



DOCTORAL THESIS

3rd Cycle Doctoral (D-LMD)

Presented by

Oussama ABDA

With a view to obtaining the doctoral diploma in 3rd Cycle Doctoral (D-LMD)

Branch: Electronics

Specialty: multidimensional analysis and machine learning for imaging

Topic

Enhancing Medical Signal Using Wavelet Transform and Artificial Intelligence

Supported, on 03 /10 / 2026, before the jury composed of:

Last and first name	Grade	Institution of affiliation	Designation
Pr. Amel Baha Houda ADAMOUMITICHE	professor	University of Djelfa	President
Pr. Hilal NAIMI	Professor	University of Djelfa	Supervisor
Dr. Mohammed DJEMAI	MCA	University of Djelfa	Examiner
Dr. Kamal SEHAIRI	MCA	Higher Teacher Training School of Laghouat	Examiner

Djelfa University, FST - 2026

DEDICATIONS

I dedicate this thesis to my beloved parents, in recognition of their sacrifices, their affection, their patience, and their unwavering support throughout my studies.

To my dear brother, Youssouf, for his continuous encouragement, understanding, and support at every step.

To all my family, for their constant support, care, and encouragement throughout my studies.

To my dear friends, for their motivation, their kindness, and their constant presence during this journey.

ACKNOWLEDGEMENTS

First and foremost, I am grateful to Allah for providing me with the strength, knowledge, and determination to accomplish this thesis.

I have received guidance and support from many individuals throughout this journey, and I wish to sincerely thank everyone who contributed, directly or indirectly, to the completion of this work.

I am especially grateful to my supervisor, **Professor NAIMI Hilal**, for her patient guidance, valuable advice, and continuous support, which have been essential to the success of this research.

I also sincerely thank **Professor MITICHE Lahcene**, head of the Laboratory, and all members of the Laboratory of Modeling, Simulation, and Optimization of Complex Systems, for their encouragement.

I would also like to express my sincere appreciation to the members of the jury, **Prof. Amel Baha Houda ADAMOU-MITICHE**, **Dr. Mohamed DJEMAI**, and **Dr. Kamel SEHAIRI**, for accepting to evaluate this work and for their valuable time, insightful comments, and constructive feedback.

My appreciation goes to all the students of the 2022/2023 promotion for their companionship, motivation, and the positive academic environment we shared.

I extend my thanks to the University of Zian Achour, Djelfa, for providing the necessary resources and support to carry out this research.

Finally, I am deeply thankful to my parents, my brother Youssouf, and my dear friends for their constant support, trust, and encouragement throughout this journey.

الملخص

تم استخدام التحويل الموجي والدكاء الاصطناعي في مجموعة واسعة من التطبيقات. في مجال الإشارات الطبية، تم توظيفهما في مهام متعددة مثل استخراج الخصائص، والتصنيف، والتقسيم. في هذا العمل، نقترح نموذجين لتصنيف أورام الدماغ انطلاقاً من صور الرنين المغناطيسي.

يعتمد النموذج الأول على مصنف SVM متعدد لتصنيف صور الرنين المغناطيسي إلى أربع فئات، وهي الورم الدبقي، والورم السحائي، والورم النخامي، وعدم وجود ورم. يجمع هذا النموذج بين التحويل الموجي غير المضمحل لاستخراج المعلومات الترددية والزمنية، يليه استخراج الخصائص باستخدام النموذج المدرب مسبقاً EfficientNet-B0. يتم تطبيق تحليل المكونات الرئيسية (PCA) لاختيار الخصائص، ويُدرَّب مصنف SVM متعدد بناءً على هذه الخصائص، مستفيداً من المعلومات الترددية والمكانية لصور الرنين المغناطيسي، مما يسمح بتحقيق دقة تصنيف عالية.

أما النموذج الثاني فيستخدم دمج تحويل الموجة الساكنة (SWT) للمعالجة المسبقة، و ResNet50V2 لاستخراج الخصائص، وشبكة LSTM للتصنيف. تبدأ مرحلة المعالجة المسبقة بتحويل SWT، حيث يتم تحليل الصور إلى عدة نطاقات ترددية لتعزيز الخصائص المرتبطة بالورم وتقليل الضجيج. ثم يتم تغيير حجم الصور إلى 224×224 بكسل لتتوافق مع نموذج ResNet50V2. يُستخدم ResNet50V2 لاستخراج الخصائص، والتي تُمرَّر بعد ذلك إلى شبكة LSTM لتصنيفها إلى أربع فئات، وهي الورم الدبقي، والورم السحائي، والورم النخامي، وعدم وجود ورم. وقد أثبتت الطريقة المقترحة فعاليتها في تحقيق دقة تصنيف عالية.

الكلمات المفتاحية: التحويل الموجي، الدكاء الاصطناعي، صور الرنين المغناطيسي، التصنيف.

Abstract

Wavelet transform and artificial intelligence have been applied to a wide range of problems. In medical signals, they have been used in many tasks such as feature extraction, classification and segmentation. In this work, we propose two models for brain tumor classification from MR images. The first model uses a multi SVM classifier to classify MR images into 4 classes (glioma, meningioma, pituitary, no tumor), the model combines the undecimated wavelet transform to extract both frequency and temporal information, followed by feature extraction using a pretrained EfficientNet-B0 model. Principal Component Analysis (PCA) is applied for features selection, The multi-SVM classifier is trained based on these features, making use of the frequency and spatial information of the MR images, allowing for high accuracy. The second model uses the integration of Stationary Wavelet Transform (SWT) for preprocessing, ResNet50V2 for feature extraction, and Long Short-Term Memory (LSTM) for classification, The preprocessing phase started with Stationary Wavelet Transform (SWT), which decomposed the images into multiple frequency bands to enhance tumor-related features while reducing noise. The images were then resized to 224x224 pixels to ensure compatibility with the ResNet50V2 model. Feature extraction was performed using ResNet50V2, the extracted features were then passed to a Long Short-Term Memory (LSTM) network for classification into four categories: Glioma, Meningioma, Pituitary, and No Tumor. this proposed method be effective in producing a high classification accuracy rate. The proposed models are evaluated using performance metrics including accuracy F1 score and recall and precision, the dataset from Kaggle, the results obtained are compared with some modern methods.

Keywords Wavelet transform, Artificial Intelligence, MR images, classification

Résumé

La transformée en ondelettes et l'intelligence artificielle ont été appliquées à un large éventail de problèmes. Dans le domaine des signaux médicaux, elles ont été utilisées pour de nombreuses tâches telles que l'extraction de caractéristiques, la classification et la segmentation. Dans ce travail, nous proposons deux modèles pour la classification des tumeurs cérébrales à partir d'images IRM.

Le premier modèle utilise un classifieur SVM multiclasse pour classer les images IRM en 4 classes (gliome, méningiome, tumeur de l'hypophyse, absence de tumeur). Ce modèle combine la transformée en ondelettes non décimée, afin d'extraire à la fois l'information fréquentielle et temporelle, suivie d'une extraction de caractéristiques à l'aide d'un modèle EfficientNet-B0 pré-entraîné. L'analyse en composantes principales (ACP) est appliquée pour la sélection des caractéristiques. Le classifieur SVM multiclasse est ensuite entraîné sur ces caractéristiques, tirant parti de l'information fréquentielle et spatiale des images IRM, ce qui permet d'obtenir une grande précision.

Le second modèle repose sur l'intégration de la transformée en ondelettes stationnaire (SWT) pour le prétraitement, de ResNet50V2 pour l'extraction de caractéristiques et d'un réseau à mémoire à long et court terme (LSTM) pour la classification. La phase de prétraitement commence par la transformée en ondelettes stationnaire (SWT), qui décompose les images en plusieurs bandes de fréquences afin de renforcer les caractéristiques liées à la tumeur tout en réduisant le bruit. Les images sont ensuite redimensionnées à 224×224 pixels afin d'assurer la compatibilité avec le modèle ResNet50V2. L'extraction de caractéristiques est réalisée avec ResNet50V2, puis les caractéristiques extraites sont transmises à un réseau LSTM pour la classification en quatre catégories : gliome, méningiome, tumeur de l'hypophyse et absence de tumeur. La méthode proposée s'avère efficace pour produire un taux de précision de classification élevé.

Les modèles proposés sont évalués à l'aide de métriques de performance, notamment l'exactitude (accuracy), le score F1, le rappel et la précision, sur un jeu de données provenant de Kaggle. Les résultats obtenus sont comparés à ceux de certaines méthodes modernes.

Mots-clés : transformée en ondelettes, intelligence artificielle, images IRM, classification

Contents

Abstract	i
List of Figures	vi
List of Tables	vii
Abbreviations	viii
Introduction	1
1 Medical Signals	4
1.1 Introduction	5
1.2 Definition of Medical Signal	5
1.3 Types of Medical Signals	5
1.3.1 Electrophysiological Signals	5
1.3.2 Medical Imaging Signals	8
1.3.3 Biomechanical Signals	11
1.4 Preprocessing and Filtering of Medical Signals	12
1.4.1 Noise Reduction	12
1.4.2 Standardization	12
1.4.3 Feature Extraction	12
1.5 Importance of Medical Signal Processing	13
1.6 Focus on MRI in Brain Tumor Detection	13
1.7 Overcoming MRI Classification Challenges	14
1.8 Conclusion	17
2 Wavelet Transforms	18
2.1 Introduction	19
2.2 Wavelet Theory	19
2.2.1 Basic Principles and Properties	19
2.3 Mother Wavelets	20
2.4 Wavelet Families and Properties	20
2.4.1 Orthogonal Wavelets	20
2.4.2 Biorthogonal Wavelets	22
2.5 Continuous Wavelet Transforms	22
2.6 Discrete Wavelet Transforms	24
2.7 Undecimated Wavelet Transforms (UWT)	25
2.8 Complex Wavelets	26
2.8.1 DT-CWT	26
2.9 Lifting wavelet transform	26

2.9.1	Split	28
2.9.2	Predict	28
2.9.3	Update	28
2.10	Applications to Medical Signals	28
2.10.1	Signal Denoising	29
2.10.2	Feature Extraction	29
2.10.3	Signal Compression and Coding	31
2.10.4	Other Applications	31
2.11	Conclusion	31
3	Artificial Intelligence: An Overview	32
3.1	Introduction	33
3.2	Artificial intelligence, machine learning, deep learning	33
3.3	Machine Learning vs Deep Learning	34
3.4	Types of ML	34
3.4.1	Supervised Learning	35
3.4.2	Unsupervised Learning	36
3.4.3	Reinforcement Learning	36
3.5	Commonly Used Machine Learning Algorithms	37
3.5.1	SVM	37
3.5.2	KNN	37
3.5.3	Decision Tree	38
3.5.4	K-means	38
3.5.5	Random Forest	39
3.5.6	Naive Bayes	39
3.5.7	Logistic Regression	40
3.5.8	Hierarchical Clustering	40
3.5.9	GMM	41
3.6	Principles of Deep Learning	41
3.6.1	ANNs	41
3.6.2	Artificial Neuron	41
3.6.3	ANN Structure	43
3.7	Training of ANN	45
3.7.1	Gradient Descent and Backpropagation	46
3.8	CNNs	46
3.8.1	Convolutional Layers	47
3.8.2	Pooling Layers	48
3.8.3	Fully Connected Layer	48
3.9	Popular CNN Architectures	49
3.9.1	VGG	49
3.9.2	EfficientNet	50
3.9.3	U-Net	51
3.9.4	ResNet	52
3.10	RNN and LSTM	53
3.11	Applications of AI	54
3.12	AI in Medical Signals	56
3.13	Challenges in AI	57
3.14	Conclusion	57

4	Application of Proposed Methods to Brain Tumor MRI Classification	58
4.1	Introduction	59
4.2	CAD Methods for Brain Disorder Detection	59
4.3	Wavelet and AI for Brain Tumor classification	60
4.4	First Proposed Model	60
4.4.1	Main Idea and Motivation	60
4.4.2	Principles and Objectives	61
4.4.3	Architecture	62
4.4.4	Database Description	62
4.4.5	Experiment Setup	63
4.4.6	Performance Measures	63
4.4.7	Results	65
4.4.8	Confusion Matrix	66
4.5	Second Proposed Model	68
4.5.1	Key concept and inspiration	68
4.5.2	Process	68
4.5.3	Results	70
4.6	Comparison and Analysis	72
4.7	Conclusion	74
	General Conclusion	76
	Bibliography	77
	Author's Publications	93

List of Figures

1.1	Example of Normal ECG signal	6
1.2	Example of a normal EEG	7
1.3	Example of sample EMG	7
1.4	Example of ultrasound imaging	8
1.5	An example of chest X-ray and computed tomography (CT) scan images	9
1.6	Brain magnetic resonance imaging (MRI)	10
1.7	X-ray image of a human brain tumor	10
1.8	Detection of brain tumor from MRI images	14
1.9	Automated Workflow for Brain Tumor Detection and Diagnosis	16
2.1	Morlet wavelet	23
2.2	A Discrete Wavelet Transform (DWT) model of 3 levels decomposition	24
2.3	The stationary wavelet transforms using multi-level decomposition	25
2.4	- (a) 1D complex wavelet, (b) filter bank analysis for 1-D DTCWT, (c) synthesis filter bank for 1-D DTCWT	27
2.5	Forward lifting wavelet transform	28
3.1	Hierarchical Relationship Between Artificial Intelligence, Machine Learning, and Deep Learning	34
3.2	Difference Between Machine Learning and Deep Learning	35
3.3	Overview of Machine Learning Types	35
3.4	Reinforcement learning	36
3.5	Classification using support vector machines	37
3.6	The Biological Neuron	42
3.7	The formal neuron	42
3.8	ANN structure	44
3.9	Sigmoid activation function	44
3.10	Tanh activation function	45
3.11	CNN architecture	47
3.12	An illustrated representation of the convolution operation	47
3.13	Visual representation of the max-pooling operation	49
3.14	Vgg16 architecture	50
3.15	Architecture of EfficientNetB0	51
3.16	U-Net architecture	52
3.17	ResNet50 model architecture	53
3.18	Working principle of LSTM architecture	54
3.19	Applications of Artificial Intelligence	56
4.1	Workflow of the first proposed system	63
4.2	The Basic Structure of a Confusion Matrix	65
4.3	Performance metrics of the first proposed model	66

4.4	Confusion matrix for first proposed model	67
4.5	Normalized confusion matrix for first proposed model	67
4.6	Block diagram for second proposed method	69
4.7	Performance metrics for second proposed model	70
4.8	Confusion matrix for second proposed method	71

List of Tables

1.1	Overview of AI Techniques in MRI Images Processing	15
2.1	Comparison of Advanced Denoising Techniques for Medical Images Using Wavelet Transforms	30
3.1	Hebb's Rule	43
3.2	AI Techniques and Applications in Medical Signal Processing and Diagnosis	56
4.1	Dataset Distribution by Tumor Class	63
4.2	Performance metrics for the first proposed method	66
4.3	Dataset Composition Used in This Study	69
4.4	Comparison of Brain Tumor Classification Methods	72

Abbreviations

AI Artificial Intelligence

MRI Magnetic Resonance Imaging

SVM Support Vector Machine

ANN Artificial Neural Network

BWT Biorthogonal Wavelet Transform

CAD Computer-Aided Diagnosis

CT Computed Tomography

DL Deep Learning

DWT Discrete Wavelet Transform

ECG Electrocardiogram

EEG Electroencephalogram

EMG Electromyogram

GAN Generative Adversarial Network

GMM Gaussian Mixture Model

KNN K-Nearest Neighbors

LSTM Long Short-Term Memory

ML Machine Learning

MRI Magnetic Resonance Imaging

PCA Principal Component Analysis

PET Positron Emission Tomography

RNN Recurrent Neural Network

ReLU Rectified Linear Unit

SVM Support Vector Machine

SWT Stationary Wavelet Transform

UWT Undecimated Wavelet Transform

ViT Vision Transformer

Introduction

Medical signals are data representing the biological and physiological activity of the human body [1]. They are recorded for diagnosis and monitoring purposes. These signals are essential for non-invasive medical analysis and play a central role in healthcare and clinical decision-making [2]. These signals include electrocardiograms (ECGs), which record the electrical activity of the heart [3]; electroencephalograms (EEGs) which measure brain activity [4]; and electromyograms (EMGs) which capture electrical signals generated by skeletal muscles[5]. Other important examples include respiratory signals, which are used to assess breathing rate and volume [6], and blood pressure signals which monitor the force of blood flow against vessel walls [7]. Furthermore, medical imaging signals, such as those generated by magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound, provide visual information revealing the body's internal structures. These signals aid in precise diagnosis and therapy planning [8]. In the field of medical signals one problem is the classification of MRI scans to detect brain tumors, Early detection is crucial for treatment planning [9]. Due to the large amount of data, manual diagnosis by specialists is time-consuming and prone to human error[10]. Researchers in this field have used various artificial intelligence (AI) approaches, including machine learning (ML) and deep learning (DL), to develop numerous computer-aided diagnostic (CAD) systems [11]. These methods face many challenges, such as the availability of data and feature extraction from MRI images [12]. Traditional ML methods rely on handcrafted features, which may not capture all relevant information and require domain expertise [13]. Although DL techniques automate feature extraction, they often require substantial computer resources and large amounts of data [14]. In this thesis, we address these challenges by combining wavelet transforms, which are a powerful mathematical tool for analyzing signals [15] with different AI techniques, including ML and DL. This thesis focuses on finding an effective way to combine wavelet transforms and artificial intelligence techniques to improve the analysis of medical signals, with a particular focus on classifying MRI images for brain tumor detection. The aim is to improve feature extraction and overcome the limitations of traditional methods by combining the undecimated wavelet transform (UWT) with deep learning models such as EfficientNet-B0. The study examines the advantages of using wavelet transforms to extract and classify features in medical data, demonstrating their ability to capture frequency and spatial information. The efficiency of the proposed strategy is evaluated using a medical image dataset and is then compared with existing methods. Finally, we identify potential areas for development and explore the pros and cons of the proposed method in medical signal processing. The thesis aims to improve the accuracy and efficiency of medical signal processing by incorporating wavelet analysis into deep learning systems. The study will evaluate the proposed method using medical image datasets and compare its results with those of the most advanced methods available today. It also assesses the method's pros and cons and suggests potential improvements and optimizations.

In this study, we present two models for classifying brain tumors using MRI scans. The first is based on a multi-SVM classifier, which leverages the undecimated wavelet transform (UWT) to extract both frequency and temporal information from MRI images. This is followed by feature extraction using a pretrained EfficientNet-B0 model, and Principal Component Analysis (PCA) is applied for feature selection. The multi-SVM classifier utilizes these selected features to classify the MRI images into four categories: glioma, meningioma, pituitary tumor, and no tumor. The second model combines Stationary Wavelet Transform (SWT) for preprocessing, ResNet50V2 for feature extraction, and Long Short-Term Memory (LSTM) for classification. SWT decomposes images into mul-

multiple frequency bands to enhance tumor-related features and reduce noise. The images are then resized to 224×224 pixels for compatibility with ResNet50V2, which extracts relevant features. These features are fed into an LSTM network to classify MRI images into four categories: Glioma, Meningioma, Pituitary, and No Tumor.

This thesis consists of four chapters organized as follows:

1. The first chapter is devoted to the general concepts of medical signals. It presents the different types of medical signals, including electrophysiological, imaging, and biomechanical signals. This chapter also covers key preprocessing techniques such as noise reduction, standardization, and feature extraction. Special attention is given to MRI signals in the context of brain tumor detection, with a focus on the use of artificial intelligence methods for image classification and diagnosis.

2. The second chapter introduces wavelet transforms. It explains wavelet theory, key properties, and types of wavelets such as orthogonal, biorthogonal, and complex. The chapter describes continuous, discrete, undecimated and lifting wavelet transforms. It also covers practical applications including denoising, feature extraction, compression, segmentation, and classification of medical signals like ECG, EEG, and MRI.

3. The third chapter is devoted to an overview of artificial intelligence. It presents the main concepts and techniques of AI, including machine learning, deep learning, natural language processing, and computer vision. The chapter traces the development of AI and highlights its wide applications across various sectors. The second part focuses on AI's role in medical signal processing, covering its use in disease diagnosis, signal analysis, patient monitoring, and prediction.

4. The fourth chapter focuses on Application of Proposed Methods to Brain Tumor MRI Classification. It introduces two models that combine wavelet transforms with deep learning. The first model uses UWT with EfficientNet-B0, PCA, and multi-SVM for classifying MRI images into four tumor types. The second model applies SWT for denoising, ResNet50V2 for feature extraction, and LSTM for classification. Both models aim to improve accuracy by combining wavelet transform with machine learning and deep learning techniques.

Finally, a conclusion concludes this thesis.

Chapter 1

Medical Signals

1.1 Introduction

Medical signals are essential to contemporary healthcare because they make it possible to analyze and interpret a variety of physiological data for diagnosis, monitoring, and treatment [16]. These medical signals include electrocardiograms (ECGs), electroencephalograms (EEGs), electromyograms (EMGs), and magnetic resonance imaging (MRI). It is important to extract insights and rational frameworks that can help healthcare professionals make informed decisions [17].

Advances in technology allow us to capture and interpret medical signals, which has allowed us to develop sophisticated diagnostic tools and systems. The main objective is to increase the capabilities of healthcare professionals by automating repetitive operations and supplying more precise data for clinical decision-making, rather than replacing them. The practical applications of advanced imaging and signaling techniques have increased, and further advances in theory and technology are required [18].

The challenge in using medical signals is to capture and accurately interpret complex physiological data to yield meaningful insights [19]. The primary goal is to provide a comprehensive overview of medical parameters, divided into electrophysiological parameters, medical imaging parameters, and biomechanical parameters. and highlights important preprocessing and filtering procedures necessary for successful analysis, ensuring the reliability and usefulness of the data [20].

1.2 Definition of Medical Signal

Several researchers have defined medical signals and emphasized their importance in healthcare. For example, Akay in 1994 defined a medical signal as: "a representation of a physiological parameter, captured over time, to monitor and diagnose the health status of an individual". Similarly, Enderle et al. in 2000 described medical signals as: "biological signals that are collected and interpreted using various techniques to understand and evaluate the functioning of different body systems" [21].

These definitions underscore the objectives and challenges associated with medical signals. Transmission of medical data that represent diverse physiological processes involving noise and artifacts is contained in medical signals. The analysis of these signals aims to provide information on the status and function of the body by assisting in the diagnosis, monitoring, and treatment of certain medical disorders.

1.3 Types of Medical Signals

Medical signals play a key role in diagnosing and managing physical conditions. These signals can be broadly classified as electrophysiological, medical imaging or biomechanical. Each category comprises specific indicators that offer detailed insights into different aspects of human health [22].

1.3.1 Electrophysiological Signals

Electrophysiological signals are a subset of medical signals that measure the electrical activity generated by physiological processes [23]. These signals are essential for under-

standing the functioning of the heart, brain, muscles, and other organs. Common types of electrophysiological signals include:

Electrocardiogram (ECG)

An electrocardiogram (ECG) records the electrical activity of the heart over time by applying electrodes to the skin. This essential cardiology instrument is used to identify heart problems such as myocardial infarction and arrhythmias. The ECG signal provides information on the heart's rhythm, rate and electrical conduction pathways . Figure 1.1. presents an example of a normal ECG signal[24].

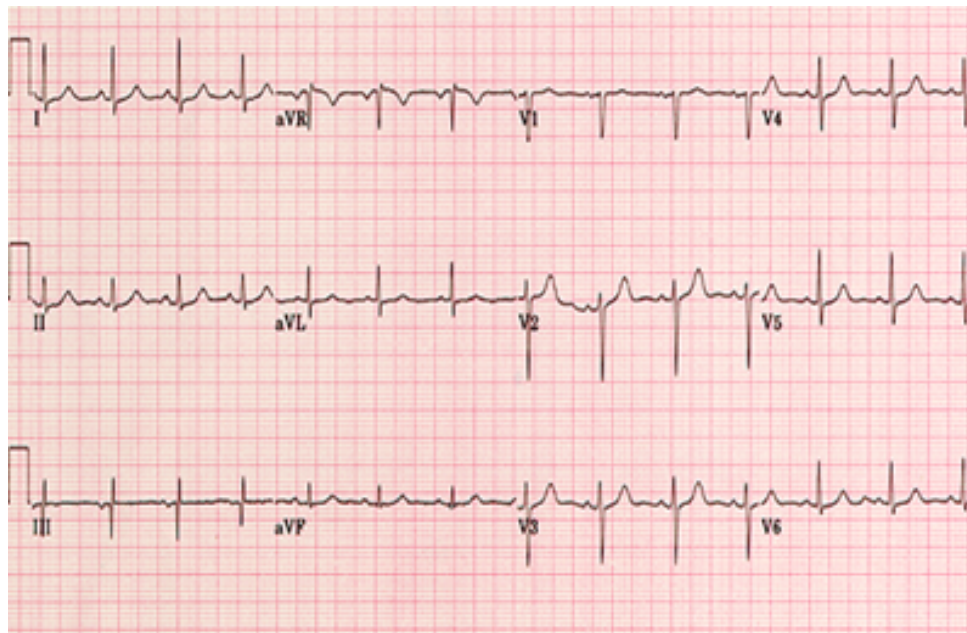


Figure 1.1: Example of Normal ECG signal [24]

Electroencephalogram (EEG)

Electroencephalography (EEG) is one method for measuring the electrical activity of the brain. EEG is often used to identify and track neurological problems, such as brain damage, sleep disorders, and epilepsy. Electrodes inserted into the scalp measure spontaneous electrical activity and receive EEG data. Figure 1.2 shows an example of a normal EEG [25].

