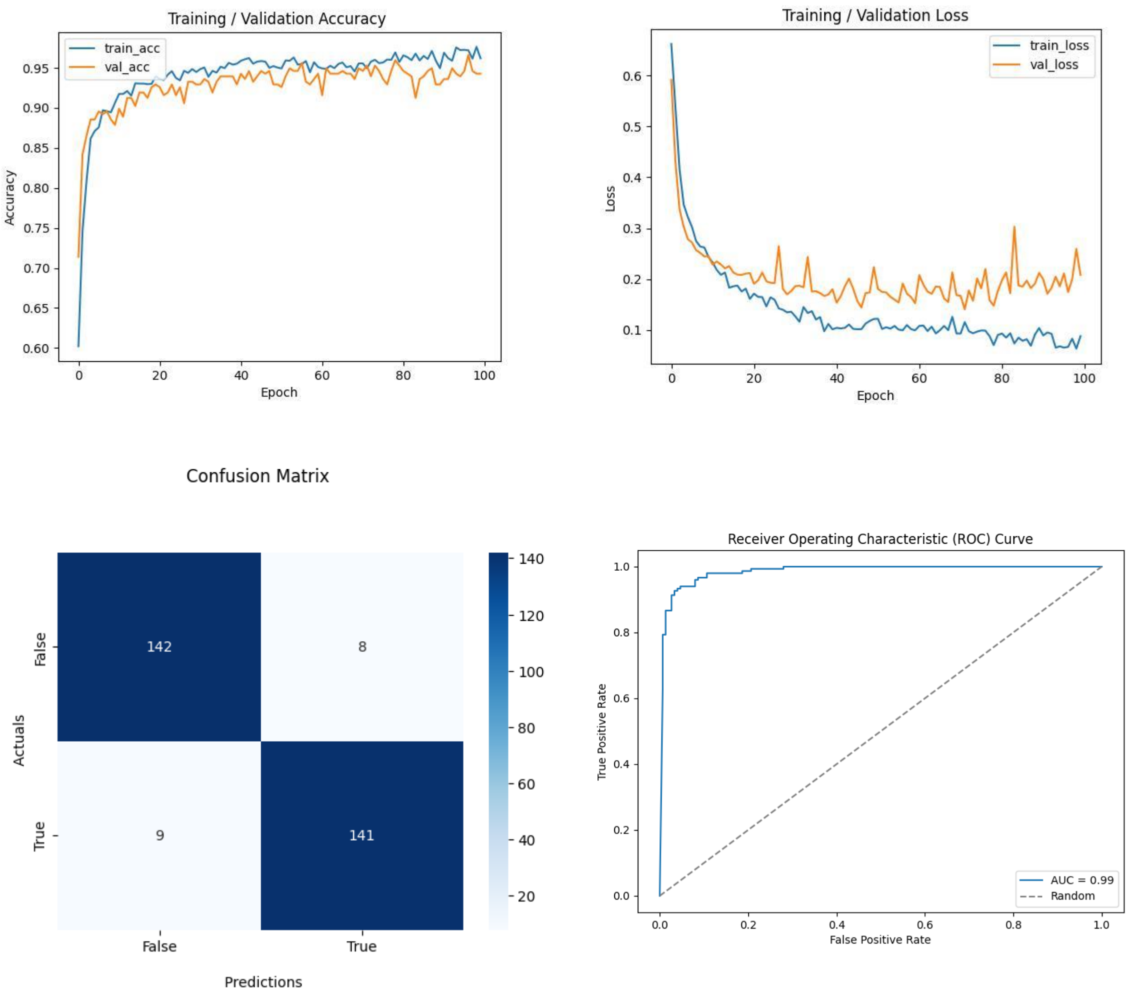
In the extended learning phase of 100 epochs, the VGGNet16 architecture demonstrated a better training performance. The table that follows (Table 15) provides an overview of the outcomes that were achieved.

Table 15
VGGNet16 Performance Evaluation Summary II.

Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	99.18
	Loss	02.10
Validation	Accuracy	98.56
	Loss	04.25
Testing	Accuracy	94.33
	Loss	20.60

During the training phase, the VGGNet16 hits an accuracy of 99.18% and a low loss of 02.10. While maintaining a high validation accuracy of 98.56%, the model exhibited a slight increase in validation loss to 04.25, indicating robust generalization capabilities. However, during testing, the accuracy slightly decreased to 94.33%, accompanied by an elevated loss of 20.60. In addition to assessing accuracy and loss metrics, a confusion matrix was used to undertake an in-depth evaluation of the VGGNet16 model's performance (See Figure 21).

Figure 21
VGGNet16 Performance Evaluation II.



VGGNet19

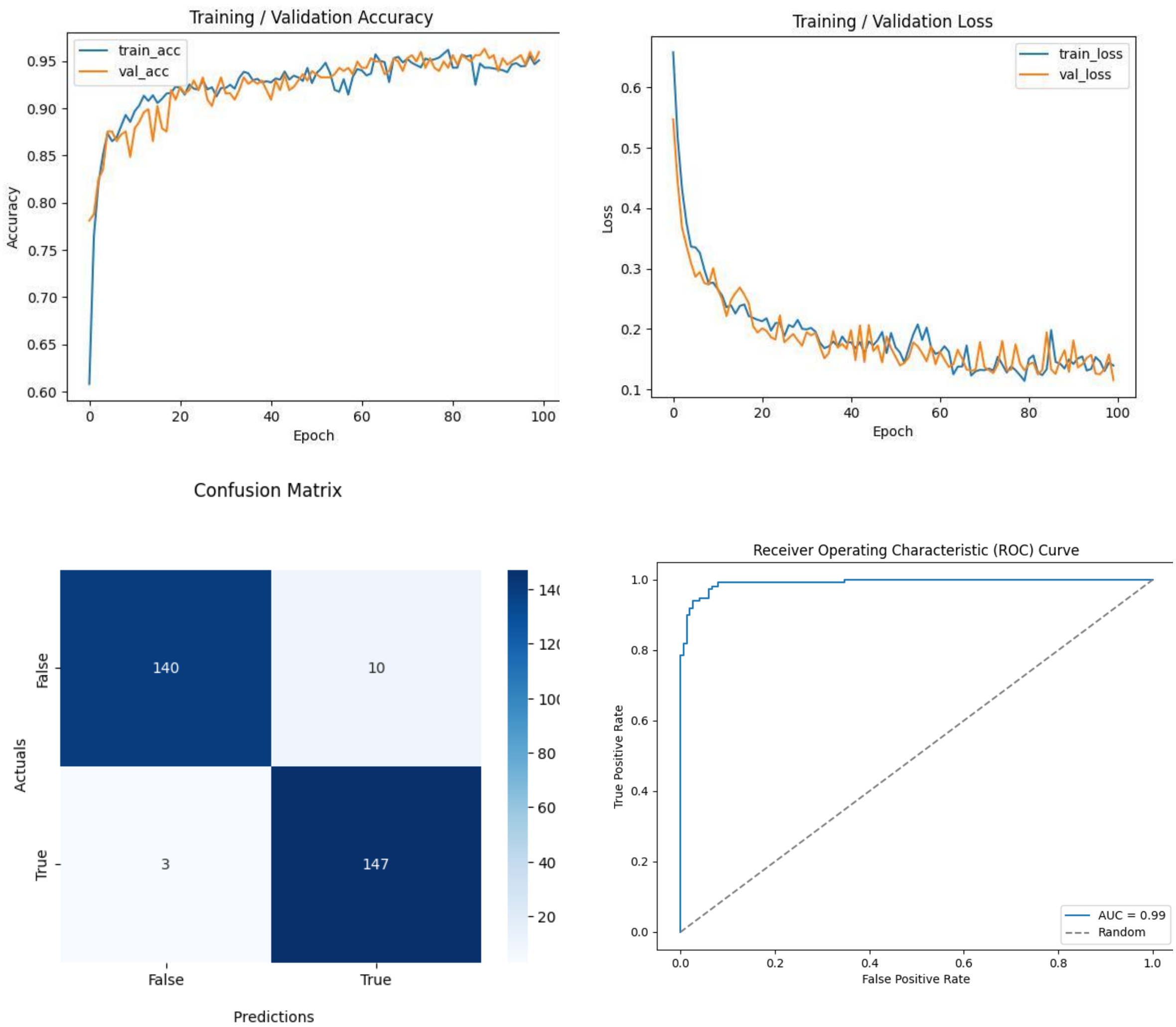
In the extended learning phase of 100 epochs, the VGGNet19 architecture shows a higher training performance. The table below (Table 16) provides an overview of the outcomes that were achieved.

Table 16
VGG19 Performance Evaluation Summary II.

Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	98.93
	Loss	04.95
Validation	Accuracy	98.67
	Loss	05.59
Testing	Accuracy	95.67
	Loss	11.80

The model learnt effectively from the training data, as evidenced by a high training accuracy of 98.93%. On the training set, the model's predictions appear to be reasonably confident, based on the training loss of 04.95%. Although it is not as high as the training accuracy, the validation accuracy is still rather good at 98.67%. The model shows good generalization to previously encountered data, as seen by the validation loss value of 05.59%, which is consistent with the training loss. The testing accuracy is lower than the training and validation accuracies, at 95.67%. The testing loss of 11.80 is significantly lower than the testing loss of other baseline models, although it is larger than the training and validation losses. In addition to assessing accuracy and loss metrics, a comprehensive evaluation of the VGGNet19 model's performance was conducted using a confusion matrix (See Figure 22).

Figure 22
VGGNet19 Performance Evaluation II.



Modified DenseNet1.0

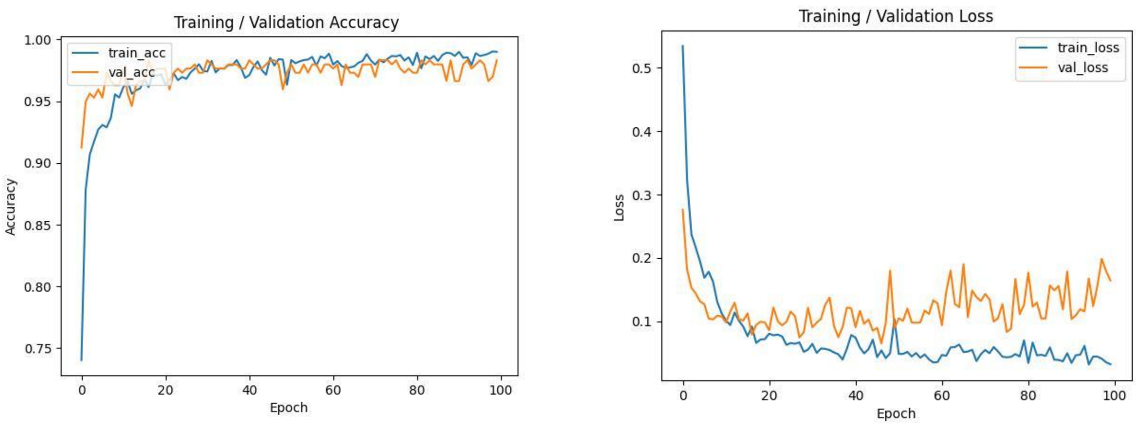
In the extended learning phase with 100 epochs, the Modified DenseNet1.0 architecture showcased remarkable performance across training, validation, and testing sets (See Table 17).

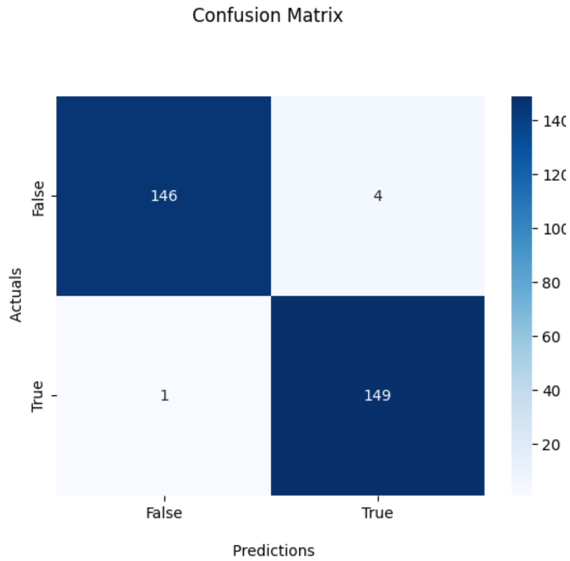
Table 17
Modified DenseNet1.0 Performance Evaluation Summary II.

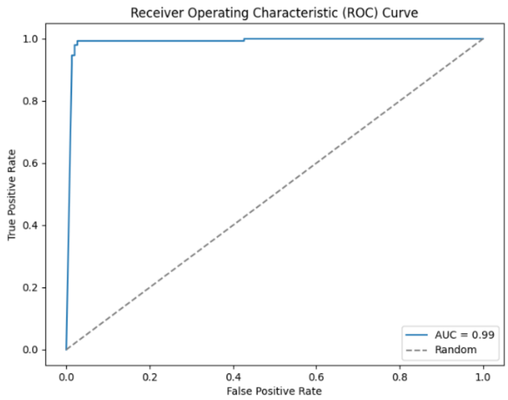
Learning Phase	Evaluation Metrics	Value (%)
Training	Accuracy	99.93
	Loss	0.39
Validation	Accuracy	99.78
	Loss	02.08
Testing	Accuracy	98.33
	Loss	16.28

Modified DenseNet1.0, our proposed model, demonstrates exceptional performance across various evaluation metrics. Our model achieves an outstanding training accuracy of 99.93% and a low loss of 0.39, exhibiting how well it learns from the training batch of data. Modified DenseNet1.0 maintains a high accuracy of 99.78% during the validation phase, closely aligning with the training performance (See Figure 23), and a low loss of 02.08%, suggesting robust generalization to unseen data. The testing phase further validates the model's capabilities with a remarkable accuracy of 98.33%, although accompanied by a slightly high loss value of 16.28.

Figure 23
Modified DenseNet1.0 Performance Evaluation II.







This overall performance underscores the effectiveness of Modified DenseNet1.0 in achieving accurate classifications and generalizing well to new, independent datasets, making it a promising model for the targeted task. Continued monitoring and future adjustments such as advanced image processing and ensemble learning will surely enhance its performance.

The model performs outstandingly, according to the evaluation findings based on precision, recall, and F1-score. With a 98% overall accuracy rate, the model successfully distinguished between images that had been classified as tumorous and those that were not. Consistently outstanding precision and recall levels point to a well-balanced model. The model achieves 99% precision and recall for non-tumorous images and 97% precision and 99% recall for the tumorous ones. The F1-score value of 98% for both classes, which combines recall and precision, shows our exceptional model's efficacy. There are just 4 erroneous positives and 1 false negative in the confusion matrix, which shows how rarely the model misclassifies data. These findings collectively imply that the model does a very good job of correctly classifying images. Overall, these results suggest that the model performs exceptionally well in precisely recognizing images of both classes, making it a robust solution for the given task.

