People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University of Amar Telidji - Laghouat
Faculty Of Science and Technology

Department of Electronics
Specialty: Instrumentation

Final-year project
(Master's degree)

Classification of medical images using Deep Learning

Presented by:

Ibrahim Hireche

Supervised by:

Zitouni Abdelkader

Academic Year: 2024 / 2025

Dedication

" To my dear parents, for their invaluable sacrifices, unwavering love, tenderness, constant support, and prayers throughout my studies,

To my brothers and sisters, for their continuous encouragement and steadfast moral support,

To all my family and friends, for their presence and invaluable support throughout my academic journey,

May this work be the fulfillment of your heartfelt wishes and the fruit of your unyielding support,

Thank you from the bottom of my heart."

Acknowledgments

First and foremost, I would like to express my deep gratitude to Allah for granting me the strength, patience and perseverance to complete this work. This achievement is the culmination of many years of dedication and hard work.

I am deeply honored to extend my sincere thanks and heartfelt appreciation to my supervisor, Mr. Zitouni, for his invaluable support, guidance, and the time he dedicated to supervising me throughout this project. His expertise and encouragement have been instrumental in shaping the success of this work.

I would also like to express my gratitude to all of the professors at Amar Teledji University, whose dedication to teaching and knowledge sharing has greatly enriched my academic journey over the past five years.

My sincere thanks go to the staff members of the Department of Science and Technology for their support and assistance during my studies.

I would like to extend my appreciation to the members of the jury for their time and consideration in reviewing my work. Their insightful feedback and constructive criticism have been greatly appreciated.

Finally, I am grateful to everyone who has contributed, directly or indirectly, to the successful completion of this thesis. Your support, encouragement, and guidance have been invaluable throughout this process.

Abstract

This report explores medical image classification through Convolutional Neural Networks (CNNs), integrating theoretical foundations with practical application. It begins with a review of essential AI and deep learning principles, with particular emphasis on CNN architecture and its effectiveness in feature extraction. For the practical component, a DenseNet-121 model is implemented to classify two distinct medical imaging datasets: chest X-rays (COVID-19, pneumonia, and normal cases) and brain MRIs (three tumour types and normal). Following preprocessing and training, the model demonstrated excellent performance, achieving an accuracy and F1-score both exceeding 0.99. These results highlight the robustness of CNNs in medical image analysis and their potential in supporting clinical diagnostic processes.

Keywords: Artificial intelligence, Machine learning, Deep learning, Classification, Convolutional neural network.

Résumé

Ce rapport explore la classification d'images médicales à l'aide des réseaux de neurones convolutifs (CNN), en combinant des fondements théoriques solides avec une mise en œuvre pratique. Il commence par une revue des concepts essentiels de l'intelligence artificielle et de l'apprentissage profond, en mettant l'accent sur l'architecture des CNN et leur efficacité dans l'extraction de caractéristiques. Dans la partie pratique, un modèle DenseNet-121 est implémenté pour classifier deux jeux de données d'imagerie médicale : des radiographies thoraciques (cas de COVID-19, pneumonie et normaux) et des IRM cérébrales (trois types de tumeurs et cas normaux). Après le prétraitement et l'entraînement, le modèle a démontré d'excellentes performances, atteignant une précision et un score F1 supérieurs à 0,99. Ces résultats mettent en évidence la robustesse des CNN dans l'analyse d'images médicales et leur potentiel pour soutenir les processus de diagnostic clinique.

Mots clés: Intelligence artificielle, Apprentissage automatique, Apprentissage profond, Classification, Réseau de neurones convolutifs.

ملخص

يستعرض هذا التقرير تصنيف الصور الطبية باستخدام الشبكات العصبية الالتفافية (CNNs)، جامعًا بين الأسس النظرية والتطبيق العملي. يبدأ التقرير بمراجعة المبادئ الأساسية للذكاء الاصطناعي والتعلم العميق، مع التركيز بشكل خاص على بنية الشبكات العصبية الالتفافية وفعاليتها في استخلاص السمات. في الجانب العملي، يُطبّق نموذج DenseNet-121 لتصنيف مجموعتين مختلفتين من بيانات التصوير الطبي: صور الأشعة السينية للصدر (حالات كوفيد-19، الالتهاب الرئوي و الحالات الطبيعية) وصور الرنين المغناطيسي للدماغ (ثلاثة أنواع من الأورام والحالات الطبيعية). بعد المعالجة المسبقة والتدريب، أظهر النموذج أداءً ممتازًا، محققًا دقةً و قيمة F1 تجاوزت 0.99. تُبرز هذه النتائج قدرة الشبكات العصبية الالتفافية على تحليل الصور الطبية وإمكاناتها في دعم عمليات التشخيص السريري.

كلمات مفتاحية : الذكاء الاصطناعي، التعلم الآلي، التعلم العميق، التصنيف، الشبكة العصبية الالتفافية.

Contents

Dedication I

Acknowledgments II

Abstract III

Résumé IV

ملخص V

1 Introduction **1**

2 Basic Principles and Concepts of Artificial Intelligence **3**

2.1 Introduction 3

2.2 Core principals of Artificial Intelligence 4

2.3 Definition 4

2.4 Applications of Artificial Intelligence 4

2.4.1 Healthcare 4

2.4.2 Finance 5

2.4.3 Autonomous Vehicles 5

2.4.4 Retail 6

2.4.5 Agriculture 7

2.4.6 Manufacturing 7

2.5 Limitations of Artificial Intelligence 8

2.6 Conclusion 9

3 Foundational Tools and Algorithms for Artificial Intelligence **10**

3.1 Introduction 10

3.2 Machine Learning 10

3.2.1 Definition 10

3.2.2 Types of Machine Learning 11
3.3 Deep Learning 13
3.3.1 Definition 13
3.3.2 Artificial Neural Network 14
3.3.3 Artificial Neuron 14
3.4 Architectures of Deep Learning Neural Networks 15
3.4.1 Feedforward Neural Network 15
3.4.2 Convolutional Neural Network 15
3.4.3 Recurrent Neural Network 15
3.4.4 Long Short-Term Memory Network and Gated Recurrent Unit 15
3.4.5 Auto-encoders 16
3.4.6 Generative Adversarial Network 16
3.4.7 Transformers 16
3.5 Conclusion 16

4 Image Classification Using Convolutional Neural Networks 17
4.1 Introduction 17
4.2 Definition 17
4.3 Biological inspiration Behind CNNs 18
4.4 Convolutional Neural Network Architecture 18
4.4.1 Feature Extraction Part 19
4.4.2 Classification Part 21
4.5 Classification Assessment 22
4.5.1 Classification Accuracy and its Limitations 22
4.5.2 Classification Rate/Accuracy 23
4.5.3 Recall 24
4.5.4 Precision 24
4.5.5 F-measure (F1-score) 24
4.6 Conclusion 24

5 Implementation of CNN-Based Classification on medical images 26
5.1 Introduction 26
5.2 Tools and Implementation Environment 26
5.2.1 Hardware Setup 26

5.2.2 Python Programming Language 27
5.2.3 Libraries and Frameworks 27
5.3 Chest X-ray Classification 28
5.3.1 Data Description and Analysis 28
5.3.2 Model Development and Training 30
5.3.3 Visualisation of Results 32
5.3.4 Model Performance and Result Analysis 34
5.4 Brain Tumour Classification 35
5.4.1 Data Description and Analysis 35
5.4.2 Model Development and Training 38
5.4.3 Visualisation of Results 38
5.4.4 Model Performance and Result Analysis 39
5.5 Conclusion 41

General Conclusion 43

List of Figures

2.1 AI-powered detection system highlighting critical lung issues in a chest X-ray . . 5

2.2 **Waymo** formerly known as the **Google Self-Driving Car Project** 6

2.3 **HiCart**'s AI-Powered Shopping Cart. 6

2.4 **UAV** agriculture drone sprayer 7

2.5 Electron Microscope Imaging of Silicon Wafer Defects powered by **ExtractAI** . 8

3.1 Relationship between AI, ML, and Deep Learning 13

3.2 Structure of a typical Artificial Neural Network 14

3.3 Mathematical Structure of a Artificial Neuron 15

4.1 Example of a Convolutional Operation 19

4.2 Common Activation Functions in CNNs. 20

4.3 Example of Pooling operations 20

4.4 A pooled feature map being flattened and inputted into a fully connected layer . 21

4.5 Class Selection in CNNs for single and multi output 21

4.6 General structure of a Convolutional Neural Network (CNN) 22

5.1 Examples of Chest X-Rays from healthy patients 29

5.2 Examples of Chest X-Rays from patients with Pneumonia 29

5.3 Examples of Chest X-Rays from patients with COVID-19 30

5.4 Image before preprocessing and after applying resizing and data augmentation . 31

5.5 A schematic illustration of the DenseNet-121 architecture 31

5.6 Training and validation loss per epoch 33

5.7 Training and validation accuracy per epoch 33

5.8 Confusion Matrix 33

5.9 Examples of Brain MRI Scans from Healthy Individuals 36

5.10 Examples of Brain MRI Scans from patients with Glioma 37

5.11 Examples of Brain MRI Scans from patients with Meningioma 37

5.12 Examples of Brain MRI Scans from patients with Pituitary tumour 38

5.13 Training and validation loss per epoch 38
5.14 Training and validation accuracy per epoch 39
5.15 Confusion Matrix 39

List of Tables

4.1 Confusion matrix for binary classification 23

5.1 Hardware configuration used in Google Colab 27

5.2 Training Configuration Summary 32

5.3 Classification metrics derived from the confusion matrix 34

5.4 Classification metrics derived from the confusion matrix 40

List of Acronyms

AI Artificial Intelligence

ML Machine Learning

DL Deep Learning

ANN Artificial Neural Network

FNN Feedforward Neural Network

CNN Convolutional Neural Network

RNN Recurrent Neural Network

PCA Principal Component Analysis

S3VM Semi-supervised Support Vector Machines

DQN Deep Q Network

PPO Proximal Policy Optimization

CPC Contrastive Predictive Coding

BERT Bidirectional Encoder Representations from Transformers

k-NN k-Nearest Neighbors

SVM Support Vector Machines

LSTM Long Short-Term Memory Network

GRU Gated Recurrent Unit

GAN Generative Adversarial Network

NLP Natural Language Processing

Chapter 1

Introduction

In recent decades, artificial intelligence (AI) has transitioned from a conceptual ambition to a transformative force across numerous industries, including healthcare. AI refers to the capability of machines to perform tasks that typically require human intelligence, such as reasoning, learning, and problem-solving [1]. This technological evolution is not merely a breakthrough but a continuation of computer science, empowered by the proliferation of algorithms, exponential growth in computational power, and the availability of vast datasets interconnected through high-speed networks.

A critical subfield of AI is machine learning (ML), which enables systems to identify patterns in data and improve performance without being explicitly programmed [2]. Within ML, deep learning (DL) has garnered significant attention for its ability to automatically extract high-level features from raw data using multi-layered artificial neural networks [3]. These architectures have demonstrated outstanding capabilities in areas such as natural language processing and, notably, computer vision.

In healthcare, one of the most impactful applications of AI—particularly DL—is in medical imaging. Deep learning techniques have achieved remarkable accuracy in image recognition tasks, enabling the classification, detection, and analysis of medical images. This capability is crucial in supporting clinical decision-making, including disease diagnosis, treatment planning, and patient monitoring. However, medical images present inherent challenges, such as inter-patient variability, low contrast between healthy and pathological tissues, and the need for highly precise interpretation [9].

Given these complexities, this report focuses on the classification of medical images using deep learning techniques. The goal is to design a good and reliable automatic classification model capable of accurately identifying pathological patterns in medical imagery.

The objectives of this project are as follows :

- Explore the theoretical underpinnings of AI, Machine Learning, and Deep Learning, with a particular focus on their applications in image analysis.

- Implement a deep learning model capable of classifying medical images with a high degree of accuracy.
- Evaluate the model's performance using standard metrics and assess its limitations.
- Analyse and discuss the results, drawing conclusions and suggesting potential future improvements.

The structure of this report is organized into two major parts:

Theoretical Part: This section offers a comprehensive overview of the fundamental principles of artificial intelligence (AI), including its applications and illustrative examples in the first chapter. The second chapter delves into some of the most widely used AI algorithms, focusing on Machine Learning and Deep Learning. The topics covered include the main types of machine learning, supervised, unsupervised, reinforcement learning...The third chapter offers a focused study on one of the prominent deep learning architectures: Convolutional Neural Network (CNN).

Practical Part: This section presents the implementation phase of the project, where the theoretical concepts discussed earlier are applied in practice. It begins with an introduction to the dataset used for training and evaluation, followed by a detailed explanation of the preprocessing steps applied to prepare the data. The chapter then outlines the development of the CNN model, including its architecture, training procedure, and hyperparameter settings. The model's performance is assessed using standard evaluation metrics, and the results are analysed to highlight both its strengths and limitations. This hands-on approach demonstrates how AI techniques, particularly CNNs, can be effectively used to solve real-world classification tasks.

Chapter 2

Basic Principles and Concepts of Artificial Intelligence

2.1 Introduction

In recent years, AI has emerged as one of the most transformative and rapidly advancing technologies of the modern era. Driven by significant progress in computational power, the availability of large datasets, and breakthroughs in algorithms, AI has shifted from a theoretical research area to a practical tool integrated into numerous aspects of daily life. Today, AI technologies underpin systems ranging from virtual assistants and recommendation engines to advanced medical diagnostics and autonomous vehicles.