Electromyogram (EMG)

An electromyogram (EMG) measures the electrical activity produced by skeletal muscle tissue. It is used to diagnose neuromuscular disorders, assess muscle function and examine muscle activation patterns. EMG indicators are collected using surface or intramuscular electrodes that detect the electrical potentials generated during muscle contraction [27]. Figure 1.3 shows a typical EMG signal.

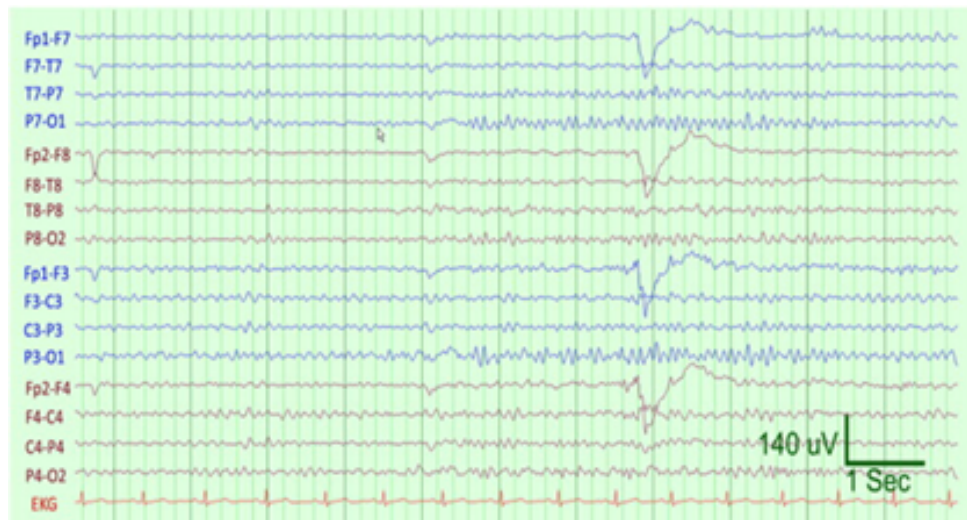


Figure 1.2: Example of a normal EEG [26]

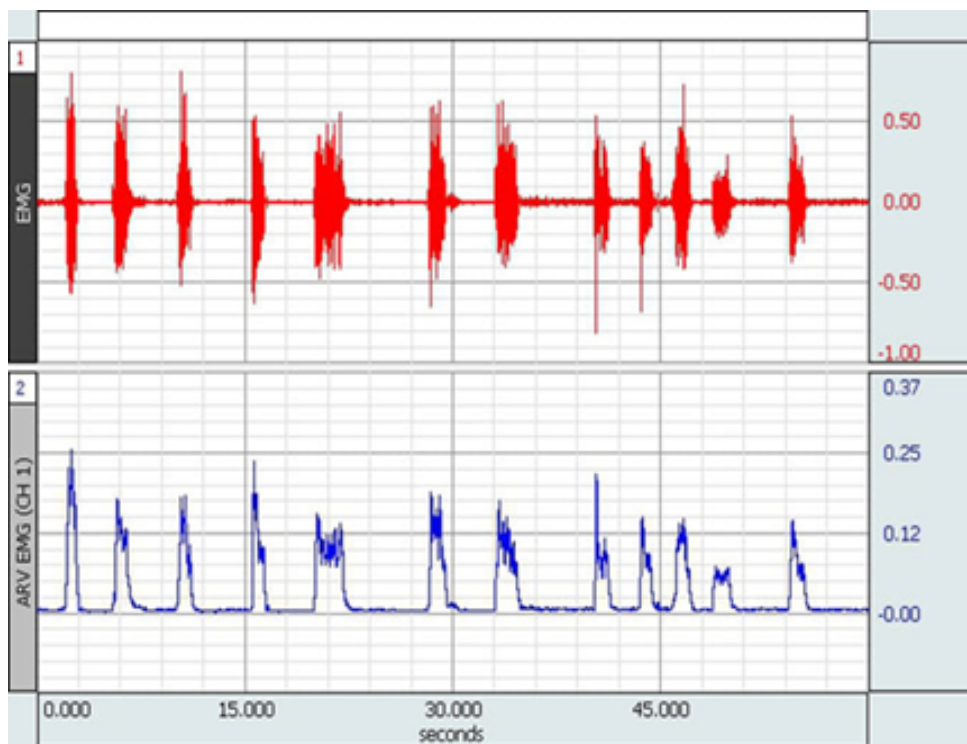


Figure 1.3: Example of sample EMG [28]

Other Electrophysiological Signals

In addition to the ECG, EEG and EMG, other types of electrophysiological signal include the electrooculogram (EOG), which measures eye movements, and the electrogastrogram (EGG), which monitors stomach activity. These signals offer valuable insights into specific physiological functions and have a variety of diagnostic and research applications.

1.3.2 Medical Imaging Signals

Medical imaging indicators are a vital class of medical signals that provide a visual depiction of the inside of the body. These signals enable clinicians to visualize bodily systems and processes, which is essential for diagnosing and monitoring various medical conditions [29]. Key types of medical imaging signals include:

Ultrasound Imaging

ultrasound imaging, also known as sonography, uses high-frequency sound waves to create images of internal body structures. This technique is widely used in obstetrics to monitor fetal development, as well as in cardiology, urology and musculoskeletal diagnostics. Ultrasound signals are generated and received using a transducer, which converts the echoes into real-time images displayed on a screen [30]. figure 1.4 represent an example of ultrasound imaging.

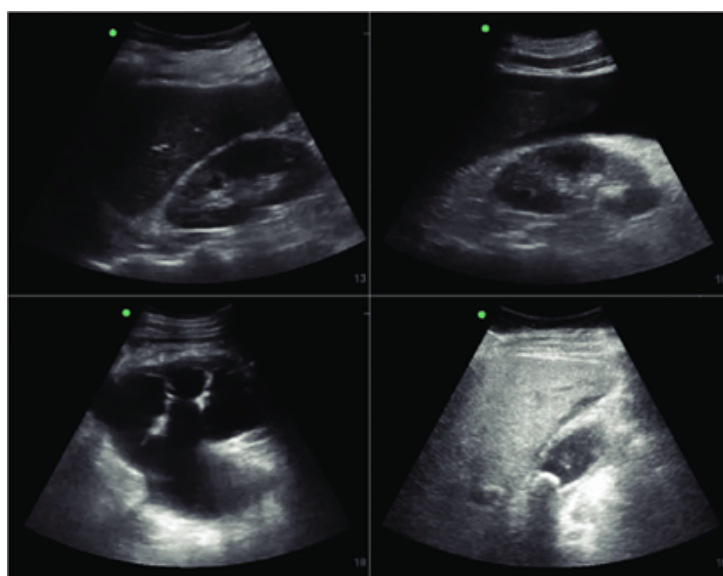


Figure 1.4: Example of ultrasound imaging [31]

Computed Tomography (CT)

Computed tomography (CT) uses X-rays to produce detailed, cross-sectional images of the body. A computer reconstructs these images from several X-ray measurements collected from various angles around the patient to provide a thorough perspective of the inside anatomy. CT scans are very helpful in diagnosing traumatic injuries and diseases, including cancer and heart disease. [32]. Figure 1.5 present An illustration of computed tomography (CT) scan with chest x-ray pictures.

Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) uses the principles of nuclear magnetic resonance (NMR) to provide comprehensive anatomical and functional information without the need for invasive procedures. Unlike X-ray or CT scans, MRI does not involve ionising radiation, making it a safer alternative for medical imaging. MRI works by exploiting the magnetic properties of hydrogen protons present in the body's tissues. When a

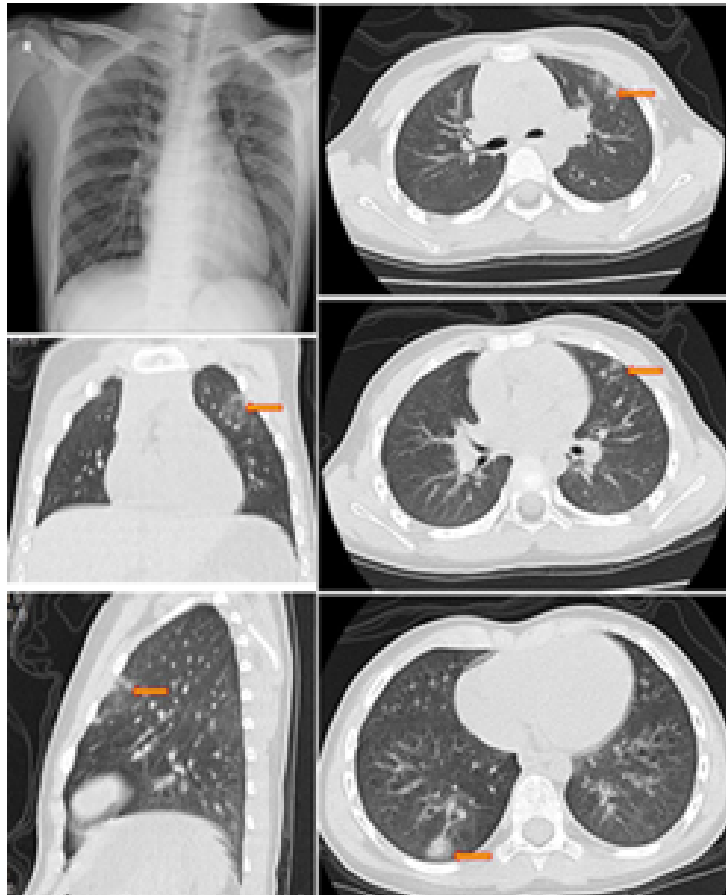


Figure 1.5: An example of chest X-ray and computed tomography (CT) scan images [33]

strong magnetic field is applied, the protons align themselves within the field. Radiofrequency (RF) pulses are then used to disrupt the protons' equilibrium. As the protons realign with the magnetic field, they emit radiofrequency (RF) signals, which can be detected and processed to create images of the body's internal structures. MRI signals are complex and vary depending on the materials and imaging sequences used. The free induction decay (FID) signal represents protons' radiofrequency decay upon recombination with the magnetic field and is the primary signal of interest. Image reconstruction uses the frequency-domain MR signal, produced by applying a Fourier transform to the time-domain FID signal [34]. MRI is widely used in various medical applications, including Neurological imaging MRI is particularly useful for imaging the brain, spinal cord, and other neurological structures, enabling the detection of tumors, hemorrhages, and other abnormalities. Figure 1.6 shows a brain magnetic resonance imaging (MRI) scan. Physicians can use a brain MRI scan to detect diseases such as bleeding, edema, developmental anomalies, tumors, infections, inflammation, injury or stroke damage, and blood vessel issues. Doctors can also use an MRI to determine the source of seizures or headaches.

X-ray

X-rays are widely used to identify a variety of medical issues, such as fractures, infections, arthritis, dental problems, osteoporosis and bone tumours. They are primarily used to examine bones and joints. In some cases, X-rays can also help to detect abnormalities in soft tissue involving internal organs. However, for brain-related assessments, X-rays are

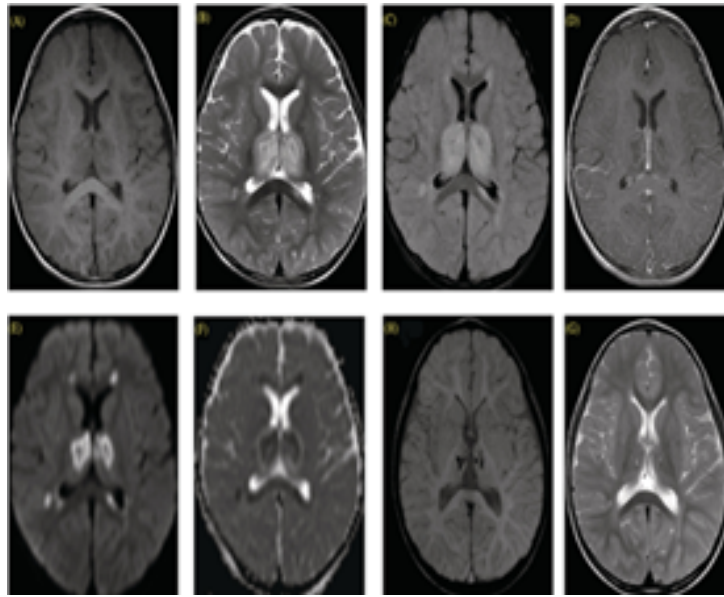


Figure 1.6: Brain magnetic resonance imaging (MRI) [35]

less favoured than CT or MRI scans. This is because X-rays have limited resolution with regard to soft tissues, making them inadequate for clearly visualising brain structures or detecting tumours[36]. Figure 1.7 shows an example of a brain X-ray with a tumour.



Figure 1.7: X-ray image of a human brain tumor [37]

X-rays involve exposure to ionising radiation, raising safety concerns, particularly with frequent use. In contrast, MRI and CT scans provide superior soft tissue contrast and are safer for repeated imaging, especially MRI scans, as they do not involve radiation. Therefore, X-rays are mainly used in urgent or trauma cases where fast general imaging is required. For more precise brain imaging, MRI and CT are usually preferred.

Other Modalities Imaging

One nuclear medicine imaging method that provides insight into bodily functions and metabolism is positron emission tomography (PET). Glucose is the most commonly used radioactive tracer for PET scans. After being injected into the patient's body, this radioactive material decays, releasing positrons that eventually combine with the body's electrons to create gamma rays. PET scanner detects these gamma rays and uses them to create images that display the state of the body's various tissues and organs. PET is especially helpful in oncology because it can identify regions of elevated metabolic activity, which is suggestive of tumor growth. PET aids in the detection of malignant cells. PET is also used in neurology to identify alterations in brain metabolism, such as those underlying Alzheimer's disease, as well as in cardiology to evaluate blood flow to the heart.[38].

1.3.3 Biomechanical Signals

Biomechanical signals are defined as data recorded from various physiological processes and physical interactions within the body. These signals are extremely useful for explaining and understanding the mechanical properties and functional behaviour of biological tissues and systems. This section will provide an overview of the most common types of biomechanical signal, how they are acquired and their importance in medical and research applications[39].

Pressure Signals

Pressure signals are measurements of the force exerted on or acting on biological tissues. In most cases, these signals are detected using pressure sensors or transducers. These signals are important for determining the forces involved in many physiological processes, such as blood flow, tissue compression and joint loading. Understanding pressure signals can help to identify conditions such as hypertension and tissue health, as well as the impact of external forces on the body[40].

Deformation Signals

Deformation signals denote changes in the shape or structure of tissues when subjected to applied forces or internal stresses. They are conventionally measured using imaging techniques or strain gauges, making them very important in the study of tissue mechanics, wound healing and changes caused by disease processes that alter tissue properties. Therefore, analysing deformation signals can help explain how tissues adapt to mechanical stresses and may aid in diagnosing and treating musculoskeletal disorders[41].

Movement Signals

Movement signals correspond to the dynamics of body movements in terms of velocity, acceleration and displacement. These signals can typically be captured by motion sensors, accelerometers, or gyroscopes. These signals are crucial for analysing motor functions, gait patterns, and the impact of physical activity on health. Movement signals provide useful information for rehabilitation processes, ergonomic assessments, and the study of motor control mechanisms[42].

1.4 Preprocessing and Filtering of Medical Signals

Medical signal analysis comprises two vital steps: preprocessing and filtering. These steps are designed to improve signal quality, remove noise, and extract relevant features for better diagnosis, monitoring and treatment in medical applications.

1.4.1 Noise Reduction

Medical signals are often tainted by various types of noise, which can originate from internal physiological processes or external sources such as electromagnetic interference. To improve the signal-to-noise ratio (SNR) and preserve pertinent information, noise cancellation techniques are employed [43].

The techniques of noise reduction from medical signals include :

Digital filtering uses low-pass, high-pass, band-pass, and notch filters to eliminate specific frequency components associated with noise.

Adaptive Filtering Employs algorithms that adjust filter parameters dynamically to remove noise while retaining the original signal characteristics.

Wavelet denoising uses the wavelet transform to split data into various scales, selectively eliminating noise components.

1.4.2 Standardization

Standardisation ensures that medical signals can be compared across different sessions, equipment and patients. This involves normalising the amplitude, adjusting for baseline drift and compensating for variations in signal acquisition conditions[44]. Methods of standardisation include:

Baseline Correction Removes any constant or slowly varying offset in the signal.

Amplitude Normalization Adjusts the signal amplitude to a standard range, facilitating comparison and analysis.

Time Alignment Synchronizes signals recorded at different times or from different devices to a common reference.

1.4.3 Feature Extraction

Feature extraction is the process of identifying and measuring the specific aspects of a medical signal that are relevant for analysis. These characteristics may include detecting abnormalities, classifying signal patterns or monitoring changes over time [3]. Many techniques are used for feature extraction.

Time-Domain Analysis Involves calculating statistical measures such as mean, variance, and skewness of the signal.

Frequency-Domain Analysis Uses Fourier transform to identify dominant frequencies, power spectrum, and harmonic components.

The wavelet transform provides a multi-resolution analysis of the signal, concurrently recording frequency and temporal information. Effective preprocessing and filtering are essential to ensure that medical signals are accurately processed, producing more dependable diagnoses and improved patient outcomes.

1.5 Importance of Medical Signal Processing

Modern healthcare relies heavily on medical signal processing. This technology assists physicians in diagnosis, monitoring, and treatment planning by enabling the extraction of useful information from unprocessed physiological and imaging data.

- It improves the quality and clarity of signals by removing noise and artifacts.
- It enhances the visibility of clinically relevant features, such as tumor boundaries in MRI or abnormal rhythms in ECG.
- It supports early diagnosis through automated detection of patterns or abnormalities.
- It allows continuous monitoring of patients, enabling real-time alerts and timely interventions.
- Examining patient data over time helps with tailored therapy.

In medical imaging, processing techniques like filtering, segmentation, feature extraction, and classification help identify and characterize conditions such as tumors, lesions, or organ deformities. In physiological signals, such as ECG or EEG, processing techniques detect arrhythmias, epileptic seizures, or brain activity patterns. Additionally, medical signal processing enables the incorporation of machine learning and artificial intelligence, enhancing decision-making efficiency and accuracy. These algorithms can identify intricate patterns, learn from large datasets, and produce precise results for therapeutic applications. Without efficient signal processing, medical data could be inaccurate, incomplete, or useless. As the volume and complexity of medical data increase, robust signal processing becomes crucial for turning it into useful insights. Processing medical signals is more than just a technical step. It forms the basis for developing healthcare systems that are accurate, intelligent, and easily accessible.

1.6 Focus on MRI in Brain Tumor Detection

A brain tumor is an abnormal growth of cells in the brain or surrounding tissues. These tumors may be malignant (cancerous) or benign (non-cancerous). Because tumors press against or invade surrounding tissues, they can interfere with normal brain function. The symptoms depend on the tumor's location and size. Early detection of a brain tumour increases the chances of successful treatment, This results in enhanced survival rates and better outcomes. Timely intervention can prevent the growth or spread of the tumour to other parts of the brain, thereby reducing the risk of serious neurological damage. It also enables doctors to use less aggressive treatments with fewer side effects, thus improving patients' quality of life. Magnetic resonance imaging (MRI) is the most effective and widely used technique for diagnosing brain tumors because of its exceptional capacity to provide finely detailed images of brain regions. MRI scans use powerful magnetic fields and radio waves to create images of the brain, helping to detect abnormal growths and determine their size, location, and impact on surrounding tissues. Compared to other imaging methods, such as CT scans, this noninvasive method is safer for repeated use because it does not involve ionizing radiation. Furthermore, advanced MRI techniques, such as functional MRI (fMRI), diffusion tensor imaging (DTI), and contrast-enhanced MRI, can provide essential information about the type and grade of tumors, which is crucial for developing treatment plans and predicting outcomes [9].Figure 1.8 represents the clinical process of brain MRI analysis, where a medical professional examines a series of brain scan slices for diagnostic evaluation.



Figure 1.8: Detection of brain tumor from MRI images [9]

DTI (diffusion tensor imaging) is a sophisticated MRI method that helps visualize white matter pathways by mapping the diffusion of water molecules across the brain. This outlines the tumor's relationship to significant brain structures, which is crucial for surgical planning to prevent harm to vital neurological connections. Another extremely effective MRI method is functional magnetic resonance imaging (fMRI), which can identify variations in blood flow and assess brain activity. Since the functional regions of the brain involved in speech, movement, and cognition have been mapped, fMRI has proven highly helpful in diagnosing and treating brain tumors. If clinicians understand how tumors affect essential brain functions, they can design interventions that maximize tumor eradication while maintaining these functions. This understanding is essential for the early identification of brain tumors and subsequent treatment. Clinicians find this modality appealing because it accurately and noninvasively conveys excellent image quality. Advanced MRI techniques significantly increase the diagnostic accuracy of the test by offering details on tumor characterization and treatment planning that are essential to the overall outcome and patients' quality of life. The table 1.1 provide overview of AI Techniques in MRI Images Processing:

1.7 Overcoming MRI Classification Challenges

Classifying MRI images of brain tumors is a challenging task for radiologists. Above all, the characteristics of tumors vary in shape, size, location, and appearance, so classification becomes quite a challenge when the tumors have vague or irregular boundaries. In such scenarios, where the complexity is very high, high expertise is demanded; on the other hand, even highly skilled radiologists are sometimes confused about different tumor types. Moreover, the diagnosis from an MRI image is rather subjective and may vary from radiologist to radiologist for the same images. The subjectivity is further complicated by the fact that the task of image analysis by radiologists is really time-consuming, whereby a single image has to be scrutinized and thus fatigues are prone to error. The major challenges are detecting subtle abnormalities, mostly in early-stage tumors. Further, the

Table 1.1: Overview of AI Techniques in MRI Images Processing

AI Techniques	Description	Application in MRI Images Processing
Machine Learning (ML) [45]	Algorithms that learn from data to make predictions or decisions.	- Tumor classification - Tissue segmentation - Anomaly detection
Support Vector Machines (SVM) [46]	A supervised ML algorithm used for classification by finding the optimal hyperplane separating data classes.	- Classifying brain tumors - Differentiating between healthy and diseased tissues
Random Forests [47]	An ensemble learning method that combines multiple decision trees for improved accuracy.	- Feature selection - Multi-class classification of brain abnormalities
K-Nearest Neighbors (KNN) [48]	A simple ML algorithm that classifies data points based on the majority class of their nearest neighbors.	- Tumor classification - Detecting anomalies in MRI images
Deep Learning (DL) [13]	Advanced ML using neural networks with multiple layers to model complex data patterns.	- Automatic tumor detection - Image segmentation - MRI image enhancement
Convolutional Neural Networks (CNNs) [49]	Neural networks designed for image data, excelling in feature extraction from spatial hierarchies.	- Brain tumor segmentation - MRI image classification - Feature extraction
Recurrent Neural Networks (RNNs) [50]	Neural networks tailored for sequential data, capturing temporal dependencies.	- Analysis of time-series MRI data - Sequence prediction in functional MRI
Autoencoders [51]	Unsupervised learning models that learn efficient representations of data, often used for dimensionality reduction.	- Denoising MRI images - Feature extraction - Anomaly detection
Hybrid Models [52]	Combines different AI techniques to leverage their strengths, enhancing performance.	- CNN-SVM for improved MRI classification - Wavelet-CNN for noise reduction and feature extraction
Reinforcement Learning (RL) [53]	AI that learns to make sequences of decisions to maximize a reward function.	- Optimization of MRI scanning protocols - Adaptive image enhancement
Explainable AI (XAI) [54]	Techniques that make AI models interpretable and transparent, crucial for clinical adoption.	- Understanding model predictions in MRI diagnostics - Highlighting important regions in MRI images

variability in the MRI acquisition parameters, like the differences in the imaging protocols and the settings of the machines, can alter the quality of the image and thus further complicate the precise classification. Based on these challenges that were discussed above, the need for Computer-Aided Diagnosis in developing a system for the precise classification of the MRI images was seen. Such a Computer-Aided Diagnosis system enhances accuracy and consistency because it analyzes the MRI images with the help of advanced algorithms and machine learning approaches, doing away with the subjective factors in human interpretation. These systems are able to pick up very subtle abnormalities that might easily be missed by the naked eye, such as the discovery of early-stage tumors and subtle variations indicative of malignancy. In addition, CAD systems help to increase the effectiveness of their work by processing huge amounts of images in a short period of time, which is of utmost importance in clinical settings where timely diagnosis is critical. They also help to support radiologists with a second opinion, and thus facilitate more informed decisions for those cases that involve complexities. CAD systems standardize diagnoses through consistent algorithms for different patients, making them invariable across radiologists and institutions. In addition, CAD systems could even include advanced imaging methods capable of revealing an integral overview of tumors and enhancing the treatment plan through techniques based upon machine learning. On the whole, CAD systems have emerged as the indispensable tools in mitigating challenges encountered by radiologists during MRI image classification, and this goes on to improve patient outcome. The figure 1.9 represents the automated pipeline for brain tumor detection and diagnosis from MRI images.