Tables 18 and 19 provide a thorough summary of the second phase of experiment findings and a comparative analysis of the CNN model's performance on various classification approaches.

Table 18
Selected Baseline Models Learning Evaluation II

Learning phase Epochs=100		Training		Validation		Testing	
Evaluation Metrics %		Accuracy	Loss	Accuracy	Loss	Accuracy	Loss
Classification Architectures	ResNet50	88.15	30.85	88.41	30.78	87.33	33.66
	VGGNet19	98.93	04.95	98.67	05.59	95.67	11.80
	VGGNet16	99.18	02.10	98.56	04.25	94.33	20.60
	Modified DenseNet1.0	**99.93**	**0.39**	**99.78**	**02.08**	**98.33**	**16.28**

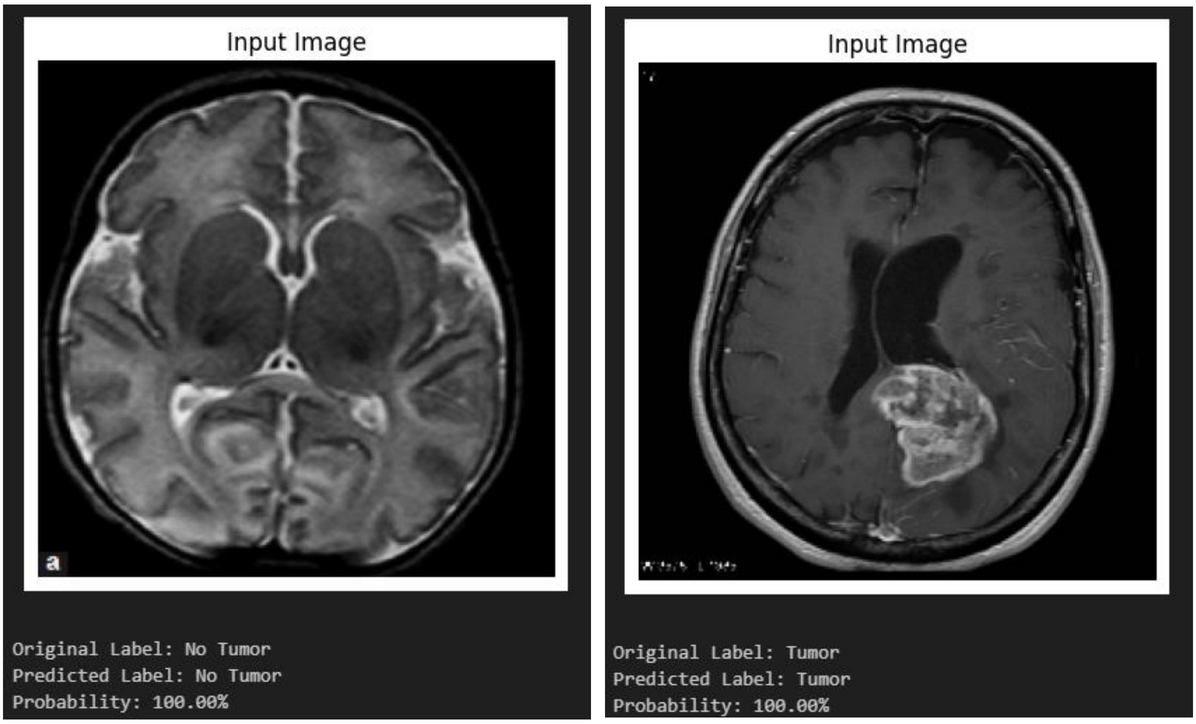
Table 19
Selected Baseline Models Evaluation metrics on 100 epochs.

Evaluation Metrics % Epochs=100		Class	Precision	Recall	F1-score	Sensitivity	Specificity	AUC
Classification Architectures	VGGNET16	Yes	95	94	94	94	94.67	99
		No	94	95	94			
	VGGNET19	Yes	94	98	96	98	93.33	99
		No	98	93	96			
	ResNet50	Yes	90	84	87	84	90.67	94
		No	85	91	88			
	Modified DenseNet1.0	**Yes**	**97**	**99**	**98**	**99.33**	**97.33**	**99**
		No	**99**	**97**	**98**			

Testing Process

We tested the obtained Model versions on unseen data which is an MRI Image of the brain that has no Tumour with the indication of the probability percentage of a Tumour's existence or non-existence. We got the following results presented in Figure 24.

Figure 24
One-Simple prediction on unseen data using targeted images.



We have also run a group test of 15 samples and we have obtained a very satisfying results compared to other baseline models tested on the same group test.

Figure 25
15-Simples predictions on unseen data using testing data [:15]

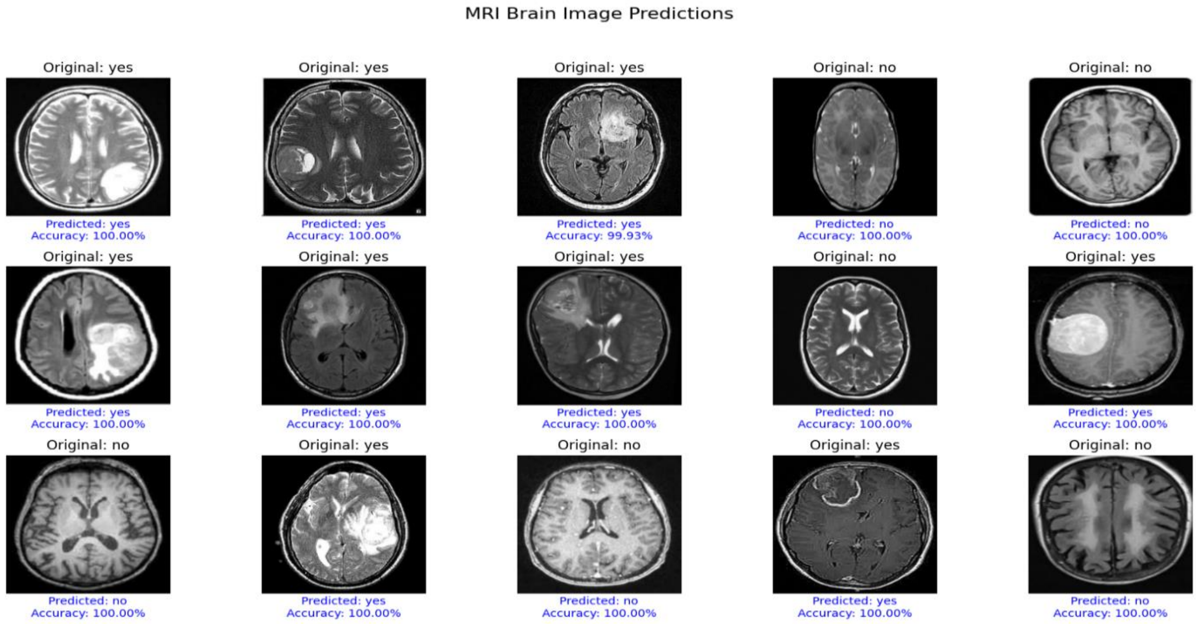
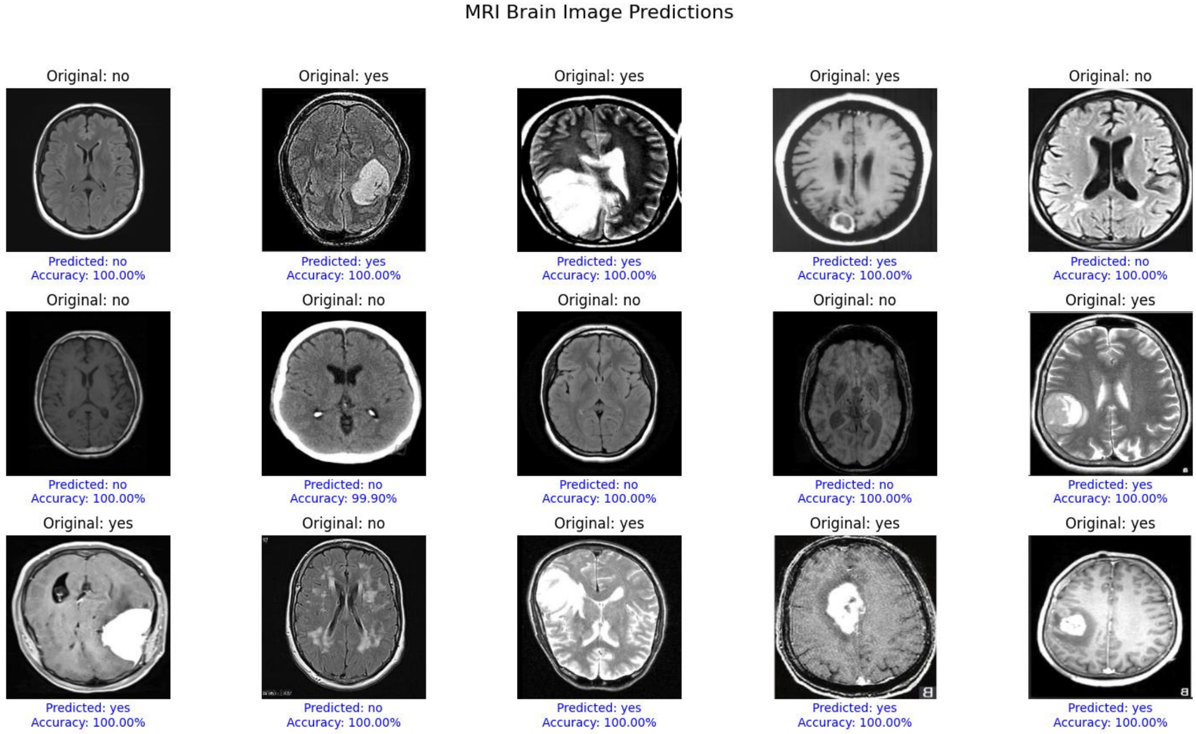


Figure 26

15-Simples predictions on unseen data using testing data [Random]



We have then run a group test of 20 samples and we have obtained a very satisfying results compared to other baseline models tested on the same group test.

Figure 27
20-Simples predictions on unseen data using testing data [:20]

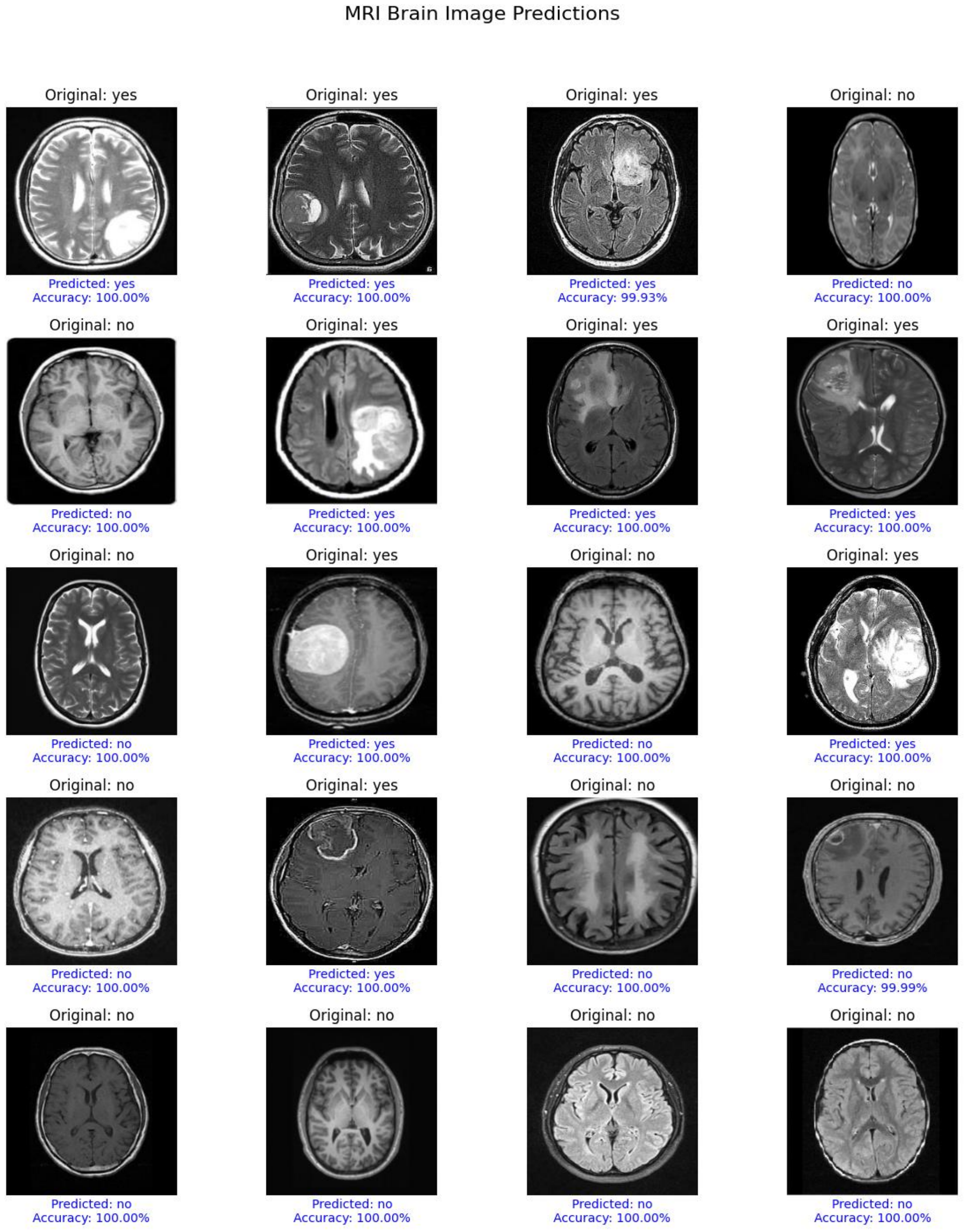
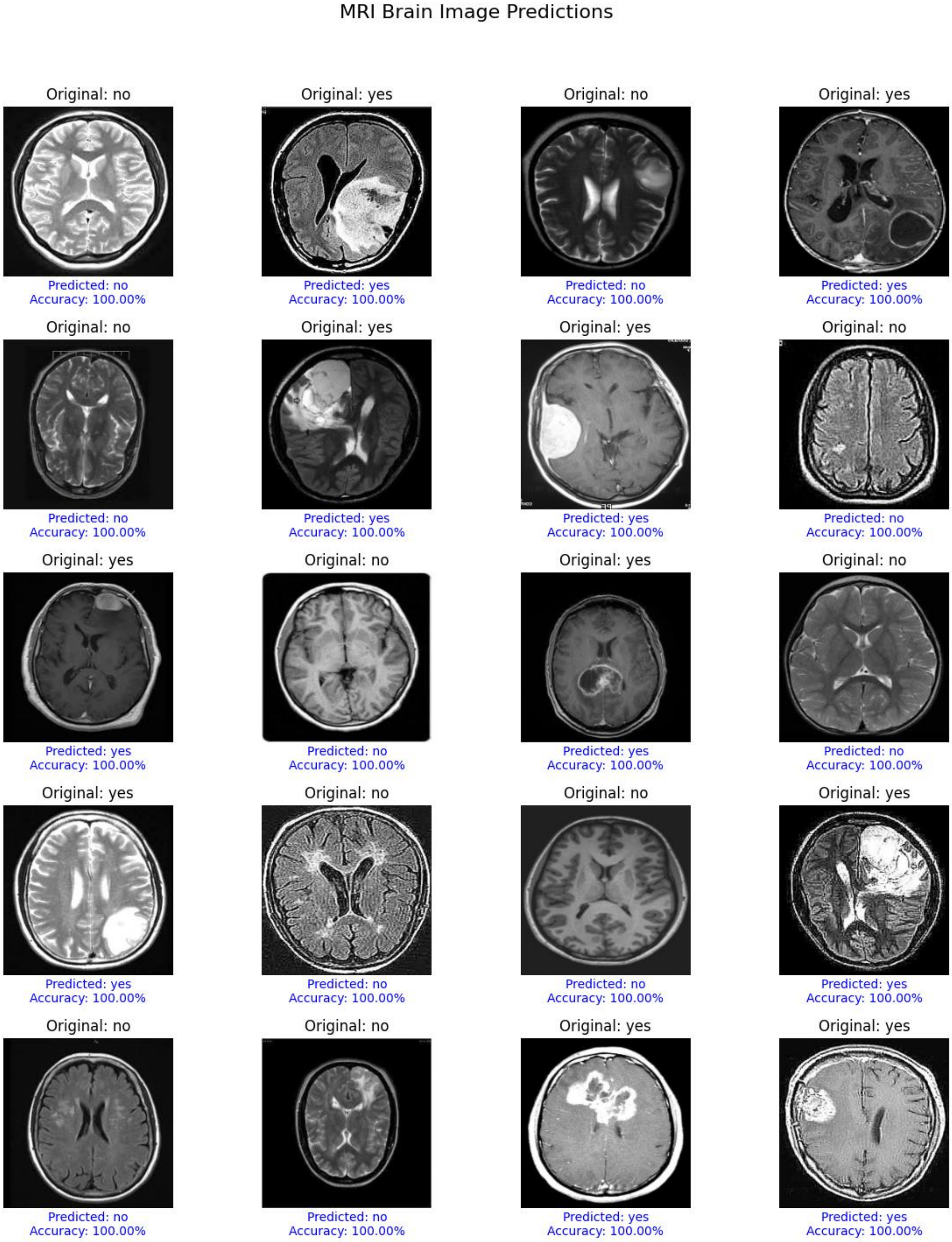