The growing relevance of AI is not only due to its technical capabilities but also to its ability to enhance efficiency, decision-making, and innovation across multiple sectors. In industries such as healthcare, finance, education, transportation, and entertainment, AI has introduced new opportunities and reshaped traditional practices. Its integration into society has reached a point where many AI-driven processes operate seamlessly in the background, becoming almost invisible to the average user while significantly impacting their experiences.

Given this widespread adoption, understanding the core concepts of AI has become essential, particularly in specialized fields like medical imaging where AI's contributions are redefining diagnostic practices.

This first chapter of the report will present a comprehensive overview of Artificial Intelligence by:

- Defining AI and explaining its fundamental principles.
- Exploring its applications across multiple fields.
- Highlighting key advantages that make AI a powerful tool.
- Discussing some important limitations and challenges that remain.

2.2 Core principals of Artificial Intelligence

Artificial Intelligence (AI) is founded on several key principles that enable machines to replicate aspects of human intelligence. These principles include perception, where AI systems gather and interpret information from their environment [3]; reasoning and decision-making, which allow AI to process data logically and solve problems [4]; and learning, through which systems improve over time by identifying patterns in data [5]. Adaptation ensures AI systems adjust their behaviour to dynamic environments [6], while autonomy allows them to perform tasks with minimal human intervention. Finally, interaction and communication, particularly through Natural Language Processing, enable AI to effectively engage with humans [7]. Together, these principles form the foundation for the wide-ranging applications of AI across modern industries.

2.3 Definition

Artificial Intelligence is defined as the field that is focused on creating systems capable of performing tasks that typically require human intelligence. These tasks include decision-making, problem-solving, learning, and perception. AI encompasses both the development of algorithms that allow machines to learn from data and the creation of systems that can adapt to new information without human intervention. AI systems can range from narrow applications, such as facial recognition and recommendation algorithms, to more general capabilities like autonomous vehicles. The ultimate goal of AI research is to create machines that can simulate human-like cognition and even surpass human capabilities in certain areas [1][8].

2.4 Applications of Artificial Intelligence

Artificial intelligence has become a cornerstone of modern technological advancements, affecting almost every aspect of society. From healthcare to finance, AI applications are reshaping industries by automating complex processes, making data-driven decisions, and providing innovative solutions. In the following sections, we explore various fields where AI is making a significant impact, highlighting its transformative effects and some examples of its applications.

2.4.1 Healthcare

AI in healthcare is revolutionizing diagnostic procedures, personalized treatment plans, and medical research. Machine learning and deep learning models are employed for tasks like disease detection, medical imaging, and predicting patient outcomes. In medical imaging, AI systems can automatically detect conditions such as tumours in radiology scans with accuracy comparable to human experts. For example, AI algorithms are used in detecting breast cancer from mammograms, significantly improving early detection rates [9]. Moreover, AI is employed in drug discovery and personalized treatment recommendations, making healthcare more efficient and tailored to individual patients.

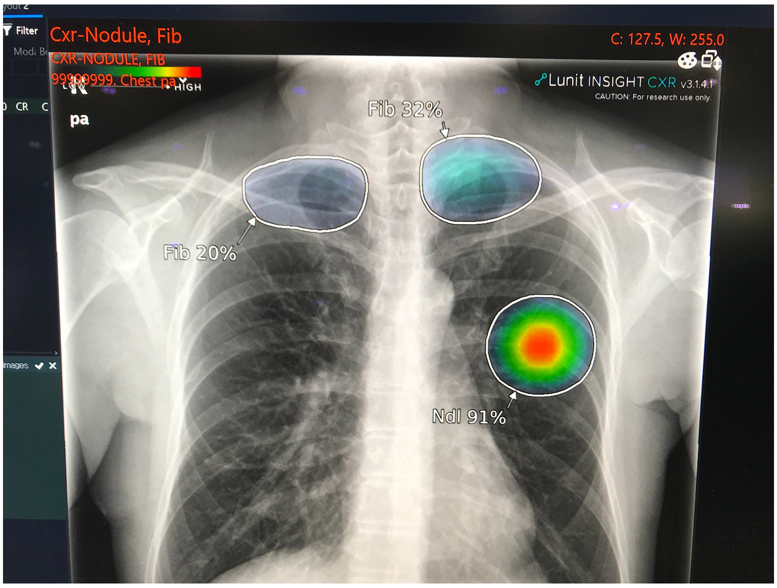


Figure 2.1: AI-powered detection system highlighting critical lung issues in a chest X-ray [47]

2.4.2 Finance

AI plays a crucial role in the financial sector by enhancing decision-making processes and improving operational efficiency. Algorithmic trading uses machine learning to predict market movements and execute trades at high speed and accuracy, helping financial institutions achieve better returns. In fraud detection, AI models analyse transaction patterns to identify potentially fraudulent activities. For instance, AI-driven tools are used to detect unusual spending patterns on credit cards, preventing fraud [10]. Furthermore, AI applications in **credit scoring** and risk assessment assist banks in making more informed lending decisions, enhancing financial inclusion and reducing risk.

2.4.3 Autonomous Vehicles

Autonomous vehicles represent one of the most exciting applications of AI, where AI systems make real-time decisions based on sensor data to navigate roads safely. Self-driving cars use AI to interpret data from cameras, radar, and **LIDAR** sensors, allowing them to understand their surroundings and make decisions such as stopping for pedestrians, navigating traffic, and avoiding obstacles. AI models also help in path planning and route optimization. For example, Tesla's Autopilot system utilizes deep learning for recognizing road signs and making safe driving decisions [11]. This technology holds the potential to reduce accidents and transform transportation systems globally.

Figure 2.2: **Waymo** formerly known as the **Google Self-Driving Car Project** [48]

2.4.4 Retail

In the retail sector, AI enhances customer experiences and optimizes supply chain management. Recommendation systems, which suggest products based on user behaviour and preferences, are one of the most prominent AI applications. These systems power platforms like Amazon and Netflix, where algorithms analyse past purchases or viewed content to suggest similar products [7]. AI is also used in inventory management, where machine learning models predict demand patterns, ensuring efficient stock levels and minimizing waste. Additionally, AI-driven chatbots and virtual assistants are employed for customer support, providing 24/7 assistance and improving customer satisfaction.

Figure 2.3: **HiCart**'s AI-Powered Shopping Cart [49]

2.4.5 Agriculture

AI is transforming agriculture by improving productivity, sustainability, and resource management. AI applications in precision farming include predicting crop yields, monitoring soil health, and optimizing irrigation. Machine learning models can analyse satellite data and sensor readings to detect diseases in crops early, allowing farmers to take preventative measures before a problem spreads. AI-powered drones and robots are also used for tasks like planting seeds, harvesting crops, and spraying fertilizers, reducing labour costs and increasing operational efficiency [1]. These applications contribute to more sustainable farming practices and better food security.

Figure 2.4: **UAV** agriculture drone sprayer [50]

2.4.6 Manufacturing

The manufacturing industry is leveraging AI to optimize production processes, improve quality control, and reduce costs. AI is used in predictive maintenance, where machine learning algorithms analyse data from factory equipment to predict fAIlures before they occur, reducing downtime and extending the life of machinery. Additionally, AI in supply chain optimization helps companies forecast demand, manage inventories, and streamline logistics, resulting in cost savings. Robotics and automation powered by AI are also revolutionizing manufacturing, enabling 24/7 production without compromising quality [12]. These advancements lead to more efficient and flexible manufacturing systems.

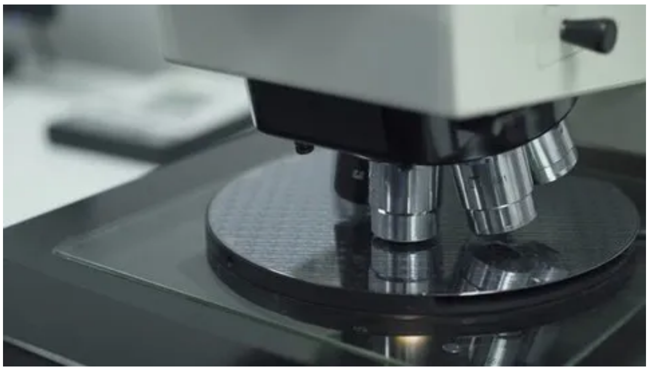


Figure 2.5: Electron Microscope Imaging of Silicon Wafer Defects powered by **ExtractAI** [51]

2.5 Limitations of Artificial Intelligence

While AI offers numerous applications and use cases across various fields, it is not without its challenges and inherent costs. In this section, we outline some of its key limitations.

- **Data Dependency and Quality:** AI systems, especially those based on machine learning and deep learning, require large amounts of high-quality, labelled data to perform effectively. Poor or biased datasets can lead to inaccurate or unfair outcomes. For instance, in medical imaging, an AI trained on unbalanced data may fail to generalize across diverse patient populations [13].
- **Lack of Explain-ability (Black-Box Models):** Many advanced AI models, particularly deep neural networks, operate as ”black boxes,” meaning their internal decision-making processes are not transparent. This lack of interpretability poses challenges in critical fields like healthcare and finance where understanding the rationale behind a decision is essential [14].
- **Generalization Challenges:** AI models often struggle to generalize beyond the specific conditions of their training datasets. They may perform well in controlled environments but fAIl in real-world, unpredictable scenarios [15]. This limits their reliability and deployment across dynamic settings.
- **Ethical and Bias Issues:** AI can inadvertently perpetuate or even amplify social biases present in training data, leading to discriminatory outcomes in areas such as hiring, law enforcement, or healthcare [16]. Ethical concerns also arise around privacy, surveillance, and consent.
- **High Computational and Resource Costs:** Training and deploying state-of-the-art AI models can be computationally intensive, requiring significant energy and financial resources. This environmental and economic cost restricts access to AI technologies, especially for smaller organizations [17].

- **Vulnerability to Adversarial Attacks:** AI models, particularly those used in security-critical applications like autonomous driving or medical diagnostics, can be vulnerable to adversarial attacks—small, carefully crafted perturbations that can mislead the AI into making wrong decisions [18].

2.6 Conclusion

Artificial Intelligence has swiftly evolved from a theoretical framework to a transformative force deeply integrated into diverse sectors of modern society. In this chapter, we examined the foundational principles of AI, highlighting its central objective of replicating intelligent behaviour. We also reviewed how AI applications are revolutionizing various fields: enhancing diagnostic precision in healthcare, strengthening financial security through fraud detection, enabling the development of autonomous vehicles, personalizing customer experiences in retail, promoting sustainable agricultural practices, and improving operational efficiency in manufacturing. While these examples underscore AI's substantial contributions, they also reveal critical limitations, including ethical challenges, inherent biases, and technical difficulties related to transparency and generalization.

Chapter 3

Foundational Tools and Algorithms for Artificial Intelligence

3.1 Introduction

Solving tasks such as object recognition in images or speech recognition, although seemingly simple, has proven extremely challenging when approached using traditional manual methods. This difficulty arises from the inherent complexity and variability of real-world data, such as sound recordings or image pixels. For machines, an image is merely a matrix of numbers representing pixel intensities, and a sound signal is a sequence of numerical values corresponding to variations in air pressure over time. Thanks to advances in artificial intelligence, particularly machine learning, machines have gained the ability to interpret these numerical patterns — transcribing sound signals into sequences of words and identifying objects like cats in images, even as their appearance and surrounding context vary significantly.

In this chapter, we explore the key tools and algorithms that underpin the development of artificial intelligence. We begin by examining the subfield known as Machine Learning, followed by a detailed discussion of its main types — with a particular focus on Deep Learning. We delve into the concept of artificial neural networks, their various architectures, and the principles that govern their operation, to provide a solid foundation for understanding deep learning. Finally, we outline how deep learning models function, explore their primary domains of application, and conclude the chapter with practical and illustrative examples.

3.2 Machine Learning

3.2.1 Definition

Machine Learning (ML) is a subfield of artificial intelligence that focuses on enabling systems to learn patterns from data and improve their performance on specific tasks without being explicitly programmed.

As defined by Tom M. Mitchell:

> "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."[2]

3.2.2 Types of Machine Learning

Supervised Learning

In supervised learning, the model is trained on labelled data, meaning that each input is paired with the correct output (target). The goal is for the model to learn the mapping from input to output, so it can predict the output for new, unseen data [19].

Common Algorithms used:

- Linear Regression
- Logistic Regression
- Support Vector Machines (SVM)
- Decision Trees
- k-Nearest Neighbors (k-NN)

Example: Email spam detection, where emails are labelled as **"spam"** or **"not spam"**.

Unsupervised Learning

In unsupervised learning, the model is trained on unlabeled data. The goal is to find patterns or structures within the data without any predefined outputs [20]. It is typically used for clustering, dimensionality reduction, and anomaly detection.

- Common Algorithms used:

- K-means Clustering
- Hierarchical Clustering
- Principal Component Analysis (PCA)
- Auto-encoders

Example: Customer segmentation in marketing, where the model groups customers based on purchasing behaviour without predefined labels.

Semi-supervised Learning

Semi-supervised learning lies between supervised and unsupervised learning. The model is trained on a small amount of labelled data and a large amount of unlabelled data. This method is useful when labelling data is expensive or time-consuming [21].