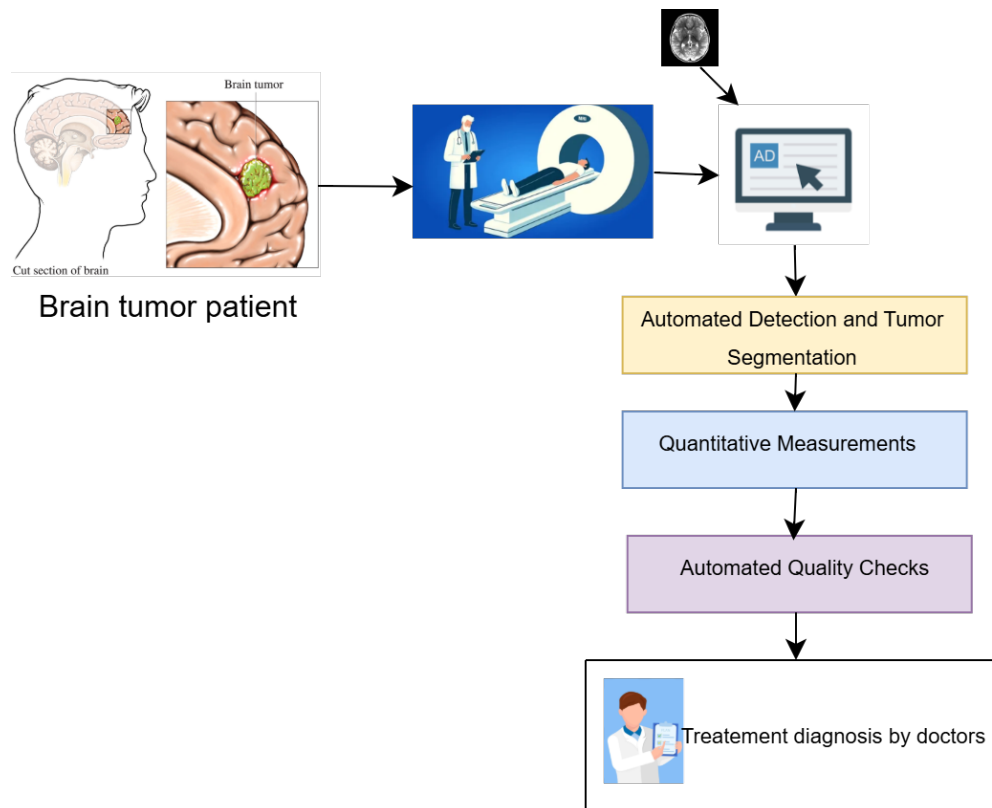


Figure 1.9: Automated Workflow for Brain Tumor Detection and Diagnosis [55]

1.8 Conclusion

In this chapter, we provided an explanation of medical signals, including their definitions and types, such as electrophysiological, imaging and biomechanical signals. We then discussed some significant pre-processing methods that enhance the quality and reliability of these signals, such as noise cancellation, standardisation and feature extraction techniques. We paid particular attention to MRI images, especially in the context of brain tumour detection, investigating their preprocessing steps in detail. Finally, we provided an overview of AI techniques that are increasingly being applied to MRI images for brain tumour classification. This will provide a basic understanding, and subsequent chapters will cover more advanced, state-of-the-art developments in medical signal processing and analysis.

Chapter 2

Wavelet Transforms

2.1 Introduction

In the latter half of the 20th century, the wavelet transform emerged as a powerful alternative. In contrast, the Fourier transform only operates in the frequency domain, the wavelet transform provides time and frequency localisation [56]. This dual capability enables the effective analysis of transient and continuous components within a signal. Consequently, wavelet transforms are particularly well-suited to medical signal processing, including ECG, EEG and MRI [57]. This chapter begins with the basic theory of wavelets, including definitions and key properties. It explains the differences between continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The discussion then moves on to cover various wavelet families, such as orthogonal, biorthogonal and complex wavelets, highlighting their unique characteristics and suitability for different types of medical signals.

Having established the theoretical foundation, the chapter goes on to explore the practical applications of wavelet transforms in medical signal processing. Key use cases include signal denoising, feature extraction, time-frequency analysis and data compression. These applications demonstrate the advantages of wavelet methods over traditional approaches, offering higher precision, greater efficiency and better adaptability to the nature of medical data.

2.2 Wavelet Theory

The wavelet transform is based on wavelet theory, which is the mathematical study of signal analysis at different resolutions [58]. Unlike conventional Fourier transforms, which only offer frequency information, wavelet transforms record a signal's time and frequency properties. This dual capability is essential when analyzing non-stationary signals, whose characteristics vary over time. The equation below represents the wavelet transform of a signal $x(t)$, using a scaled and translated version of a mother wavelet $\psi(t)$:

$$W(s, \tau) = \int_{-\infty}^{\infty} x(t) \cdot \frac{1}{\sqrt{s}} \psi^* \left(\frac{t - \tau}{s} \right) dt \quad (2.1)$$

In this expression, s denotes the scale parameter controlling frequency resolution, τ is the translation parameter for time localization, and $\psi^*(t)$ is the complex conjugate of the mother wavelet.

2.2.1 Basic Principles and Properties

Wavelets are small oscillations and, in both time and frequency, localisation is possible. The wavelet transform works by decomposing a signal into a series of basic functions, known as wavelets, which can effectively represent the various characteristics of a signal at different scales. Each wavelet is obtained from a prototype function known as the 'mother wavelet' [59]. The key features of wavelets are:

Localization: The wavelets are localized either in time or in frequency, so one can analyze a transient, short-lived phenomenon within a signal.

Multiresolution: Wavelets can be adapted to analyze signals in multiple resolutions so that they are able to do both coarse and fine analysis.

Vanishing Moments: Wavelets having vanishing moments are much powerful in representing the polynomial trends in signals; that would be helpful in solving the problem of removing noise.

2.3 Mother Wavelets

The mother wavelet (ψ) is a core function from which all other wavelets are derived[60]. It is a mathematically defined function and generally selected according to the particular application needs. The mother wavelet has to be oscillatory and have finite energy. Some most commonly useful examples of mother wavelet include the Haar wavelet, Daubechies wavelets, and the Morlet wavelet.

Through scaling and translation, wavelets are created from a single fundamental wavelet, also known as the "mother" wavelet function, denoted by the symbol $\psi(t)$. The mother wavelet can be expressed mathematically as follows:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (2.2)$$

where:

- a is the scale parameter (dilation or compression),
- b is the translation parameter (shifting),
- t is time,
- $\psi(t)$ is the mother wavelet.

2.4 Wavelet Families and Properties

Various wavelet families that have proven to be particularly significant are integrated into digital image processing [61].

2.4.1 Orthogonal Wavelets

Orthogonal Wavelets are wavelets which are fitted for signal decomposition and reconstruction processes, both without losing any information. One of the important features of orthogonal wavelets is that the functions of wavelets and scaling are orthogonal; their inner product is zero. That property ensures the stability and invertibility of the mathematical processes for decomposition and reconstruction [62].

Haar Wavelets

The earliest and most basic wavelet in the family is the Haar wavelet. Functions obtained by sampling Haar functions in two different locations are known as discrete Haar wavelets. These Haar functions can be efficiently represented in matrix form, with each row representing a Haar function. The Haar wavelet appears to be a step function and is highly discontinuous. The Daubechies db1 wavelet is equivalent to the Haar wavelet. According to the Haar wavelet theorem, all functions are continuous on the interval $[0,$

1] and can be expressed as a convergent series with respect to the system's components [63]. The mathematical definition of the original Haar wavelet can be defined as follows:

$$\text{Haar}(0, X) = 1, \quad \text{pour } X \in [0, 1] \quad (2.3)$$

$$\text{Haar}(1, X) = \begin{cases} 1, & \text{for } X \in [0, \frac{1}{2}] \\ -1, & \text{for } X \in [\frac{1}{2}, 1] \end{cases} \quad (2.4)$$

where X is the discrete point of the function computation interval.

The Haar matrix of all ranges can be recursively constructed using:

$$M(n) = \begin{bmatrix} M(n-1) \\ 2^{(n-1)/2} D(n-1) \end{bmatrix} \quad \text{with } M(0) = 1 \quad (2.5)$$

Here:

- $M(n)$ represents the matrix of discrete Haar functions of degree 2^n ,
- $D(n-1)$ denotes the identity matrix of degree 2^{n-1} scaled appropriately,
- K is the Kronecker (tensor) product.

Additionally:

$$M(n) \neq M(n)^T \quad \text{for } n > 1 \quad \text{and} \quad M(n)^{-1} = 2^{-n} \cdot M(n)^T$$

The Haar wavelet functions are asymmetric, and their elements take values of 1, -1, or 0. The product of these functions can be expressed in terms of powers of $\sqrt{2}$. These discrete and orthogonal Haar functions are defined within the interval $[0, 1]$.

Daubechies Wavelets

This is an important family of wavelets thanks to their compact support and smoothness properties. Daubechies wavelets are defined by the number of vanishing moments, which determine their ability to represent polynomial trends in a signal. Daubechies-4 db4 has four vanishing moments, offering a good balance between smoothness and compactness. Consequently, it is widely used in the analysis of smooth and transient signals.

Each Daubechies wavelet function ψ_n is compactly supported. The index number n is also connected to the number of vanishing moments[64].

$$\int_{-\infty}^{+\infty} x^u \psi_n(x) dx = 0, \quad 0 \leq u \leq n \quad (2.6)$$

Another crucial characteristic of Daubechies wavelets is their ability to expand their regularity linearly with their support width. With the exception of db1, the equivalent of the Haar wavelet, Daubechies wavelets are not symmetric or asymmetric with respect to any axis. Each Daubechies wavelet satisfies the admissibility requirement, ensuring safe reconstruction.

2.4.2 Biorthogonal Wavelets

Unlike orthogonal wavelets, which restrict themselves to their scaling functions for orthogonality, biorthogonal wavelets extend this feature so that each is orthogonal with possibly different wavelet and scaling functions [65]. This relaxation of the strict principles of orthogonality allows wavelets to be designed with desirable properties such as symmetry and linear phase, which are important for certain signal processing tasks [66].

A significant advantage of biorthogonal wavelets is their ability to achieve symmetry and linear phase [67], which is typically challenging or even impossible with purely orthogonal wavelets. This ensures that the wavelets do not introduce phase distortions, which is very important for image and signal processing applications that concern the original structure of the signal [68]. Conversely orthogonal wavelets are quite asymmetric and therefore not very effective for tasks where visual or structural integrity must be preserved.

The biorthogonal wavelet design allows for two filter banks, a decomposition filter bank and a reconstruction filter bank, which do not necessarily have to be identical [69]. This is especially crucial in real-world situations, such as denoising, where maintaining the signal's phase and structure during decomposition and reconstruction is essential. These wavelets have many applications in medical signal processing, such as image compression and denoising, where maintaining data integrity requires highly precise results.

The Cohen-Daubechies-Feauveau (CDF) wavelet, a popular family of biorthogonal wavelets, is widely used in signal and image processing applications. Image compression in JPEG 2000 is a common example. Processing medical images such as MRIs and CT scans requires the ability to represent both sharp and smooth signal fluctuations, a key benefit of biorthogonal wavelets [70].

Advantages of Biorthogonal Wavelets

Symmetry: Symmetry is key for many practical applications, such as medical imaging, in which phase distortion can cause visual artifacts.

Linear Phase: The phase should be linear and preserved, since it holds the signal structure, especially for medical images in edge detection and segmentation.

Efficient Reconstruction: With separate wavelet and scaling functions for decomposition and reconstruction, biorthogonal wavelets can ensure signals to be reconstructed efficiently after processing without loss of critical information.

In a nutshell, the biorthogonal wavelet transform is more flexible than orthogonal wavelets, hence fit very well for medical signal processing tasks in which the preservation of the signal integrity along with accurate reconstruction is highly critical [67].

2.5 Continuous Wavelet Transforms

The wavelet transform offers greater flexibility when selecting the analysis function than the Fourier transform does. Unlike the Fourier transform, it does not require the use of sine waves, but rather allows the use of compact oscillating waveforms. A wavelet function, denoted by the symbol

$\psi(t)$, is used to effectively separate different frequencies. This function combines the analysis pattern and the windowing process. A common example is the Morlet wavelet, which is illustrated in the figure 2.1, Various wavelet functions have been developed for the continuous wavelet transform (CoWT), each with unique properties[71].

A function $\psi(t)$ is considered a wavelet when it meets certain mathematical conditions.

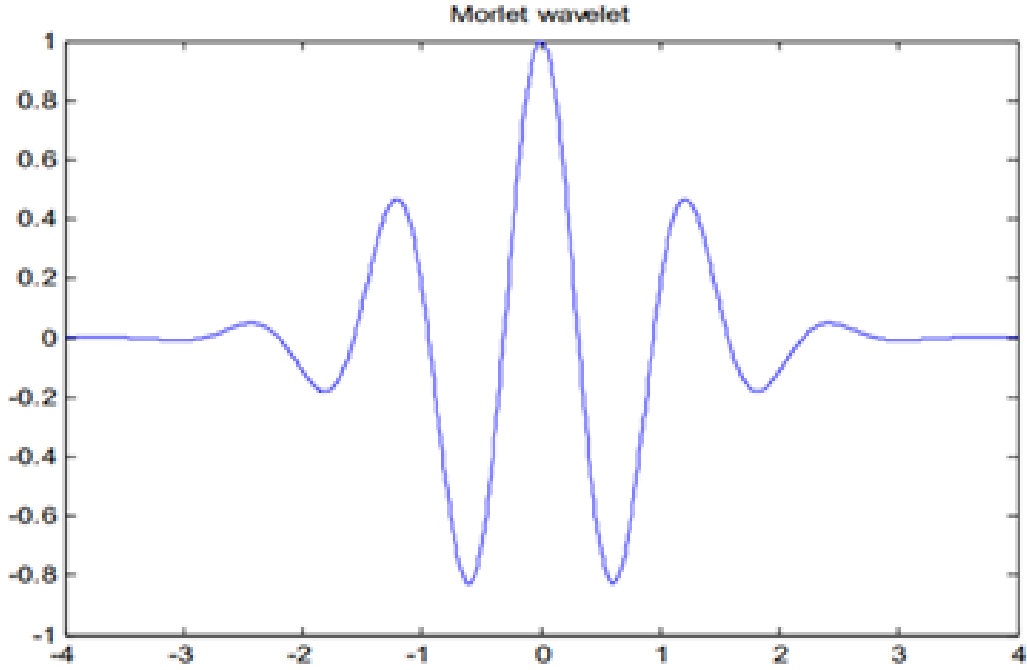


Figure 2.1: Morlet wavelet [72]

$$E = \int |\psi(t)|^2 dt < \infty \quad (2.7)$$

The energy, denoted by E , must be finite and corresponds to the square root of the magnitude of the analyzing function $\psi(t)$.

If $\psi(f)$ represents the Fourier transform of the wavelet $\psi(t)$, the following requirement must be satisfied.

$$C_\psi = \int_0^\infty \frac{|\psi(f)|^2}{f} df < \infty \quad (2.8)$$

The continuous wavelet transform is defined as follows:

$$X_{WT}(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} X(t) \psi^* \left(\frac{t - \tau}{s} \right) dt \quad (2.9)$$

The transformed signal, denoted as $X_{WT}(\tau, s)$, depends on the scale parameter s and the translation parameter τ . The original wavelet, known as the mother wavelet, is used in its complex conjugate form when dealing with complex wavelets. To maintain consistent energy across different scales, the wavelet coefficients are normalized by dividing by $\frac{1}{\sqrt{|s|}}$.

This ensures equal energy distribution for wavelets at all scales.

Like the Fourier Transform (FT) and Short-Time Fourier Transform (STFT), the Wavelet Transform (WT) also has an inverse. This is called the Inverse Continuous Wavelet Transform (ICOWT), and it is defined by a specific mathematical expression.

$$X(t) = \frac{1}{C_\psi^2} \iint_{-\infty}^{+\infty} X_{WT}(\tau, s) \cdot \frac{1}{s^2} \psi \left(\frac{t - \tau}{s} \right) d\tau ds \quad (2.10)$$

2.6 Discrete Wavelet Transforms

The Discrete Wavelet Transform (DWT) is a method used to analyze signals and images by breaking them down into discrete frequency bands. DWT has been used in several domains, including medical image denoising and signal and image processing [73].

Signal analysis can be carried out at different frequency resolutions using the discrete wavelet transform (DWT). In particular, DWT exhibits modest frequency resolution but remarkable time resolution when working with high frequencies. Conversely, DWT lacks time resolution but provides significant frequency resolution for low frequencies. To perform time-frequency analysis using DWT, a signal x is subjected to a number of filtering processes using high-pass and low-pass filters. A half-band low-pass filter eliminates all frequencies above half of the signal's highest frequency, and a half-band high-pass filter eliminates all frequencies below half of the signal's highest frequency [73].

Initially, the signal x samples are subjected to a concurrent decomposition process via convolution with low- and high-pass filters characterised by impulse responses g and h , respectively [74]. This procedure yields the following outcomes:

$$Y_{\text{low}}[n] = \sum_{k=-\infty}^{\infty} x[k] g[2n - k] \quad (2.11)$$

$$Y_{\text{high}}[n] = \sum_{k=-\infty}^{\infty} x[k] h[2n - k] \quad (2.12)$$

The values $Y_{\text{low}}[n]$ and $Y_{\text{high}}[n]$ denote the outcomes derived from the low-pass and high-pass filters, respectively, subsequent to a down sampling procedure featuring a reduction factor of 2.

Currently, only half as many samples are used in the signal representation as in the original signal. Consequently, the time resolution is cut in half. However, the frequency resolution is doubled because the frequency spectrum now only covers half of what it did before. The results obtained from the high-pass and low-pass filters are referred to as detail and approximation coefficients, respectively. If more decomposition is needed to extract the desired frequency spectrum from the input signal, the decomposition approach can be used again. With each further decomposition step, the signal separates into low-frequency and high-frequency bands. Consequently, the time resolution is cut in half, and frequency resolution is doubled in comparison to the previous level due to the filtering and sub-sampling process. As illustrated in Figure 2.2, this process forms a three-level filter bank.

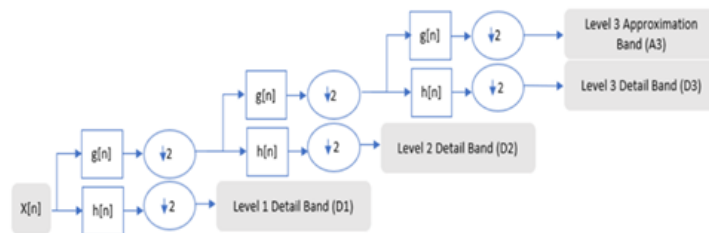


Figure 2.2: A Discrete Wavelet Transform (DWT) model of 3 levels decomposition [75]

When creating a new signal using the inverse discrete wavelet transform (IDWT)—the opposite of the DWT—the approximation and detail coefficients derived from signal decomposition are essential. To rebuild the new signal, the signals at each level must first be upsampled and multiplied by two, then sent through high-pass and low-pass filters [74].

2.7 Undecimated Wavelet Transforms (UWT)

Also known as stationary wavelet transform the primary constraint of the Continuous Wavelet Transform (CWT) is the intricate nature arising from numerous parameter selections, which render its practical implementation less straightforward [76]. Additionally, due to its computational demands, the CWT is most viable when both the original signal and the mother wavelet have explicit mathematical forms. Conversely, the Stationary Wavelet Transform (SWT) conducts discrete wavelet decomposition without down sampling. Upon decomposition, the resultant length of output doubles the input length due to the filter bank structure. This permits the preservation of information lost in coefficient downsampling. The scale and shift components in continuous wavelet decomposition change constantly over time. At the discrete wavelet decomposition field, $b \in R$ and $a \in R^+ \setminus 0$, the factors of scale and shift exhibit variation across distinct numerical values. ($a = 2^j$, $b = k \cdot 2^j$, $(j, k) \in \mathbb{Z}^2$) In the context of the decomposition utilizing stationary wavelet transform (SWT), the parameter j denotes the level of decomposition, while k represents the discrete time position.

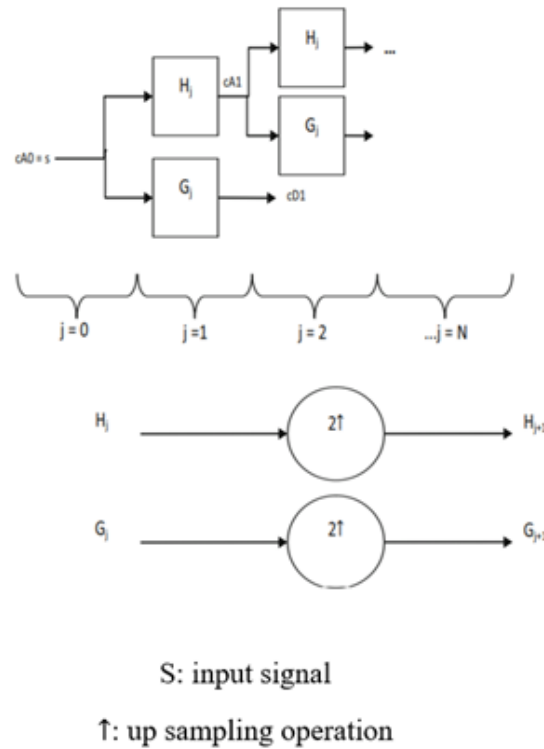


Figure 2.3: The stationary wavelet transforms using multi-level decomposition [77]

Figure 2.3 visually illustrates the sequential process of SWT decomposition at different levels. This method partitions signals into low-pass components and other details.

The process of signal decomposition involves separating a signal into two parts: a frequency component and a higher-frequency component. Approximation coefficients come from low-pass filtering, while detail coefficients come from high-pass filtering. To reconstruct a signal that has undergone decomposition at level j , one must combine the approximation coefficients and detail coefficients from level $j + 1$. This method enables the reconstruction of the original signal with a high degree of accuracy.

$$cA_j = cA_{j+1} + cD_{j+1} \quad (2.13)$$

In this context, the cA_j 's denote the approximation coefficients at level j , whereas the cA_{j+1} 's signify the approximation coefficients at level $j + 1$. The cD_{j+1} 's, on the other hand, pertain to the detail coefficients at level $j + 1$ [78].

2.8 Complex Wavelets

Complex wavelets are an extension of traditional wavelet transforms, using complex-valued wavelets to provide richer information particularly regarding phase and amplitude — than real-valued wavelets. They are employed in signal processing and various scientific disciplines for tasks such as image analysis, time-frequency analysis and noise reduction[59].

2.8.1 DT-CWT

The Dual-Tree Complex Wavelet Transform (DTCWT) is an improved form of the Discrete Wavelet Transform (DWT). Unlike the DWT, the DTCWT adds directional sensitivity and approximate shift invariance in high-dimensional spaces. Thanks to approximate shift invariance in the DTCWT, identical textures produce consistent response features regardless of where they are sampled. In contrast, the DWT may yield varied responses for the same texture when sampled from different places. Furthermore, the DTCWT's directional sensitivity enhances its ability to process diagonal features more effectively than the DWT. This enables the DTCWT to depict geometric features without creating checkerboard-like patterns, making it useful for denoising images [79].

Two critically sampled DWTs are applied to the same dataset simultaneously in order to perform the dual-tree complex wavelet transform (DTCWT) on a given signal, $x(n)$. Additionally, using the same filters for the higher and lower trees is not beneficial. Consequently, the filters in this setup are designed specifically. The sub-bands from the lower tree comprise the imaginary part of the complex wavelet transform, while the sub-bands from the higher DWT comprise the real part. This transformation provides shift invariance and expands by a factor of two [80]. Figure 2.4 illustrates the structure of the complex wavelet transform. (a) shows the real and imaginary parts of a 1D complex wavelet. (b) shows the decomposition tree for the real and imaginary components. (c) shows the corresponding reconstruction tree.