Figure 28
20-Simples predictions on unseen data using testing data [Random]

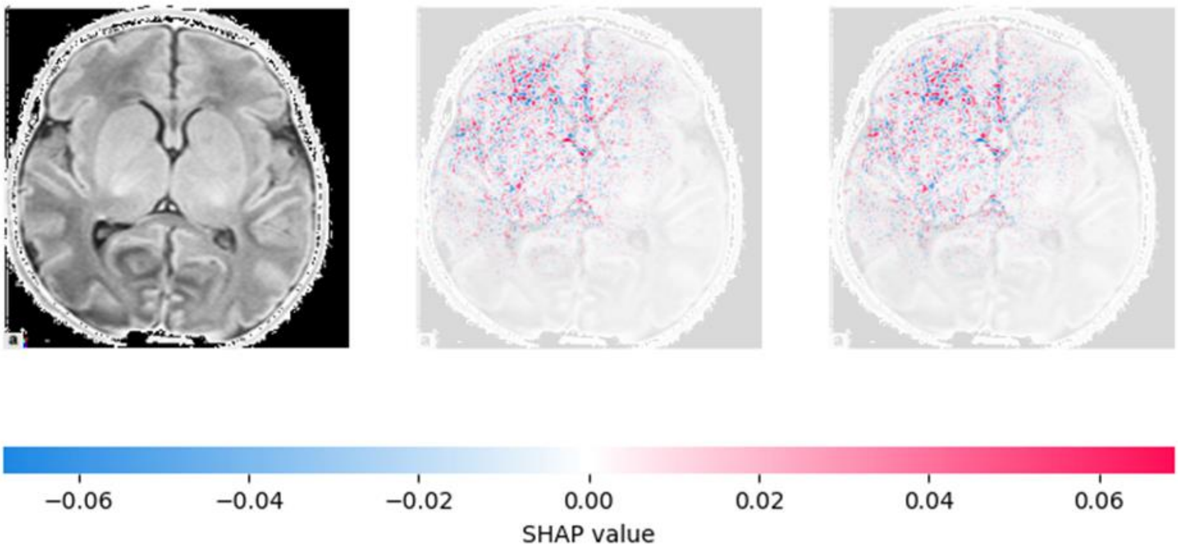


Explainability Process

Based on these results, we have decided to carry on our research utilizing mainly the results obtained from Modified DenseNet1.0 and adding explainability to it using the Shapely values technique paving the road towards our upcoming version Modified DenseNet2.0. ML models are powerful but hard to interpret. However, SHAP values are a promising tool to help understand how the model features impact predictions.

The graphs of SHAP values provide essential details about how comprehensible the predictions made by our model are. These visuals highlight how every pixel in the input MRI image contributes to the final prediction produced by our model, "Modified DenseNet1.0". We have generated an initial demonstration of what to expect from shapely value-based explanation for our model's prediction testing on tumorous and non-tumorous brain MRI images (See Figures 29 to 32) with two focused areas: Shapely values and brain-focused Image plots for more visual explainability. The graphs that follow show how each feature contributes to the overall prediction obtained by testing an unseen data input to our ongoing under development "Modified DenseNet2.0".

Figure 29
Non-Tumorous Shapely Values-focused Image Plot.



Brain-focused Image Plots provide an overlay of the Shapley values on the original MRI image. This combined result allows for a direct visual correlation between the image features and their respective Shapley values. Areas with more intense coloration signify regions with greater influence on the model's decision. This graph serves as a comprehensive representation, aiding in the identification of anatomical structures or abnormalities that contribute significantly to the predictive outcome.

Figure 30
Non-Tumorous Brain-focused Image Plot.

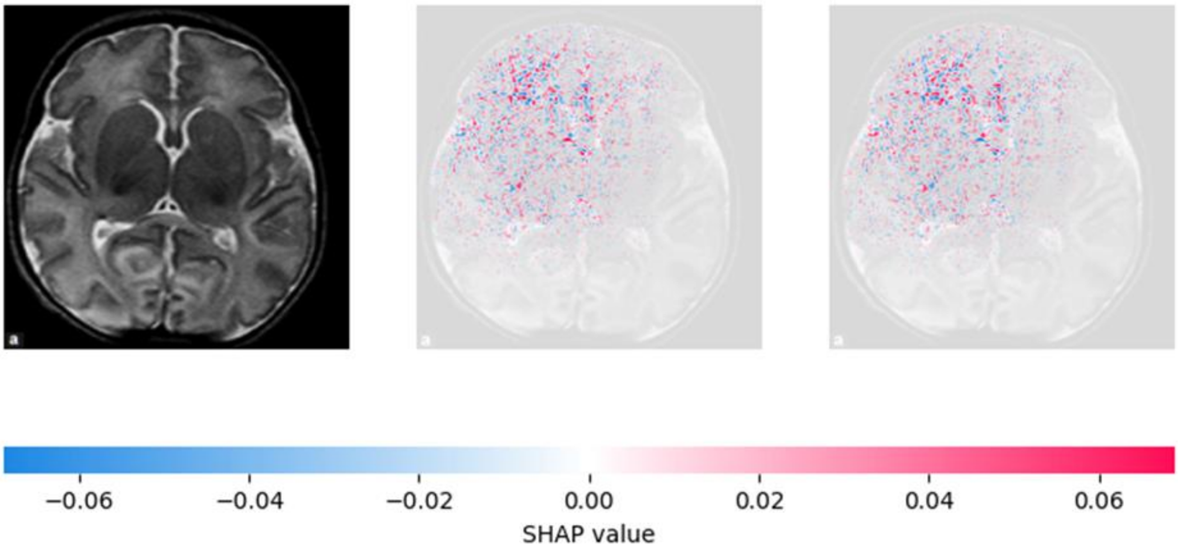
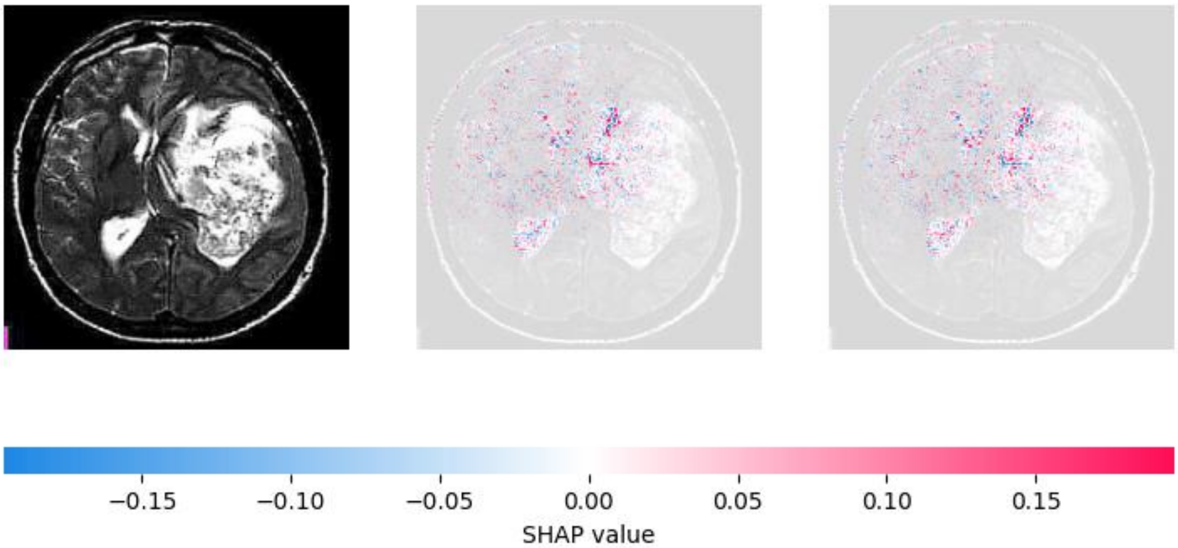
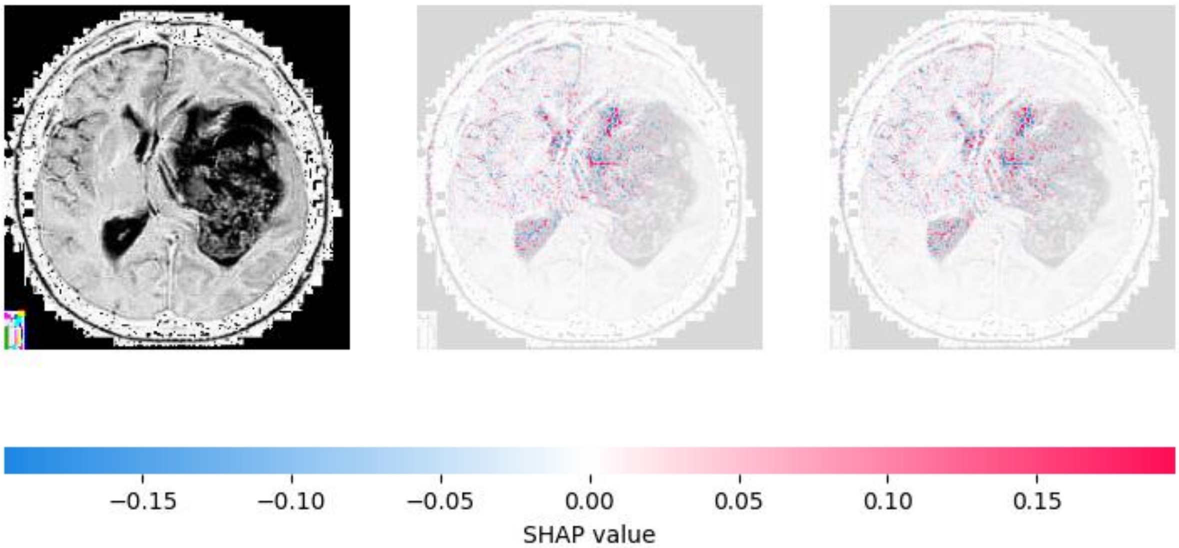


Figure 31
Tumorous Shapely Values-focused Image Plot.



Shapely Values-focused Image Plots illustrate the individual pixel contributions to the model's prediction. Each pixel's Shapley value is depicted, representing its impact on the prediction. Areas with higher absolute Shapley values suggest a greater influence on the model's decision. Interpretation of this graph involves scrutinizing regions where Shapley values peak or trough, offering an explainability of the image features most involved in the model's decision-making process.

Figure 32
Tumorous Brain-focused Image Plot.



Model deployment

As part of our model's performance demonstration, we have deployed Modified DenseNet1.0 to the health web application run by Near East University Health AI-IoT platform.

Once we load the Health AI-IoT application and login successfully, we click on "Image Diagnosis" from the dashboard. Next, we will get various options to check. We select "Brain Cancer" button. Then, we get the first web page that will allow the user to input their MRI image through a file browser feature (See Figure 33). The next step is to click on "Detect Tumor" button (See Figure 34).

Figure 33
Uploading Images through files browser.