- Common Algorithms used:
  - Semi-supervised Support Vector Machines (S3VM)
  - Semi-supervised Learning with Graphs

Example: Image recognition where only a few images are labelled, but many more are unlabelled.

Reinforcement Learning

In reinforcement learning, the model learns by interacting with an environment and receiving feedback in the form of rewards or penalties. The goal is to learn a policy that maximizes the cumulative reward over time [22].

- Common Algorithms used:
  - Q-learning
  - Deep Q Network (DQN)
  - Policy Gradient Methods
  - Proximal Policy Optimization (PPO)

Example: A self-driving car learning to navigate a road based on rewards for reaching the destination and penalties for collisions.

Self-supervised Learning

Self-supervised learning is a special type of unsupervised learning where the model generates its own labels from the data. It is often used in Natural Language Processing (NLP) and computer vision [23].

- Common Algorithms used:
  - Contrastive Predictive Coding (CPC)
  - Bidirectional Encoder Representations from Transformers (BERT)

Example: Language models like **GPT**, which predict the next word in a sentence based on context.

Transfer Learning

Transfer learning is a technique where a model trained on one task is reused as the starting point for a model on a second task. This is particularly useful when there is a lack of sufficient labelled data for the new task but abundant data for a related one.

- Common Algorithms used:
  - Fine-tuning pre-trained neural networks
  - Domain adaptation techniques

Example: Using a pre-trained image classification model, such as one trained on **ImageNet**, to classify medical images with a smaller dataset.

3.3 Deep Learning

3.3.1 Definition

Deep learning is a subfield of machine learning that focuses on learning data representations through architectures composed of multiple layers of artificial neural networks. These networks are designed to automatically learn hierarchical features from raw input data without the need for extensive manual feature engineering [3]. Deep learning excels in tasks such as image and speech recognition, natural language processing, and complex decision-making.

From the concepts explored earlier, the hierarchical relationship between the three classes (Artificial Intelligence, Machine Learning and Deep Learning) can be illustrated as follows:

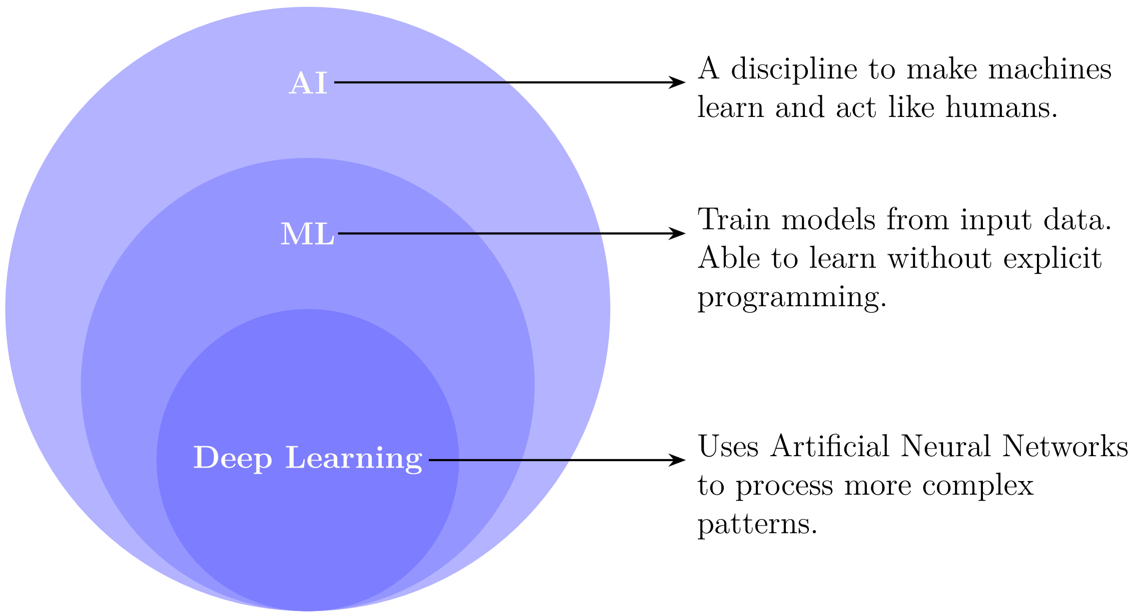


Figure 3.1: Relationship between AI, ML, and Deep Learning

3.3.2 Artificial Neural Network

Deep learning relies on a sequence of layers composed of non-linear processing units known as neurons to extract or transform data features. The output of each layer serves as the input to the next, forming what is known as an **artificial neural network**. These networks are inspired by the structure and function of biological neurons in the human brain. An Artificial Neural Network (ANN) consists of numerous interconnected artificial neurons, and the depth of the network increases with the number of layers and neurons involved. A deep learning model is typically structured as a series of layers:

- **Input layer:** Receives the raw data (e.g., pixel values from an image).
- **Hidden layers:** Multiple layers where neurons are connected with weights; these layers extract features of increasing complexity.
- **Output layer:** Produces the final result (e.g., class label or probability).

Each connection between neurons has an associated weight, and learning consists of updating these weights based on a loss function using optimisation techniques like back-propagation and gradient descent [5].

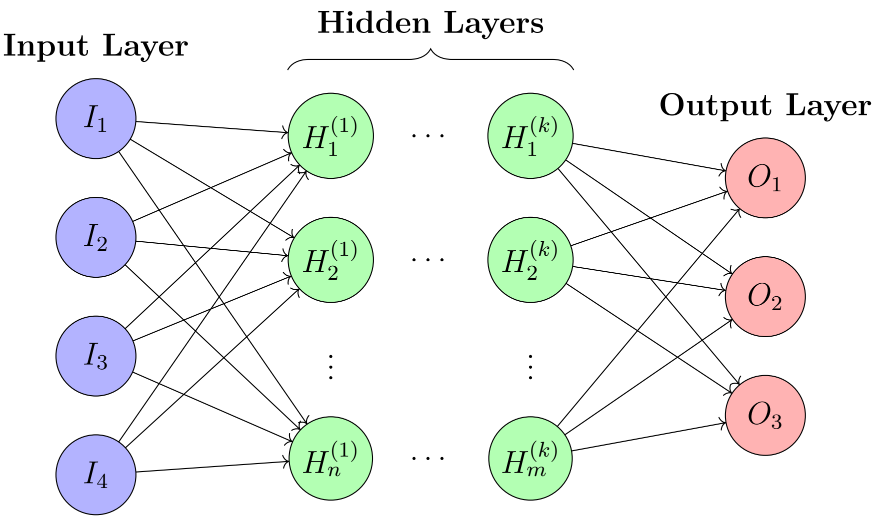


Figure 3.2: Structure of a typical Artificial Neural Network

3.3.3 Artificial Neuron

An artificial neuron, which serves as the fundamental building block of an artificial neural network, takes multiple inputs $x_1, x_2, ..., x_n$ each multiplied by an associated weight $w_1, w_2, ..., w_n$, sums them along with a bias b, and passes the result through an activation function f. This process can be expressed as:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right)$$

f introduces non-linearity, enabling the network to learn complex patterns. Common choices include **sigmoid**, **ReLU**, and **tanh**. During training, the weights and bias are adjusted using **back-propagation** and an optimisation algorithm such as **stochastic gradient descent** to minimise prediction error [3, 5, 2].

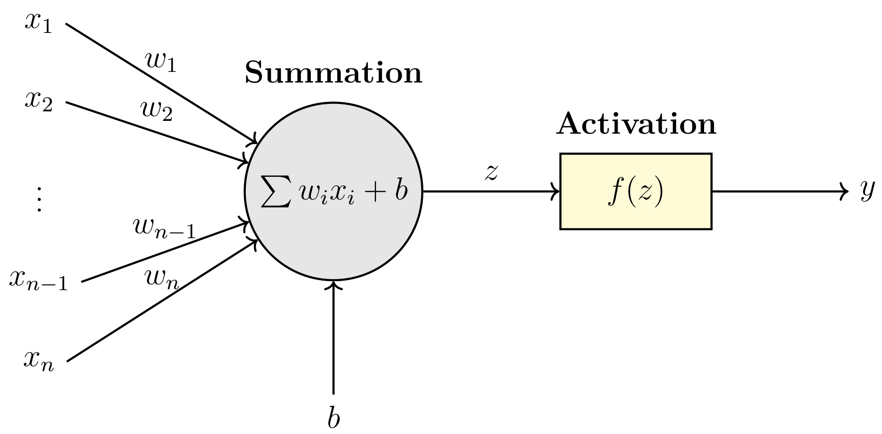


Figure 3.3: Mathematical Structure of a Artificial Neuron

3.4 Architectures of Deep Learning Neural Networks

3.4.1 Feedforward Neural Network

The simplest form of neural networks where information moves in only one direction forward from input nodes, through hidden layers, to output nodes. They are commonly used for tasks such as classification and regression.

3.4.2 Convolutional Neural Network

Designed specifically for processing grid-like data such as images. They use convolutional layers to automatically detect spatial features like edges, textures, and shapes, making them ideal for image and video analysis [5].

3.4.3 Recurrent Neural Network

Specialized for sequential data such as time series or natural language. They have loops that allow information to persist through steps, enabling the network to learn from past inputs.

3.4.4 Long Short-Term Memory Network and Gated Recurrent Unit

Advanced types of Recurrent Neural Network (RNN)s that address the **vanishing gradient problem**. They are better at learning long-term dependencies in sequences, making them effective in language modelling and translation.

3.4.5 Auto-encoders

Unsupervised neural networks that learn to compress data into a lower-dimensional representation and then reconstruct the original input. They are often used for noise reduction and dimensionality reduction.

3.4.6 Generative Adversarial Network

Comprise two competing networks, a generator and a discriminator. The generator creates synthetic data, while the discriminator evaluates them against real data. This setup is used to generate realistic images, text, or audio [3].

3.4.7 Transformers

A powerful architecture introduced for handling sequential data without recurrence. Transformers rely on self-attention mechanisms to process entire sequences in parallel, making them the backbone of modern language models such as **BERT** and **GPT** [24].

3.5 Conclusion

In this chapter, we have explored the foundational tools and algorithms that enable artificial intelligence systems to learn and make decisions. We began by introducing machine learning, defining its primary categories, supervised, unsupervised, and reinforcement learning, each offering distinct approaches to extract patterns from data. We then dived into deep learning, a subset of machine learning that uses artificial neural networks to model complex representations. Special emphasis was placed on the architecture and functioning of artificial neurons, which serve as the fundamental building blocks of deep neural networks. Various types of deep learning architectures were also discussed, each tailored to specific tasks and data types. Through these explorations, it became clear that the strength of AI lies not only in its computational power but also in the flexibility and adaptability of its underlying algorithms. This theoretical foundation paves the way for the next chapter, where we apply these methods in the context of medical image classification using Convolutional Neural Networks.

Chapter 4

Image Classification Using Convolutional Neural Networks

4.1 Introduction

Convolutional Neural Networks (CNNs) have become the cornerstone of deep learning models for image classification tasks, owing to their exceptional ability to automatically learn and extract hierarchical features from raw pixel data. Unlike traditional machine learning models that rely on manual feature extraction, CNNs use layers of convolutional operations to learn relevant features, making them highly efficient for visual recognition tasks such as object detection, facial recognition, and medical image analysis. Inspired by the human visual system, CNNs capture complex patterns and structures in images, which are crucial for accurate classification [5].

CNNs have demonstrated exceptional performance in a variety of applications, particularly in image recognition and classification. They have proven effective in tasks such as detecting objects in photographs, recognizing faces, and identifying abnormalities in medical images. Their ability to learn intricate patterns in images is key to their success, enabling accurate and reliable feature recognition, which is essential for decision-making and analysis across various domains.

In this chapter, we explore CNNs in detail, focusing on their architecture and key components. We will discuss the fundamental principles behind each aspect of CNNs and their crucial role in efficient image classification.

4.2 Definition

A Convolutional Neural Network is a type of deep learning model specifically designed to process structured data such as images, text or audio. CNNs are a subclass of feedforward neural networks that learn to automatically extract hierarchical features from raw input data through convolutional layers. These layers use filters (or kernels) that optimize the extraction of low- and high-level features, such as edges and textures, from images. CNNs are particularly powerful for visual recognition tasks as they are capable of capturing complex patterns and structures in images, which is essential for accurate classification. Unlike traditional machine learning models, which require manual feature engineering, CNNs leverage the convolution operation to learn relevant features directly from raw data. This ability to learn from data, rather than relying on predefined rules, makes CNNs the de facto standard for deep learning-based approaches to computer vision and image processing [3, 5].

A convolutional neural network (CNN) is a type of feedforward neural network that learns features via filter (or kernel) optimization. This type of deep learning network has been applied to process and make predictions from many different types of data, including text, images, and audio [5]. Convolution-based networks are the de-facto standard in deep learning-based approaches to computer vision and image processing [5].

4.3 Biological inspiration Behind CNNs

The architecture of CNNs is inspired by the structure and functioning of the human visual system[39]. In the human brain, the visual cortex processes visual information through a hierarchical structure, where simple features such as edges and lines are first detected and then progressively combined to recognize more complex patterns, objects, and scenes. Similarly, CNNs are designed to process images through layers that progressively extract more complex features.

The initial layers of a CNN function like the early stages of visual processing in the human brain, detecting simple features such as edges and corners. As we move deeper into the network, more complex features like textures, patterns, and shapes are identified. The final layers of the CNN combine these features to make a decision or classification, similar to how the human brain processes visual data to identify and interpret objects in the environment.