2.9 Lifting wavelet transform

In the traditional DWT, the wavelet transform is carried out through a sequence of filtering operations derived from the Fourier transform, where low-pass and high-pass filters

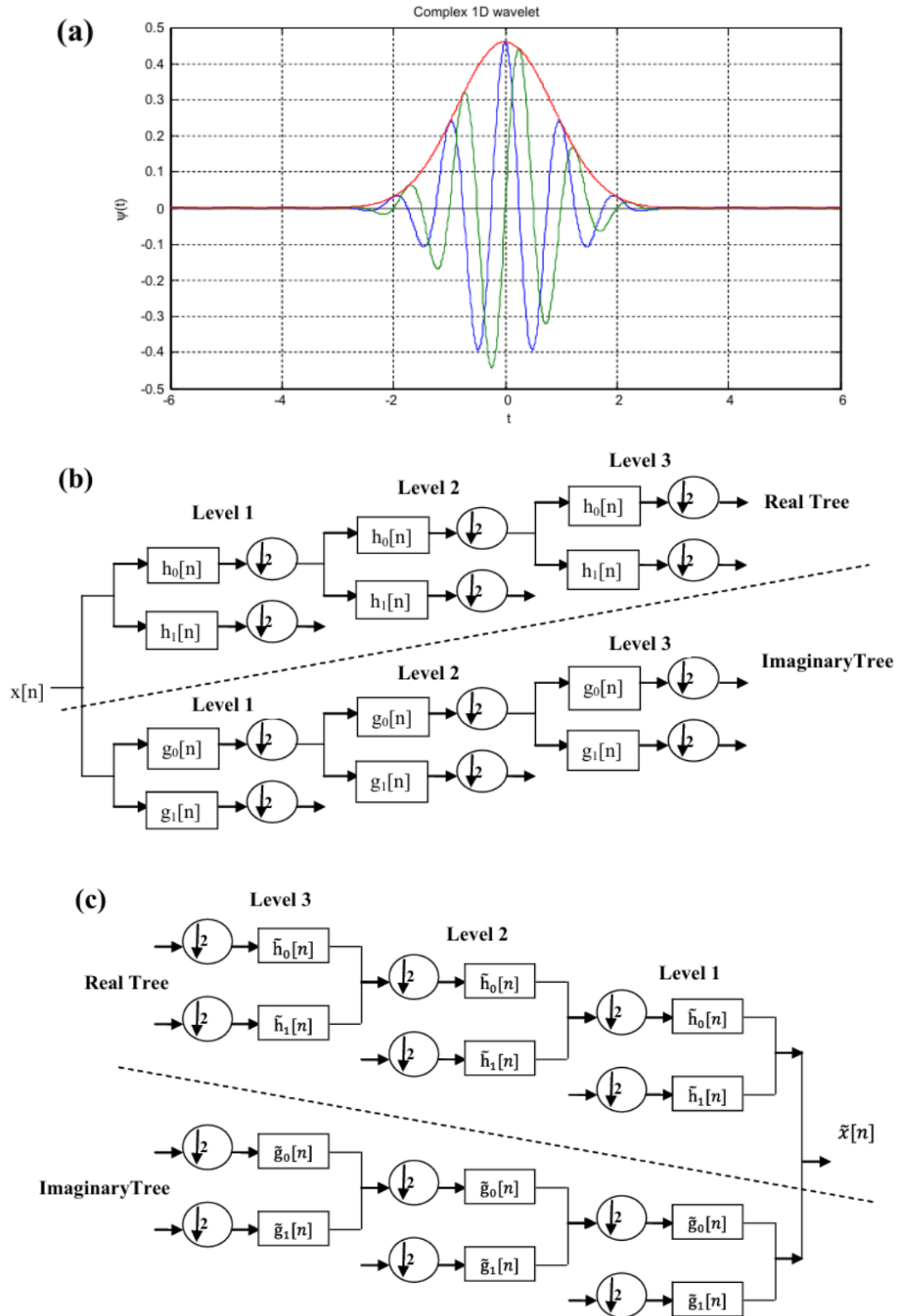


Figure 2.4: - (a) 1D complex wavelet, (b) filter bank analysis for 1-D DTCWT, (c) synthesis filter bank for 1-D DTCWT [80]

combined with down-sampling generate the approximation and detail sub-bands. In contrast, the LWT is a variant of DWT that applies the lifting scheme, which decomposes

the signal (or image) through three steps: split, predict, and update, without relying on the Fourier transform [81]. Known as the second generation of wavelets, LWT provides a more flexible and efficient approach, reducing computational cost and memory usage. Its implementation follows three successive steps 2.5, described as follows:

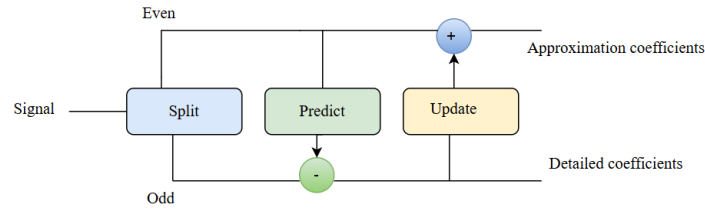


Figure 2.5: Forward lifting wavelet transform [82]

2.9.1 Split

In the split stage, also known as the lazy wavelet, the signal is separated into two subsets, typically odd and even samples.

2.9.2 Predict

In this step, the value of a pixel at an odd position (X_o) is estimated using the values of its neighboring pixels at even positions (X_e). The difference between the predicted and actual values of the odd sample is then stored in its position, forming the detail band (D_n). In regions with gradual intensity changes, predicted values closely match the actual ones, resulting in detail coefficients near zero. Conversely, areas with significant intensity variations produce higher detail coefficients, indicating stronger local differences. This property is leveraged in the proposed texture mask. The prediction step can be expressed as:

$$Y_{2n+1} = X_o - PREDICT(X_e) \quad (2.14)$$

2.9.3 Update

To preserve the overall structure of the signal while reducing its energy, the update step computes the average of adjacent odd samples and assigns this value to the corresponding even sample. Since pixel intensities are not perfectly linear, the even sample must also be adjusted using the difference obtained in the prediction stage. The resulting signal is referred to as the approximation band, as it retains the main structure of the original signal but with half the number of samples:

$$Y_{2n} = X_e + UPDATE(Y_{2n+1}) \quad (2.15)$$

2.10 Applications to Medical Signals

The wavelet transform has emerged as a powerful tool for analysing medical signals, which often exhibit complex and non-stationary characteristics [83]. Thanks to its ability

to localise both time and frequency information, the wavelet transform offers significant advantages over traditional Fourier methods. Its applications extend across various domains of medical signal processing, providing efficient techniques for denoising, feature extraction, compression and classification. The following subsections will cover key areas in which the wavelet transform has proven particularly useful [84].

2.10.1 Signal Denoising

Although noise removal is a widely recognised application of wavelets[85], it tends to be more prevalent in image processing than in signal processing. Denoising involves adjusting the intensity of specific frequency channels to reduce or suppress noise. If the signal is confined to a specific frequency range, noise can be reduced by deactivating all frequency channels outside this range. In cases where random noise is present, eliminating minor fluctuations within the signal can help with denoising. This can be achieved by setting the coefficients of the highest frequency channel to zero, thereby minimising small variations in the signal.

Wavelet for image denoising Wavelet transforms are increasingly recognized for their effectiveness in image denoising[86], leveraging their ability to capture both spatial and frequency information. Various methodologies have emerged, each enhancing the denoising process while preserving image quality, the table 2.1 summarizes some of the methodologies.

2.10.2 Feature Extraction

Feature extraction is an essential phase in the evaluation and categorization of medical signals. The wavelet transform enables meaningful features to be extracted from time- and frequency-dependent signals or MRI images [89]. By decomposing signals or images into various scales, wavelet transforms capture transient and localised features that are often missed by traditional Fourier-based methods. Feature extraction using wavelet transforms is a powerful technique for extracting features from medical images. This technique is ideal for analyzing non-stationary signals because it effectively captures time and frequency data. It is essential for training effective classifiers because it produces compact and informative representations of the original data. Wavelet-transformed features preserve critical textural and structural details at various resolutions, which is crucial for distinguishing between tumor types when classifying brain tumors from MRI images. These extracted features increase the effectiveness and precision of classification models by decreasing dimensionality while maintaining discriminative qualities. Wavelet-based features improve a model's ability to identify optimal decision boundaries in high-dimensional feature spaces when used to train classifiers, such as Support Vector Machines (SVMs). SVMs in particular benefit from the feature separability of the wavelet transform, which helps them better identify intricate patterns. Similarly, because wavelet-based features enhance the signal-to-noise ratio, making them more resilient to irrelevant data, other classifiers, such as Random Forest, k-Nearest Neighbors, and Multilayer Perceptron (MLP), also perform better when trained on them. Overall, the integration of wavelet transform for feature extraction lays a strong foundation for reliable classification by emphasizing the most relevant and localized features from MRI data, supporting the development of accurate and generalizable diagnostic systems.

Table 2.1: Comparison of Advanced Denoising Techniques for Medical Images Using Wavelet Transforms

Technique	Description	Advantage
Dual Tree Complex Wavelet and Shrinkage with the Wiener Filter Technique [87]	This approach uses either hard or soft thresholding operators of dual tree complex wavelet transform for the denoising of medical images.	Provides a better balance between smoothness and accuracy than the discrete wavelet transforms (DWT) and is less redundant than the stationary wavelet transforms (SWT). Evaluated using PSNR and SSIM.
Lifting Dual Tree Complex Wavelet Transform (LDTTCWT) with Hybrid Thresholding and Wiener Filter [80]	The approach involves two steps: estimating the noise level of the image and then using LDTTCWT for denoising.	Enhances the quality of medical images by preserving data-carrying structures like surfaces and edges.
Hybrid Thresholding with Wiener Filter Technique using Wavelet Transforms [88]	The image is decomposed using DWT, then hybrid thresholding (Bayes Shrink) and Wiener filter are applied to reduce noise and improve image quality.	Improves image quality and reduces noise level. Evaluated using SSIM and PSNR.
Image Denoising Network Based on Sub-band Information Sharing Using Dual-tree Complex Wavelet [79]	The network combines spatial and transform domains through dual-tree complex wavelet transform to capture structured and time-frequency localized features.	Effectively recovers complex texture regions. Addresses difficulties in previous denoising methods.

2.10.3 Signal Compression and Coding

Medical imaging data, such as magnetic resonance imaging (MRI) or computed tomography (CT) scans, generates vast amounts of information, necessitating efficient compression techniques for storage and transmission [90]. The wavelet transform's ability to condense signal energy into a small set of significant coefficients makes it an effective tool for signal compression. By converting the signal into the wavelet domain, where only the most important information is preserved, wavelet-based compression, also referred to as wavelet coding, reduces the amount of data. Medical imaging standards like JPEG 2000 extensively use this technique because it minimizes image quality degradation while drastically reducing file size. Wavelet-based compression is essential for quickly and effectively transmitting medical images over networks for telemedicine and remote diagnostics without sacrificing diagnostic precision.

2.10.4 Other Applications

In addition to denoising, feature extraction and compression, the wavelet transform is used for other advanced tasks, such as data fusion, image segmentation and classification [91]. In medical image fusion [92], For instance, wavelets are used to merge several imaging modalities, such as CT and MRI scans, into a single image while maintaining the complementary information from each source. Improving contrast and detail raises the standard of diagnostic imaging. Wavelet transforms identify edges and regions of interest at various scales for medical image segmentation, enabling the precise identification of diseased or anatomical features. This is useful when segmenting brain tumors [92], for instance, wavelets can effectively capture subtle texture differences between normal and abnormal tissue, aiding accurate diagnosis and treatment planning. Wavelet-based features have also been employed in classification tasks such as tumour detection and grading [93]. Leveraging the wavelet transform's ability to extract multi-scale features enables machine learning models to distinguish between different types of tumours or disease states with greater accuracy. This approach has been used to classify breast cancer, lung nodules, and neurological disorders using MRI, CT, and other imaging modalities.

2.11 Conclusion

This chapter explores the application of wavelet theory. It introduces key concepts, such as continuous and discrete wavelet transforms and various wavelet families. The chapter highlights the flexibility and robustness of the theory for analyzing complex medical signals, such as electrocardiograms (ECGs), electroencephalograms (EEGs), and magnetic resonance imaging (MRI) scans. The wavelet transform enhances the precision and efficacy of medical signal processing tasks, and this chapter discusses its practical applications. For the MRI-based brain tumor classification task, we utilized the stationary wavelet transform (SWT), also referred to as the undecimated wavelet transform (UWT), to improve feature extraction while maintaining spatial integrity. By eliminating the downsampling step present in conventional wavelet transforms, SWT/UWT preserves fine spatial structures and high-frequency components at multiple scales. This characteristic leads to a more accurate representation of tumor regions and enhances classification robustness. Continued progress in wavelet-based techniques is anticipated to expand their impact in medical image analysis and automated diagnostic systems.

Chapter 3

Artificial Intelligence: An Overview

3.1 Introduction

Artificial intelligence (AI), a subfield of computer science, seeks to create machines that can carry out operations such as learning, analyzing, predicting, and making decisions, which often require human intellect [94]. Some of the main areas that fall under artificial intelligence (AI) include natural language processing (NLP), computer vision, deep learning (DL), and machine learning (ML) [95]. Since its early beginnings, AI has evolved significantly. It started with simple attempts to simulate basic human behavior and progressed to complex systems capable of self-learning and achieving performance that often surpasses human abilities in specific applications [96]. Today, AI is widely applied across many vital sectors, including smart transportation and autonomous driving, financial services and market analysis, cybersecurity and threat detection, industrial production management, and healthcare [97]. In the medical field specifically, AI has become a powerful tool to support physicians and improve healthcare quality [98]. It is used for medical signal analysis (such as ECG, EEG, MRI, CT, and X-Ray), automated disease diagnosis, patient monitoring through vital signals, and predicting disease progression based on long-term clinical data [99]. Due to the complex and dynamic nature of medical signals, AI techniques have become essential for accurately and efficiently extracting meaningful information from them [100]. This chapter is divided into two main parts: the first part presents general concepts and techniques of artificial intelligence, while the second part focuses on its specific applications in medical signal processing.

3.2 Artificial intelligence, machine learning, deep learning

Artificial intelligence (AI) is the simulation of human intelligence in computers that are designed to understand, learn, and solve problems like humans do. AI includes several methods, such as computer vision, natural language processing, and machine learning. AI systems can process large volumes of data to identify trends, reach conclusions, and carry out activities independently. These systems learn from experience and improve over time without explicit programming. AI applications are found in various sectors, including robotics, healthcare, and finance [101].

Machine Learning (ML), a type of artificial intelligence, allows computers to learn from data and make judgments or predictions without explicit programming. ML involves creating algorithms that can recognize trends in data, improve over time, and adapt to new information. ML methods include reinforcement learning, supervised learning, and unsupervised learning. Applications such as image classification, recommendation systems, driverless cars, and speech recognition extensively use these models. Machine learning's main advantage is its ability to generalize and learn from experience [102].

Deep Learning (DL) a branch of machine learning, is centered on multi-layered neural networks, also known as "deep neural networks" [103]. It processes vast volumes of data and automatically recognizes intricate patterns by simulating the composition and operations of the human brain. These models excel at tasks such as autonomous driving, natural language processing, and image and audio recognition. While these models require substantial data and processing power, they can achieve greater accuracy than traditional machine learning techniques. Two important architectures are recurrent neural networks (RNNs) and convolutional neural networks (CNNs).

The figure 3.1 represents illustrates the relationship between three core concepts in modern computational intelligence: Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL).

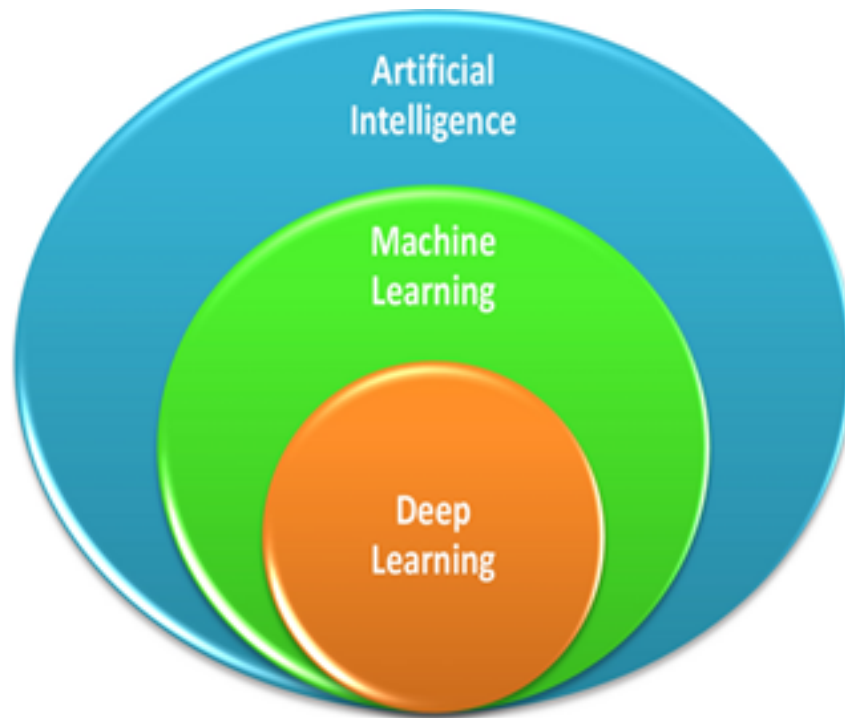


Figure 3.1: Hierarchical Relationship Between Artificial Intelligence, Machine Learning, and Deep Learning [104]

3.3 Machine Learning vs Deep Learning

Although machine learning (ML) and deep learning (DL) are both subfields of artificial intelligence, they differ in terms of feature processing, data requirements, and complexity. ML models that work well with smaller datasets, such as Support Vector Machines (SVMs), k-Nearest Neighbors (KNNs), and Decision Trees, typically require manual feature extraction. DL, on the other hand, is appropriate for large-scale, high-dimensional datasets because it automatically extracts and learns features from raw data using deep neural networks, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs). By learning hierarchical representations directly from the data, DL eliminates the need for domain expertise, which is often required for feature engineering in ML. However, DL models require longer training times and more processing power [105].

The figure 3.2 represent this difference between ML and DL:

3.4 Types of ML

Machine learning (ML) can be divided into three primary categories based on how the model learns from data such as provided in figure 3.3:

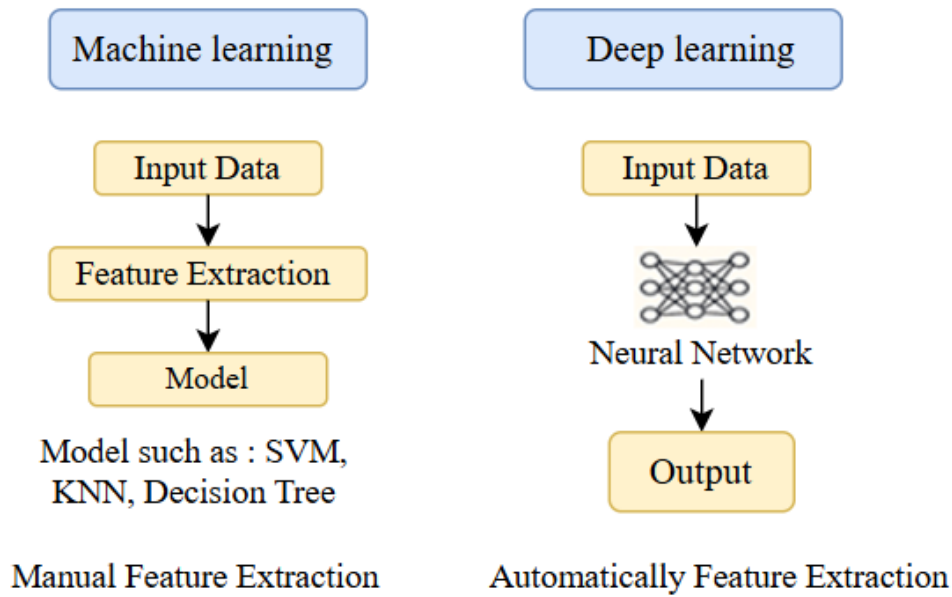


Figure 3.2: Difference Between Machine Learning and Deep Learning [105]

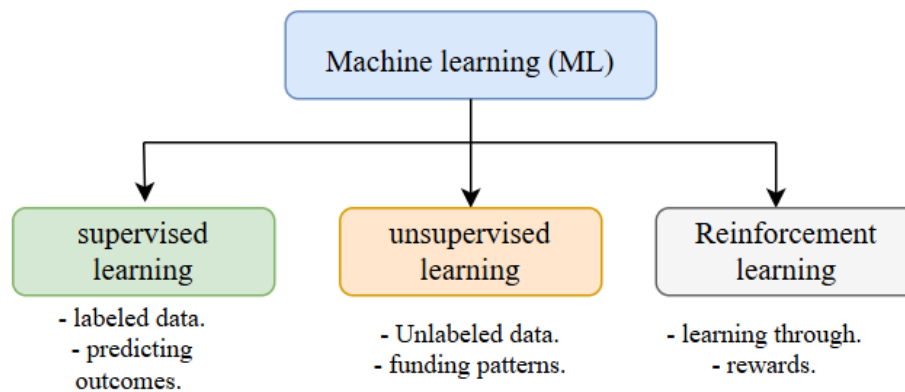


Figure 3.3: Overview of Machine Learning Types [106]

3.4.1 Supervised Learning

In supervised learning, a subset of machine learning, a labeled dataset comprising input data and matching target outputs is used to train the model. The objective is to minimize the discrepancy between the expected and actual values in order to learn the mapping between the inputs and outputs. Supervised learning consists of two primary tasks: regression, which produces continuous outputs, and classification, which produces categorical outputs [107].

o **Classification:** is a supervised learning problem with a categorical output variable. The objective is to use patterns discovered from labeled training data to categorize new data points into one of many predefined classes. For example, the goal could be to determine whether an MRI scan shows a pituitary tumor, meningioma, glioma, or no tumor.[108].

o **Regression:** is a supervised learning job with a continuous output variable. The objective is to understand the relationship between the input features and use that knowledge to predict the output value. For instance: One example is using information from MRI scans to estimate the size of a brain tumor in millimeters.[109].

3.4.2 Unsupervised Learning

Unsupervised learning is one type of machine learning technique in which the model is trained on data without explicitly labeled results. Unlike supervised learning, which depends on input-output pairs, unsupervised learning looks for hidden structures, patterns, or correlations in the data. Unsupervised learning algorithms are particularly helpful for tasks such as clustering, dimensionality reduction, and anomaly detection because their main objective is to find underlying groupings or representations. Clustering, which groups data points according to their similarities, is one of the most popular applications of unsupervised learning. Well-known clustering techniques include K-means, DBSCAN, and hierarchical clustering. Without knowing the labels beforehand, these methods divide the data into discrete clusters, allowing for the identification of innate patterns in the dataset [110].

3.4.3 Reinforcement Learning

Through interaction with an environment, an agent can learn to make consecutive judgments through reinforcement learning, a type of machine learning. To optimize long-term cumulative rewards, the agent performs actions, receives input in the form of rewards or penalties, and adjusts its behavior accordingly. Learning occurs through trial and error, guided by the results of prior actions[111].Figure 3.4 illustrates Reinforcement learning the process in which an agent uses actions to interact with its surroundings, Based on the environment's response, the agent receives rewards and updates its strategy through observed states.

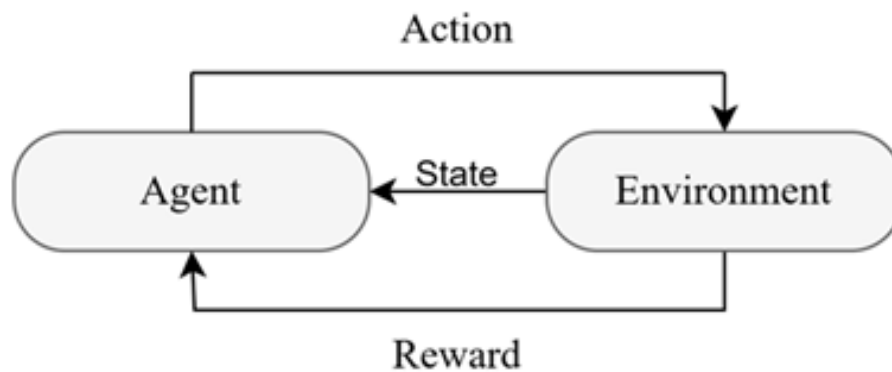


Figure 3.4: Reinforcement learning [111]

3.5 Commonly Used Machine Learning Algorithms

3.5.1 SVM

The Support Vector Machine (SVM) is a supervised machine learning approach used for classification and regression applications. Created by Vladimir Vapnik in the 1990s, SVM is renowned for its solid theoretical underpinnings, adaptability, and ease of use, even for those with no background in data mining [112]. The main objective of the algorithm is to find a hyperplane that successfully divides classes. Although there are a number of hyperplanes available, the optimal hyperplane maximizes the margin—the distance between the closest points of each class—to improve the classification of subsequent data points. The hyperplane is defined by these "support vectors," which are the closest points, as shown in Figure 3.5; There are two types of SVM classifiers: linear and nonlinear. A linear SVM model finds a straight hyperplane that optimizes the distance to the closest points of each class. This model assumes that the data is linearly separable. However, real-world data is often more dispersed, reducing the effectiveness of linear separation. Vapnik developed the "kernel trick" to address this issue, converting data into a higher-dimensional space that allows for nonlinear classification. SVMs have several advantages: they can handle high-dimensional feature spaces, they have superior classification accuracy, and they are less prone to overfitting. However, selecting the correct kernel is crucial because an incorrect one can lead to errors.