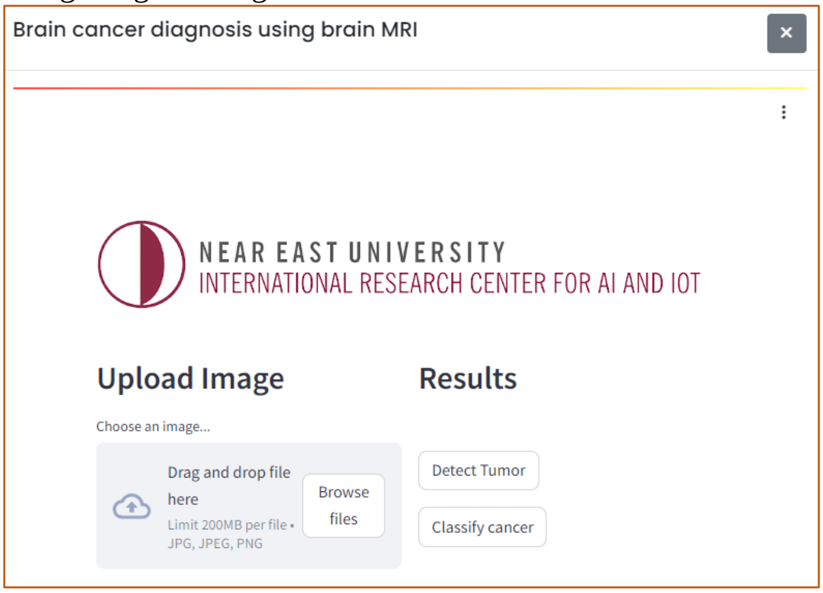
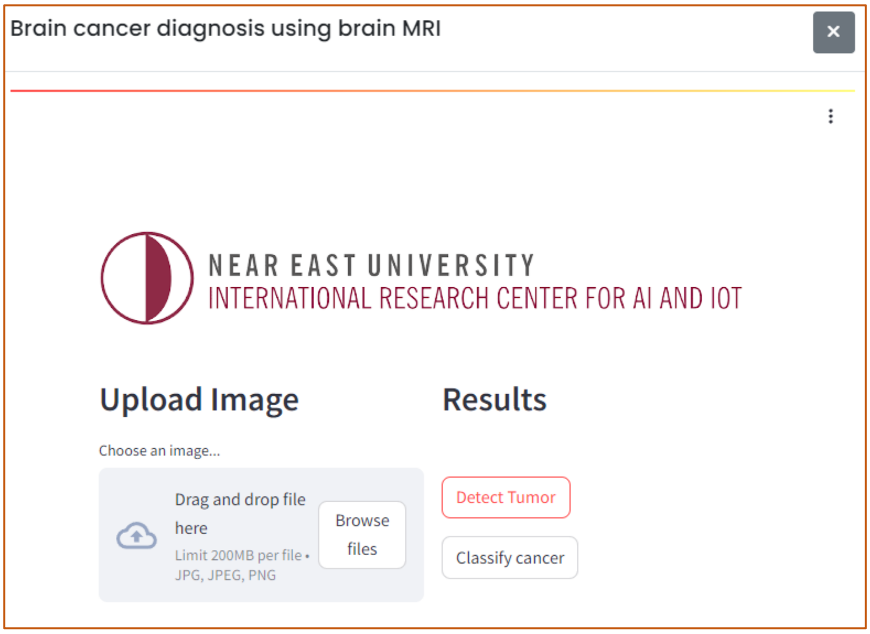


Figure 34
“Detect Tumor” Button.



Once the input image is uploaded to the system, we press on “Detect Tumor” button and wait for the results (See Figure 35 and Figure 36).

Figure 35

Image uploaded through files browser.

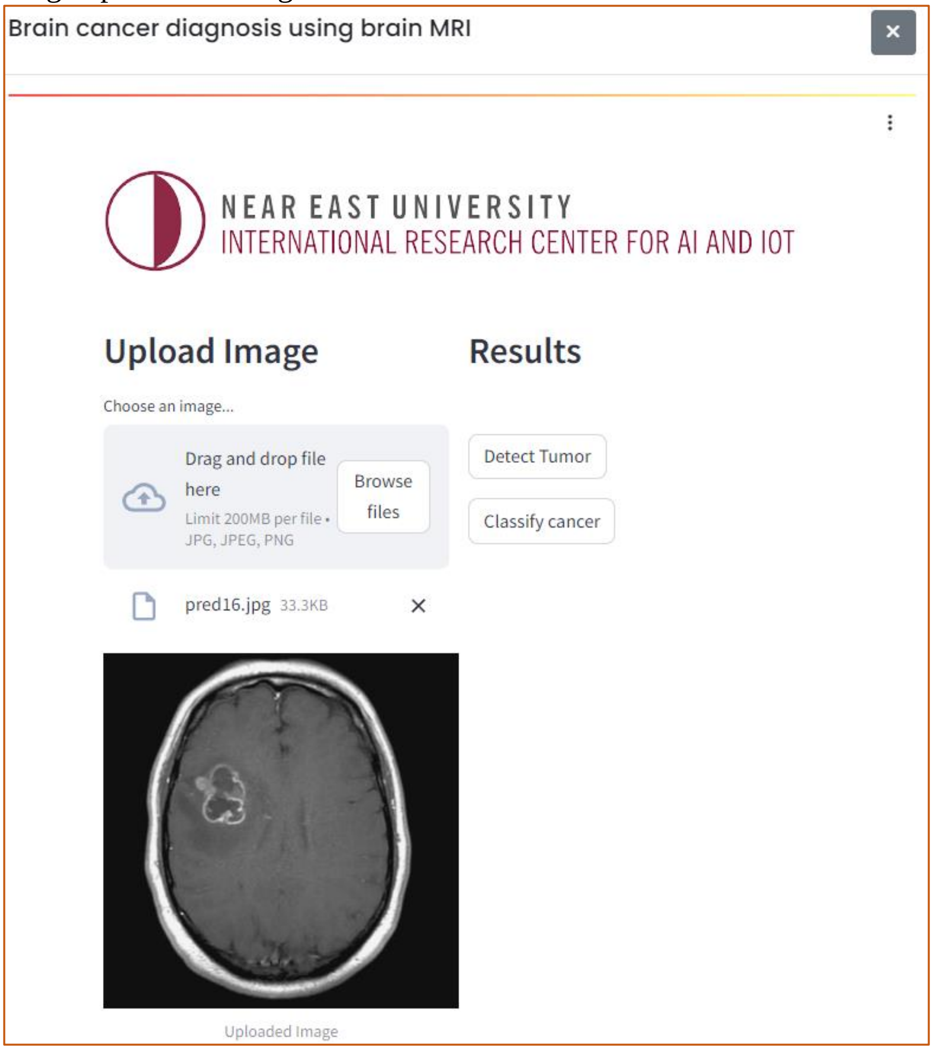
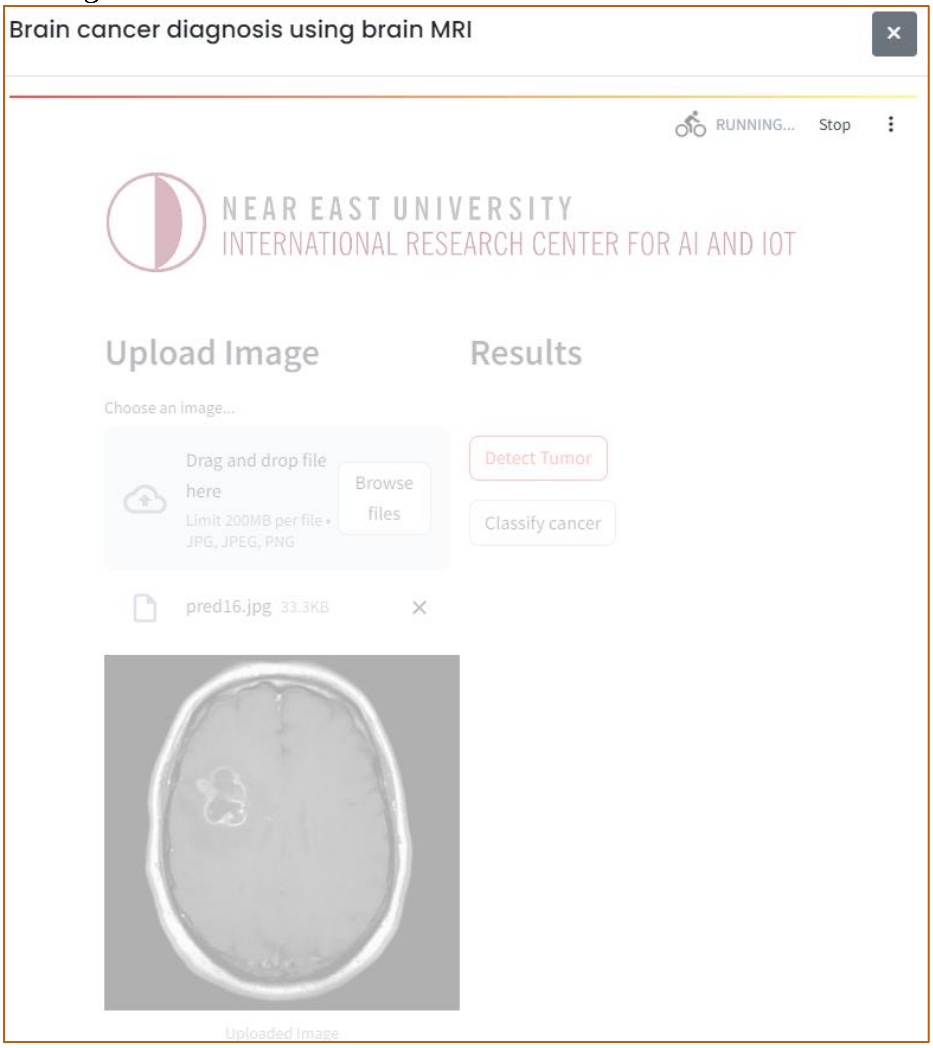


Figure 36
Waiting for Tumour Detection results.



The results will be out in few seconds, as demonstrated in the following figures (See Figure 37 and Figure 38).

Figure 37 demonstrates a prediction test conducted over unseen MRI image input by the user. The output shows detection of Tumour in the MRI image with a confidence probability of 100%.

Figure 37
Tumour Detection results.

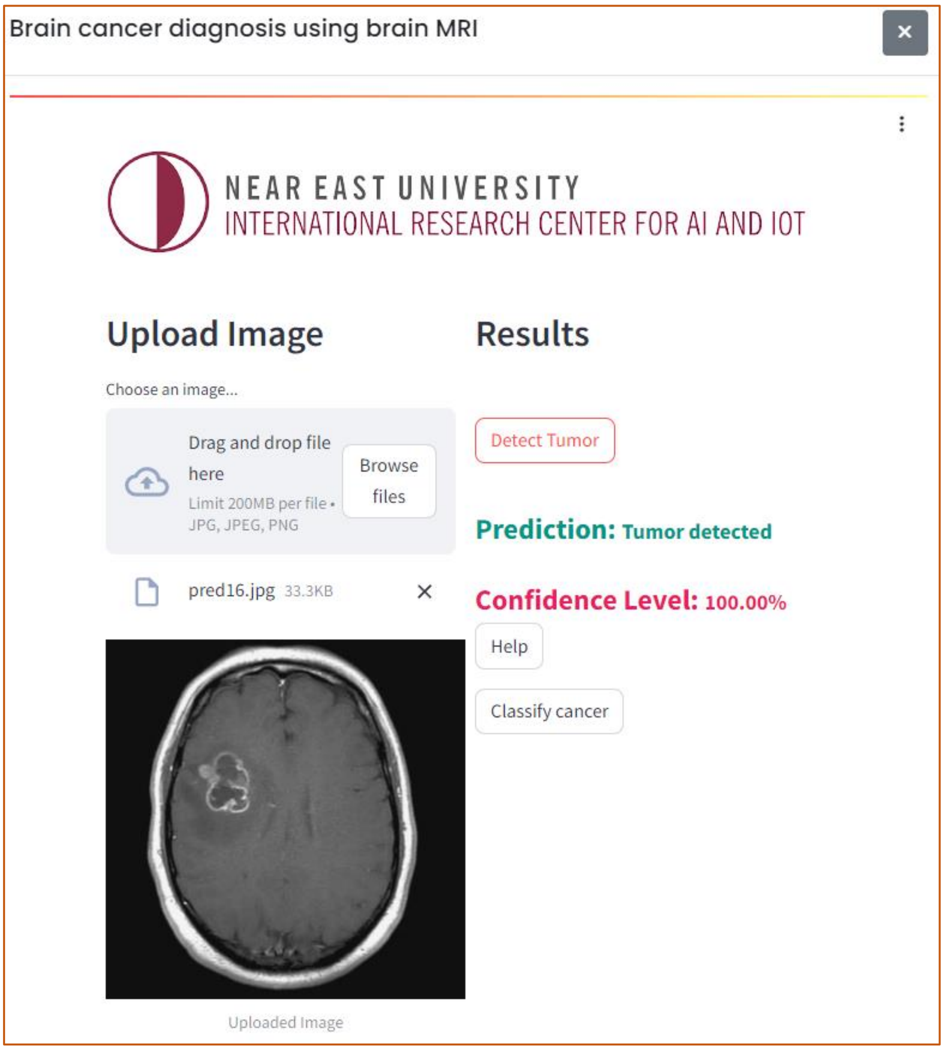
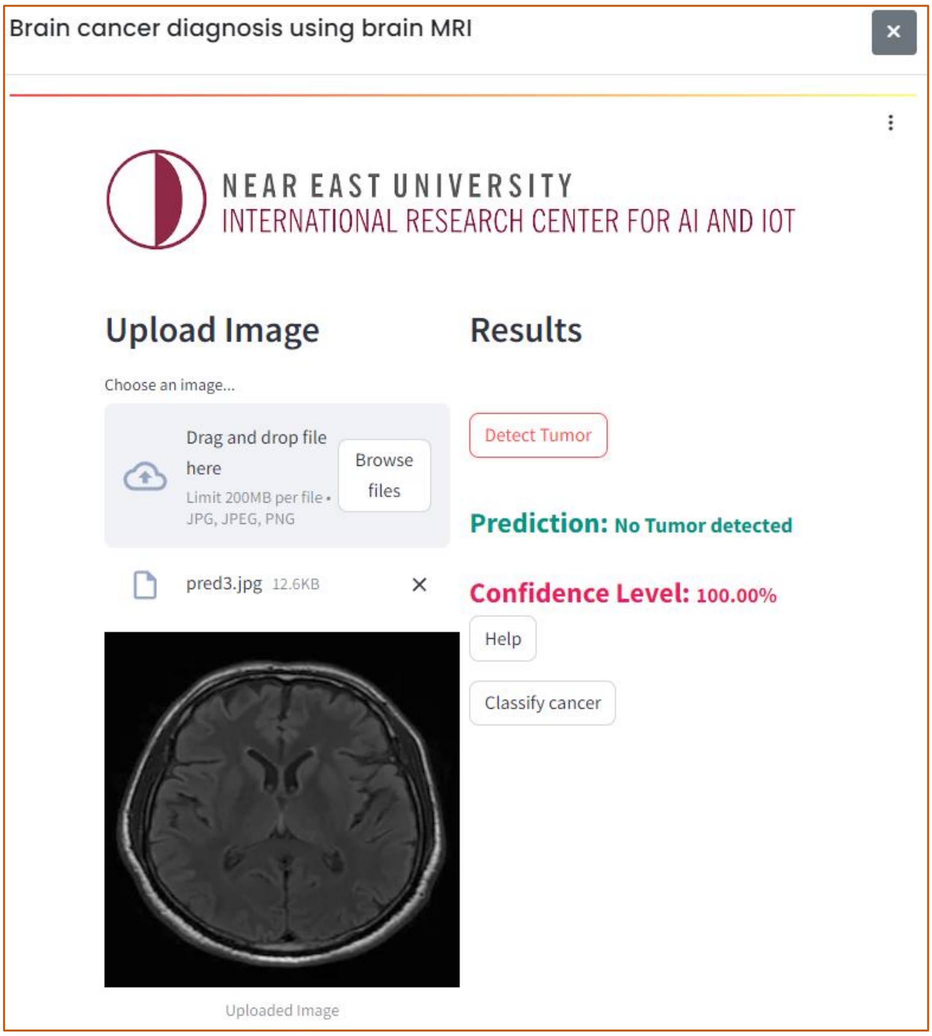


Figure 38 demonstrates a prediction test conducted over another unseen MRI image input by the user. The output shows no detection of Tumour in the MRI image with a confidence probability of 100% as well.