This biologically inspired design enables CNNs to efficiently handle image recognition tasks by mimicking the hierarchical way in which humans process visual stimuli, allowing them to extract increasingly sophisticated features from raw input data. This inspiration from the visual system has been one of the key factors behind the success of CNNs in computer vision tasks, such as object detection, facial recognition, and medical image analysis.

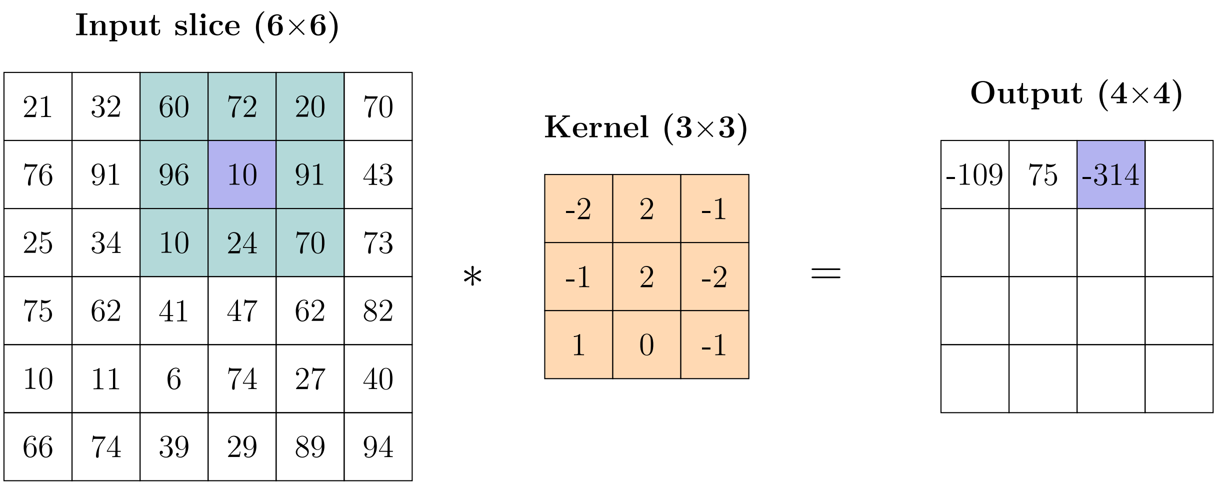
4.4 Convolutional Neural Network Architecture

Convolutional neural networks are structured hierarchically, designed to progressively extract increasingly complex features from raw input data, typically images. The architecture can be divided into two main parts: the **feature extraction part** and the **classification part**. Each part comprises several distinct layers, each serving specific functions essential for effective and accurate image classification.

4.4.1 Feature Extraction Part

Convolutional Layers:

Convolutional layers are the cornerstone of a CNN, responsible for extracting features from the input data. They perform the convolution operation, where a filter (or kernel) slides over the input image to compute dot products, creating feature maps. This allows CNNs to capture spatial hierarchies and local dependencies in the data.[25]



sample calculation:

$(60 \times -2) + (72 \times 2) + (20 \times -1) +$

$(96 \times -1) + (10 \times 2) + (91 \times -2) +$

$(10 \times 1) + (24 \times 0) + (70 \times -1) = \boxed{\mathbf{-314}}$

Figure 4.1: Example of a Convolutional Operation

Activation Functions:

The activation function introduces non-linearity to the model, enabling it to learn complex patterns. Typically, the Rectified Linear Unit (ReLU) is used, but other activation functions like sigmoid or tanh can be employed depending on the specific use case.[26]

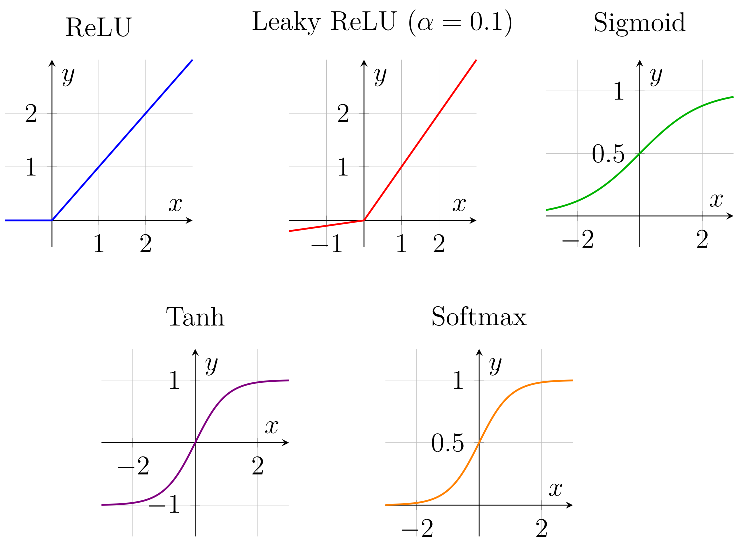


Figure 4.2: Common Activation Functions in CNNs.

Pooling Layers:

Pooling layers reduce the spatial dimensions (width and height) of the input, helping reduce computation and memory usage while preserving essential features. Common types of pooling include max pooling, where the maximum value from each patch is taken, and average pooling, which calculates the average of the patch.[27]

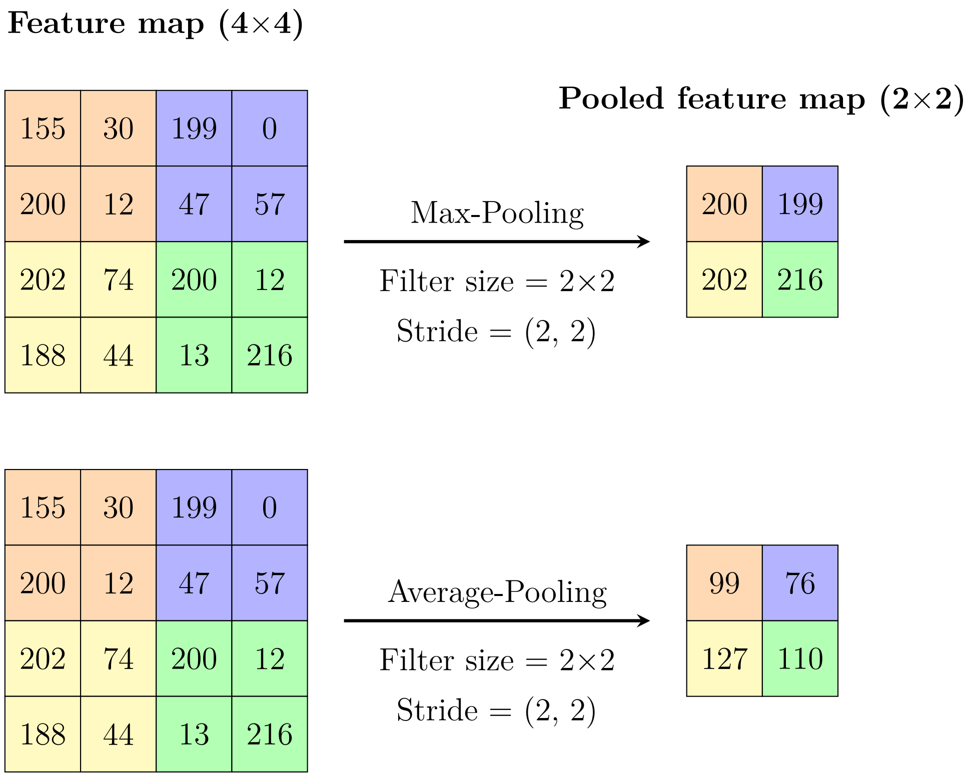


Figure 4.3: Example of Pooling operations

4.4.2 Classification Part

Fully Connected Layers:

Fully connected layers are traditional layers found in neural networks, where each neuron is connected to every neuron in the previous layer. These layers are used after feature extraction to make final predictions.[28]

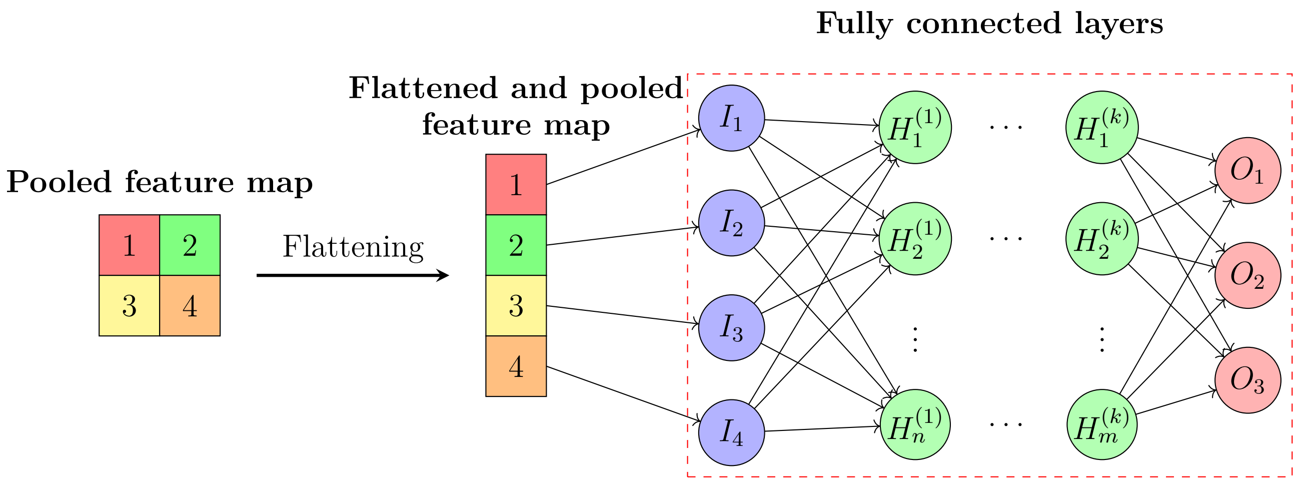


Figure 4.4: A pooled feature map being flattened and inputted into a fully connected layer

Output Layer:

The output layer in a CNN is responsible for producing the final classification or prediction. The type of output layer depends on the task: for binary classification, a sigmoid activation function is typically used, while for multi-class classification, a softmax activation function is employed. [3]

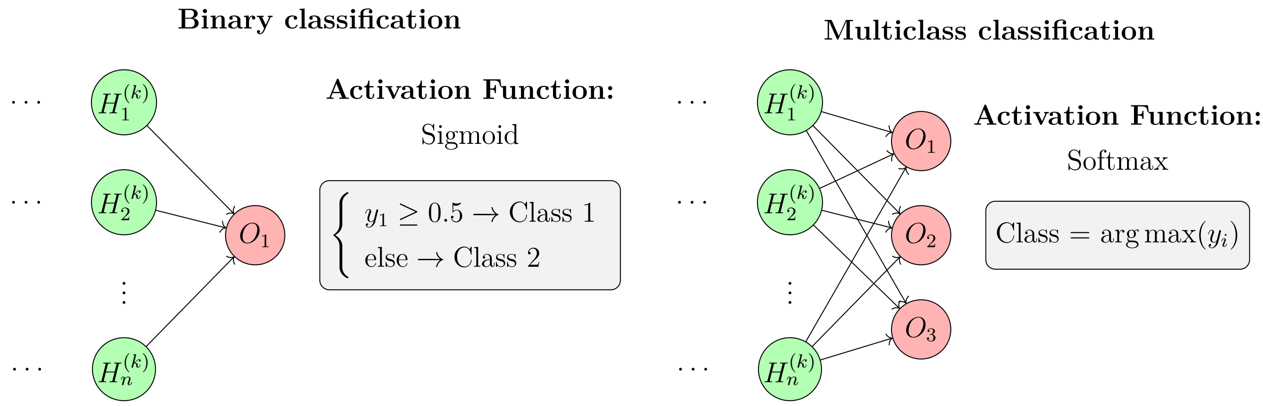


Figure 4.5: Class Selection in CNNs for single and multi output

By integrating all these fundamental components, each contributing a specific role in feature extraction and classification, we arrive at the complete structure of a Convolutional Neural Network:

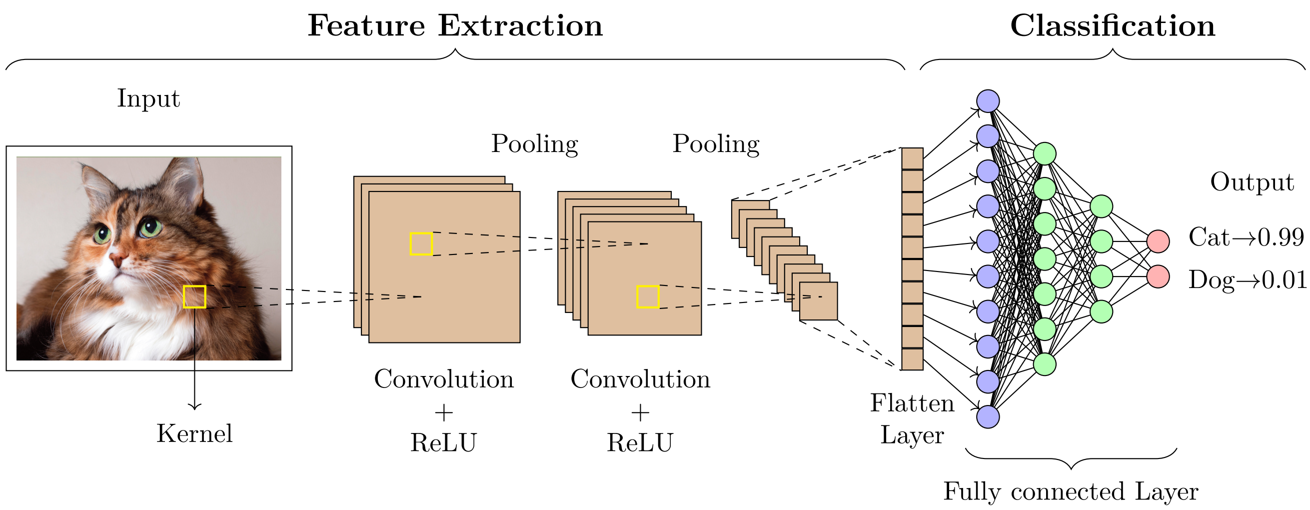


Figure 4.6: General structure of a Convolutional Neural Network (CNN)

4.5 Classification Assessment

In order to evaluate classification systems and compare their robustness for a given application, a classic approach consists in using the confusion matrix. A confusion matrix is a technique for summarizing the performance of a classification algorithm. Calculating a confusion matrix can give us a better idea of what our classification model is getting right and what types of errors it is making [38].