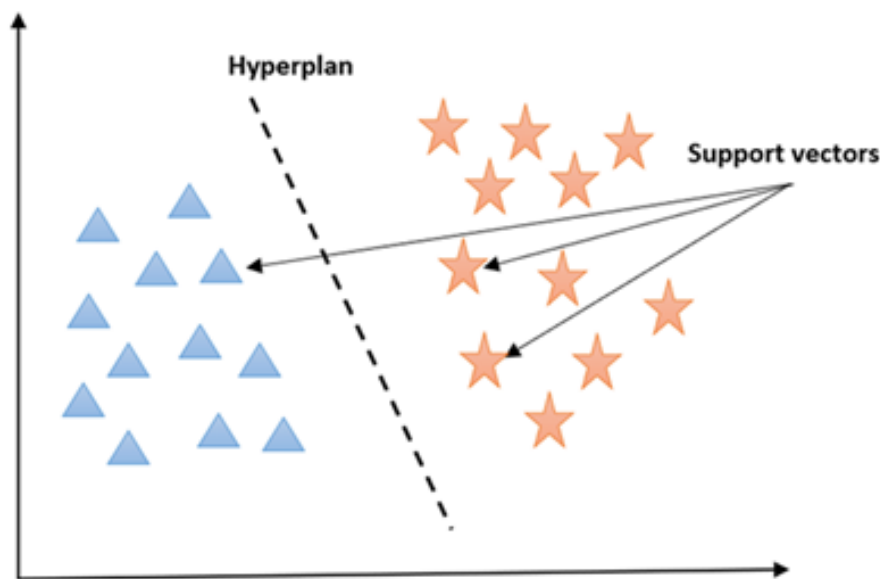


Figure 3.5: Classification using support vector machines [113]

3.5.2 KNN

K-Nearest Neighbors (KNN) is a popular supervised machine learning technique that can handle both classification and regression tasks. This non-parametric approach is often called a "lazy learner" because it doesn't create an explicit model during training. Instead, a new data point is assigned a class or value based on its similarity to the k closest samples from the training set [114]. The KNN process involves the following steps:

- * K value selection: Choosing the value of K, which determines how many neighbors will be considered. This is a key factor in the model's accuracy.

- * Distance calculation: Calculating the distance between the new data point and each point in the training set using metrics such as Euclidean, Manhattan, or Minkowski.

- * Identifying K nearest neighbors: Selecting the K training samples with the smallest distance to the new input.

- * Prediction: For classification, the algorithm assigns the most frequent class among the K neighbors. For regression, it computes the average value of the neighbors.

3.5.3 Decision Tree

Decision trees (DT), a popular technique in non-parametric supervised learning, can be applied to both regression and classification problems. Various DT algorithms exist, with the most notable being CART, ID3, and C4.5. A DT consists of several key components: the root node represents the complete dataset; internal decision nodes correspond to attributes used for splitting the data; branches denote the outcomes of those splits; and leaf nodes, which have no further divisions, indicate the final class labels or values[115]. DT operates based on the following steps:

- * Selecting decision attributes: The algorithm identifies the most effective attributes to split the dataset by minimizing impurity, often using Gini index or entropy

- * Splitting the data: New decision nodes are created by splitting the dataset based on the selected attributes.

- * Building the tree: This procedure is carried out recursively on every additional node until a halting condition is satisfied.

- * Assigning class labels: A label is applied to each leaf node. This label indicates the most common class at that node in terms of classification. In regression, it is the mean of the target values.

- * Making predictions: The built tree is then used to forecast results for additional inputs by following the path from root to leaf depending on feature values.

3.5.4 K-means

K-means clustering is one of the most popular unsupervised machine learning methods. It groups data points into clusters according to their similarities. This clustering technique is based on centroids and reduces the distance between each cluster's centroid and its associated data points. The centroid is the average position of the points within a cluster. In this approach, each data point is paired with the cluster whose centroid is closest to it. K-means iteratively refines the cluster assignments until the centroids stabilize. A user-specified number of clusters is represented by "k." [116]. The k-means algorithm follows these steps:

- * Choose k sites at random to serve as the initial centroids.

- * Using a distance metric, assign each data point to the closest centroid.

- * By averaging the locations of every point in each cluster, recalculate the centroids.

- * Continue doing this until the centroids stop changing.

K-means clustering is widely used because it is easy to use and computationally effective, making it appropriate for large datasets. However, it has several shortcomings. It cannot identify clusters with complex shapes, requires the user to predetermine the number of clusters, and produces results that can change based on where the initial centroids

are placed. Additionally, k-means may not work well when clusters contain different sizes or densities.

3.5.5 Random Forest

Random Forest (RF) is an ensemble learning method used for both classification and regression tasks. It constructs multiple decision trees by training each one on a random subset of the data using the bagging approach. The final prediction is made by combining the outputs from all trees in the ensemble. This model typically achieves higher accuracy than using a single decision tree [117]. The steps of the RF algorithm are as follows:

- * Bootstrap sampling: The training set is split into several random subsets, each used to train a different decision tree.

- * Random feature selection: A random subset of features is selected for consideration at each split in the decision trees.

- * Tree construction: Each decision tree is built using its corresponding bootstrap sample and selected features.

- * Final prediction: For classification, the predicted class is the one most commonly chosen by the trees (majority voting). For regression, the prediction is the average of the outputs from all trees. Random Forest (RF) is an ensemble learning method used for both classification and regression tasks. It constructs multiple decision trees by training each one on a random subset of the data using the bagging approach. The final prediction is made by combining the outputs from all trees in the ensemble. This model typically achieves higher accuracy than using a single decision tree.

The steps of the RF algorithm are as follows:

- * Bootstrap sampling: The training set is split into several random subsets, each used to train a different decision tree.

- * Random feature selection: A random subset of features is selected for consideration at each split in the decision trees.

- * Tree construction: Each decision tree is built using its corresponding bootstrap sample and selected features.

- * Final prediction: For classification, the predicted class is the one most commonly chosen by the trees (majority voting). For regression, the prediction is the average of the outputs from all trees.

3.5.6 Naive Bayes

The Bayes theorem forms the basis of the probabilistic supervised learning algorithm Naïve Bayes (NB), which uses the assumption that features are conditionally independent to simplify calculations. Despite this presumption, NB is effective in real-world applications, including spam detection, text classification, and binary and multi-class issues [118].

Main concepts of the NB algorithm:

- Bayes' theorem: Computes the probability of a class A given observed features B using:

$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B)} \quad (3.1)$$

where:

- $P(A | B)$: posterior probability of class A given feature B
- $P(B | A)$: likelihood of B given A
- $P(A)$: prior probability of class A
- $P(B)$: prior probability of feature B

Naïve assumption: Assumes that all features are conditionally independent given the class. For features B_1 to B_n :

$$P(A | B_1, B_2, \dots, B_n) = P(A) \prod_{i=1}^n P(B_i | A) \quad (3.2)$$

Prediction: The class with the highest posterior probability is selected using the Maximum A Posteriori (MAP) decision rule:

$$\text{MAP} = \arg \max_{A_k \in A} P(A_k | x_{\text{new}}) \quad (3.3)$$

3.5.7 Logistic Regression

Logistic Regression (LR) is a probabilistic supervised learning algorithm mainly used for binary and linear classification, despite its name suggesting a regression approach. It performs well when the data classes are linearly separable. LR estimates the probability of a data point belonging to a class using the Sigmoid (logistic) function[119]. Key components of the LR algorithm include:

* Linear combination: A weighted sum of the input features is calculated to form a value z , where

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (3.4)$$

Here, β_0 is the intercept, and $\beta_1, \dots, \beta_n$ are the feature weights.

Logistic function (Sigmoid): This function maps the value of z into a probability between 0 and 1 using the following formula:

$$g(z) = \frac{1}{1 + \exp(-z)} \quad (3.5)$$

Prediction: The model calculates the probability for each input and classifies the instance based on a threshold (commonly 0.5).

3.5.8 Hierarchical Clustering

Hierarchical clustering produces a tree-like structure by periodically combining or dividing data points according to similarity. This technique creates a binary tree, which provides a clear visual representation of the clustering procedure by gradually grouping related points [120]. The process consists of the following:

1. Begin by considering each point as a distinct cluster.
2. Update the proximity matrix by merging the two nearest clusters and determining the shortest distance between them.
3. Continue doing this until a single final cluster is formed from all the points.

This method produces dendrograms, which graphically depict the cluster structure. It is also simple to use. However, even if the merge is poor, the mechanism cannot undo the decision once the points have been aggregated. This method is slower than techniques like k-means and can be difficult to understand when deciding how many clusters to use for large datasets.

3.5.9 GMM

In unsupervised learning, a Gaussian mixture model (GMM) is a probabilistic method of clustering. In this model, every data point is assumed to originate from a composite of many Gaussian (normal) distributions [121]. GMM is based on two main concepts:

- *Gaussian distribution: a continuous probability distribution shaped like a bell curve.

- * Mixture mode: a statistical model that represents data as coming from a mix of several probability distributions.

The GMM uses these principles to model data by estimating the parameters of each Gaussian component and the proportion of each in the dataset. Its mathematical formulation follows from this assumption.

3.6 Principles of Deep Learning

Deep learning models and comprehends intricate patterns in data using multi-layered artificial neural networks, allowing traits to be extracted automatically without human involvement. Deep learning is often used for autonomous systems, natural language processing, and image recognition [122].

3.6.1 ANNs

The fundamental building blocks of deep learning are artificial neural networks (ANNs). ANNs consist of interconnected layers of nodes, commonly referred to as neurons, which process data and identify patterns through experience. ANNs are modeled after the composition and operation of the human brain [123]. These networks can generalize and anticipate new data because they can handle complex, nonlinear data properties. In artificial neural networks (ANNs), learning occurs through interconnections between artificial neurons across multiple layers, similar to the way learning occurs in the human nervous system. In humans, learning involves slight changes in synaptic connections between neurons. ANNs are modeled after biological neural networks [124].

Neurons are the basic building blocks of the human nervous system. These cells are specialized for transmitting brain signals and processing electrical impulses. Each of the billions of neurons in the human brain is connected to approximately 10,000 others. Neurons have an ideal structure for their purpose. The four primary parts of a typical neuron are the axon, dendrites, cell body, and synapses Figure 3.6 provides a depiction of the structure of a biological neuron[125].

3.6.2 Artificial Neuron

A synthetic neuron functions similarly to a biological neuron. After processing weighted inputs with an activation function, it generates an output. Since neurons are organized into layers, the network can be trained to recognize patterns and convert inputs into

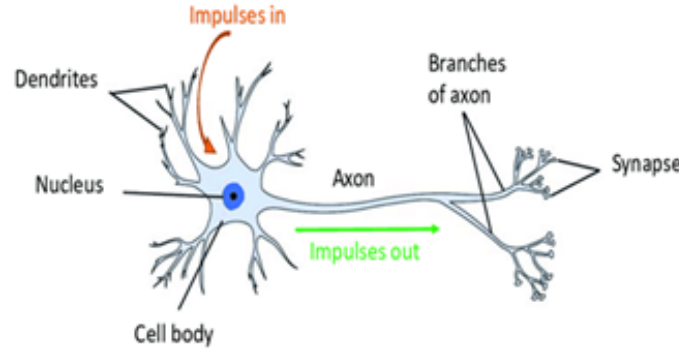


Figure 3.6: The Biological Neuron[125]

outputs. McCulloch and Pitts presented the first formal neuron model in the 1940s, marking the beginning of neuron modeling [126]. This mathematical model triggers a response when the weighted total of the input signals exceeds a predetermined threshold. Figure 3.7 depicts the series of actions performed by this neuron. However, until 1949 when Hebb proposed a learning concept without offering specific equations, this model lacked a learning mechanism.

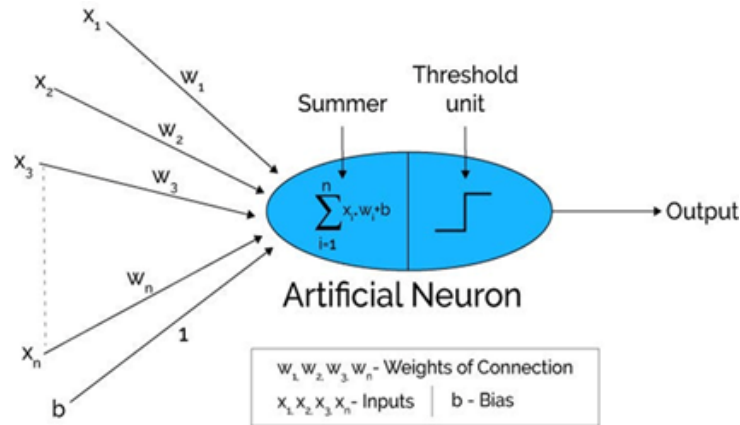


Figure 3.7: The formal neuron [127]

The output value (Z) is derived from the sum of the inputs (X_n), each multiplied by their respective weights (W_n), followed by applying an activation function (F) to this weighted sum. This process can be mathematically expressed as:

$$Z = F(S) = F\left(\sum_{n=1}^N W_n X_n + b\right) \quad (3.6)$$

Although this simplified model cannot fully capture the intricacy of biological neurons, it makes it possible to investigate the connections between neurons in more intricate networks. Inspired by biological processes, the Hebbian learning rule states that when two neurons fire simultaneously, their connection strengthens. Since weights between neurons are modified by repeated exposure to samples, this process can be considered either supervised or unsupervised learning. According to the rule first proposed by Hebb in 1949, when neuron A frequently activates neuron B, one or both neurons change in a way

that strengthens the connection. The simultaneous firing of presynaptic and postsynaptic neurons determines the change in connection strength, or weight, as shown in Table 3.1:

Table 3.1: Hebb's Rule

A_i	A_j	ΔW_{ij}
0	0	–
0	1	0
1	0	0
1	1	+

The weight change that was carried out is represented by ΔW_{ij} . The following equation can be used to model this rule in the context of formal neurons:

$$W_{ij}(t + \Delta t) = W_{ij}(t) + \mu \Delta W_{ij}(t) \quad (3.7)$$

With $W_{ij}(t)$ representing the coactivity between two neurons, $W_{ij}(t + \Delta t)$ denotes the updated weight, $W_{ij}(t)$ is the previous weight, and μ is the learning rate.

3.6.3 ANN Structure

An artificial neural network (ANN) consists of three primary layer types: the input layer, one or more hidden layers, and the output layer. These layers are arranged sequentially. In a feedforward neural network, also referred to as a multilayer perceptron (MLP), data moves from input to output in a single direction [128].

- **Input Layer** This is the first layer of the network. Each node represents a feature or variable from the input dataset. The number of features equals the number of input nodes. Each node sends its value to the following layer without any processing.

- **Hidden Layer(s)** One or more hidden layers are located between the input and output layers. Each hidden layer neuron receives input from every neuron in the layer above and applies weights to the results. Then, it runs the results through an activation function, such as sigmoid or ReLU. These layers identify and extract patterns from the input data. The network's ability to learn and generalize depends heavily on the number of hidden layers and neurons per layer.

- **Output Layer** This is the network's final layer. It generates the expected output using the patterns that the hidden layers have learned. The type of task determines the number of neurons in this layer. In multi-class classification, for instance, each output neuron produces a probability score corresponding to a class.

Figure 3.8 show the structure of ANN

Activation Function

A neural network's activation function uses the weighted sum of its inputs to determine whether to activate a neuron. Adding non-linearity to the model allows the network to recognize intricate patterns [130]. Common activation functions include:

Sigmoid

The equation of the sigmoid function is:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.8)$$

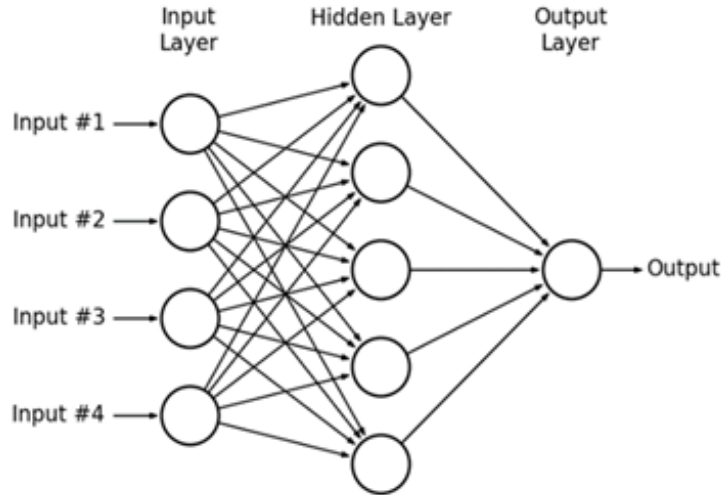


Figure 3.8: ANN structure [129]

Figure 3.9 illustrates the activation function of the sigmoid, which exhibits non-linear characteristics and generates continuous outputs. Unlike the threshold function that produces discrete outputs in the range $[0, 1]$, the sigmoid function can map inputs into an analog form. However, it suffers from the vanishing gradient problem, where large inputs are compressed between 0 and 1, significantly reducing their derivatives and leading to poor results. To address this, activation functions like ReLU are used, which avoid the issue of small derivatives [131].

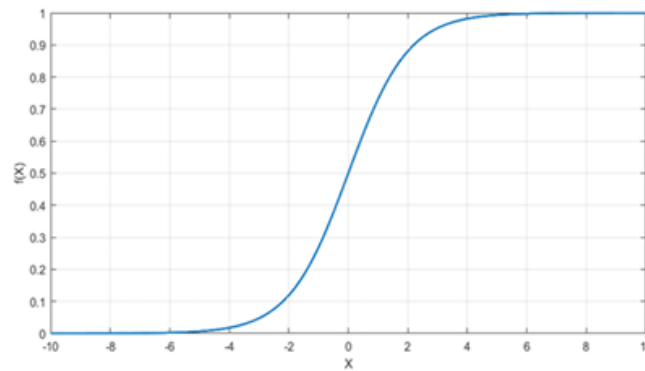


Figure 3.9: Sigmoid activation function [131]

ReLU

Currently, the most popular non-linear activation function is the Rectified Linear Unit (ReLU) [132], ranging from 0 to infinity. It converts all negative values to zero. Mathematically, it can be expressed as:

$$F(x) = \max(0, x), \quad F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases} \quad (3.9)$$

Tanh activation function

The Hyperbolic Tangent (Tanh) function is another activation function that provides advantages comparable to those of the sigmoid function [133]. Its key feature is that the output range spans from -1 to 1. As shown in Figure 3.10, the Tanh function is both continuous and differentiable. The definition of the output is as follows:

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (3.10)$$

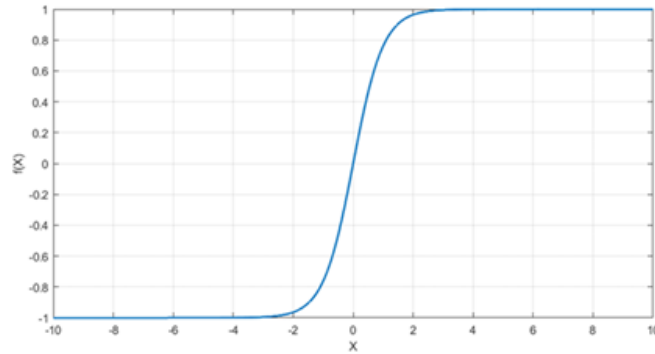


Figure 3.10: Tanh activation function [133]

Softmax

Essentially, the Softmax function combines multiple sigmoid functions. The Softmax function is intended for multiclass classification problems, while the sigmoid function produces values ranging from 0 to 1 that reflect probabilities in binary classifications. It provides a probability for every data point across all classes. The output layer of a network designed for multiclass classification contains an equal number of neurons as there are target classes [134].

$$\text{Softmax}(z_i e^x) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (3.11)$$

Where:

- z_i is the input to the Softmax function (the raw output from the network for class i),
- e^{z_i} is the exponential of the input,
- The denominator is the sum of the exponentials of all inputs, ensuring that the output values sum to 1.

3.7 Training of ANN

In order to minimize the loss (or error) function and improve task performance, the weights and biases of a neural network must be adjusted during training. This procedure uses input-output data pairs to direct changes. However, it is difficult to determine the precise minimum of a loss function since it is often complex and nonlinear. Instead, the

model's parameters are updated based on the function's gradient using gradient descent, a popular optimization technique that reduces error iteratively [135].

3.7.1 Gradient Descent and Backpropagation

To minimize prediction error when training an artificial neural network, its internal parameters, such as weights and biases, must be improved. Two important methods in this process are backpropagation and gradient descent [136].

Gradient Descent

Gradient descent is one iterative optimization technique. It determines the gradient of the error function in relation to the model's parameters. To minimize the error, it updates the parameters in the opposite direction of the gradient. The procedure continues until the error function converges or hits a minimal threshold. Before the iterations begin, the network's parameters are usually initialized randomly [137].

Backpropagation

The conventional technique for applying gradient descent to neural networks is backpropagation [138]. Doing the following allows the network to modify its internal weights and biases:

Forward Pass: Input data travels throughout the network from the input layer to the output layer. Weights, biases, and activation functions are applied at each layer to transform the input data, resulting in an output.

Error Calculation: After generating the output, we calculate the difference between the actual target (ground truth) and the expected output. This discrepancy is referred to as the "mistake."

Backward Pass: The error spreads backward through the network. This method determines the contribution of each bias and weight to the inaccuracy of each layer.

Weight Update: Gradient descent is used to update weights and biases in order to lower the total error.

Iteration: This cycle forward pass, error calculation, backward pass, and weight update is repeated over many iterations. With each iteration, the network improves its performance by learning from the data.

3.8 CNNs

The Convolutional Neural Network (CNN) model is one type of deep learning model that works especially well for processing grid-like data, such as photographs. The ability of CNNs to automatically extract spatial and hierarchical characteristics from unprocessed input has made them standard in computer vision applications, including classification, segmentation, and detection. CNNs require less preprocessing than conventional classification methods. During training, CNNs automatically extract valuable characteristics from data, eliminating the need for manually created filters. Each layer of a typical CNN uses a differentiable function to convert input activations into output activations [139], as shown in Figure 3.11. The core elements of a CNN include:

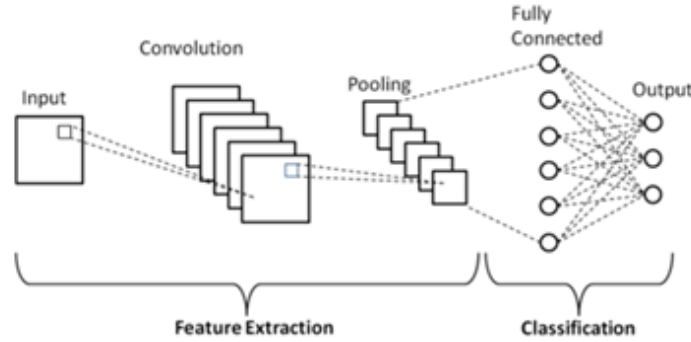


Figure 3.11: CNN architecture [140]

3.8.1 Convolutional Layers

A convolutional layer is the central component of a CNN. This layer uses a collection of learnable filters, or kernels, which are smaller than the original image. Through training, these filters acquire the ability to recognize particular features. Each filter creates an activation map by moving geographically over the image and performing element-wise multiplication (dot product) between the filter and the appropriate input region. As shown in Figure 3.12, An activation map is created by applying a filter to a section of an image, such as the red-highlighted area, and repeating this process throughout the image. The output of the convolutional layer consists of stacked activation maps from various filters. Each neuron in an activation map has the same filter weights and represents a small part of the input. Thanks to its local connectedness, the network can identify patterns unique to local areas. While later layers identify complex features, such as structures or objects, early levels identify simple patterns, such as edges [141].