Figure 38

No-Tumor Detection results.



CHAPTER V

Discussion

Recently, several models, algorithms, and frameworks were presented to conduct different types of classification, binary, and multiple classes. The proposed architecture AlexNet achieved an accuracy of 80% whereas GoogleNet achieved a slightly higher accuracy of 86%, coming to VGGNet16 which has achieved 86.1 %. DenseNets utilized in our research reported an approximate training accuracy of 94.37% on the CIFAR-100 dataset and 97.05% on the CIFAR-10.

Using medical MRI images in our experiment and testing on two different datasets one that is balanced and small in size, and the other one is imbalanced and large we can confidently conclude that data augmentation is an effective solution when it comes to covering lack of data issue. It was sufficient to enhance the accuracy of our model reaching 99.56% with the first dataset trained on 100 epochs. (See Table 4 and Table 5).

We used the second dataset excluding data augmentation due to data sufficiency and trained the proposed model Modified DenseNet1.0, achieving 99.93% accuracy which we believe provided reliable and stable results based on the predictions mentioned previously and the confusion matrix results that surpassed the other baseline models' out noticeably on an overall observation regardless the high accuracy showcased by a first-class winner.

We stress on our model's output that achieved the best training accuracy trained on 100 epochs and the best testing loss accuracy trained on 50 epochs. In fact, we examine the intricate interplay between training and testing epochs in our model development. The achievement of optimal training accuracy after 100 epochs underscores the model's adeptness at assimilating complex patterns within the training data, demonstrating its proficiency through iterative learning cycles. The strategic decision to conclude training at this juncture aligns with the goal of ensuring a comprehensive grasp of the training dataset's complexity. Simultaneously, the realization of optimal testing loss accuracy after 50 epochs highlights the model's capabilities in generalizing to new, unseen data, reflecting its ability to recognize unseen data effectively. This dualistic strategy, combining extensive learning from the training dataset and judicious evaluation on a distinct dataset, enhances the model's overall efficacy and adaptability. The nuanced choice of different epoch counts for training and testing serves as a deliberate measure to supress the risk of overfitting, achieving in a well-balanced model capable of both precision and generalization.

Figure 39

Training accuracy comparison of Modified DenseNet1.0 and other baseline models, Epochs=100.

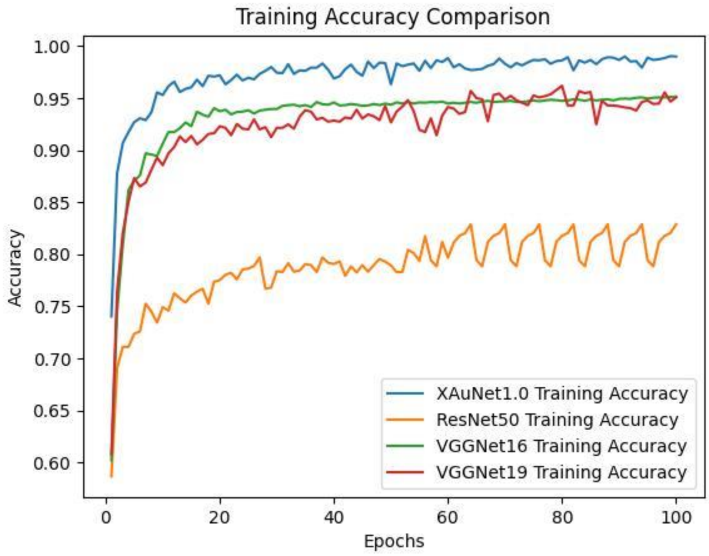


Figure 40

Training loss comparison of Modified DenseNet1.0 and other baseline models, Epochs=100.

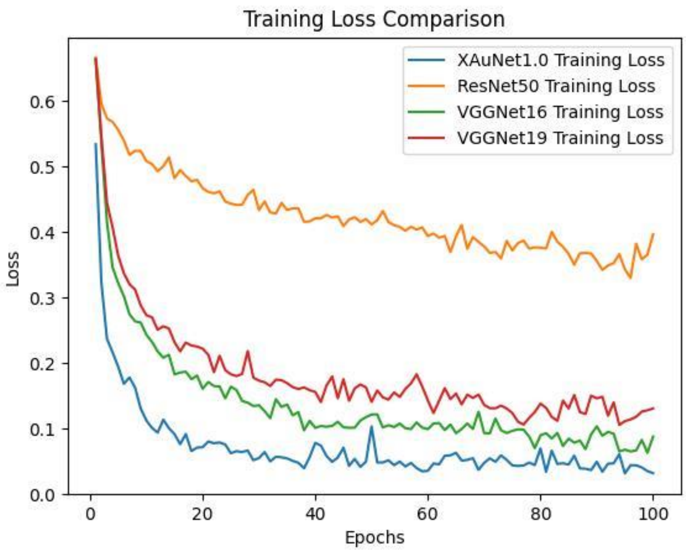


Figure 41

Training accuracy comparison of Modified DenseNet1.0 and other baseline models, Epochs=50.

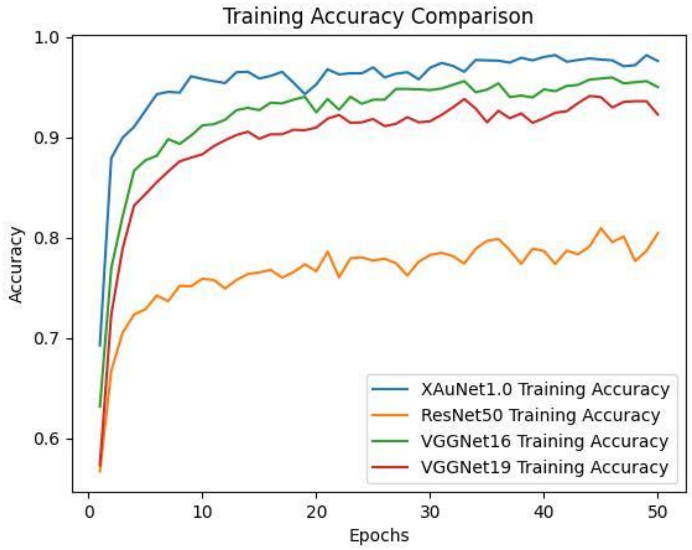
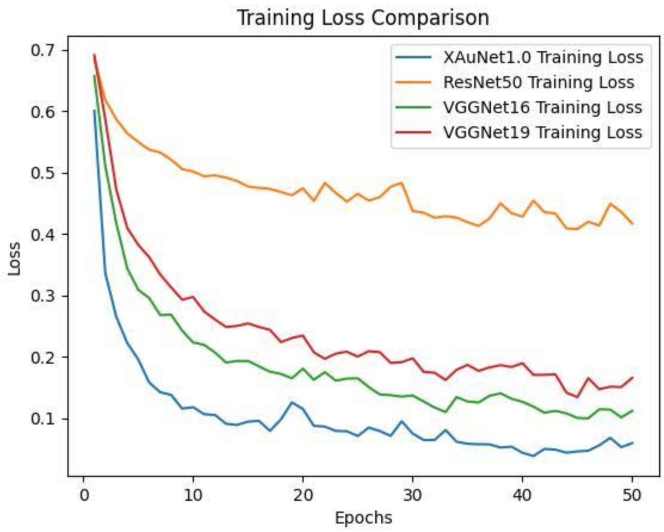


Figure 42

Training loss comparison of Modified DenseNet1.0 and other baseline models, Epochs=50.



For more clarity and interpretability, we introduce our upcoming explained model "Modified DenseNet2.0" trained on 3000 samples of pre-processed MRI images of tumorous and non-tumorous brain states using DenseNets and shapely values for explainability and visual explanations. The second objective of this research is to demonstrate to healthcare professionals how the model we recommend could be explained while retaining high accuracy. Consequently, the implemented classifier yielded an exceptional accuracy of 99.93% in contrast to the most recent approaches, as indicated in the subsequent table (Table 20).

Table 20

Comparison of our proposed model and other Deep Learning-based models in previous studies.

Author	Method	Accuracy
Ge et al., 2020.	GAN	88.82 %
Vimal Kurup et al., 2020	PNN-CNN	90%
Dehkordi et al., 2022	CNN	97.4%
Haq et al., 2022	CNN	97.96%
Zeineldin et al., 2022	ResNet-50	98.62 %
Our proposed model Modified DenseNet1.0	**DenseNets**	**99.93 %**

Anomalies or abnormalities in the brain that significantly influence the model's prediction are likely to manifest as pronounced peaks or troughs in the shapely values-focused graph (See Figures 29 and 31). Whereas the brain-focused graph offers a more intuitive visualization, allowing healthcare professionals to directly relate model predictions to specific regions of interest within the brain (See Figures 30 and 32).

It is essential to acknowledge that while SHAP values provide valuable explainability, they are not exhaustive in capturing the full complexity of neural network decision-making. Variability in interpretations may arise, and the model's predictions should always be cross-referenced with clinical expertise for comprehensive diagnosis.

In summary, the SHAP values graphs aim to strengthen the transparency of classification model's decision-making process, offering a nuanced understanding of the features influencing predictions. These visualizations not only provide insights for model developers but also hold the potential to facilitate communication between AI practitioners and healthcare professionals in the context of brain Tumour classifications and predictions.

CHAPTER VI

Conclusion and Recommendations

Our work embraces a long term journey to explore the depths of eXplainable Artificial Intelligence (XAI) for brain tumour related classification studies, utilizing a deep transfer learning approach based on the DenseNets' architecture. Our primary objective was to develop a model capable of conducting a binary classification task applied on brain tumour images obtained from MRI scans, emphasizing transparency and credibility in the decision-making process through in-depth analysis. Leveraging Shapley values for explainability, our objective is to establish a connection between the Deep Learning (DL) model's mysterious core and the requirement for trust, credibility and reliability in its predictions, especially in the context of medical diagnosis.

Our proposed model demonstrated remarkable performance in binary classification, achieving an accuracy of 99.70% and 99.93% on two distinct epochs values (50 and 100) with varying characteristics. Notably, the model's stability and reliability were evaluated employing several performance measures, including f1-score, recall, specificity, sensitivity, and accuracy, and the Area Under the Receiver Operating Characteristic Curve. The superior performance on a balanced dataset (dataset II) underscored the importance of data balance in training robust and reliable models for medical image classification tasks.

The integration of Shapley's values into our model allows for a complete awareness of the process of making decisions. By showcasing the involvement of each feature to the model's predictions, we aim to enhance transparency and build trust in the model's outcomes. The visualized explanations provided by Shapley values not only serve as means of interpretation but also contribute to the legitimacy of our model in the medical industry.

Considering the observed advantages of our proposed model Modified DenseNet1.0 on dataset II, which boasts a balanced composition, we propose its deployment for real-world tests and live usage. Cloud platforms such as Microsoft Azure, Heroku, IBM Cloud, and EC2 Amazon Web Services offer convenient avenues for deploying ML models. This would facilitate users worldwide in testing our model by simply uploading a brain image to a server, thereby expanding the accessibility and impact of our research. It is essential to outline potential avenues for future research. Expanding the dataset diversity, investigating the model's generalizability to different populations, and exploring explainability and interpretability techniques are avenues that require exploration (Modified DenseNet2.0). In fact, scenarios where individual users have to trust and comprehend the decisions made by AI systems, and when the stakes are high, are where XAI really shines. This is especially relevant in fields like healthcare, finance, criminal justice, and any domain where AI systems impact human lives and well-being (Benyamina et al., 2022). Moreover, continuous collaboration with medical professionals and the integration of evolving technologies will contribute to the ongoing enhancement of AI-assisted medical diagnosis systems to address difficulties related to manual evaluation of multiple MRI images generated in a clinic, a hospital, a major medical centre, a teaching hospital, a tertiary care hospital, academic medical centres and oncology centres. Consequently, the need for more accurate computer-based tumour detection techniques, notably if we are dealing with IoT health care systems. In spite of that, it is critical to recognize the limitations and moral consideration related to deploying AI models in the medical domain. While our model demonstrates high accuracy and robustness, it is imperative to emphasize that it should never substitute professional medical advice or diagnosis. A collaborative approach, wherein AI serves as a decision-making aid for healthcare professionals, is the ideal scenario. The significance of continual oversight by medical specialists cannot be overstated, ensuring that the AI system augments human capabilities rather than replacing them.

As we conclude this thesis, it is essential to outline potential avenues for future research. Expanding the dataset diversity, investigating the model's generalizability to different populations, and exploring additional explainability techniques are avenues that warrant exploration. Moreover, continuous collaboration with medical professionals and the integration of evolving technologies will contribute to the ongoing enhancement of AI-assisted medical diagnosis systems.

Our study has successfully trained a densely connected DL model for binary brain tumour classification, showcasing the efficacy of the DenseNet-based transfer learning approach. The incorporation of explainability through Shapley values adds a layer of transparency, addressing a critical aspect in the deployment of AI models in medical settings and Internet of Things accessible globally by pubic and professional users. While our findings are promising, they mark a stepping stone in the broader landscape of AI-assisted medical diagnosis, emphasizing the need for responsible integration and ongoing research.

References

Cooper, S B, and J Van Leeuwen. The Selected Works of A.M. Turing: His Work and Impact. Waltham, Ma, Elsevier, 2012.

Oded Maimon, & Lior Rokach. (2010). *Data Mining and Knowledge Discovery Handbook*. Springer Us.

Han, J., Pei, J., & Kamber, M. (2006). *Data Mining, Southeast Asia Edition*. Elsevier.

Aggarwal, C. C. (2016). *Data mining: the textbook*. Springer.