4.5.1 Classification Accuracy and its Limitations

Classification accuracy is the ratio of correct predictions to total predictions made.

$$\text{Classification accuracy} = \frac{\text{Correct predictions}}{\text{Total predictions}} \times 100$$

Classification accuracy can also easily be turned into a misclassification rate or error rate by inverting the value, such as:

$$\text{Misclassification rate} = \left(1 - \frac{\text{Correct predictions}}{\text{Total predictions}}\right) \times 100$$

The main problem with classification accuracy is that it hides the detail we need to better understand the performance of our classification model. For instance, with 3 or more classes we may get a classification accuracy of 80%, but we don't know if that is because all classes are being predicted equally well or whether one or two classes are being neglected by the model. But thankfully we can avoid this last problem by using a confusion matrix [38].

The confusion matrix shows the way in which our classification model is confused when it makes predictions. It gives us insight not only into the errors being made by our classifier but more importantly the types of errors that are being made. It is this breakdown that overcomes the limitation of using classification accuracy alone. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key of the confusion matrix [38].

It is a two-dimensional matrix, indexed in one dimension by the true class of an object and in the other by the class that the classifier assigns. For binary classification, the confusion matrix is a 2 × 2 table that displays the four possible combinations of predicted and actual class labels [38] (as shown in Table 4.1):

	Class 1 Predicted	Class 2 Predicted
Class 1 Actual	**TP**	**FN**
Class 2 Actual	**FP**	**TN**

Table 4.1: Confusion matrix for binary classification

With:

- **Class 1:** Positive
- **Class 2:** Negative
- **Positive (P):** Observation is positive (for example: is an apple).
- **Negative (N):** Observation is not positive (for example: is not an apple).
- **True Positive (TP):** Observation is positive, and is predicted to be positive.
- **False Negative (FN):** Observation is positive, but is predicted negative.
- **True Negative (TN):** Observation is negative, and is predicted to be negative.
- **False Positive (FP):** Observation is negative, but is predicted positive.

4.5.2 Classification Rate/Accuracy

Classification Rate or Accuracy is given by the relation:

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

However, there are problems with accuracy. It assumes equal costs for both kinds of errors. A 99% accuracy can be excellent, good, mediocre, poor or terrible depending upon the problem [38].

4.5.3 Recall

Recall can be defined as the ratio of the total number of correctly classified positive examples divide to the total number of positive examples. High Recall indicates the class is correctly recognized (small number of FN) [38]. Recall is given by the relation:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

4.5.4 Precision

To get the value of precision we divide the total number of correctly classified positive examples by the total number of predicted positive examples. High Precision indicates an example labelled as positive is indeed positive (small number of FP) [38]. Precision is given by the relation:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

High recall, low precision: This means that most of the positive examples are correctly recognized (low FN) but there are a lot of false positives.

Low recall, high precision: This shows that we miss a lot of positive examples (high FN) but those we predict as positive are indeed positive (low FP).

4.5.5 F-measure (F1-score)

Since we have two measures (Precision and Recall), it is helpful to have a measurement that represents both of them. We calculate an F-measure which uses Harmonic Mean in place of Arithmetic Mean as it punishes more the extreme values. The F-Measure will always be nearer to the smaller value of Precision or Recall [38].

$$\text{F-measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

4.6 Conclusion

In this chapter, we presented a comprehensive overview of Convolutional Neural Networks (CNNs), a cornerstone of modern deep learning techniques in image classification. We began with a definition of CNNs and discussed their biological inspiration, particularly the human visual cortex. We then examined the essential components that form the architecture of CNNs, including convolutional layers, activation functions, pooling layers, fully connected layers, and output layers, each playing a crucial role in enabling hierarchical feature extraction and robust classification.

Additionally, we introduced key performance metrics used to evaluate classification models, such as accuracy, precision, recall and F-measure. These indices provide deeper insights into the model's predictive capabilities.

This foundation sets the stage for the next chapter, where we will explore how CNNs are applied in image classification tasks, the methodologies used, and a detailed evaluation of model performance using these metrics.

Chapter 5

Implementation of CNN-Based Classification on medical images

5.1 Introduction

This chapter provides a detailed walkthrough of the complete process involved in developing an image classification model using CNNs. It begins with an overview of the tools, libraries, and computing environment employed for implementation, followed by an analysis of the dataset, its structure, content, and suitability for the classification task.

Next, we outline the preprocessing steps required to prepare the data, including image resizing, normalisation, and other transformations aimed at improving model performance. The architecture of the CNN is then presented, with emphasis on the design rationale behind layer selection and activation functions.

The model is trained using a well-defined set of hyperparameters, and its performance is assessed using standard evaluation metrics such as accuracy, precision, recall, and loss. These results are complemented by visualisations to better interpret training behaviour and classification performance. The chapter concludes with a discussion of the outcomes, highlighting the model's strengths and limitations.

5.2 Tools and Implementation Environment

5.2.1 Hardware Setup

The training and experimentation for the CNN model were conducted using **Google Colab**, a cloud-based platform that provides free access to computational resources, including GPUs. This setup eliminates the need for powerful local hardware while ensuring efficient training times.[41]

During implementation, Colab provided access to the following resources:

Component	**Specification**
Processor	Intel(R) Xeon(R) CPU @ 2.30GHz
RAM	15 GB
GPU	NVIDIA Tesla T4
Operating System	Ubuntu-based virtual environment
Disk space	112.6 GB

Table 5.1: Hardware configuration used in Google Colab

Using Google Colab allowed for seamless integration with Python libraries and facilitated the training of deep learning models without the constraints of local computational limitations.

5.2.2 Python Programming Language

Python was selected as the programming language for this project due to its simplicity, readability, and the vast ecosystem of tools it offers for scientific computing and machine learning. Its clean syntax and extensive support libraries enable rapid prototyping and ease of implementation, which is essential for developing and testing deep learning models efficiently [29]. Moreover, Python's integration with popular machine learning and deep learning frameworks has made it the de facto standard language in the AI research community [30]. Its ability to handle tasks ranging from data preprocessing to model evaluation within a single, cohesive environment further solidifies its suitability for end-to-end machine learning pipelines [31].

5.2.3 Libraries and Frameworks

The implementation of this image classification model was supported by a set of powerful libraries and frameworks, each serving specific functions ranging from data manipulation to model building and evaluation.

TensorFlow and Keras: For constructing the CNN, TensorFlow was used, a versatile framework designed for deep learning tasks. It supports large-scale machine learning applications and is widely adopted in both research and production environments [32, 42]. Keras, which is integrated with TensorFlow, was utilised as the high-level API for designing the model due to its user-friendly interface that facilitates rapid prototyping of deep learning models [40, 43]. Keras allows for the easy definition of layers, activation functions, and optimisers, significantly streamlining the process of model creation. This combination of TensorFlow and Keras ensures both flexibility and ease of use.

NumPy: a fundamental library for numerical computations in Python, providing efficient data structures like arrays that are essential for storing and manipulating image data [33]. It was used throughout the project for operations such as data normalisation and image transformations, which are crucial steps in the preprocessing pipeline. NumPy's optimised performance ensures that large datasets can be handled efficiently, making it an ideal choice for high-volume image classification tasks.

Pandas: a data manipulation and analysis library that facilitates working with structured data [34]. It was primarily used for organising and processing the dataset, particularly in the preparation stages before feeding the data into the CNN. The ability to handle large datasets with ease and support for operations like data merging and reshaping made it a vital tool for data preprocessing and analysis.

Matplotlib and Seaborn: For visualising the training process and evaluating model performance, Matplotlib and Seaborn were employed [35, 36]. Matplotlib is a versatile library for creating static, animated, and interactive plots. It was used to generate training and validation accuracy/loss graphs, which helped in monitoring model convergence. Seaborn, built on top of Matplotlib, was used to create more advanced statistical plots, such as confusion matrices, which offer a detailed view of classification performance.

Scikit-learn: a widely-used open-source machine learning library in Python, Scikit-learn provided essential tools for evaluating model performance and preprocessing the dataset. It was employed to perform stratified dataset splitting, ensuring proportional class representation across training, validation, and test sets. For evaluation, Scikit-learn's utilities such as the confusion matrix and classification report were used to compute and visualise some performance metrics including accuracy, and loss values. [37, 44].

5.3 Chest X-ray Classification

5.3.1 Data Description and Analysis

The dataset used in this study is titled *COVID, Pneumonia & Normal Chest X-ray Images* and is publicly available on Kaggle[45]. It consists of Chest X-ray (CXR) images organised into three distinct categories: **COVID-19**, **Normal**, and **Pneumonia**, each stored in separate subfolders. This directory structure is ideal for supervised learning in medical image classification.

Dataset Composition

- **COVID-19:** 1,626 images
- **Normal:** 1,802 images
- **Pneumonia:** 1,800 images
- **Total:** 5,228 images

All images are preprocessed and resized to 256×256 pixels and saved in PNG format.

Data Sources and Credibility

The images have been collected from reliable and publicly accessible medical image repositories:

- Eurorad
- Radiopaedia
- P. Mooney, Chest X-ray Pneumonia Datasets

Visual and Clinical Characteristics

- **Normal:** Refers to chest X-ray images of individuals with healthy lungs. These images do not show any signs of infection, inflammation, or abnormalities in the lung fields, pleura, or surrounding structures.

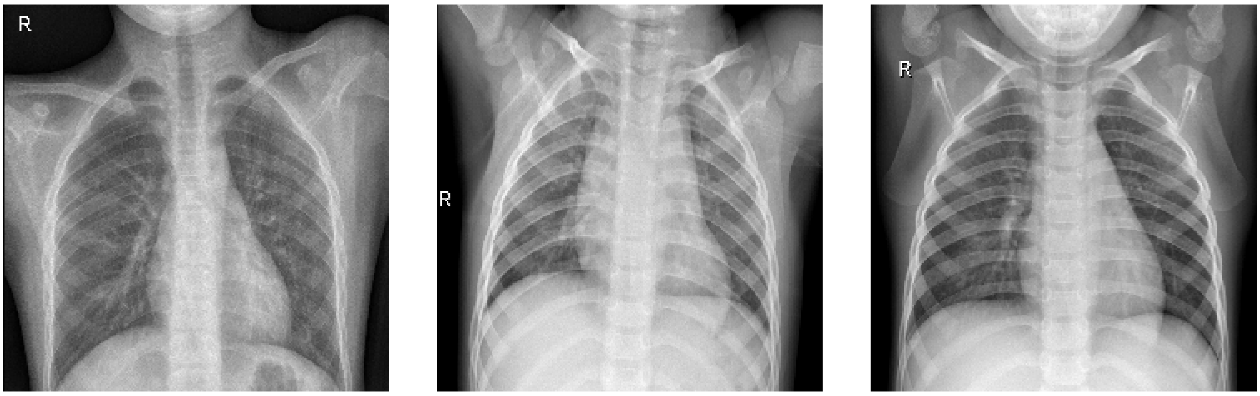


Figure 5.1: Examples of Chest X-Rays from healthy patients

- **Pneumonia:** An infection that inflames the air sacs in one or both lungs, which may fill with fluid or pus. Pneumonia can be bacterial, viral, or fungal in origin. Chest X-rays may show localized or diffuse opacities, commonly affecting a single lung lobe (lobar pneumonia) or scattered areas (bronchopneumonia).