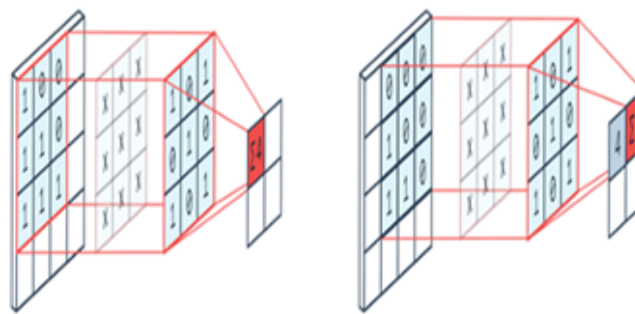


Figure 3.12: An illustrated representation of the convolution operation [141]

Key elements to consider when constructing a convolutional layer include:

Kernel (Filter): A kernel is a compact matrix that scans the input (such as an image) to detect specific patterns. At each position, it multiplies the corresponding input region element-wise and sums the result. These filters are trainable and designed to recognize features like edges or textures. **Feature Map:** This is the output generated by applying a kernel to the input. Each kernel produces a distinct feature map that highlights certain

patterns or details in the original data.

Stride: Stride defines the number of steps the kernel moves across the input in both horizontal and vertical directions. A small stride retains more detail, while a larger stride reduces the size of the resulting feature map. Common values are 1, 2, or 3.

Padding: Padding adds extra borders around the input before convolution begins, helping to preserve the spatial dimensions. Common options are “valid” (no padding) and “same” (padding added to keep output size equal to input). The size of the output volume can be calculated using:

$$W_{\text{out}} = \frac{W - K + 2P}{S} + 1 \quad (3.12)$$

Where:

- W = input width or height,
- K = kernel size,
- S = stride,
- P = padding,
- D_n = number of filters.

This results in an output dimension of $W_{\text{out}} \times W_{\text{out}} \times D_n$.

3.8.2 Pooling Layers

Another important CNN process is pooling, which non-linearly reduces the spatial dimensions of feature maps. Although there are several non-linear techniques, max pooling is the most popular. In this method, the input is divided into smaller rectangular portions, and the highest value is chosen from each piece. This operation is applied distinctly to every depth channel and spatial location. A standard pooling layer reduces the spatial dimension by 75% using a 2x2 window with a stride of 2 in both height and width. It selects the highest value from a set of four in each phase. The depth remains constant [142]. For instance, applying this on a 4x4 input using a 2x2 filter results in a 2x2 output, as illustrated in Figure 3.13

3.8.3 Fully Connected Layer

In neural networks, fully connected layers, also known as dense layers, are typically located at the end. Unlike convolutional and pooling layers, which deal with particular input regions, these layers connect all of the neurons in the current layer to all of the neurons in the preceding layer [144]. here’s how they function:

Neuron response: All of the neurons in the previous layer send messages to each neuron in the dense layer. The dense layer uses an activation function to interpret these signals and generate an output showing its reaction.

Output tasks: These layers are frequently used in tasks like regression or classification because they connect extracted features to predefined output labels or numerical values.

Final result: The output from the final dense layer is often sent through a softmax algorithm for classification purposes. However, the results of regression may not change.

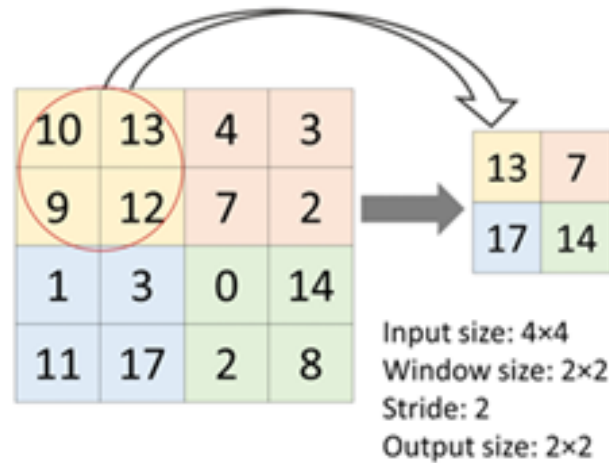


Figure 3.13: Visual representation of the max-pooling operation [143]

3.9 Popular CNN Architectures

Several CNN architectures have become standard in deep learning research due to their performance, scalability, and adaptability to different tasks. These architectures are widely used in medical imaging applications, especially for classification, detection, and segmentation of abnormalities in MRI, CT, and other modalities. Below are some of the most commonly used models:

3.9.1 VGG

The Visual Geometry Group at Oxford University, under the direction of Andrew Zisserman and Karen Simonyan, created the deep convolutional neural network architecture known as VGG (Visual Geometry Group). In their 2014 study, "Very Deep Convolutional Networks for Large-Scale Image Recognition," they presented this architecture, which was essential to the development of deep learning for computer vision problems.

The architecture of VGG

VGG is a deep but simple architecture that focuses on simplicity while increasing depth. It uses small (3x3) convolution filters throughout the network, allowing deeper networks to learn complex representations from images[145]. Similar to early convolutional networks, the architecture consists of a stack of convolutional layers followed by fully connected layers. However, it is much deeper. The network reduces spatial dimensions while preserving crucial features for subsequent layers by employing a max pooling layer after several convolutional layers. Similar to a traditional multi-layer perceptron, the output is flattened and sent through fully connected layers. A softmax activation function is used in the final layer for classification tasks.

VGG introduced a variety of models according to the number of layers: VGG11 has eight convolutional layers and three fully connected layers; VGG13 has ten convolutional layers and three fully connected layers; VGG16 has thirteen convolutional layers and three fully connected layers; and VGG19 has sixteen and nineteen weight layers. These models were commended for their ability to perform challenging classification tasks. The

figure 3.14 represent vgg 16 architecture.

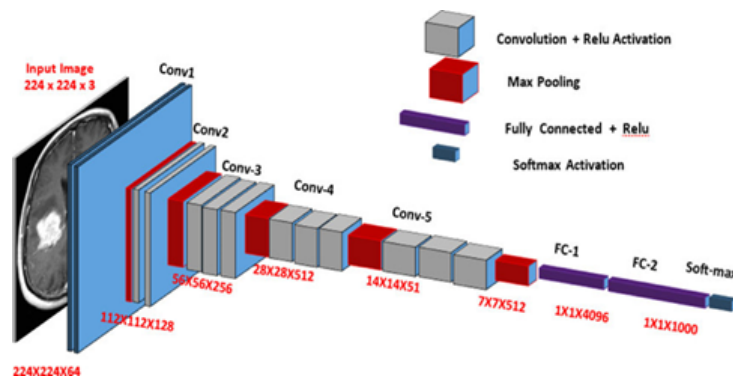


Figure 3.14: Vgg16 architecture [146]

VGG16 consists of three fully connected layers with 4,096 and 1,000 neurons, five max pooling layers, 13 convolutional layers, and a softmax layer for class probabilities. The input image usually has a pixel size of 224 x 224. The output is classified using a softmax layer.

VGG's architecture is uniform, using 3x3 filters across all layers, simplifying design and expanding network depth. Its deep representations learn detailed, hierarchical data, making it powerful for image classification. VGG models are widely used for transfer learning, fine-tuning pre-trained weights for various computer vision problems.

3.9.2 EfficientNet

EfficientNet is a series of convolutional neural network (CNN) designs that seeks to combine computational efficiency with excellent accuracy. Created by Mingxing Tan and Quoc V. Le, EfficientNet introduced a novel scaling technique called compound scaling in 2019. This technique balances the trade-offs between model depth, width, and resolution. Compared to earlier designs, such as ResNet and VGG, EfficientNet is renowned for its ability to achieve state-of-the-art accuracy in image classification with fewer parameters and lower computational cost.

EfficientNet balances the three main parameters—depth (number of layers), width (number of channels), and resolution (input image size)—using a compound scaling technique. MobileNetV2's inverted bottleneck residual blocks (MBConv), which reduce computing costs, form the basis of the architecture. The compound scaling technique ensures that these parameters are selected to optimize efficiency within a specified computational budget by equally adjusting depth, width, and resolution across several model versions (EfficientNet-B0 to B7) [147].

EfficientNet comes in multiple versions (B0 to B7), with B0 being the baseline model and B7 being the largest, most powerful model. The EfficientNet-B0 architecture serves as the base upon which other versions (B1 to B7) are scaled up. The compound scaling method allows for different models to operate on higher resolution input images. For example, EfficientNet-B1 uses an input resolution of 240x240, 18 layers, and slightly more filters per layer, while EfficientNet-B7 uses a 600x600 input resolution, with significantly more layers and filters.

The figure 3.15 represent the architecture of EfficientNetB0, it uses a compound scaling method, which uniformly scales depth, width, and resolution across the network, as

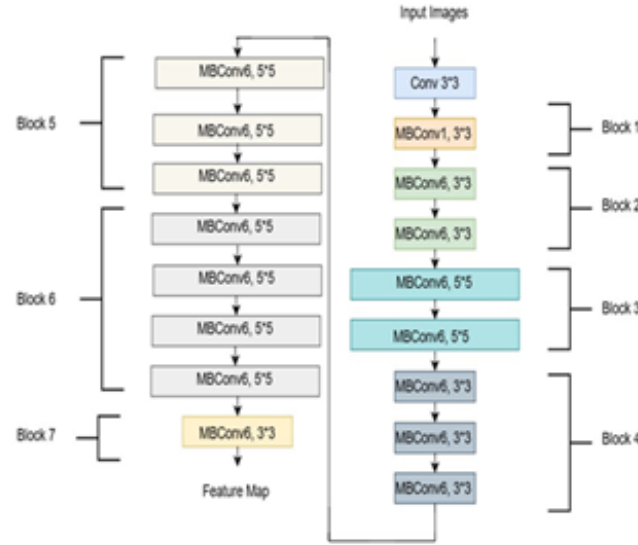


Figure 3.15: Architecture of EfficientNetB0 [148]

illustrated in the diagram where each block (MBConv) has specific kernel sizes and expansion factors. EfficientNetB0 is known for its ability to achieve high accuracy while keeping computational costs low.

3.9.3 U-Net

The U-Net, a convolutional neural network (CNN) architecture, was created specifically for biomedical image segmentation. It was pioneered by Olaf Ronneberger [149]. Due to its effectiveness in generating accurate, pixel-level predictions, the model developed by Philipp Fischer and Thomas Brox in 2015 has grown to become one of the most popular models for segmentation tasks. U-Net's main characteristic is its U-shaped architecture, which uses symmetric encoder-decoder paths to provide high-resolution localization and context.

The U-Net architecture is a deep learning model with a U-shaped design that gradually reduces spatial resolution to capture context. The U shape comes from its two primary components: the encoder (the contracting path) and the decoder (the expanding path). To reduce spatial dimensions, the encoder uses convolutional layers, 2x2 max pooling layers, and ReLU activation. As the network becomes deeper, the model can capture more complex characteristics because each downsampling step doubles the number of feature channels.

To restore spatial resolution, the decoder upsamples while mirroring the encoder. Thanks to its up-convolution layers and skip connections, the model can use both low-level and high-level features. These connections concatenate relevant feature maps from the encoder path to the decoder. These connections enhance the segmentation output by preserving spatial information lost during downsampling. Applications of the U-Net architecture include autonomous driving, image-to-image translation, biomedical image segmentation, and agriculture. Among its primary features are skip connections between the encoder and decoder, which maintain spatial details and increase the precision of segmentation borders.

The figure 3.16 represent the architecture of U-Net

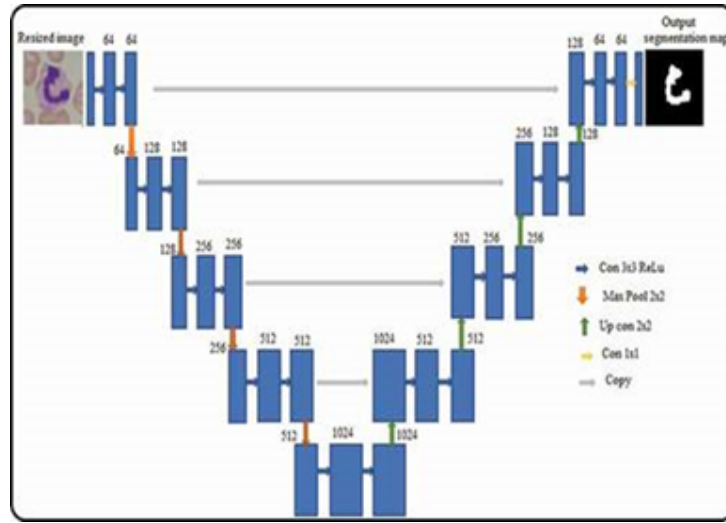


Figure 3.16: U-Net architecture [150]

3.9.4 ResNet

Residual Neural Networks (ResNet), a type of deep learning architecture, were created to address the vanishing or exploding gradient issue that arises in very deep neural networks [151]. Thanks to the introduction of the residual learning idea by ResNet, deeper networks can be taught efficiently. The 2015 proposal of this architecture by Kaiming He et al. transformed deep learning, particularly for image classification problems.

The key concepts of ResNet are:

Residual Functions: Residual learning is based on the hypothesis that learning the "residual" (the difference between the input and the desired output) is easier than learning the entire function. Rather than learning a direct mapping, $H(x)$, the network learns a residual mapping, $F(x) = H(x) - x$, where x is the input and $H(x)$ is the desired output. The function can then be reformulated as $H(x) = F(x) + x$, which simplifies the learning process. This optimizes convergence and improves performance in deep networks.

Identity Shortcut Connections: ResNet uses skip connections, also known as shortcut connections. In this method, a layer's input is directly added to its output. This allows data to move through the network unaltered by intermediate layers. This can be stated mathematically as follows:

$$y = F(x, \{w_i\}) + x \quad (3.13)$$

Here, $F(x, \{w_i\})$ represents the residual mapping to be learned, and the input x is directly added via the shortcut connection. This technique allows the network to learn modifications to the input rather than completely new transformations.

The numbers represent the total number of layers in the network. There are multiple versions of ResNet, including ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. To decrease the number of parameters and computational cost, deeper variants (ResNet-50 and beyond) replace fundamental residual blocks with bottleneck blocks, Figure 3.17 represents the ResNet50 model architecture, highlighting its five main stages

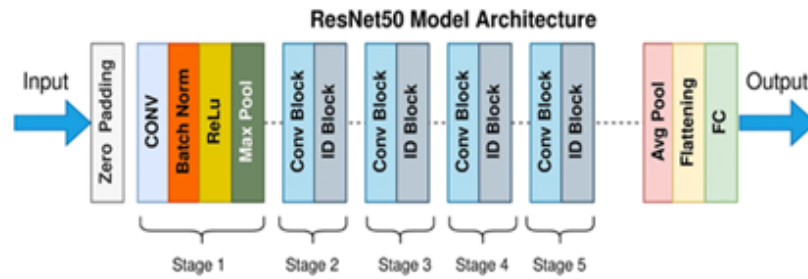


Figure 3.17: ResNet50 model architecture [152]

composed of convolutional and identity blocks. The model begins with preprocessing layers and ends with average pooling, flattening, and a fully connected output layer.

Thanks to residual learning, deep networks like ResNet can achieve state-of-the-art performance on various tasks, such as segmentation, object detection, and image classification. ResNet overcomes the limitations of traditional deep networks, which struggled with convergence issues and performance degradation, by enabling them to learn simpler residual functions. ResNets use residual learning to build networks with hundreds or thousands of layers, improving training time and optimization. They prevent degradation and ensure deeper layers don't deteriorate performance by learning identity mappings when necessary.

3.10 RNN and LSTM

A recurrent neural network (RNN) is a type of deep neural network that is trained on sequential or time-series data. It can be used to develop a machine learning (ML) model that generates sequential predictions or inferences inputs[153], RNNs can be used to solve problems such as sentiment analysis, speech recognition, image captioning, natural language processing (NLP), and language translation, as well as other ordinal or temporal issues.

RNNs are made up of neurons, which are data processing nodes that cooperate to carry out intricate tasks. There are three types of neurons: input, output, and hidden. After the input layer receives the data to be processed, the output layer delivers the outcome. Data processing, analysis, and prediction occur in the hidden layer.

Long Short-Term Memory (LSTM)

The LSTM architecture was introduced to address the issue of long-term memory retention in RNNs. The working mechanism of the LSTM architecture differs from CNNs[154], which handle input data segments independently. While this method works well for uncorrelated data, it falls short when dealing with correlated data, such as time series. Figure 3.18 show working principal of LSTM architecture.

"T-time" refers to the length of the input sequence. The four gates that comprise the LSTM architecture—the input gate, forget gate, update gate, and output gate—operate as artificial neural networks. In other words, t-time denotes the LSTM's time step.

Input gate: It regulates the amount of fresh input that should be kept in the cell state. At each time step, it determines which elements of the input data are significant enough to be added to the cell state.

Forget gate: It decides how much of the prior cell state data should be kept. This

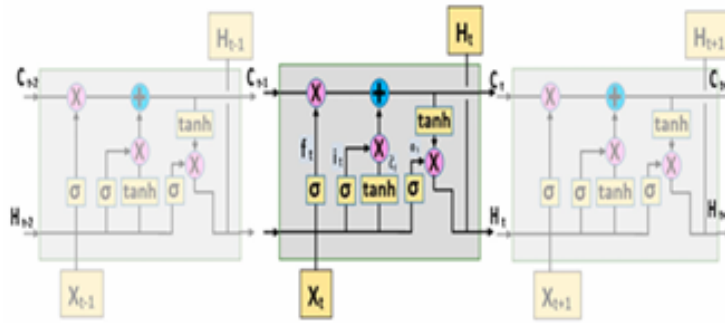


Figure 3.18: Working principle of LSTM architecture [155]

allows the network to eliminate outdated or irrelevant information by determining which aspects of the previous time step should be "forgotten" or deleted.

Update Gate (often referred to as the Cell State Update): It updates the cell state by combining the input and forget gates. To maintain pertinent knowledge, it combines new input data with previously stored data.

Output Gate: It uses the updated cell state and input to determine the appropriate end output for the current time step. This gate determines the information sent to the subsequent time step and, eventually, to the final output layer.

The precise location in the input sequence is referred to as "time step t " or "t-time." Each time step denotes a point in the sequence, such as a value in a time series or a word in a phrase. LSTM processes inputs successively across time. Consequently, the LSTM modifies its internal cell state at each time step by using the gates, the previous cell state, and the current input. Due to their capacity to store and update data over several time steps, LSTMs can learn patterns across lengthy sequences.

3.11 Applications of AI

Artificial intelligence (AI) has revolutionized a number of fields by allowing machines to carry out tasks that normally require human intelligence [101]. Some of these tasks include learning, thinking, problem solving, pattern identification, and decision making. The main areas where AI is being used are as follows:

Healthcare and Medical Diagnosis

- Assists in disease diagnosis using medical imaging (e.g., MRI, CT scans).
- Detects anomalies such as tumors, fractures, or lesions using deep learning.
- AI models predict disease progression and patient outcomes.
- Chatbots and virtual assistants support patient interaction and monitoring.

Signal and Image Processing

- Enhances image quality and removes noise using autoencoders or GANs.

- Extracts meaningful features from biomedical signals (EEG, ECG) for classification.
- Supports segmentation and detection tasks in medical images.

Natural Language Processing (NLP)

- Enables language translation, text summarization, and sentiment analysis.
- Powers intelligent assistants (e.g., Siri, Alexa) and automated chat systems.
- Processes and classifies medical records and literature.

Education and E-Learning

- Adapts content delivery based on learner performance.
- Automates grading and feedback.
- Detects learning difficulties through behavioral data analysis.

Finance and Business

- Automates fraud detection using anomaly detection algorithms.
- Predicts market trends and customer behavior.
- Improves decision-making in investment and risk analysis.

Robotics and Automation

- Powers autonomous robots for manufacturing, logistics, and surgery.
- Integrates vision and motion planning for real-time decision-making.
- Enhances precision in minimally invasive surgeries.

Smart Systems and Internet of Things (IoT)

- Manages data from smart devices for energy optimization, safety, and automation.
- Enables predictive maintenance in smart factories.
- Integrates with smart homes, cities, and healthcare systems.

The different applications of AI are illustrated in [Figure 3.19](#).



Figure 3.19: Applications of Artificial Intelligence [101]

3.12 AI in Medical Signals

Artificial intelligence (AI) has greatly impacted the analysis and interpretation of medical signals, improving patient care and diagnostic accuracy. For example, in electrocardiograms (ECGs), AI algorithms are used for feature extraction and classification to detect arrhythmias and other heart disorders [156], in Electroencephalogram (EEG) AI helps in identifying patterns related to neurological disorders, such as epilepsy, by analyzing complex brain wave data [157], in Electromyogram (EMG) AI techniques are applied to assess muscle health and diagnose neuromuscular diseases [158] the table 3.2 two summarize the application of AI techniques in medical signals.

Table 3.2: AI Techniques and Applications in Medical Signal Processing and Diagnosis

Application	AI Technique Used	Medical Signals
Diagnosis and monitoring [159]	Prediction, classification	ECG, EEG, EMG
Feature extraction [160]	Machine learning, deep learning	High-dimensional physiological data
Medical imaging [161]	Image segmentation, pattern recognition, image classification	MRI, CT, Ultrasound
Pain assessment [162]	Time-frequency analysis, classification	EMG, ECG
Mental health [163]	Cognitive state analysis	EEG, ECG
Explainable AI [99]	Model interpretability	Various biomedical signals

The use of AI methods, particularly Explainable AI (XAI) and deep learning models, has shown great promise in improving the accuracy and clarity of brain tumor classification. However, significant obstacles remain before AI can be used for this purpose, including ethical issues and the need for validation and approval from medical professionals.

3.13 Challenges in AI

A number of significant obstacles limit the reliability and scalability of artificial intelligence (AI) across disciplines [164]. Data-related issues include poor data quality, high labeling costs, and imbalanced datasets, all of which affect model accuracy and fairness. On the technical side, problems like overfitting, lack of interpretability, and high computational requirements hinder model performance and deployment [165]. Ethical issues like prejudice, privacy infringement, and responsibility seriously call into question trust and safety in AI systems. Deployment obstacles include domain adaptation issues, limitations in real-time processing, and performance deterioration over time due to model drift. Addressing these challenges is critical for building robust, ethical, and adaptable AI systems [166].

3.14 Conclusion

In recent years, artificial intelligence (AI) tools have significantly improved medical diagnostics. These technologies have outperformed traditional machine learning techniques in tasks including segmentation, detection, and classification. This chapter outlines the goals, key methods, and applications of AI in the medical field.

The following chapter will discuss the integration of AI methods with the wavelet transform to improve medical signals. Our focus is on using this hybrid technology to classify brain tumors from MRI data. Our approach employs wavelet-based processing alongside the EfficientNetB0 and ResNet50V2 models to extract features, thereby enhancing accuracy, facilitating early detection, and aiding in clinical decision-making.

Chapter 4

Application of Proposed Methods to Brain Tumor MRI Classification

4.1 Introduction

Classifying brain tumors is crucial for medical diagnosis and treatment planning. Magnetic resonance imaging (MRI) is commonly used to identify brain malignancies because it offers fine-grained anatomical and soft tissue contrast [167]. However, due to the varying sizes, forms, and aggressiveness of brain tumors, they are still difficult to classify accurately [168]. Traditional classification methods often struggle with these complexities, leading to the need for more advanced techniques that combine feature extraction and deep learning models. The wavelet transform has proven to be a useful tool for analyzing medical images and signals, as it enables multi-resolution representation of signals. By decomposing MR images into different frequency components, wavelet transform enhances important structural details that aid in classification. However, In particular, Convolutional Neural Networks (CNNs) have demonstrated significant efficacy in automated feature extraction and categorization using deep learning models. One way to improve the classification accuracy of brain tumors is to combine wavelet transform and deep learning.

This chapter presents two proposed models for MRI-based brain tumor classification. The first model integrates Undecimated Wavelet Transform (UWT) with EfficientNet-B0, followed by Principal Component Analysis (PCA) for dimensionality reduction and multi-SVM for classification of MRI scans into four classes: glioma, meningioma, pituitary and no tumor. While the second model is based on Stationary Wavelet Transform (SWT) for noise reduction, ResNet50V2 for deep feature extraction, and Long Short-Term Memory (LSTM) for classification. This model enhances classification accuracy by leveraging the noise-filtering capability of SWT, the deep feature representation of ResNet50V2, and the sequential learning capability of LSTM. These models were tested on MRI Kaggle dataset.

4.2 CAD Methods for Brain Disorder Detection

Computer-aided diagnosis (CAD) systems are crucial in medical imaging because they greatly aid in identifying and examining brain problems [169]. These systems employ a variety of techniques, from sophisticated deep learning models to more traditional machine learning methods, to interpret medical images and support clinical decisions. Their main goals are to boost diagnostic precision, lower the chances of human error, and facilitate the early recognition of brain conditions, leading to better patient care. For brain-related disorders, CAD systems are designed to handle various classification tasks, each addressing specific clinical requirements [170]. These may involve binary classification (normal vs. abnormal scans) or multi-label classification when multiple conditions exist in a single case. Each classification type introduces unique challenges, requiring specific strategies to manage the complexity of brain imaging data.