Del Campo Barraza, S. M. (2017). Unsupervised Feature Learning Applied to Condition Monitoring. In https://www.ltu.se/. Luleå University of Technology, Graphic Production. https://www.diva-portal.org/smash/get/diva2:1090081/FULLTEXT01.pdf

Chatterjee, S., Sciarra, A., Dünnwald, M., Tummala, P., Agrawal, S., Aishwarya Jauhari, Kalra, A., Steffen Oeltze-Jafra, Speck, O., & Nürnberger, A. (2023). StRegA: Unsupervised Anomaly Detection in Brain MRIs using Compact Context-encoding Variational Autoencoder. ArXiv (Cornell University). https://doi.org/10.58530/2022/0172

Iqbal, H., Khalid, U., Hua, J., & Chen, C. (2023). Unsupervised Anomaly Detection in Medical Images Using Masked Diffusion Model. ArXiv. /abs/2305.19867

Iromi R Paranavithana, Thashika D Rupasinghe, and Daniel D Prior, "Unsupervised Learning and Market Basket Analysis in Market Segmentation," Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2021, 7-9 July, 2021, London, U.K., pp122-127

Joppe, D. (2016). OVERNIGHT PLEASURE TRAVELLERS' ACTIVITIES IN CANADA: USING ASSOCIATION RULE MINING TO DISCOVER RELATIONSHIPS. In Scholar Works. Travel and Tourism Research Association: Advancing Tourism Research Globally. 21.

Solan, Z., Horn, D., Eytan Ruppin, & Edelman, S. (2005). Unsupervised learning of natural languages. Proceedings of the National Academy of Sciences of the United States of America, 102(33), 11629–11634. https://doi.org/10.1073/pnas.0409746102

Chen, Y., Mancini, M., Zhu, X., & Akata, Z. (2022). Semi-Supervised and Unsupervised Deep Visual Learning: A Survey. IEEE Transactions on

Pattern Analysis and Machine Intelligence, 1–23. https://doi.org/10.1109/tpami.2022.3201576

François-LavetV. (2018). *An introduction to deep reinforcement learning*. NOW.

Sutton, R. S., & Barto, A. G. (2014). *Reinforcement Learning: An Introduction*.

Gamble, C., & Gao, J. (2018, August 17). Safety-first AI for autonomous data centre cooling and industrial control. Google DeepMind; Google. https://deepmind.google/discover/blog/safety-first-ai-for-autonomous-data-centre-cooling-and-industrial-control/

Paulus, R., Xiong, C., & Socher, R. (2017). A Deep Reinforced Model for Abstractive Summarization. ArXiv. /abs/1705.04304

Choi, E., Hewlett, D., Google, J., Polosukhin, I., Lacoste, A., & Berant, J. (n.d.). Coarse-to-Fine Question Answering for Long Documents. Retrieved January 1, 2024, from https://www.cs.utexas.edu/~eunsol/files/papers/acl17eunsol.pdf

Grissom II, A., He, H., Boyd-Graber, J., Morgan, J., & Daume III, H. (2014). Don't Until the Final Verb Wait: Reinforcement Learning for Simultaneous Machine Translation. In Empirical Methods in Natural Language Processing. https://users.umiacs.umd.edu/~jbg/docs/2014_emnlp_simtrans.pdf

Yu, C., Liu, J., & Nemati, S. (2020). Reinforcement Learning in Healthcare: A Survey. https://arxiv.org/pdf/1908.08796.pdf

Al Sallab, A. A., & Rashwan, M. A. (2012). *Self-learning machines using Deep Networks* (Vol. 5). International Journal of Computer Information Systems and Industrial Management Applications.

KröseB., & Van, P. (1996). *An introduction to neural networks*. University Of Amsterdam.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation* (2nd ed.). Upper Saddle River, N.J.: Prentice Hall.

Nwadiugwu, M. C. (2020). Neural Networks, Artificial Intelligence and the Computational Brain. *ArXiv*. /abs/2101.08635

Mehlig, B. (2021). *Machine Learning with Neural Networks, An Introduction for Scientists and Engineers*. Cambridge University Press.

Omondi, A. R., & Rajapakse, J. C. (2006). *FPGA Implementations of Neural Networks*. Springer Science & Business Media.

Dkhichi, F., & Oukarfi, B. (2014). Neural Network Training By Gradient Descent Algorithms: Application on the Solar Cell. *International Journal of Innovative Research in Science, Engineering and Technology*, *03*(08), 15696–15702. https://doi.org/10.15680/ijirset.2014.0308084

Dimililer, K. (2013). Backpropagation Neural Network Implementation for Medical Image Compression. *Journal of Applied Mathematics*, *2013*, 1–8. https://doi.org/10.1155/2013/453098

Qi, X., Chen, G., Li, Y., Cheng, X., & Li, C. (2019). Applying Neural-Network-Based Machine Learning to Additive Manufacturing: Current Applications, Challenges, and Future Perspectives. *Engineering*, *5*(4), 721–729. https://doi.org/10.1016/j.eng.2019.04.012

Vié, A. (2020). Qualities, Challenges and Future of Genetic Algorithms. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3726035

Zhang, B., Wu, Y., Lu, J., & Du, K.-L. (2011). Evolutionary Computation and Its Applications in Neural and Fuzzy Systems. *Applied Computational Intelligence and Soft Computing*, *2011*, 1–20. https://doi.org/10.1155/2011/938240

Baldominos, A., Saez, Y., & Isasi, P. (2019). Hybridizing Evolutionary Computation and Deep Neural Networks: An Approach to Handwriting Recognition Using Committees and Transfer Learning. *Complexity*, *2019*, 1–16. https://doi.org/10.1155/2019/2952304

Roberts, Lawrence G. *Machine Perception of Three-Dimensional Solids*. 1963. M.I.T. Press, Cambridge, MA., 1965, pp. pp.159-197, engr.case.edu/merat_francis/EECS%20490%20F06/References/Papers/Roberts%203D%20Perception.pdf. Accessed 1 Jan. 2024.

American Cancer Society. (2021). *American Cancer Society*. Cancer.org; American Cancer Society. https://www.cancer.org/

Cancer.net. (2019). Cancer.net. https://www.cancer.net/

Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*(10), 464–470. https://doi.org/10.1016/j.tics.2006.08.004

Bromberger, S., Chicago, & London. (1992). *On What We Know We Don't Know Explanation, Theory, Linguistics, and How Questions Shape Them*.

http://web.stanford.edu/group/cslipublications/cslipublications/bromberger-corpus/On-What-We-Know-We-Dont-Know.pdf

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining Explanations: An Overview of Interpretability of Machine Learning. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*. https://doi.org/10.1109/dsaa.2018.00018

Kanerva, O. (2019). *Evaluating explainable AI models for convolutional neural networks with proxy tasks*. Www.semanticscholar.org.

Randolph Mayes, G. (n.d.). *Internet Encyclopedia of Philosophy | An encyclopedia of philosophy articles written by professional philosophers.* Internet Encyclopedia of Philosophy. Retrieved April 12, 2023, from https://iep.utm.edu/

Kulicki, P. (2020). Aristotle's Syllogistic as a Deductive System. *Axioms*, *9*(2), 56. https://doi.org/10.3390/axioms9020056

Bacon, S. (1620). *Novum Organum*. P. F. Collier. (Bacon, 1620)

Simon, H. A., & Newell, A. (1971). Human problem solving: The state of the theory in 1970. *American Psychologist*, *26*(2), 145–159. https://doi.org/10.1037/h0030806

Kind, P. M. (2013). Establishing Assessment Scales Using a Novel Disciplinary Rationale for Scientific Reasoning. *Journal of Research in Science Teaching*, *50*(5), 530–560. https://doi.org/10.1002/tea.21086

Dunbar K, Fugelsang J. Scientific Thinking and Reasoning. InHolyoak KJ, Morrison RG, The Cambridge Handbook of Thinking and Reasoning. Cambridge University Press; 2005:705,720 (Dunbar and Fugelsang, 2005)

Miller, Tim. "Explanation in Artificial Intelligence: Insights from the Social Sciences." *Artificial Intelligence*, vol. 267, Feb. 2019, pp. 1–38, https://doi.org/10.1016/j.artint.2018.07.007.

Hempel, C. G., & Oppenheim, P. (1948). "Studies in the Logic of Explanation." *Philosophy of Science, 15*.

Kim, M.-Y., Atakishiyev, S., Babiker, H. K. B., Farruque, N., Goebel, R., Zaïane, O. R., Motallebi, M.-H., Rabelo, J., Syed, T., Yao, H., & Chun, P. (2021). A Multi-Component Framework for the Analysis and Design of Explainable Artificial Intelligence. *Machine Learning and Knowledge Extraction*, *3*(4), 900–921. https://doi.org/10.3390/make3040045

Ozsoz, M., Mubarak, A., Said, Z., Aliyu, R., Al Turjman, F., & Serte, S. (2021). Deep learning-based feature extraction coupled with multi-class SVM for COVID-19 detection in the IoT era. *International Journal of Nanotechnology*, *1*(1), 1. https://doi.org/10.1504/ijnt.2021.10040115

Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., Siddiqui, T., & Mellit, A. (2023). A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope. Sustainability, 15(7), 5930. https://doi.org/10.3390/su15075930

Altman, N. S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. The American Statistician, 46(3), 175–185. https://doi.org/10.1080/00031305.1992.10475879

Akella, P., Kumar, R., & Al-Turjman, F. (2023). A novel hybrid model for automatic diabetic retinopathy grading and multi-lesion recognition method based on SRCNN & YOLOv3. *International Journal of Nanotechnology*, *20*(5/6/7/8/9/10), 615–643. https://doi.org/10.1504/ijnt.2023.134022

Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. ArXiv.org. https://arxiv.org/abs/1409.1556

Pathak, M. K., Pavthawala, M., Patel, M. N., Malek, D., Shah, V. J., & Vaidya, B. (2019). Classification of Brain Tumor Using Convolutional Neural Network. *3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 128–132.

Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., & Baik, S. W. (2019). Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *Journal of Computational Science*, *30*, 174–182. https://doi.org/10.1016/j.jocs.2018.12.003

Ayadi, W., Elhamzi, W., Charfi, I., & Atri, M. (2021). Deep CNN for Brain Tumor Classification. *Neural Processing Letters*, *53*(1), 671–700. https://doi.org/10.1007/s11063-020-10398-2

Pashaei, A., Sajedi, H., & Jazayeri, N. (2018). *Brain Tumor Classification via Convolutional Neural Network and Extreme Learning Machines*.

Abiwinanda, Nyoman, et al. "Brain Tumor Classification Using Convolutional Neural Network." *IFMBE Proceedings*, 30 May 2018, pp. 183–189, www.springerprofessional.de/en/brain-tumor-classification-using-convolutional-neural-network/15802612, https://doi.org/10.1007/978-981-10-9035-6_33.

Naseer, A., Yasir, T., Azhar, A., Shakeel, T., & Zafar, K. (2021). Computer-Aided Brain Tumor Diagnosis: Performance Evaluation of Deep Learner CNN Using Augmented Brain MRI. *International Journal of Biomedical Imaging*, *2021*, 1–11. https://doi.org/10.1155/2021/5513500

Alanazi, M. F., Ali, M. U., Hussain, S. J., Zafar, A., Mohatram, M., Irfan, M., AlRuwaili, R., Alruwaili, M., Ali, N. H., & Albarrak, A. M. (2022). Brain Tumor/Mass Classification Framework Using Magnetic-Resonance-Imaging-Based Isolated and Developed Transfer Deep-Learning Model. *Sensors*, *22*(1), 372. https://doi.org/10.3390/s22010372

Bacanin, N., Bezdan, T., Venkatachalam, K., & Al-Turjman, F. (2021). Optimized convolutional neural network by firefly algorithm for magnetic resonance image classification of glioma brain tumor grade. *Journal of Real-Time Image Processing*. https://doi.org/10.1007/s11554-021-01106-x

Jati, G. K., Manurung, R., & Suyanto. (2013, January 1). *13 - Discrete Firefly Algorithm for Traveling Salesman Problem: A New Movement Scheme* (X.-S. Yang, Z. Cui, R. Xiao, A. H. Gandomi, & M. Karamanoglu, Eds.). ScienceDirect; Elsevier. https://www.sciencedirect.com/

Tang, Z., Chuang, K. V., DeCarli, C., Jin, L.-W., Beckett, L., Keiser, M. J., & Dugger, B. N. (2019). Interpretable classification of Alzheimer's disease pathologies with a convolutional neural network pipeline. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-10212-1

McKenzie, A. T., Marx, G. A., Koenigsberg, D., Sawyer, M., Iida, M. A., Walker, J. M., Richardson, T. E., Campanella, G., Attems, J., McKee, A. C., Stein, T. D., Fuchs, T. J., White, C. L., Vonsattel, J.-P., Teich, A. F., Gearing, M., Glass, J., Troncoso, J. C., Frosch, M. P., & Hyman, B. T. (2022). Interpretable deep learning of myelin histopathology in age-related cognitive impairment. *Acta Neuropathologica Communications*, *10*(1). https://doi.org/10.1186/s40478-022-01425-5

Zhang, Q., Wu, Y., & Zhu, S.-C. (2018). *Interpretable Convolutional Neural Networks*. https://arxiv.org/pdf/1710.00935

Tomita, N., Abdollahi, B., Wei, J., Ren, B., Suriawinata, A., & Hassanpour, S. (2019). Attention-Based Deep Neural Networks for Detection of Cancerous and Precancerous Esophagus Tissue on Histopathological Slides. *JAMA Network Open*, *2*(11), e1914645. https://doi.org/10.1001/jamanetworkopen.2019.14645

Apicella, A., Isgrò, F., Prevete, R., & Tamburrini, G. (2020). Middle-Level Features for the Explanation of Classification Systems by Sparse Dictionary Methods. *International Journal of Neural Systems*, *30*(08), 2050040. https://doi.org/10.1142/s0129065720500409

Duffy, G., Jain, I., He, B., & Ouyang, D. (2022). Interpretable deep learning prediction of 3d assessment of cardiac function. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, *27*(27), 231–241. https://pubmed.ncbi.nlm.nih.gov/34890152/

Hu, S., Teng, F., Huang, L., Yan, J., & Zhang, H. (2021). An explainable CNN approach for medical codes prediction from clinical text. *BMC Medical Informatics and Decision Making*, *21*(S9). https://doi.org/10.1186/s12911-021-01615-6

Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P.-A. (2019). Accurate and interpretable evaluation of surgical skills from kinematic data using fully convolutional neural networks. *International Journal of Computer Assisted Radiology and Surgery*, *14*(9), 1611–1617. https://doi.org/10.1007/s11548-019-02039-4

Sattarzadeh, S., Sudhakar, M., Lem, A., Mehryar, S., Plataniotis, K. N., Jang, J., Kim, H., Jeong, Y., Lee, S., & Bae, K. (2021). Explaining Convolutional Neural Networks through Attribution-Based Input Sampling and Block-Wise Feature Aggregation. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(13), 11639–11647. https://doi.org/10.1609/aaai.v35i13.17384

Oviedo, F., Ren, Z., Sun, S., Settens, C., Liu, Z., Hartono, N. T. P., Ramasamy, S., DeCost, B. L., Tian, S. I. P., Romano, G., Gilad Kusne, A., & Buonassisi, T. (2019). Fast and interpretable classification of small X-ray diffraction datasets using data augmentation and deep neural networks. *Npj Computational Materials*, *5*(1). https://doi.org/10.1038/s41524-019-0196-x

Ancona, M., Ceolini, E., Öztireli, C., & Gross, M. (2018). Towards better understanding of gradient-based attribution methods for Deep Neural Networks. *ArXiv: 1711.06104 [Cs, Stat]*, *4*. https://arxiv.org/abs/1711.06104

Tosun, A. B., Pullara, F., Becich, M. J., Taylor, D. L., Fine, J. L., & Chennubhotla, S. C. (2020). Explainable AI (xAI) for Anatomic Pathology. *Advances in Anatomic Pathology*, *27*(4), 241–250. https://doi.org/10.1097/pap.0000000000000264

Spinner, T., Schlegel, U., Schäfer, H., & El-Assady, M. (2020). explAIner: A Visual Analytics Framework for Interactive and Explainable Machine Learning. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 1064–1074. https://doi.org/10.1109/TVCG.2019.2934629

Jha, S. S. (2022). An Overview on the Explainability of Cyber-Physical Systems. *The International FLAIRS Conference Proceedings*, *35*. https://doi.org/10.32473/flairs.v35i.130646

Haag, F., Hopf, K., Vasconcelos, P. M., & Staake, T. (2022). Augmented Cross-Selling Through Explainable AI—A Case from Energy Retailing. *ECIS 2022 Research Papers*.