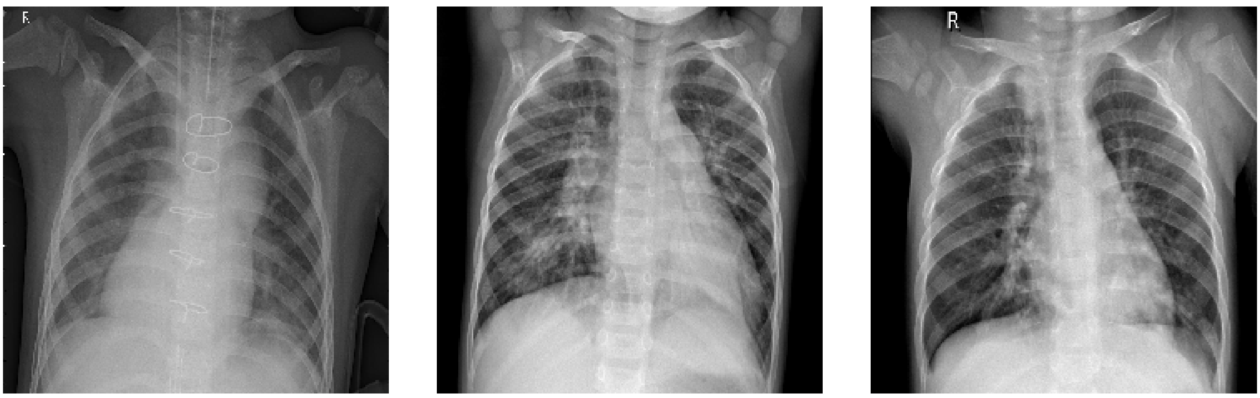


Figure 5.2: Examples of Chest X-Rays from patients with Pneumonia

- **COVID-19:** Caused by the **SARS-CoV-2** virus, COVID-19 can lead to viral pneumonia. In chest X-rays, COVID-19 often presents as bilateral, peripheral ground-glass opacities and consolidations, particularly in the lower lobes. These patterns are typically more diffuse and symmetric compared to other types of pneumonia.

Figure 5.3: Examples of Chest X-Rays from patients with COVID-19

Dataset Balance and Observations

The dataset is relatively balanced across all three categories, reducing the need for heavy rebalancing or data augmentation. However, clinical similarities between pneumonia and COVID-19 X-rays may pose challenges in model generalisation and accuracy.

5.3.2 Model Development and Training

This section details the entire model development pipeline, from preparing the input data to configuring and training the **DenseNet-121** model. It includes data preprocessing strategies, architectural choices, and training configuration adapted to the classification of chest X-ray images.

Data Preprocessing

Subsequently, pixel values were normalised to the $[0, 1]$ range by dividing each value by 255, a common preprocessing step that facilitates faster convergence during training.And to improve the model's ability to generalise and to mitigate the risk of overfitting, data augmentation techniques such as horizontal flipping, random rotations, and zooming were applied during the training phase using Keras' `ImageDataGenerator`.

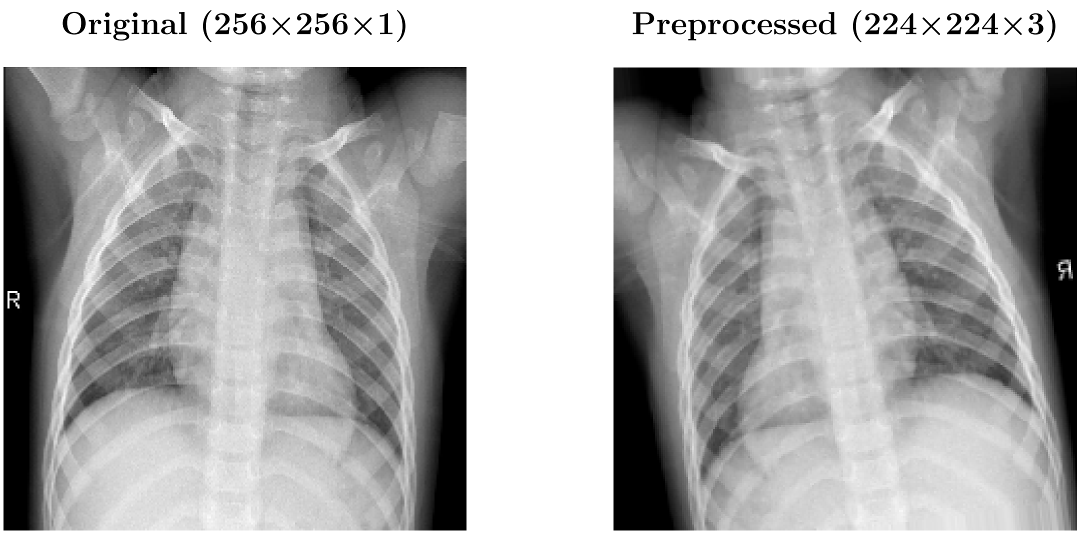


Figure 5.4: Image before preprocessing and after applying resizing and data augmentation

Model Architecture

For this project, we utilised **DenseNet-121**, a pre-trained convolutional neural network known for its efficient feature propagation and compact parameter count. The original final classification layer of DenseNet-121, which is designed for 1000 ImageNet classes, was replaced with a fully connected layer tailored to our three target classes (COVID-19, Pneumonia, Normal). A softmax activation function was applied to output class probabilities.

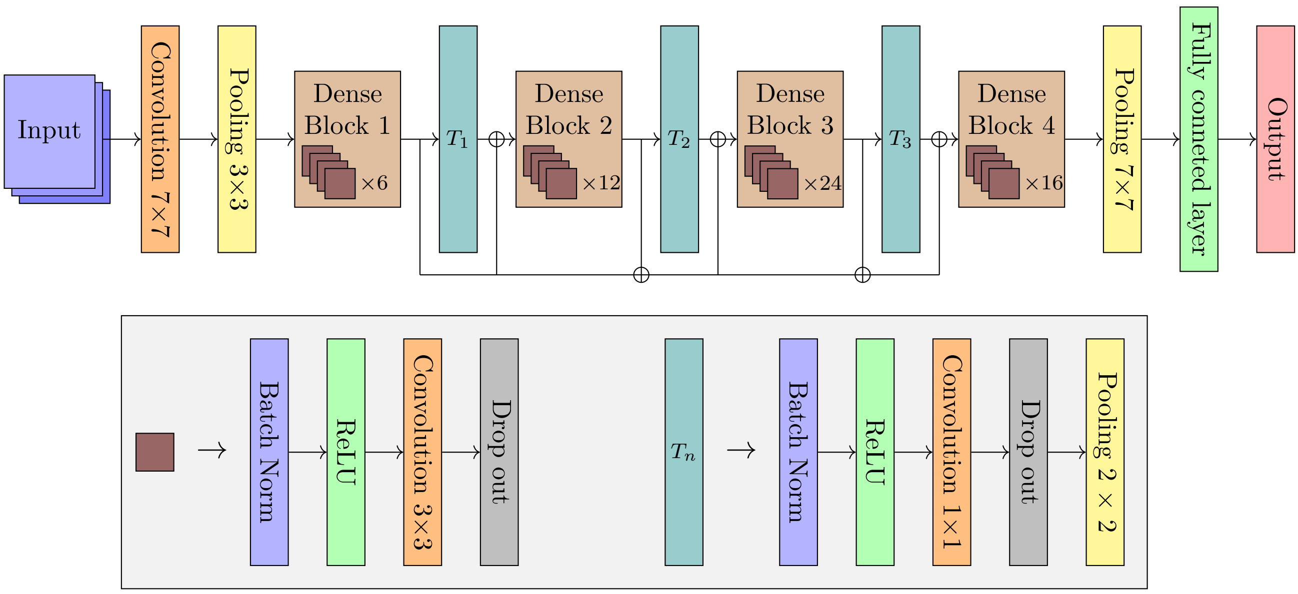


Figure 5.5: A schematic illustration of the DenseNet-121 architecture [52]

Training Configuration

Parameter	Value / Description
Optimizer	Adam
Learning Rate	0.0001
Batch Size	32
Epochs	50
Loss Function	Categorical Cross-Entropy
Data Split	Training: 70% , Validation: 15%, Test: 15%

Table 5.2: Training Configuration Summary

- **Optimizer (Adam):** An adaptive learning rate optimisation algorithm that combines the benefits of AdaGrad and RMSProp. It computes individual learning rates for different parameters from estimates of first and second moments of the gradients.
- **Learning Rate:** Determines the step size during each update of the model weights. A smaller values allows for more stable and gradual convergence.
- **Batch Size:** The number of samples processed before the model's internal parameters are updated.
- **Epoch:** One epoch is a complete pass through the entire training dataset.
- **Loss Function (Categorical Cross-Entropy):** A loss function commonly used in multi-class classification problems. It measures the dissimilarity between the predicted probability distribution and the true class labels.

To optimise training efficiency and reduce the risk of overfitting, early stopping was implemented to halt training when the validation performance ceased to improve over a defined number of epochs, thus preventing unnecessary computation. Additionally, a learning rate scheduler was used to dynamically reduce the learning rate when the model's performance plateaued, allowing finer adjustments during later training stages for better convergence.

5.3.3 Visualisation of Results

Training and validation accuracy and loss were plotted over the course of the training epochs to monitor the model's learning dynamics. These visualisations help in diagnosing training issues such as overfitting or underfitting. Furthermore, the confusion matrix offers a visual summary of the model's classification performance across the three classes: COVID-19, pneumonia, and normal.

Figure 5.6: Training and validation loss per epoch

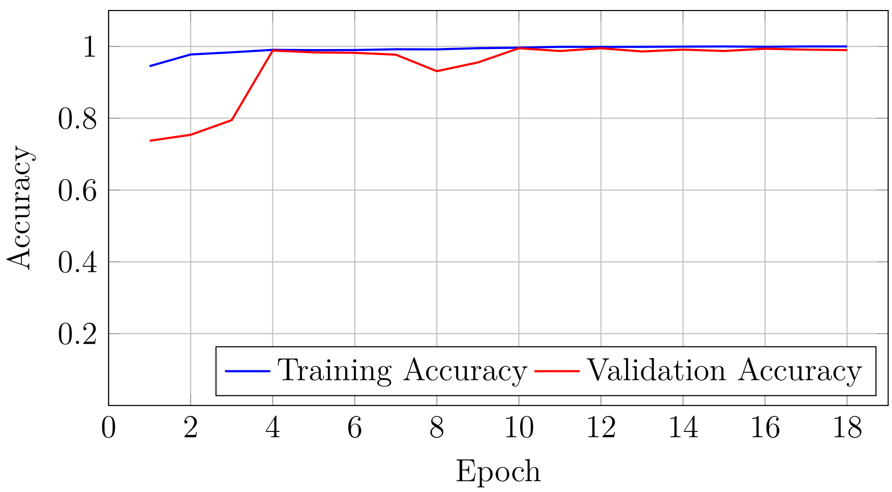


Figure 5.7: Training and validation accuracy per epoch

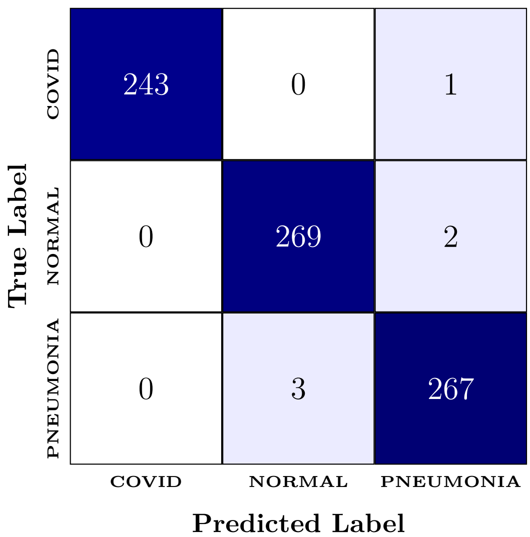


Figure 5.8: Confusion Matrix

5.3.4 Model Performance and Result Analysis

To evaluate the effectiveness of the DenseNet121 model on the chest X-ray classification task, its performance was analysed through visual representations of training dynamics and classification outcomes.

Training and Validation Performance

Figure 5.6 and 5.7 shows the evolution of training and validation loss and accuracy respectively over the course of training. The training loss decreases steadily and stabilises at a low value, indicating that the model is learning effectively from the training data. The validation loss initially fluctuates but quickly converges and remains consistently low after around the fifth epoch. These trends suggest that the model avoids overfitting and maintains generalisation to unseen data.

The accuracy plots confirm this observation. Training accuracy reaches near-perfect levels within the first few epochs and remains stable throughout. Meanwhile, validation accuracy starts lower but improves rapidly, surpassing 95% by epoch four and closely tracking the training accuracy thereafter. The narrow gap between training and validation accuracy implies that the model generalises well, with no signs of underfitting or overfitting.

Together, these results demonstrate that the model has learned discriminative features effectively and is well-optimised for the classification task.

Confusion Matrix Analysis

From the confusion matrix illustrated in Figure 5.8, the following classification performance indicators were derived:

Class	Precision	Recall	F1-score
COVID	1.0000	0.9959	0.9979
NORMAL	0.9890	0.9926	0.9908
PNEUMONIA	0.9889	0.9889	0.9889
Macro Average	0.9926	0.9925	0.9925
Weighted Average	0.9924	0.9924	0.9924
Overall Accuracy	0.9924		

Table 5.3: Classification metrics derived from the confusion matrix

The confusion matrix provides a detailed view of the model's classification performance across the three target categories: **COVID**, **NORMAL**, and **PNEUMONIA**. From this matrix, essential evaluation metrics such as precision, recall, F1-score, and overall accuracy have been computed and summarised in Table 5.3.

The model achieved an overall classification accuracy of **99.24%**, indicating highly reliable predictions on the test set. Breaking down the performance by class:

- **COVID:** The model achieved a perfect precision score of **1.0000**, meaning there were no false positives when identifying COVID cases. The recall of **0.9959** shows that only a small number of COVID cases were missed. The resulting F1-score of **0.9979** confirms excellent balance between sensitivity and specificity.