CAD techniques have evolved over time. Initial systems depended on manually engineered features, where expert knowledge guided the extraction of relevant image attributes. With the rise of deep learning [171], especially convolutional neural networks (CNNs), the focus has shifted to automated feature learning, resulting in higher performance—especially when handling large-scale datasets. Whether based on traditional methods or deep learning, the effectiveness of CAD systems depends on the chosen approach and its adaptability to medical imaging tasks. While CAD systems using artificial intelligence, especially deep learning, have achieved notable success in brain disease classification, they still face several limitations [172]. Deep learning models require large amounts of labeled

data for effective training, which is often difficult to obtain in the medical field. These models can also act as “black boxes,” offering limited interpretability of how decisions are made—an issue that poses challenges in clinical settings where transparency is essential. Moreover, noise, intensity variations, and irrelevant background information in MRI images can reduce the accuracy and reliability of AI models[173].

To address these limitations, integrating AI with signal processing techniques like the wavelet transform has shown promising results. The wavelet transform excels at decomposing images into multi-scale and multi-resolution representations, capturing both spatial and frequency information. This preprocessing step enhances important features and reduces noise, making the input more meaningful for AI-based models. By combining wavelet transform with deep learning, classification systems can achieve better performance, robustness, and interpretability. The next section presents this hybrid approach, focusing on the use of integrated wavelet-AI techniques for brain tumor classification.

4.3 Wavelet and AI for Brain Tumor classification

The wavelet transform captures both spatial and frequency information. Tumor identification relies on minor image features, such as edges, textures, and localized anomalies, all of which are highlighted by this dual depiction. These aspects are frequently overlooked when raw photos are used directly as input for AI models.

When paired with artificial intelligence (AI) tools such as convolutional neural networks (CNNs), support vector machines (SVMs), or multi-layer perceptrons (MLPs), wavelet-transformed images function as an enhanced input source. They supply cleaner, more organized data that highlights tumor-relevant traits, thereby improving model learning. Studies have shown that, for small and imbalanced datasets, integrating wavelet-transformed images improves generalization, reduces false positives, and increases classification accuracy. The classification process becomes more reliable and interpretable when the wavelet transform is used as an input phase because it ensures that the AI models focus on the most informative areas of the image. For this reason, the wavelet-AI combination is a useful and efficient method for diagnosing brain tumors in real-world applications.

4.4 First Proposed Model

4.4.1 Main Idea and Motivation

Traditional brain tumor classification methods struggle with accurately distinguishing between different tumor types due to challenges in feature extraction, computational efficiency, and robustness. Conventional deep learning models, such as CNNs, require extensive labeled datasets and often fail to generalize well due to overfitting. Additionally, handcrafted feature extraction methods lack adaptability to complex medical imaging variations. To address these limitations, we propose a hybrid approach that integrates the Undecimated Wavelet Transform (UWT) and EfficientNet-B0 for feature extraction. UWT captures both spatial and frequency information without down sampling, preserving critical tumor details. EfficientNet-B0, a lightweight deep learning model, extracts high-level features from the original and wavelet-transformed images. Principal Component Analysis (PCA) is applied to reduce dimensionality while maintaining essential

information, followed by a multi-SVM classifier to ensure precise categorization of brain tumors into four classes: glioma, meningioma, pituitary tumor, and no tumor.

4.4.2 Principles and Objectives

The first proposed model is designed to enhance the classification of brain tumors by leveraging the advantages of wavelet transform and deep learning techniques. The principles guiding this model are:

- **Multi-Resolution Analysis:** The Undecimated Wavelet Transform (UWT) allows the decomposition of MRI images into multiple frequency bands without down-sampling. This ensures that important spatial and frequency-domain information is retained, which is crucial for identifying tumor characteristics.
- **Robust Feature Extraction:** EfficientNet-B0, a deep learning model optimized for computational efficiency and accuracy, extracts high-level spatial features from the original and wavelet-transformed images, enabling improved tumor classification.
- **Dimensionality Reduction:** Principal Component Analysis (PCA) is applied to eliminate redundant information while preserving critical features. This step reduces computational complexity and prevents overfitting.
- **Effective Classification:** A multi-class Support Vector Machine (SVM) classifier is used for final tumor classification. Multi-SVM is known for its strong generalization ability, making it suitable for medical image classification tasks where inter-class variations are high.

Objectives

The key objectives of this model are:

- **Improve Classification Accuracy:** By integrating UWT with deep learning, the model aims to enhance feature representation, leading to higher classification performance.
- **Preserve Structural Details:** The use of UWT ensures that no spatial information is lost during decomposition, which is essential for precise tumor classification.
- **Reduce Computational Overhead:** PCA helps in reducing the feature set without compromising important details, making the classification process more efficient.
- **Enhance Generalization:** Multi-SVM provides a robust classification framework that can generalize well across different MRI scans, reducing overfitting issues common in deep learning models.
- **Validate on Real-World Data:** The model is tested on the Kaggle MRI dataset, ensuring its applicability in practical medical scenarios.

The proposed model aims to provide a reliable, efficient and accurate method for brain tumour classification by adhering to these principles and objectives, thereby addressing the limitations of traditional and standalone deep learning approaches.

4.4.3 Architecture

The first proposed system follows these steps as provided in Figure 4.1[174]:

- **Dataset Splitting:** A stratified split is applied, allocating 80% of the dataset for training and 20% for testing.
- **Wavelet Transform:** The Undecimated Wavelet Transform (UWT) is used to extract both frequency and time-domain features. This step produces four sets of coefficients:
 - Approximation coefficients (cA)
 - Horizontal detail coefficients (cH)
 - Vertical detail coefficients (cV)
 - Diagonal detail coefficients (cD)
- **Feature Extraction:** EfficientNet-B0 extracts features from both the original MRI images and their corresponding UWT coefficients (cA, cH, cV, cD), generating 5000 features per image.
- **Feature Selection:** Principal Component Analysis (PCA) is applied to select the most relevant features, reducing dimensionality while preserving essential information.
- **Classification:** The selected features are used as inputs for a multi-SVM classifier to categorize MRI images into four classes:
 - Glioma
 - Meningioma
 - No Tumor
 - Pituitary

This approach enhances classification performance by leveraging both spatial and frequency-domain features.

4.4.4 Database Description

The dataset used in this study was compiled by Masoud Nickparvar [175] and combines three sources: fgshare [176], the SARTAJ dataset [177], and another unspecified dataset. It includes 5712 MRI images classified into four categories: Meningioma (1339 images), Glioma (1321 images), No Tumor (1595 images), and Pituitary Tumor (1457 images), as shown in Table 4.1. Each category includes images from different MRI views—axial, coronal, and sagittal which provide a detailed understanding of the structure and characteristics of the tumors.

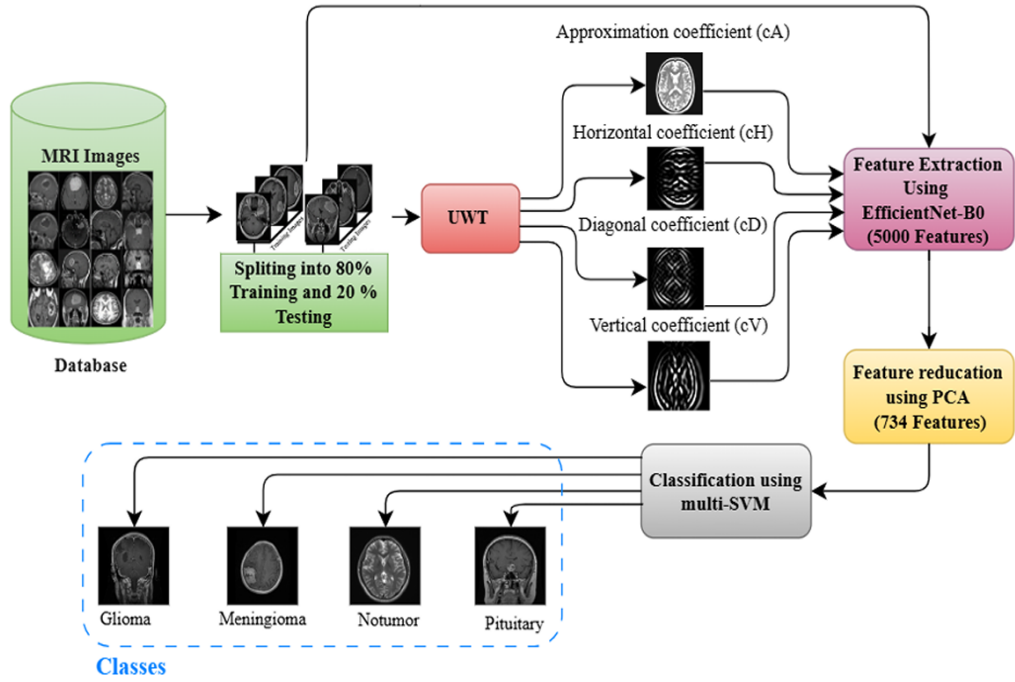


Figure 4.1: Workflow of the first proposed system [174]

Table 4.1: Dataset Distribution by Tumor Class

Class	Number of Images
Meningioma	1339
Glioma	1321
No Tumor	1595
Pituitary	1457
Total	5712

4.4.5 Experiment Setup

The following hardware and software configurations were used for the experiments:

- **Processor:** Intel Core i7-8650U, 4 cores / 8 threads, base 1.9 GHz, turbo up to 4.2 GHz.
- **Memory:** 8 GB SSD RAM.
- **GPU:** Integrated Intel UHD Graphics 620.
- **Operating System:** Windows 11.
- **Software:** MATLAB R2023a.

4.4.6 Performance Measures

It is crucial to evaluate the classifier after training to confirm its applicability, assess its accuracy, and determine its efficiency. There are a variety of ways to evaluate a classifier.

Accuracy

The simplest indicator for assessing classification is accuracy, which shows the percentage of correctly classified cases relative to the total number of cases. However, when working with unbalanced datasets, accuracy can be misleading despite providing a general sense of the model's performance.

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} \quad (4.1)$$

Where:

- TN = True Negatives
- TP = True Positives
- FN = False Negatives
- FP = False Positives

True negatives (TN) denote accurate negative forecasts, while true positives (TP) denote accurate positive predictions. False negatives (FN) are inaccurate negative forecasts, and false positives (FP) are inaccurate positive predictions. These metrics are used to evaluate the efficacy of categorization models.

Precision

Precision, also known as positive predictive value (PPV), is the percentage of correct positive predictions made by a model. Since it emphasizes the importance of precise positive predictions, precision is particularly important when false positives carry significant costs.

$$\text{Precision} = \frac{TP}{FP + TP} \quad (4.2)$$

Recall

Recall, also known as sensitivity, is the percentage of true positive cases that are correctly identified as such. In medical diagnostics, where a false negative could have serious consequences if a positive case goes undetected, this statistic is crucial.

$$\text{Recall} = \frac{TP}{FN + TP} \quad (4.3)$$

F1 Score

The F1 score is a single statistic that combines precision and recall by taking their harmonic mean. Since it considers both precision and recall, it is useful for working with imbalanced datasets.

$$\text{F1 Score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (4.4)$$

Specificity

The percentage of true negatives among all negative instances is known as specificity or the True Negative Rate (TNR). This statistic is essential when reducing the expense of false positives is a top concern.

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4.5)$$

Confusion Matrix

The confusion matrix provides a comprehensive picture of the model's performance by displaying the counts of true positives, true negatives, false positives, and false negatives. Because it breaks down the model's predictions and identifies misclassifications, the matrix is an invaluable error detection tool.

The figure 4.2 depicts the fundamental structure of the confusion matrix:

		True Class	
		Positive	Negative
Predicated class	Positive	TP	FP
	Negative	FN	TN

Figure 4.2: The Basic Structure of a Confusion Matrix

4.4.7 Results

The suggested approach begins with the classification of a dataset containing 5,712 MRI images into four groups: pituitary, meningioma, glioma, and no tumor. 80% of the dataset is used for training and 20% for testing. A combination of techniques, including UWT, EfficientNetB0, PCA, and multi-SVM, is applied. The coefficients are extracted using the UWT from both the training and testing data which, along with the original images, serve as inputs for EfficientNetB0. This CNN model is chosen for its strong feature extraction capabilities in computer vision tasks. PCA selects the most relevant features, ensuring optimal performance. Multi-SVM is utilized for classification because it is so good at handling multi-class problems. Performance is assessed using F1-score, recall, accuracy, and precision.

The main objective of this project is to create and evaluate a high-performance brain tumor classification model that uses cutting-edge methodologies to obtain accurate and

dependable tumor detection from MRI images. Table 4.2 and Figure 4.3 show the performance metrics of the suggested model.

Table 4.2: Performance metrics for the first proposed method

Accuracy	F1-score	Specificity	Precision	Recall
0.9750	0.9748	0.9915	0.9755	0.9746

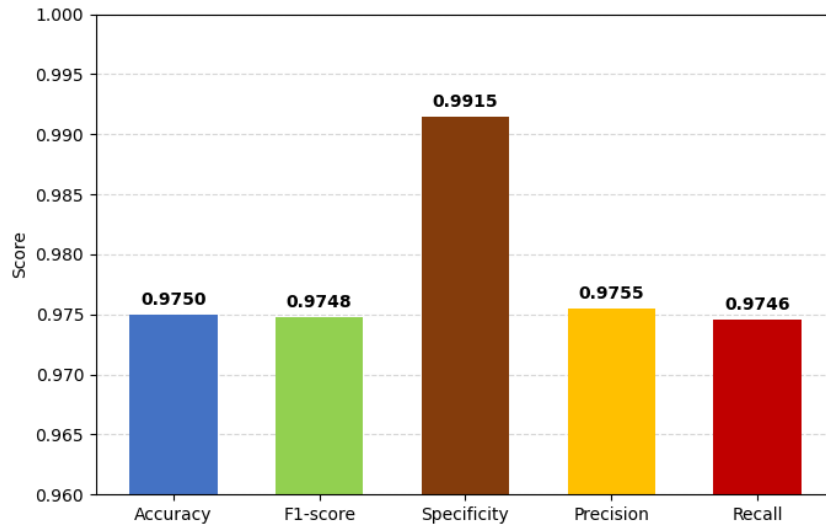


Figure 4.3: Performance metrics of the first proposed model [174]

Table 4.2 and Figure 4.3 present the results of the first proposed method, which achieved an accuracy of 97.50

The obtained results demonstrate the effectiveness of the proposed feature extraction strategy for accurate brain tumor classification. The integration of UWT, EfficientNet-B0, PCA, and multi-SVM enables the model to exploit complementary information from the original MRI images and their wavelet coefficients while maintaining a compact feature representation. EfficientNet-B0 provides discriminative deep features, whereas PCA reduces the dimensionality of the extracted feature space before classification. The combination of these components contributes to the strong classification performance achieved by the proposed method, as reflected by its high accuracy, precision, recall, F1-score, and specificity.

Key implementation challenges, such as data privacy, integration with medical workflows, and model interpretability, are addressed through local deployment, user-friendly interfaces, and clear selection criteria. Continuous learning mechanisms ensure the model remains relevant with advancements in medical knowledge and practice.

4.4.8 Confusion Matrix

A confusion matrix was created to evaluate the performance of the suggested brain tumor categorization model, as shown in Figures 4.4 and 4.5. These matrices provide an in-depth breakdown of the model's classification accuracy for each of the four tumor types: glioma, meningioma, pituitary, and no tumor.

Figure 4.5 presents the confusion matrix with absolute values, showing that the model correctly classified 240 out of 250 glioma cases (96%), 278 out of 284 meningioma cases

True Class	glioma	240	10		
	meningioma	1	278	5	
	notumor		3	311	
	pituitary	2	8		284
		glioma	meningioma	notumor	pituitary
		Predict class			

Figure 4.4: Confusion matrix for first proposed model [174]

True Class	glioma	0.96	0.04		
	meningioma	0.0035	0.9789	0.0176	
	notumor		0.0096	0.9904	
	pituitary	0.0068	0.0272		0.9659
		glioma	meningioma	notumor	pituitary
		Predict class			

Figure 4.5: Normalized confusion matrix for first proposed model [174]

(97.89%), 311 out of 314 no tumor cases (99.04%), and 284 out of 294 pituitary cases (96.56%). Misclassifications were minimal, with 10 glioma cases misclassified as meningioma, 5 meningioma cases misclassified as pituitary, and a few no tumor and pituitary cases misclassified into other categories. Figure 4 displays the normalized confusion ma-

trix, highlighting the classification performance in relative proportions: 96% for glioma, 97.89% for meningioma, 99.04% for no tumor, and 96.56% for pituitary. The model's ability to accurately distinguish between tumor and non-tumor cases is demonstrated by the low misclassification rates, which highlight the robustness of the proposed approach. The highest accuracy was observed for pituitary and non-tumor cases. These results confirm the efficacy of the classification strategy that combines multi-SVM for classification, PCA for feature selection, EfficientNet-B0 for feature extraction, and UWT. The model's potential for practical medical applications, in which accurate diagnoses are essential for clinical decision-making, is demonstrated by its high level of accuracy.

4.5 Second Proposed Model

4.5.1 Key concept and inspiration

To further improve classification accuracy and robustness, we propose a second model integrating Stationary Wavelet Transform (SWT), ResNet50V2, and Long Short-Term Memory (LSTM) networks. SWT removes noise from MRI images while preserving essential structural details, improving feature clarity before extraction. ResNet50V2 extracts high-level spatial features, capturing complex patterns associated with different tumor types. LSTM processes these extracted features to identify sequential dependencies, enhancing classification by recognizing subtle variations across tumor categories. This combination provides a robust framework for classifying MRI images into glioma, meningioma, pituitary, and no tumor categories, significantly improving classification accuracy compared to conventional methods.

4.5.2 Process

The figure 4.6 illustrates the second proposed method for brain tumor classification using MRI images[178]. The process consists of the following key steps:

MRI Dataset

The input consists of MR images containing different types of brain tumors, including glioma, meningioma, pituitary tumors, and cases with no tumor.

Dataset Description for Second Proposed Method

The dataset used in this research, made available by Masoud Nickparvar[179], It consists of 7,023 brain MRI scans, split into two subsets. The training set contains 5,712 images, and the testing set contains 1,311 images. All images fall into one of four categories: pituitary tumor, meningioma, glioma, or no tumor.

Table 4.3 shows how the MRI dataset is divided into four categories: meningioma, glioma, no tumor, and pituitary tumor. The training set contains 5,712 images: 1,339 meningioma, 1,321 glioma, 1,595 no tumor, and 1,457 pituitary tumor. The testing set has 1,311 images: 306 meningioma, 300 glioma, 405 no tumor, and 300 pituitary tumor. This distribution ensures balanced training and accurate assessment of the classification model across all classes.

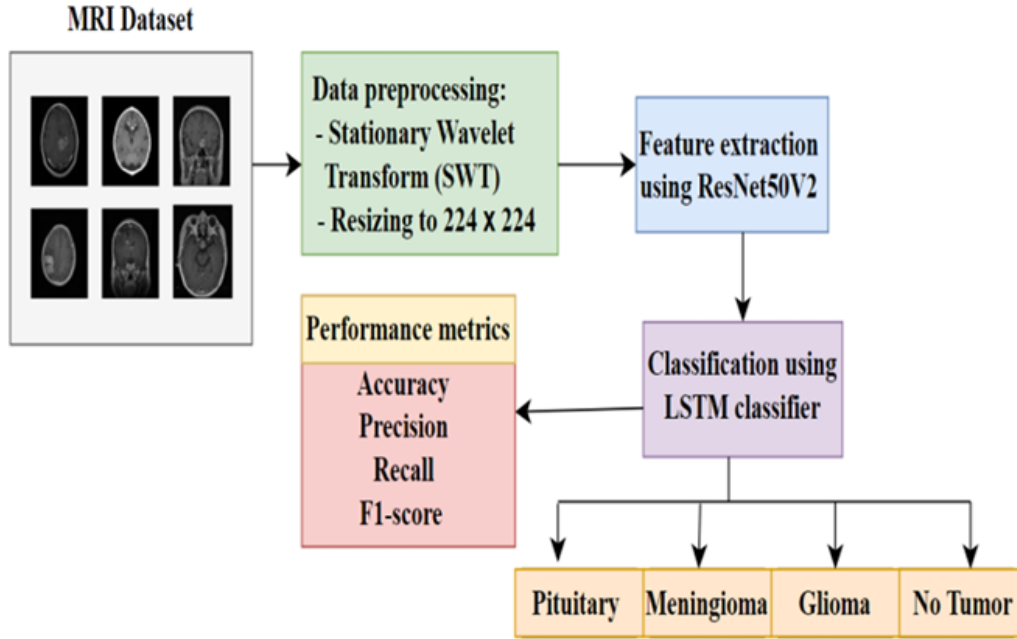


Figure 4.6: Block diagram for second proposed method [178]

Table 4.3: Dataset Composition Used in This Study

Class	Number of Training Images	Number of Testing Images
Meningioma	1339	306
Glioma	1321	300
No tumor	1595	405
Pituitary	1457	300
Total	5712	1311

Data Preprocessing

Stationary Wavelet Transform (SWT): Applied to remove noise while preserving essential structural details in the MRI images.

Resizing: All images are resized to 224×224 pixels to ensure uniform input dimensions for feature extraction.

Feature Extraction using ResNet50V2

ResNet50V2, a deep convolutional neural network, extracts high-level spatial features from the preprocessed images.

These features capture intricate patterns that help differentiate tumor types.

Classification using LSTM

The extracted features are passed through an LSTM classifier, which captures sequential dependencies in the data.

LSTM improves classification performance by identifying temporal relationships in extracted features.

The classification results are assessed using accuracy, precision, recall, and F1-score to measure the model's effectiveness. This method enhances tumor classification by leveraging SWT for noise reduction, ResNet50V2 for spatial feature extraction, and LSTM for temporal pattern recognition, leading to improved accuracy and robustness.

4.5.3 Results

This section presents the findings of the second proposed method for classifying brain tumors. This method combines long short-term memory (LSTM) for classification, stationary wavelet transform (SWT) for preprocessing, and ResNet50V2 for feature extraction. The studies used the Kaggle brain tumor dataset, which comprises 7,023 MRI images divided into four classes: glioma, meningioma, pituitary, and no tumor. The dataset was split into 5712 training images and 1311 testing images. Google Colab was used to implement the suggested strategy and make use of its computing resources for effective training and assessment. In the preprocessing phase, the first step was the Stationary Wavelet Transform (SWT), which divided the pictures into frequency bands to improve tumor-related characteristics and reduce noise. To ensure compatibility with the ResNet50V2 model, the photos were downsized to 224×224 pixels to ensure compatibility with the ResNet50V2 model. ResNet50V2, a deep learning model pre-trained on ImageNet, was then used to extract features from the MRI scans. Next, a Long Short-Term Memory (LSTM) network classified the collected features into four groups: pituitary, meningioma, glioma, and no tumor.

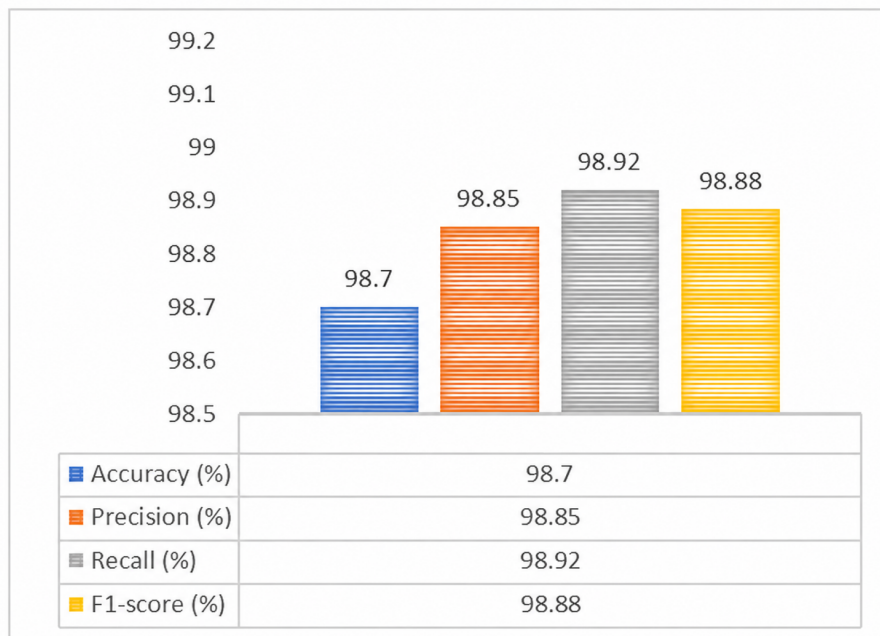


Figure 4.7: Performance metrics for second proposed model [178]

The figure 4.7 present the accuracy, precision, recall, and F1-score of the second proposed classification model. The model achieved an accuracy of 98.70%, with a precision of 98.85%, a recall of 98.92%, and an F1-score of 98.88%. These results demonstrate the model's ability to distinguish between the different types of brain tumors in MRI

images. A recall of 98.92% indicates that the model accurately detects tumor cases, while the precision of 98.85% and the F1-score of 98.88% reflect its strong overall classification performance. By limiting false-positive predictions and correctly identifying tumor cases, the model maintains consistently high values across all evaluation metrics. These results demonstrate the reliability of the proposed model for brain tumor classification.