Iliadou, E., Su, Q., Kikidis, D., Bibas, T., & Kloukinas, C. (2022). Profiling hearing aid users through big data explainable artificial intelligence techniques. *Frontiers in Neurology*, *13*. https://doi.org/10.3389/fneur.2022.933940

Lourenço das Neves, M. I. (2020). *Maria Inês Lourenço das Neves Opening the black-box of artificial intelligence predictions on clinical decision support systems*. https://run.unl.pt/bitstream/10362/126699/1/Neves_2021.pdf

Zhang, W., Dimiccoli, M., & Lim, B. Y. (2022). Debiased-CAM to mitigate image perturbations with faithful visual explanations of machine learning. *ArXiv: 2012.05567[Cs]*.

TensorBoard | TensorFlow. (2019). TensorFlow. https://www.tensorflow.org/tensorboard

Pedretti, G., Graves, C. E., Serebryakov, S., Mao, R., Sheng, X., Foltin, M., Li, C., & Strachan, J. P. (2021). Tree-based machine learning performed in-memory with memristive analog CAM. *Nature Communications*, *12*(1), 5806. https://doi.org/10.1038/s41467-021-25873-0

Eder, M., Moser, E., Holzinger, A., Jean-Quartier, C., & Jeanquartier, F. (2022). Interpretable Machine Learning with Brain Image and Survival Data. *BioMedInformatics*, *2*(3), 492–510. https://doi.org/10.3390/biomedinformatics2030031

Wäldchen, S., Pokutta, S., & Huber, F. (2022, June 28). *Training Characteristic Functions with Reinforcement Learning: XAI-methods play Connect Four*. Proceedings.mlr.press; PMLR. https://proceedings.mlr.press/v162/waldchen22a.html

Taniguchi, H., Takata, T., Takechi, M., Furukawa, A., Iwasawa, J., Kawamura, A., Taniguchi, T., & Tamura, Y. (2021). Explainable Artificial Intelligence Model for Diagnosis of Atrial Fibrillation Using Holter Electrocardiogram Waveforms. *International Heart Journal*, *62*(3), 534–539. https://doi.org/10.1536/ihj.21-094

Lu, X., Tolmachev, A., Yamamoto, T., Takeuchi, K., Okajima, S., Takebayashi, T., Maruhashi, K., & Kashima, H. (2021). Crowdsourcing Evaluation of Saliency-based XAI Methods. *ArXiv: 2107.00456 [Cs]*. https://arxiv.org/abs/2107.00456

Klein, L., El-Assady, M., & Jäger, P. F. (2022). From Correlation to Causation: Formalizing Interpretable Machine Learning as a Statistical Process. *ArXiv: 2207.04969 [Cs]*. https://arxiv.org/abs/2207.04969

Idrees, M., & Sohail, A. (2022). Explainable machine learning of the breast cancer staging for designing smart biomarker sensors. *Sensors International*, *3*, 100202. https://doi.org/10.1016/j.sintl.2022.100202

Hassan, Esraa, et al. “The Effect of Choosing Optimizer Algorithms to Improve Computer Vision Tasks: A Comparative Study.” *Multimedia Tools and Applications*, 28 Sept. 2022, https://doi.org/10.1007/s11042-022-13820-0.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Christoph Molnar. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 2022. Kindle Edition ed., vol. 2023-08-21, 21 Aug. 2023, p. 540, christophm.github.io/interpretable-ml-book/. Accessed 1 Jan. 2024.

https://plus.google.com/u/0/+Datacamp. “Learn R, Python & Data Science Online.” *Datacamp.com*, 2011, www.datacamp.com/.

Mubarak, A. S., Ameen, Z. S., & Al-Turjman, F. (2022). Effect of Gaussian filtered images on Mask RCNN in detection and segmentation of potholes in smart cities. *Mathematical Biosciences and Engineering*, *20*(1), 283–295. https://doi.org/10.3934/mbe.2023013

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. ArXiv.org. https://arxiv.org/abs/1512.03385

Huang, G., Liu, Z., & Weinberger, Kilian Q. (2016). Densely Connected Convolutional Networks. ArXiv.org. https://arxiv.org/abs/1608.06993

Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. https://proceedings.mlr.press/v37/ioffe15.pdf

Ge, C., Gu, I. Y.-H., Jakola, A. S., & Yang, J. (2020). Enlarged Training Dataset by Pairwise GANs for Molecular-Based Brain Tumor Classification. IEEE Access, 8, 22560–22570. https://doi.org/10.1109/access.2020.2969805

Vimal Kurup, R., Sowmya, V., & Soman, K. P. (2020). Effect of data pre-processing on brain tumor classification using capsulenet. In ICICCT 2019–System Reliability, Quality Control, Safety, Maintenance and Management: Applications to Electrical, Electronics and Computer Science and Engineering (pp. 110-119). Springer Singapore.

Dehkordi, A. A., Hashemi, M., Neshat, M., Mirjalili, A., & Sadiq, A. S. (2022). Brain Tumor Detection and Classification Using a New Evolutionary Convolutional Neural Network. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4292650

Haq, A. ul, Li, J. P., Khan, S., Alshara, M. A., Alotaibi, R. M., & Mawuli, C. (2022). DACBT: deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment. Scientific Reports, 12(1), 15331. https://doi.org/10.1038/s41598022194651

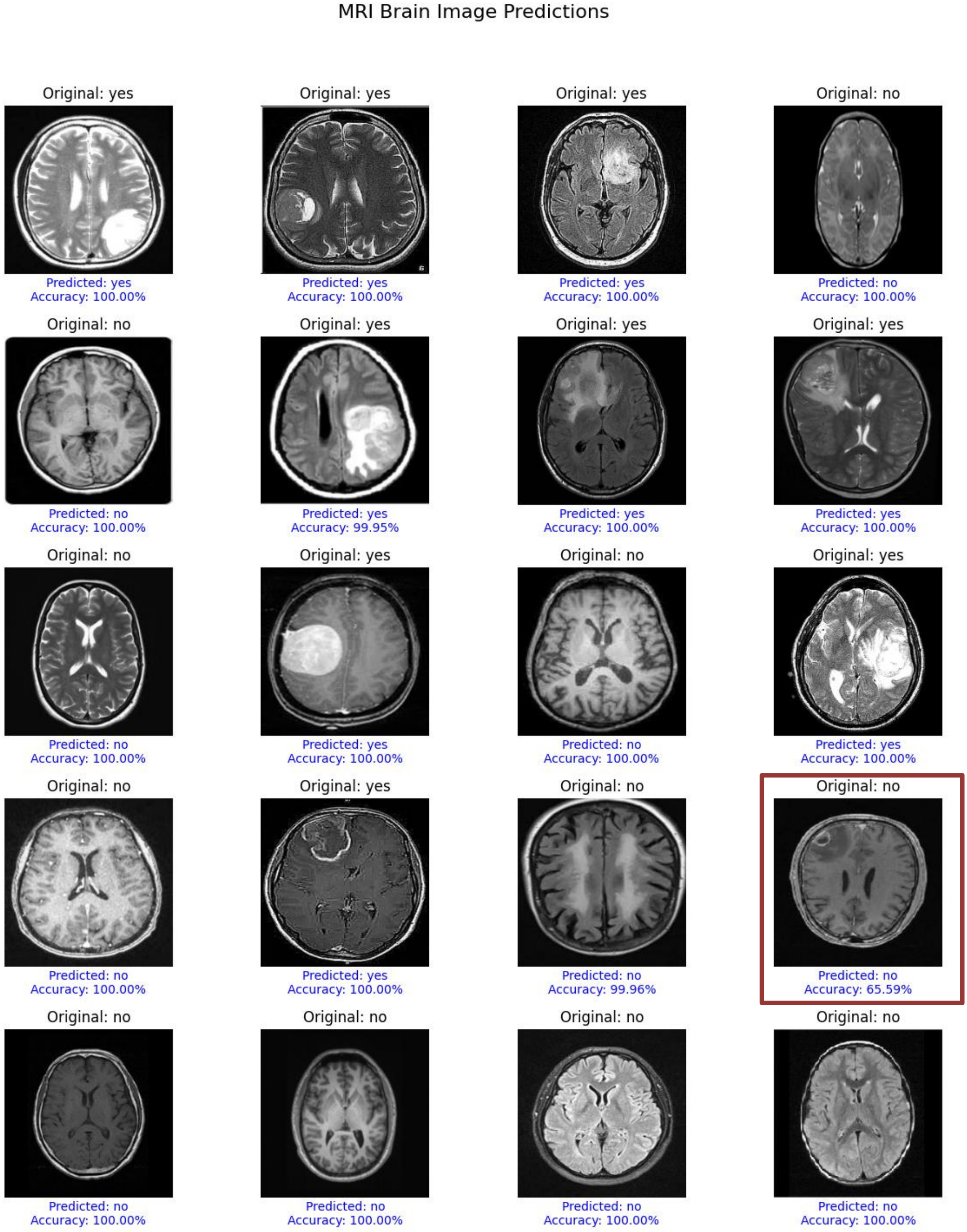
Zeineldin, R. A., Karar, M. E., Elshaer, Z., Coburger, J., Wirtz, C. R., Burgert, O., & Mathis-Ullrich, F. (2022). Explainability of deep neural networks for MRI analysis of brain tumors. International Journal of Computer Assisted Radiology and Surgery, 17(9), 1673–1683. https://doi.org/10.1007/s11548-022-02619-x

Benyamina, H., Auwalu Saleh, M., & Al-Turjman, F. (2022). Explainable Convolutional Neural Network for Brain Tumor Classification via MRI Images. IEEE. https://doi.org/10.1109/aiotcs58181.2022.00048