- **NORMAL:** The precision of **0.9890** and recall of **0.9926** reflect very few misclassifications, with the F1-score of **0.9908** affirming consistent performance on healthy X-rays.

- **PNEUMONIA:** With both precision and recall at **0.9889**, the model demonstrates uniform effectiveness in detecting pneumonia, yielding an identical F1-score.

The **macro average** and **weighted average** F1-scores are both around **0.9925**, suggesting that performance is not skewed by class imbalance. This is significant given that the dataset contains a slightly uneven distribution among classes.

The high values across all metrics confirm that the model is well-generalised and effective across all three categories, with minimal variance in performance between them. Such results indicate a overall good model.

Result Discussion

Overall, the DenseNet121 model demonstrated strong performance in the multi-class classification of chest X-ray images. The results from the confusion matrix and classification report confirmed that the model was able to distinguish effectively between the different categories, showing consistent and balanced performance across all classes.

The use of data augmentation, regularisation methods, and optimised training strategies such as early stopping and learning rate scheduling contributed to stable learning and enhanced generalisation. Although a small number of misclassifications occurred, particularly between visually similar classes like NORMAL and PNEUMONIA, the model maintained high reliability and robustness across classes.

5.4 Brain Tumour Classification

5.4.1 Data Description and Analysis

The dataset used is named *Brain tumor dataset* and is publicly available on Kaggle[46]. It consists of Brain tumour MRI images organised into four distinct categories: **Glioma**, **Meningioma**, **Pituitary** and **No Tumour**, each stored in separate subfolders.

Dataset Composition

- **Glioma:** 1621 images
- **Meningioma:** 1775 images
- **Pituitary tumour:** 1757 images
- **Normal:** 2000 images
- **Total:** 7,153 images

All images are in different pixel sizes and saved in JPG format.

Data Sources and Credibility

The images have been collected from various medical image repositories:

- Figshare
- Brain Tumor Classification (MRI)
- Br35H :: Brain Tumor Detection 2020

Visual and Clinical Characteristics

- **Normal:** Refers to brain MRI images of individuals without any detectable tumours or abnormalities. These images display a normal anatomical structure, with symmetric brain hemispheres and no signs of masses, tissue distortion, or abnormal intensities. The ventricles, grey and white matter, and surrounding areas appear well-defined and consistent.

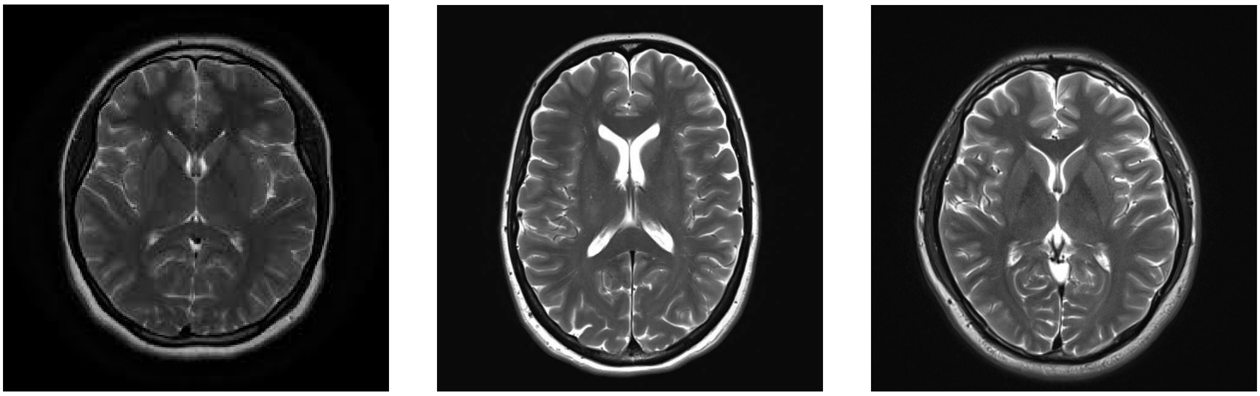


Figure 5.9: Examples of Brain MRI Scans from Healthy Individuals

- **Glioma:** A type of tumour that originates from glial cells in the brain. Gliomas tend to infiltrate the surrounding brain tissue, making their boundaries irregular and diffuse.

In MRI images, gliomas often appear as heterogeneous masses, possibly with necrotic or cystic regions. Depending on the grade, they may also exhibit surrounding oedema and cause midline shift or compression of nearby structures.

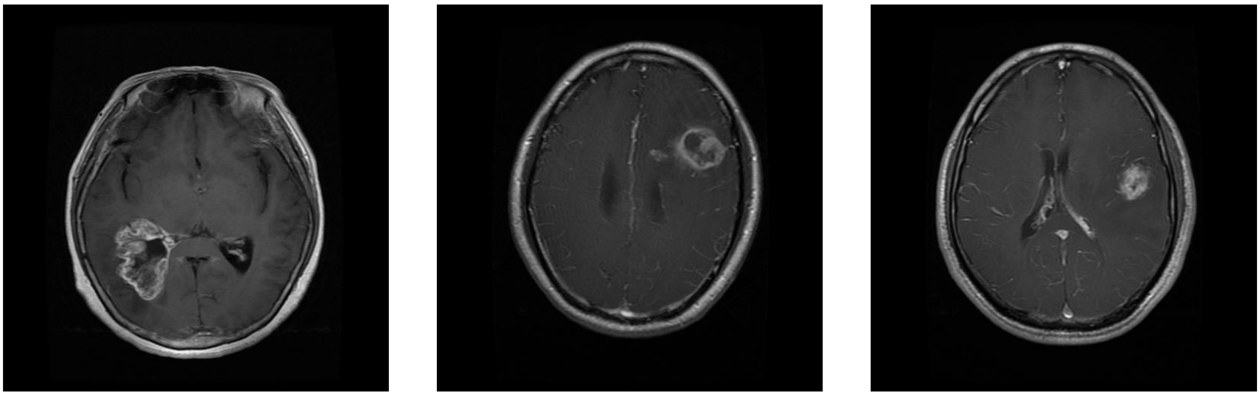


Figure 5.10: Examples of Brain MRI Scans from patients with Glioma

- **Meningioma:** Meningiomas arise from the meninges, the protective layers covering the brain and spinal cord. They are usually benign and grow slowly, forming well-circumscribed, rounded masses near the brain surface. In MRI scans, meningiomas typically appear homogenous and may cause compression of adjacent brain tissue without infiltrating it. They often enhance uniformly after contrast administration.

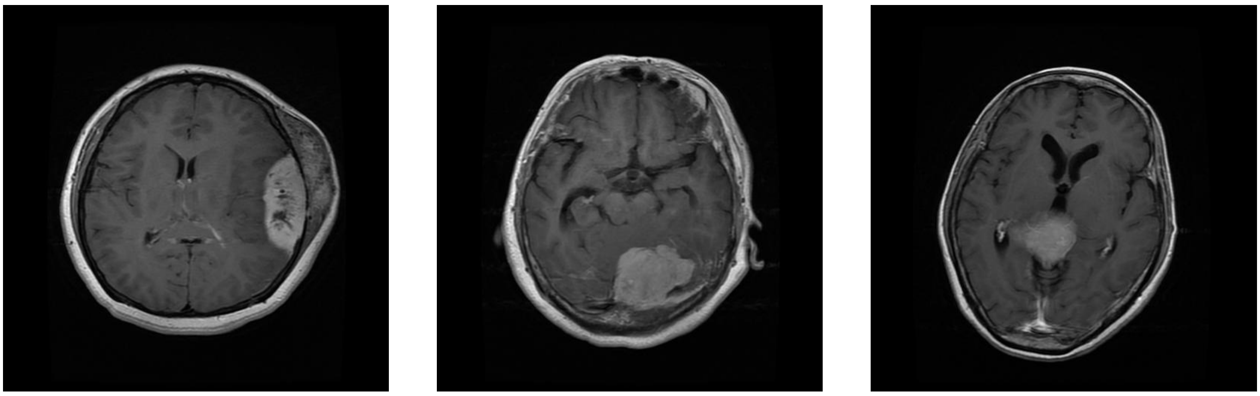


Figure 5.11: Examples of Brain MRI Scans from patients with Meningioma

- **Pituitary Tumour:** These tumours develop in the pituitary gland, located at the base of the brain. MRI images show them as distinct masses in the sella turcica region, often displacing or enlarging the gland. Larger tumours may compress the optic chiasm, potentially leading to visual disturbances. Depending on hormonal activity, these tumours may be associated with endocrine symptoms in clinical practice.

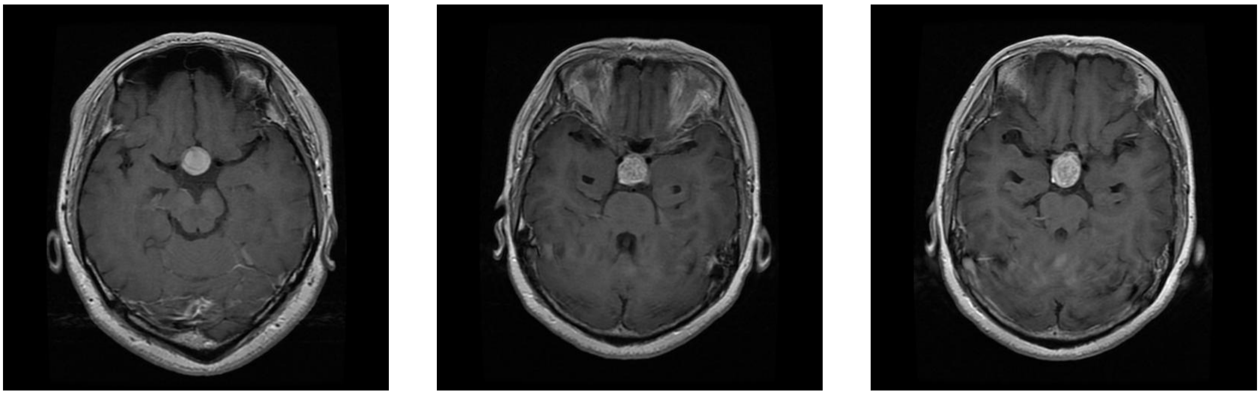


Figure 5.12: Examples of Brain MRI Scans from patients with Pituitary tumour

5.4.2 Model Development and Training

The same model development and training pipeline described in Section 5.3.2 was used for the brain tumour classification task. This includes data preprocessing, model architecture based on DenseNet121, and training configuration. Only the dataset and class labels differ, as detailed in the relevant subsections.