Appendices

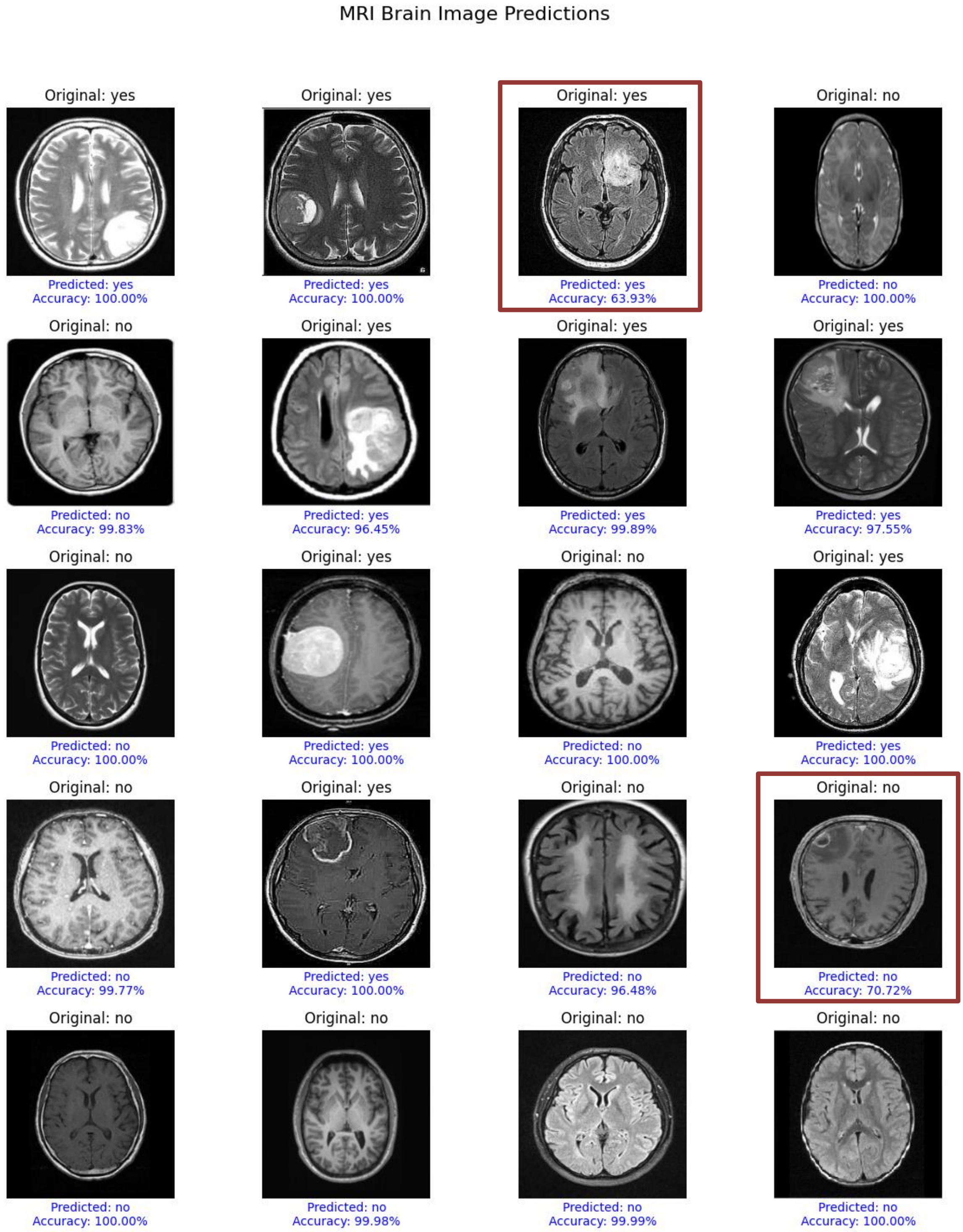
Appendix A

MRI Brain Image Predictions using VGGNet16



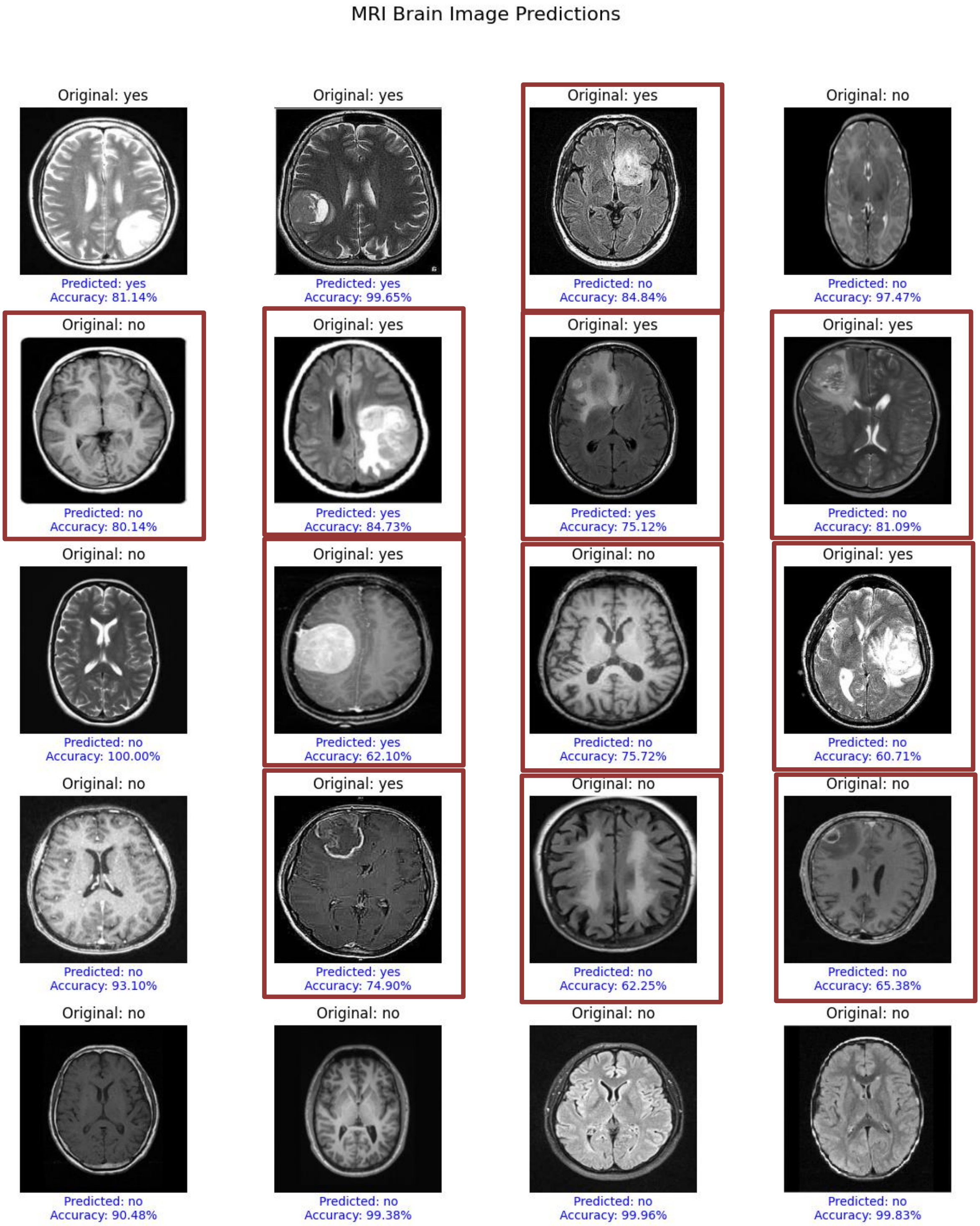
Appendix B

MRI Brain Image Predictions using VGGNet19



Appendix C

MRI Brain Image Predictions using ResNet50



Appendix D

Turnitin Similarity Report

About this page

This is your assignment inbox. To view a paper, select the paper's title. To view a Similarity Report, select the paper's Similarity Report icon in the similarity column. A ghosted icon indicates that the Similarity Report has not yet been generated.

Yakın Doğu Üniversitesi

QUICK SUBMIT | NOW VIEWING: ALL PAPERS ▾

Submit

AUTHOR	TITLE	SIMILARITY	FILE	PAPER ID	DATE
B Hadjer	ABSTRACT	0%		2314404991	07-Mar-2024
B Hadjer	CH VI	0%		2314405075	07-Mar-2024
B Hadjer	CH II	1%		2314405016	07-Mar-2024
B Hadjer	CH V	1%		2314405046	07-Mar-2024
B Hadjer	CH I	5%		2314405008	07-Mar-2024
B Hadjer	CH IV	6%		2314405039	07-Mar-2024
B Hadjer	FULL	6%		2314405084	07-Mar-2024
B Hadjer	CH III	7%		2314405026	07-Mar-2024

europass

Hadjer BENYAMINA

Nationality: Algerian **Date of birth:** 25/09/1993

Phone number: (+90) 5338832390

Email address: hadjoura.hero1993@yahoo.com

WhatsApp Messenger: +905338832390

LinkedIn: https://www.linkedin.com/in/hadjer-benyamina-6bab07ba/

Home: 62, Sehit Huseyin Amcaoglu Cad. Gonyeli., 99150 Gönyeli (Cyprus)

ABOUT ME

AI / DL and Data Science Enthusiast Researcher.

WORK EXPERIENCE

Recruitment Officer

ONCALL Interpreters & Translators [01/2024 – Current]

Email address: hadjer.benyamina@oncalleu.com | **Name of unit or department:** Recrutiment - **Business or sector:** Information and communication

- Talent acquisition within the interpreters and translators industry.
- Proficient in sourcing, screening, and selecting linguistically skilled professionals for diverse language pairs and specialized fields (Asylum, Banking and finance, Medical setting, ... etc).
- Experienced in developing and executing recruitment strategies to attract and retain highly qualified interpreters and translators.
- Demonstrated success in building and nurturing relationships with interpreters and translators.
- Proven ability to assess language proficiency, cultural competency, and subject matter expertise during candidate evaluation.
- In-depth understanding of language industry trends, including emerging technologies and best practices in interpretation and translation.
- Strong communication skills to effectively collaborate with hiring managers, project coordinators, and clients to meet language service needs.
- Proficient in managing recruitment metrics and analytics to optimize performance and drive continuous improvement in recruitment processes.
- Committed to delivering exceptional candidate experiences and facilitating successful language service delivery for clients.

Freelance Interpreter / Translator

Translators without Borders [07/2023 – Current]

Consecutive / Community Interpreting: AR <> EN
Translation: EN <> AR
FR > AR
FR > EN
Field: Legal, Medical, Asylum, Police, Social services and More.

Research And Development Officer

Oncall Europa Languages Services [19/06/2023 – 19/01/2024]

Research and Development Officer with expertise in language services and a focus on quality assurance, process optimization, and service development.
Key responsibilities include:

- Collaborating with operations to define research goals and strategies, and identifying emerging trends in language services.
- Conducting market research to understand competitor weaknesses and strengths, and ensuring successful execution of research projects.
- Conducting industry research to improve service delivery, collaborating with operations for process improvements, and forecasting potential projects based on industry newsletters.
- Collaboration and knowledge sharing among team members and departments, collaborating with external partners and experts, and participating in conferences to contribute to language service advancements.

OPI / OVI Interpreter

ONCALL Interpreters & Translators UK [17/05/2021 – 16/06/2023]

ENGLISH - ARABIC Prefered Interpreter
Tasks: Telephone and Video Interpreting services for:

- UK NHS (Hospitals and related healthcare services)
- UK Authorities (Police, City Councils...etc.)
- Asylum and Immigration applications.
- Social Services and more.

OPI / OVI Interpreter

Oncall Europa Languages Services [17/05/2021 – 16/06/2023]

ENGLISH - ARABIC Prefered Interpreter

Tasks: Telephone and Video Interpreting services for:

- EUAA (EASO)
- Asylum and Immigration applications / Registration, interviews and decision sessions.
- Resettlements programs.

Online Instructor

Utrainy.com [07/2020 – 12/2021]

- Proficient online instructor experienced in virtual teaching environments.
- Develops engaging curriculum and instructional materials for online courses.
- Facilitates interactive lessons and discussions using virtual platforms.
- Provides constructive feedback and support to students to enhance learning outcomes.
- Utilizes innovative teaching methods and technology tools to enhance online instruction.
- Creates a positive and inclusive online learning environment conducive to student success.
- Adapts teaching strategies to meet the diverse needs of learners in an online setting.
- Maintains regular communication with students and promptly addresses questions or concerns.
- Monitors student progress and assesses learning through various online assessment methods.
- Continuously seeks professional development opportunities to enhance online teaching skills.

Virtual assistant

Utrainy.com [07/2020 – 14/05/2021]

Country: Algeria

- Adapt website content for specific regions and languages.
- Manage complaint tickets for web developers and ensure timely resolution.
- Translate website content, documents, and communications accurately.

europass

- Develop comprehensive online guidelines for website usage.
- Draft GDPR-compliant text for privacy policies and data protection notices.
- Handle email correspondence, scheduling, and record-keeping.
- Update website content and ensure consistency.
- Assist in managing social media accounts and engage with followers.
- Provide basic technical support to website users and coordinate with technical teams.
- Generate reports on website performance metrics and analyze data.
- Stay updated on industry trends and identify opportunities for improvement.

Elementary School Teacher

Ministry of Education [06/09/2017 – 31/12/2020]

City: RELIZANE | **Country:** Algeria

Information Technology Help Desk

Sonelgaz [09/2015 – 10/2015]

City: RELIZANE | **Country:** Algeria

Internship Description: Administrative tasks and short visits to the Computing Center.

EDUCATION AND TRAINING

Masters in Artificial Intelligence Engineering

Near East University [02/2021 – 02/2024]

City: Lefkosa / K.K.T.C. | **Country:** Türkiye | **Website:** https://neu.edu.tr/ | **Final grade:** GPA: 3.57 | **Thesis:** Modified DenseNet1.0 for Brain Tumour Classification Via MRI Images in the IoT Era

Masters in Information Systems Engineering

Abdelhamid Ibn Badis - Université de Mostaganem [10/2015 – 06/2017]

City: Mostaganem | **Country:** Algeria | **Website:** https://www.univ-mosta.dz/universite-abdelhamid-ibn-badis-mostaganem/

Bachelor in Computer Science

Relizane University [07/2012 – 06/2015]

City: RELIZANE | **Country:** Algeria | **Website:** https://web0.univ-relizane.dz/

LANGUAGE SKILLS

Mother tongue(s): Arabic

Other language(s):

English

LISTENING C2 **READING** C2 **WRITING** C2

SPOKEN PRODUCTION C2 **SPOKEN INTERACTION** C2

French

LISTENING C1 **READING** C1 **WRITING** B2

SPOKEN PRODUCTION B2 **SPOKEN INTERACTION** B2

Russian

LISTENING A1 **READING** A1 **WRITING** A1

SPOKEN PRODUCTION A1 **SPOKEN INTERACTION** A1

Spanish

LISTENING A1 **READING** A1 **WRITING** A1

SPOKEN PRODUCTION A1 **SPOKEN INTERACTION** A1

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

DIGITAL SKILLS

Digital Skills - Test Results

Information and data literacy	**INTERMEDIATE**	Level 3 / 6
Communication and collaboration	**INTERMEDIATE**	Level 3 / 6
Digital content creation	**ADVANCED**	Level 5 / 6
Safety	**ADVANCED**	Level 5 / 6
Problem solving	**ADVANCED**	Level 5 / 6

Results from self-assessment based on The Digital Competence Framework 2.1

My Digital Skills

Python Programming Language / Deep Learning / Computer Vision / Anaconda / Jupyter Notebook / Data Visualization / Medical Imaging / Data Preprocessing / Data Science / TensorFlow / Data Analysis / Tensorflow / Matplotlib / Artificial Intelligence / Convolutional Neural Networks / OpenCV / Neural Networks and Deep Learning / Keras / Good command of Microsolft: Excel, Word, Power Point, Outlook, Teams) / Numpy / Scikit-Learn

PUBLICATIONS

[2023]
Explainable Convolutional Neural Network for Brain Tumor Classification via MRI Images doi: 10.1109/AIoTCs58181.2022.00048

H. Benyamina, A. S. Mubarak and F. Al-Turjman, 2022, AIoTCs, Nicosia, Cyprus, 2022, pp. 266-272

VOLUNTEERING

[01/2016 – 01/2020] Algeria
Proofreader https://www.dz-res.com/

HOBBIES AND INTERESTS

Writing, Reading, Drawing, Singing, Debating, Chess, Food Photography

SOCIAL AND POLITICAL ACTIVITIES

[2024 – Current]
Code Crafters Programming Club.

I won the " Programmer of the Month " title in April 2024.

europass

Link: https://drive.google.com/file/d/1xx9ubd5nIGTjSTsn62JstpgTIl8d2VcR/view?usp=sharing

RECOMMENDATIONS

Name: **Prof. Dr. Fadi Al-Turjman**
One of my professors at NEU, supervisor and advisor for 3 years.
Email: fadi.alturjman@neu.edu.tr

Name: **Asst prof Auwalu Saleh Mubarak**
One of my professors at NEU and one of my Thesis defense jury members.
Email: auwalusaleh.mubarak@neu.edu.tr

Name: **Prof.Dr. Rahib H.Abiyev**
Second email: rahib.h.abiyev@gmail.com
One of my professors at NEU during one semester.
Email: rahib.abiyev@neu.edu.tr

Name: **Said KHELIFA** | Supervisor and Professor
One of my professors at Relizane University, Algeria - from 2012 to 2015 and Supervisor.
Email: said.khelifa@univ-usto.dz

Name: **Tilsim KURTBIR** | Senior Lecturer / Interpreter
My tutor and supervisor during my one-month training period as an Interpreter at ONCALL.
Email: tkurtbir@gmail.com | **Phone number:** (+90) 5428876900

My data must be processed solely for recruitment purposes, following Article 6(1)(f) of the General Data Protection Regulation (GDPR). This information should be treated confidentially and shared only with relevant personnel involved in the recruitment process.

Nicosia - Cyprus, 19/05/2024