5.4.3 Visualisation of Results

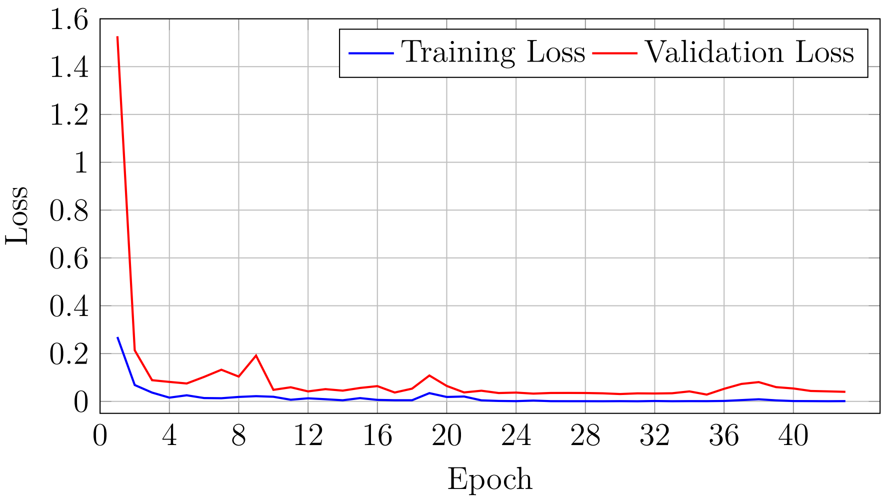


Figure 5.13: Training and validation loss per epoch

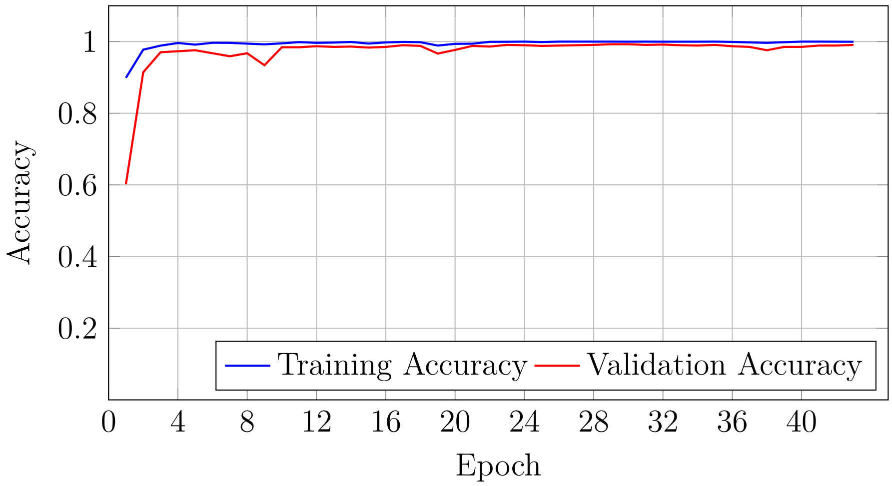


Figure 5.14: Training and validation accuracy per epoch

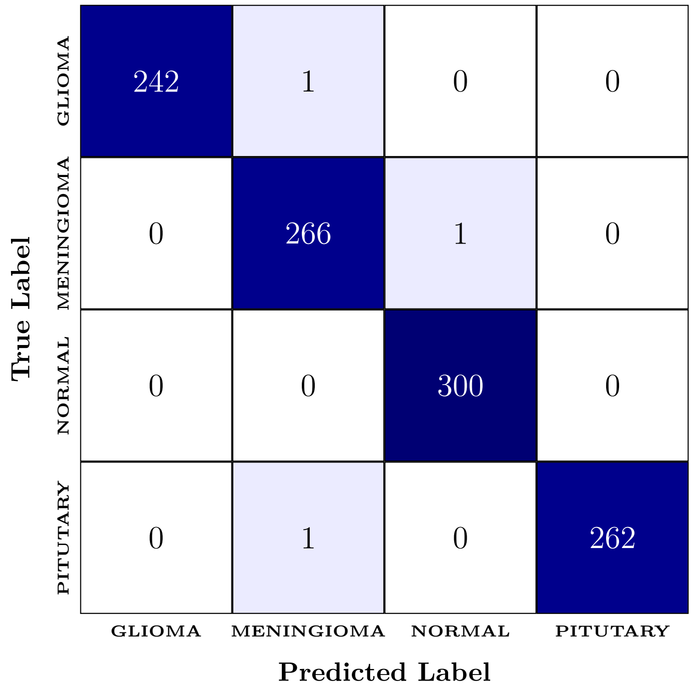


Figure 5.15: Confusion Matrix

5.4.4 Model Performance and Result Analysis

Training and Validation Performance

The training and validation performance of the DenseNet121 model is illustrated in Figures 5.13 and 5.14. The loss curves in Figure 5.13 show a sharp decrease during the initial epochs, followed by a rapid convergence for both training and validation loss. This early drop indicates that the model quickly learned to extract meaningful features from the input data. The fact that both loss curves flatten at low values suggests that the model maintained stability during training and did not overfit.

In parallel, the accuracy curves in Figure 5.14 demonstrate a rapid increase within the first few epochs, with both training and validation accuracy exceeding high levels early in the training process. The close alignment between the training and validation accuracy throughout the training cycle reflects a well-generalised model with minimal overfitting.

Together, these trends indicate that the model benefited from an effective training setup. Techniques such as data augmentation, early stopping, and learning rate scheduling likely contributed to smooth convergence and high generalisation capability. The consistent performance across both sets confirms that the model was not only able to learn discriminative features but also applied them effectively to unseen data.

Confusion Matrix Analysis

From the confusion matrix illustrated in Figure 5.15, the following classification performance indicators were derived:

Class	**Precision**	**Recall**	**F1-score**
GLIOMA	1.0000	0.9959	0.9979
MENINGIOMA	0.9925	0.9963	0.9944
NORMAL	0.9967	1.0000	0.9983
PITUITARY	1.0000	0.9962	0.9981
Macro Average	0.9973	0.9971	0.9972
Weighted Average	0.9972	0.9972	0.9972
Overall Accuracy	0.9972		

Table 5.4: Classification metrics derived from the confusion matrix

The confusion matrix provides a detailed view of the model's classification performance across the four target categories: **Glioma**, **Meningioma**, **Normal** and **Pituitary**. From this matrix, essential evaluation metrics such as precision, recall, F1-score, and overall accuracy have been computed and summarised in Table 5.4.

The results obtained from the confusion matrix reveal the model's high level of accuracy in identifying different brain tumour types. Correct classifications dominate across all categories, with only a handful of minor misclassifications occurring. This pattern strongly suggests that the model developed a nuanced understanding of the underlying visual differences between tumour types.

Pituitary and Glioma classes, in particular, were handled with a high degree of confidence. The model predicted these classes with remarkable precision, indicating minimal to no false positives. Their recall scores also show that very few relevant instances were overlooked, demonstrating reliability in detecting these tumour types.

Meningioma predictions were similarly robust. While a few samples were confused with neighbouring classes, the model maintained strong precision and recall, suggesting that it can effectively distinguish meningioma from other tumour types despite possible overlaps in imaging characteristics.

The Normal category stood out with flawless recognition. which means healthy patients are not mistakenly classified as having a tumour.

Looking at the broader metrics, both macro and weighted averages indicate highly consistent performance across classes. The close alignment of these averages implies that the model handled class imbalance gracefully and treated each tumour type with comparable effectiveness.

In summary, the model exhibited reliable and well-balanced performance throughout. Its ability to accurately classify brain MRI images across all categories, including both pathological and healthy cases, highlights its strong potential in the classification task.

Result Discussion

Overall, the DenseNet121 model exhibited excellent performance in the multi-class classification of brain MRI images. The confusion matrix and classification report demonstrated that the model was highly effective in distinguishing between tumour types: glioma, meningioma, pituitary, as well as normal (no tumour) cases, with consistently high classification accuracy across all classes.

The training pipeline, which incorporated data augmentation, regularisation techniques, and training optimisations such as early stopping and learning rate scheduling, contributed to the model's ability to learn generalisable features and avoid overfitting. While a very limited number of misclassifications were observed, mainly between tumour types with overlapping characteristics, the model maintained strong reliability and robustness. Its consistent performance across all classes further underscores its potential for supporting clinical decision-making in brain tumour diagnosis.

5.5 Conclusion

In this chapter, we implemented and evaluated a CNN-based classification system using the DenseNet-121 architecture on two medical imaging datasets: chest X-rays and brain MRIs. The model was trained and validated on both datasets following appropriate preprocessing and data augmentation, and its performance was assessed using standard metrics, including accuracy, precision, recall, and F1-score.

For the chest X-ray classification task (COVID-19, pneumonia, and normal), the model achieved an overall accuracy of 99.24%.

For the brain MRI classification task (glioma, meningioma, pituitary tumour, and normal), it achieved an overall accuracy of 99.72%.

These results demonstrate that the DenseNet-121 model, when properly configured and trained, can achieve excellent classification performance in medical imaging applications. The model showed high sensitivity and specificity across all classes, confirming its effectiveness in distinguishing between complex medical conditions.

General Conclusion

Through the development of this project, substantial knowledge and hands-on experience were gained in applying deep learning techniques to real-world medical imaging challenges. By building and evaluating classification systems for both chest X-ray and brain MRI images using the DenseNet121 architecture, we explored the complete lifecycle of a deep learning workflow. This included data preparation and preprocessing, model training, evaluation, and interpretation of results.

The model proved effective in both tasks. It successfully distinguished between COVID-19, pneumonia, and normal cases in chest radiographs, as well as accurately identified glioma, meningioma, pituitary tumours, and non-tumorous cases in brain MRI scans. Visual tools such as training and validation curves, along with confusion matrices, provided valuable insight into the model's learning dynamics and classification capabilities, highlighting both its strengths and areas where improvement is possible. Key challenges, including overfitting and class imbalance, were encountered and mitigated through strategies like data augmentation.

This project not only reinforced core theoretical concepts but also underscored the importance of proper preprocessing, careful hyperparameter tuning, and systematic evaluation for achieving robust performance in medical image analysis. Looking ahead, potential improvements include experimenting with more relevant and diverse datasets, exploring alternative CNN architectures, and fine-tuning pre-trained models for better generalisation. Ultimately, integrating the trained model into a user-friendly diagnostic application could represent a valuable step toward real-world deployment and clinical utility.

Bibliography

[1] Russell, S., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson Education.

[2] Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

[3] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

[4] Buchanan, B. G., & Shortliffe, E. H. (1984). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley.

[5] LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning. Nature*, 521(7553), 436–444.

[6] Zhang, Y., Liu, Y., & Wang, S. (2013). *Adaptive Traffic Signal Control with Deep Reinforcement Learning. Proceedings of the 22nd International Joint Conference on Artificial Intelligence.*

[7] Jurafsky, D., & Martin, J. H. (2020). *Speech and Language Processing* (3rd ed.). Prentice Hall.

[8] Haenlein, M., & Kaplan, A. (2019). *A brief history of artificial intelligence: On the past, present, and future of AI.* California Management Review, 61(4), 5-14.

[9] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van Ginneken, B., & Sánchez, C. I. (2017). *A survey on deep learning in medical image analysis.* Medical Image Analysis, 42, 60-88.

[10] Hassan, M. R., & Hussain, S. (2018). *A comprehensive survey on applications of artificial intelligence in financial markets.* Springer Nature.

[11] Goodall, N. J. (2014). *Machine learning for autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems*, 15(6), 2407-2415.

[12] Zhang, X., Yu, J., & Tan, Z. (2020). *A survey of AI applications in manufacturing.* Journal of Manufacturing Science and Engineering, 142(8), 081003.

[13] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). *A guide to deep learning in healthcare.* Nature Medicine, 25(1), 24-29.

[14] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI*. Information Fusion, 58, 82-115.

[15] Recht, B., Roelofs, R., Schmidt, L., & Shankar, V. (2019). *Do ImageNet classifiers generalize to ImageNet?* International Conference on Machine Learning (ICML).

[16] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). *A survey on bias and fairness in machine learning*. ACM Computing Surveys, 54(6), 1-35.

[17] Strubell, E., Ganesh, A., & McCallum, A. (2019). *Energy and Policy Considerations for Deep Learning in NLP*. ACL 2019.

[18] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). *Explaining and harnessing adversarial examples*. International Conference on Learning Representations (ICLR).

[19] Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.

[20] Xu, R., & Wunsch, D. (2015). *Survey of clustering algorithms*. IEEE Transactions on Neural Networks and Learning Systems, 16(3), 645-678.

[21] Chapelle, O., Schölkopf, B., & Zien, A. (2006). *Semi- Supervised Learning*. MIT Press.

[22] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.

[23] Dosovitskiy, A., & Brox, T. (2015). *Discriminative Unsupervised Feature Learning with Exemplar Convolutional Neural Networks*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(9), 1734-1747.

[24] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). *Attention is All You Need*. Advances in Neural Information Processing Systems.

[25] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). *Gradient-based learning applied to document recognition*. Proceedings of the IEEE, 86(11), 2278–2324.

[26] Nair, V., & Hinton, G. E. (2010). *Rectified Linear Units improve Restricted Boltzmann Machines* . In Proceedings of the 27th International Conference on Machine Learning (ICML-10)(pp. 807–814).

[27] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). *OverFeat: Integrated recognition, localization and detection using convolutional networks*. Proceedings of the International Conference on Learning Representations (ICLR).

[28] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.

[29] Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual.* Scotts Valley, CA: CreateSpace.

[30] Lutz, M. (2013). *Learning Python* (5th ed.). O'Reilly Media, Inc.

[31] Millman, K. J., & Aivazis, M. (2011). *Python for scientists and engineers.* Computing in Science & Engineering, 13(2), 9–12.

[32] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). *TensorFlow: A system for large-scale machine learning.* In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), 265-283.

[33] van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). *The NumPy array: A structure for efficient numerical computation.* Computing in Science & Engineering, 13(2), 22-30.

[34] McKinney, W. (2010). *Data structures for statistical computing in Python.* In Proceedings of the 9th Python in Science Conference, 51-56.

[35] Hunter, J. D. (2007). *Matplotlib: A 2D graphics environment.* Computing in Science & Engineering, 9(3), 90-95.

[36] Waskom, M., et al. (2020). *Seaborn: Statistical data visualisation.* Journal of Open Source Software, 5(49), 3021.

[37] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). *Scikit-learn: Machine learning in Python.* Journal of Machine Learning Research, 12, 2825-2830.

[38] Zitouni, A. (2020). *Image Processing Methodology and Textures Analysis for their Segmentation.* Doctorate thesis, University Amar Telidji of Laghouat, Department of Electronics.

Webographie

[39] Fukushima, K. (2007). *Neocognitron.* Scholarpedia, 2(1), 1717. http://www.scholarpedia.org/article/Neocognitron.

[40] Chollet, F., et al. (2015). *Keras*. GitHub repository. Retrieved from: https://github.com/fchollet/keras

[41] Google Colaboratory. https://colab.research.google.com

[42] TensorFlow. https://www.tensorflow.org

[43] Keras API. https://keras.io

[44] Scikit-learn. https://scikit-learn.org

[45] Sachin Kumar. (2021). *COVID19+PNEUMONIA+NORMAL Chest X-Ray Image Dataset* [Data set]. Kaggle. https://doi.org/10.34740/kaggle/dsv/1857760

[46] Tom Backert. (2024). *Brain Tumor MRI Dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/tombackert/brain-tumor-mri-data

List of Figure References

[47] HealthImaging. *Examples of FDA-cleared AI in radiology.* https://healthimaging.com/topics/artificial-intelligence/photo-gallery-examples-fda-cleared-ai-radiology

[48] Waymo. *Self-Driving Car Technology for a Reliable Ride* https://waymo.com/

[49] HiCart Corporation. *AI-Powered Shopping Carts.* https://hicartcorp.com/products/ai-powered-carts/

[50] UAV agriculture drone sprayer. DroneAssemble – FDAD-10L4X-R Agricultural Drone

[51] Opinosis Analytics Blog. *AI in Manufacturing: Examples.* https://www.opinosis-analytics.com/blog/ai-in-manufacturing-examples/

[52] DenseNet-121 architecture. ResearchGate. *A schematic illustration of the DenseNet-121 architecture*