Confusion Matrix

A confusion matrix is an important method for assessing a classification model's performance. It contrasts the expected and actual labels, offering a detailed view of how the model classifies each category. This matrix helps visualize the classifier's performance and provides insights into the types of errors it makes.

True Class	Glioma	290	8	0	2
	Meningioma	2	301	2	1
	No-tumor	0	0	405	0
	Pituitary	1	1	0	298
		Glioma	Meningioma	No-tumor	Pituitary
		Predict class			

Figure 4.8: Confusion matrix for second proposed method [178]

Figure 4.8 shows the confusion matrix for the brain tumor detection model. This matrix displays the classification results across four categories: glioma, meningioma, no tumor, and pituitary. With 290 glioma cases, 301 meningioma cases, 405 no-tumor cases, and 298 pituitary cases identified correctly, the diagonal elements demonstrate excellent performance. Very few classifications are incorrect. Eight glioma cases were mislabeled as meningioma, and two as pituitary. Two meningioma cases were misdiagnosed as glioma, two as no tumor, and one as pituitary. Each pituitary case was mislabeled as either meningioma or glioma. Notably, all 405 instances in the no-tumor group were correctly identified, achieving perfect classification. The model's ability to differentiate between different types of brain tumors and accurately identify non-tumor cases is demonstrated by its strong diagonal dominance and low misclassification rates.

4.6 Comparison and Analysis

Numerous studies on the classification of brain tumors have used publicly accessible MRI datasets, such as the Figshare and BraTS datasets. Feature extraction and classification techniques employed include convolutional neural networks (CNN), vision transformers (ViT), support vector machines (SVM), multi-layer perceptrons (MLP), and hybrid deep learning techniques. Table 4.4 contrasts the suggested models with these investigations. We note the following:

Table 4.4: Comparison of Brain Tumor Classification Methods

Author and Year	Method	Accuracy (%)	Precision (%)	F1-score (%)	Recall (%)	Sensitivity (%)	Specificity (%)
Kumar et al. [180] 2020	ResNet50 model	96.3	-	-	-	-	-
Hashemzahi et al. [181] 2020	NADE, CNN	95	-	-	-	-	-
Narmatha et al. [182] 2020	Fuzzy Brain-Storm Optimization	93.85	-	-	-	-	-
Kumar et al. [183] 2021	CNN + Residual Network + GAP	97.08	-	-	-	-	-
Nawaz et al. [184] 2022	HBTC + MLP	98.8	95.5	95.12	-	94.79	97.5
Patil et al. [185] 2022	SWT-ELOOP + SVM	-	96.66	96.66	96.67	-	-
Remzan et al. [186] 2023	ResNet50 + VGG19 + EfficientNetV2B1 + MLP	-	96.67	96.66	96.67	-	-
Babu Vimala et al. [187] 2023	Fine-tuned EfficientNet	-	92.11	92.02	95.96	-	-
Zulfiqar et al. [188] 2023	Fine-tuned EfficientNetB2	98.86	98.86	-	98.77	-	98.71
Asiri et al. [189] 2024	Swin Transformer	97	96.14	96.48	96.86	-	98.92
Alshuhail et al. [190] 2024	Modified Neural Networks + Grad-CAM	97	-	-	-	-	-
Amin et al. [191] 2025	ADE_DieT	96.09	96.10	96.09	96.09	-	-
Subba et al. [192] 2025	GoogLeNet + ABIM	97	-	-	-	-	-
First Proposed Model	UWT + EfficientNetB0 + PCA + SVM	97.5	97.75	97.48	97.86	-	99.15
Second Proposed Model	SWT + ResNet50V2 + LSTM	98.7	98.85	98.88	98.85	-	-

Kumar et al [180] proposed an approach that uses a fuzzy deformable fusion model with the Dolphin Echolocation-based Sine Cosine Algorithm (Dolphin-SCA) to preprocess MRI images prior to segmentation. Power Local Directional Pattern (Power LDP) and statistical features such as mean, variance, and skewness are employed for feature extraction. Then, Dolphin-SCA is used as the training technique for a Deep Convolutional Neural Network (Deep CNN) to classify the images. Using the BRATS and SimBRATS datasets, the method achieved a maximum accuracy of 96.3%. To classify brain tumors using MRI scans Hashemzahi et al[181]presented a hybrid model combining a convolutional neural network (CNN) and neural autoregressive distribution estimation (NADE) in their study, "Detection of Brain Tumors Using MRI Scans: A Hybrid Model." The model was trained and evaluated using 3,064 T1-weighted, contrast-enhanced images of three different types of brain tumors. The results demonstrate that the CNN-NADE hybrid excels in classification, particularly when the supply of medical images is limited. Their system achieves 95% accuracy . Narmatha et al[182] presented a fuzzy brain-storm optimization (FBSO) algorithm for brain tumor segmentation and classification by merging fuzzy logic with brain-storm optimization approaches. While the fuzzy component iterates to create an ideal network topology, the brain-storm optimization component focuses on cluster centers, though it may fall into local optima. The BRATS 2018 dataset was used to test the approach, demonstrating its effectiveness and resilience while greatly reducing segmentation time. The proposed FBSO achieves 93.85% accuracy, 94.77% precision, 95.77% sensitivity, and a 95.42% F1 score. To address the vanishing gradient and overfitting issues in deep neural networks for brain tumor classification Kumar et al[183] presented a deep network model that uses ResNet-50 and global average pooling. They used a three-tumor MRI dataset containing 3,064 images to evaluate the model's performance and compare it to other methods. The proposed method outperforms existing classification models, achieving a mean accuracy of 97.08% with data augmentation and 97.48% without. Nawaz et al[184] presented a hybrid brain tumor classification (HBTC) framework for brain tumor classification utilizing machine vision approaches. After preprocessing and segmenting the MRI dataset, the Co-occurrence Matrix (COM), Run-Length Matrix

(RLM), and gradient features are used to extract features to identify the tumor location. After selecting the nine best features, cystic, glioma, meningioma, and metastatic tumors are classified using multilayer perceptron (MLP), J48, meta bagging (MB), and random tree (RT) classifiers. With a maximum classification accuracy of 98.8%, the HBTC framework is shown to be effective, demonstrating its effectiveness and robustness, Patil et al[185] proposed a brain tumor classification method that utilizes enhanced local optimal-oriented pattern descriptors and a steerable wavelet transform for glioma, meningioma, and pituitary cancer classification. Local binary and directional pattern descriptors capture intensity details, while the Frei–Chen-based mask effectively captures edge and line subspace information. After MRI images have been broken down at various sizes, Fisher score-based feature selection is used to extract textural information and minimize dimensionality. Then, categorization is performed using a multi-class support vector machine (SVM). Experiments on a publicly accessible dataset demonstrate superior categorization performance compared to existing texture-based techniques. Remzan et al[186] created a deep learning-based automated method for classifying brain MR images to help physicians diagnose brain cancers. 80% of the 5712 MR images in the dataset are utilized for training, while the remaining 20% are used for validation. Seven pre-trained CNNs (ResNet-50, DenseNet-121, VGG-19, EfficientNetB1, EfficientNetV2B1, Inception V3, and MobileNet) are used in conjunction with Transfer Learning to extract features. A Multilayer Perceptron (MLP) is used for classification after the top-performing features from ResNet-50, VGG-19, and EfficientNetV2B1 are merged using a Features Ensemble technique. This approach outperforms state-of-the-art methods with a 95% confidence level and 96.67% accuracy (±1.04). Babu Vimala et al[187] use MRI scans and a transfer learning-based fine-tuning method with EfficientNets to classify brain tumors using deep learning. To classify glioma, meningioma, and pituitary tumors, five EfficientNet models (B0–B4) are trained and evaluated on the publicly accessible CE-MRI Figshare dataset. Models are initialized with ImageNet weights, top and fully connected layers are added, and the robustness of the models is tested via data augmentation. The model’s emphasis on tumor regions is interpreted using Grad-CAM visualization. According to the results, EfficientNetB2 outperforms other models with the best accuracy (99.06%), precision (98.73%), recall (99.13%), and F1-score (98.79%). Zulfiqar et al[188] investigated the use of MRI images and a transfer learning-based fine-tuning technique with EfficientNets for deep learning in brain tumor classification. Five EfficientNet models (B0–B4) were trained and evaluated for the classification of pituitary tumors, meningiomas, and gliomas using the CE-MRI Figshare dataset. After initializing the models with ImageNet weights, we added top and fully connected layers for classification. Numerous tests were conducted to evaluate the impact of data augmentation and the robustness of the models. Grad-CAM visualization highlights tumor regions in MRI images. EfficientNetB2 achieves the best performance with 98.86% accuracy, 98.65% precision, 98.77% recall, and 98.71% F1-score. The model is computationally efficient, lightweight, and has good generalization. In this study, an enhanced EfficientNetV2 model that incorporates a Global Attention Mechanism (GAM) and Efficient Channel Attention (ECA) improves the categorization of brain tumors in MRI data. These attention processes boost feature extraction, allowing the model to better focus on important tumor locations. Comprehensive testing on a large public dataset shows 99.76% test accuracy, outperforming earlier models. Grad-CAM visualization provides interpretability by shedding light on the model’s decision-making procedure. This study advances Computer-Aided Diagnosis (CADx) systems by increasing classification accuracy, transparency, and clinical application in brain tumor detection. Asiri et al[189]

used the Swin Transformer architecture for brain tumor image classification. This architecture minimizes computing load by using a window-based self-attention mechanism. The approach achieved 97% accuracy. To confirm its efficacy, the method was contrasted with more current models, such as Vision Transformers (ViTs). It also demonstrated good performance metrics, with an F1-score of 96.47%, specificity of 98.92%, accuracy of 96.14%, and recall of 96.86%. Alshuhail et al[190] used a Swin Transformer with a window-based self-attention mechanism for classifying brain tumors in MRI images. It achieved 97% accuracy, with strong metrics across 21 evaluation criteria, outperforming traditional CNN and ViT models. Amin et al[191] presented a multi-class brain tumor classification system using the DieT Transformer for feature extraction, PCA for dimensionality reduction, and the ADE algorithm for feature selection. The ADE-DieT model achieved an accuracy of 96.09% with a precision of 96.10%, a recall of 96.09%, and an F1-score of 96.09%. Subba et al[192] used a CNN based on GoogLeNet with an Attention-Based Inception Module combining spatial and channel attention for better feature representation. It achieved 97.62% overall accuracy.

The first proposed model in this thesis integrates undecimated wavelet transform for feature extraction, EfficientNet-B0 for deep feature extraction, principal component analysis for dimensionality reduction, and multi-SVM for classification. This combination captures both spatial and frequency-domain information, achieving 97.5 % accuracy. Compared to previous models, which primarily relied on deep CNNs or transfer learning alone, this approach enhances classification performance by incorporating handcrafted wavelet-based features and optimizing feature selection. In contrast, prior EfficientNet-based models or hybrid CNN-SVM approaches achieved lower accuracy, typically around 95% to 97%, without leveraging wavelet transform or principal component analysis. The proposed model's robustness, computational efficiency, and superior classification metrics position it as a significant advancement in MRI-based brain tumor classification.

The second proposed model in this thesis integrates stationary wavelet transform for preprocessing, ResNet50V2 for feature extraction, and a Long Short-Term Memory (LSTM) network for classification. This method enhances both spatial and frequency-based tumor features while leveraging deep learning's ability to model complex patterns. Compared to the first proposed model, which used undecimated wavelet transform, EfficientNet-B0, and multi-SVM, this model improves classification by incorporating sequential pattern learning through LSTM. Experimental results on the Kaggle dataset show an accuracy of 98.7 %, surpassing many previous deep learning approaches, including ResNet-50 and hybrid CNN-SVM models. This model offers better feature extraction and classification by addressing spatial dependencies and improving tumor representation.

4.7 Conclusion

In this chapter, we presented two models that use a combination of deep learning and wavelet transform methods to classify brain cancers based on MRI data. The first proposed model integrates Undecimated Wavelet Transform (UWT) for multi-resolution feature extraction, EfficientNet-B0 for deep feature learning, Principal Component Analysis (PCA) for dimensionality reduction, and multi-SVM for classification. This approach effectively captures both spatial and frequency-domain information, enhancing classification performance. The second proposed model employs Stationary Wavelet Transform (SWT) for noise reduction, ResNet50V2 for deep feature extraction, and Long Short-Term Memory (LSTM) for classification. This hybrid model leverages wavelet-based denoising,

deep spatial feature learning, and sequential pattern recognition to improve accuracy and robustness. Accuracy, precision, recall, and F1 score were among the performance indicators used to evaluate the two models on the Kaggle MRI dataset. The experiment's findings showed that the second model, which could incorporate spatial and sequential dependencies, had greater classification accuracy. However, the first model, with its more compact design and good performance, remains competitive. Overall, the results demonstrate the effectiveness of combining wavelet transforms and deep learning for classifying brain tumors based on MRI. With possible advancements in real-time deployment, interpretability, and integration into medical diagnostic systems, these models provide a promising foundation for further research and practical applications.

General Conclusion

This thesis explored the enhancement of medical signals using wavelet transform and artificial intelligence, focusing on MRI-based brain tumor classification. The objective was to develop efficient and accurate classification models by integrating wavelet-based preprocessing with advanced AI techniques. Two models were proposed. The first model uses the undecimated wavelet transform to capture both frequency and time-related details, then applies a pretrained EfficientNet-B0 model for feature extraction from original MRI images and its UWT coefficients. Principal Component Analysis (PCA) is used to select the most relevant features, improving efficiency while preserving important information. Finally, classification is performed using a multi-SVM classifier for accurate results. The second utilized Stationary Wavelet Transform (SWT) for preprocessing, ResNet50V2 for feature extraction, and Long Short-Term Memory (LSTM) for classification. These models aimed to improve classification accuracy while maintaining computational efficiency. The proposed models were tested on a publicly available Kaggle MRI dataset. Performance was evaluated using accuracy, precision, recall, and F1-score, and results were compared with existing state-of-the-art techniques. The findings demonstrated that the wavelet-integrated approaches provided improved classification accuracy, highlighting the effectiveness of combining frequency-domain analysis with deep learning methods.

Despite the promising results, challenges remain, such as dataset availability, computational complexity, and the need for further optimization. Future research will aim to enhance the proposed approaches by addressing current limitations. Expanding the dataset with more diverse MRI scans from multiple sources will improve model generalization. Optimizing preprocessing using adaptive or data-driven wavelet selection can capture more discriminative features. Integrating explainable AI techniques will help interpret model decisions and increase clinical trust. Combining wavelet-based features with transformer or graph neural network architectures could further improve classification accuracy.

Finally, The integration of wavelet transform with AI has shown significant potential in medical signal processing, particularly for brain tumor classification. This research contributes to the ongoing development of intelligent diagnostic tools, paving the way for more accurate and efficient medical decision-making.

Bibliography

Bibliography

- [1] R. M. Rangayyan and S. Krishnan, *Biomedical signal analysis*. John Wiley & Sons, 2024.
- [2] R. Manikandan, S. Arunprakash, R. A. Alsowail, and T. Pandiaraj, “A novel wireless sensor network deployment for monitoring and predicting abnormal actions in medical environment and patient health state,” *Alexandria Engineering Journal*, vol. 119, pp. 149–167, 2025.
- [3] A. K. Singh and S. Krishnan, “Ecg signal feature extraction trends in methods and applications,” *BioMedical Engineering OnLine*, vol. 22, no. 1, p. 22, 2023.
- [4] K. Elnaggar, M. M. El-Gayar, and M. Elmogy, “Depression detection and diagnosis based on electroencephalogram (eeg) analysis: A systematic review,” *Diagnostics*, vol. 15, no. 2, p. 210, 2025.
- [5] V. Gohel and N. Mehendale, “Review on electromyography signal acquisition and processing,” *Biophysical reviews*, vol. 12, no. 6, pp. 1361–1367, 2020.
- [6] P. Sharma, X. Hui, J. Zhou, T. B. Conroy, and E. C. Kan, “Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate,” *NPJ digital medicine*, vol. 3, no. 1, p. 98, 2020.
- [7] J. P. Mynard, A. Kondiboyina, R. Kowalski, M. M. Cheung, and J. J. Smolich, “Measurement, analysis and interpretation of pressure/flow waves in blood vessels,” *Frontiers in Physiology*, vol. 11, p. 1085, 2020.
- [8] S. Hussain, I. Mubeen, N. Ullah, S. S. U. D. Shah, B. A. Khan, M. Zahoor, R. Ullah, F. A. Khan, and M. A. Sultan, “Modern diagnostic imaging technique applications and risk factors in the medical field: a review,” *BioMed research international*, vol. 2022, no. 1, p. 5164970, 2022.
- [9] S. Solanki, U. P. Singh, S. S. Chouhan, and S. Jain, “Brain tumor detection and classification using intelligence techniques: an overview,” *IEEE Access*, vol. 11, pp. 12870–12886, 2023.
- [10] N. Rasool and J. I. Bhat, “Brain tumour detection using machine and deep learning: a systematic review,” *Multimedia tools and applications*, vol. 84, no. 13, pp. 11551–11604, 2025.
- [11] S. U. R. Khan, S. Asif, M. Zhao, W. Zou, Y. Li, and C. Xiao, “Shallowmri: a novel lightweight cnn with novel attention mechanism for multi brain tumor classification in mri images,” *Biomedical Signal Processing and Control*, vol. 111, p. 108425, 2026.

- [12] A. Kumar, S. P. Yadav, and A. Kumar, “An improved feature extraction algorithm for robust swin transformer model in high-dimensional medical image analysis,” *Computers in biology and medicine*, vol. 188, p. 109822, 2025.
- [13] S. Das and R. S. Goswami, “Advancements in brain tumor analysis: a comprehensive review of machine learning, hybrid deep learning, and transfer learning approaches for mri-based classification and segmentation,” *Multimedia Tools and Applications*, vol. 84, no. 23, pp. 26645–26682, 2025.
- [14] M. Abdullahi, O. N. Oyelade, A. F. D. Kana, M. A. Bagiwa, F. B. Abdullahi, S. B. Junaidu, I. Iliyasu, A. Ore-ofe, and H. Chiroma, “A systematic literature review of visual feature learning: deep learning techniques, applications, challenges and future directions,” *Multimedia Tools and Applications*, vol. 84, no. 19, pp. 20439–20496, 2025.
- [15] S. B. Shuvo, S. S. Alam, S. U. Ayman, A. Chakma, M. Salvi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya, “Application of wavelet transformation and artificial intelligence techniques in healthcare: A systemic review,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 15, no. 2, p. e70007, 2025.
- [16] M. Chen, D. Cui, H. Haick, and N. Tang, “Artificial intelligence-based medical sensors for healthcare system,” *Advanced Sensor Research*, vol. 3, no. 3, p. 2300009, 2024.
- [17] N. Hastings, D. Samuel, A. N. Ansari, P. Kaurani, V. S. Bhandary, P. Gautam, A. L. T. Purayil, T. Hassan, M. D. Eshwar, B. S. T. Nuthalapati, *et al.*, “The role of artificial intelligence-powered imaging in cerebrovascular accident detection,” *Cureus*, vol. 16, no. 5, 2024.
- [18] P. Gupta and M. K. Pandey, “Role of ai for smart health diagnosis and treatment,” in *Smart Medical Imaging for Diagnosis and Treatment Planning*, pp. 23–45, Chapman and Hall/CRC, 2024.
- [19] B. Nam, B. Bark, J. Lee, and I. Y. Kim, “Insightsleepnet: the interpretable and uncertainty-aware deep learning network for sleep staging using continuous photoplethysmography,” *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, p. 50, 2024.
- [20] S. Fatima, “Predictive models for early detection of chronic diseases like cancer,” *Olaoye, G*, 2024.
- [21] J. Rajeswari and M. Jagannath, “Advances in biomedical signal and image processing—a systematic review,” *Informatics in Medicine Unlocked*, vol. 8, pp. 13–19, 2017.
- [22] A. B. Naeem, B. Senapati, D. Bhuvu, A. Zaidi, A. Bhuvu, M. S. I. Sudman, and A. E. Ahmed, “Heart disease detection using feature extraction and artificial neural networks: A sensor-based approach,” *IEEE Access*, 2024.

- [23] N. Cancino-Fuentes, A. Manasanch, J. Covelo, A. Suarez-Perez, E. Fernandez, S. Matsoukis, C. Guger, X. Illa, A. Guimerà-Brunet, and M. V. Sanchez-Vives, "Recording physiological and pathological cortical activity and exogenous electric fields using graphene microtransistor arrays in vitro," *Nanoscale*, vol. 16, no. 2, pp. 664–677, 2024.
- [24] A. Joutsen, A. Cömert, E. Kaappa, K. Vanhatalo, J. Riistama, A. Vehkaoja, and H. Eskola, "Ecg signal quality in intermittent long-term dry electrode recordings with controlled motion artifacts," *Scientific Reports*, vol. 14, no. 1, p. 8882, 2024.
- [25] K. Blinowska and P. Durka, "Electroencephalography (eeg)," *Wiley encyclopedia of biomedical engineering*, vol. 10, p. 9780471740360, 2006.
- [26] P. Gélissee, S. R. Benbadis, A. Crespel, and W. O. Tatum, "Overcoming traps and pitfalls leading to misinterpretation of normal eeg variants and variation of the background activity," *Journal of Neurology*, vol. 271, pp. 3869–3878, 2024.
- [27] D. Tuken, "The efficacy and incorporation of biofeedback in medical professionals' ergonomic awareness," Master's thesis, The University of Iowa, 2024.
- [28] S. Muceli and R. Merletti, "Tutorial. frequency analysis of the surface emg signal: Best practices," *Journal of Electromyography and Kinesiology*, vol. 79, p. 102937, 2024.
- [29] A. Q. Wang, B. K. Karaman, H. Kim, J. Rosenthal, R. Saluja, S. I. Young, and M. R. Sabuncu, "A framework for interpretability in machine learning for medical imaging," *IEEE Access*, 2024.
- [30] Z. Jiang, S. E. Salcudean, and N. Navab, "Robotic ultrasound imaging: State-of-the-art and future perspectives," *Medical image analysis*, vol. 89, p. 102878, 2023.
- [31] Y. Wang *et al.*, "Deep learning in medical ultrasound imaging: A review," *Medical Image Analysis*, 2024.
- [32] I. Rozylo-Kalinowska and K. Orhan, eds., *Imaging of the Temporomandibular Joint*. Cham: Springer, 2019.
- [33] J.-Y. Ro *et al.*, "Artificial intelligence in chest radiography and computed tomography," *European Journal of Radiology*, 2024.
- [34] M. Dewey, T. Schink, and C. F. Dewey, "Claustrophobia during magnetic resonance imaging: cohort study in over 55,000 patients," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 26, no. 5, pp. 1322–1327, 2007.
- [35] L. Winter, J. Periquito, C. Kolbitsch, R. Pellicer-Guridi, R. G. Nunes, M. Hauer, L. Broche, and T. O'Reilly, "Open-source magnetic resonance imaging: Improving access, science, and education through global collaboration," *NMR in Biomedicine*, vol. 37, no. 7, p. e5052, 2024.
- [36] F. Van Gelderen and F. Van Gelderen, *Understanding X-rays*. Springer, 2004.

- [37] M. Castillo, “History and evolution of brain tumor imaging: Insights through radiology,” *Radiology*, vol. 273, no. 2S, pp. S111–S125, 2014.
- [38] J. Trotter, A. R. Pantel, B.-K. K. Teo, F. E. Escorcía, T. Li, D. A. Pryma, and N. K. Taunk, “Positron emission tomography (pet)/computed tomography (ct) imaging in radiation therapy treatment planning: a review of pet imaging tracers and methods to incorporate pet/ct,” *Advances in radiation oncology*, vol. 8, no. 5, p. 101212, 2023.
- [39] C. J. Clemente and T. J. Dick, “How scaling approaches can reveal fundamental principles in physiology and biomechanics,” *Journal of Experimental Biology*, vol. 226, no. 7, p. jeb245310, 2023.
- [40] T. G. La, K. C. T. Nguyen, N. Kaipatur, E. H. Lou, P. W. Major, and L. H. Le, “Investigating transducer–tissue interface pressure for soft tissue stress–strain behavior and the effects on echoic intensities in ultrasound imaging of periodontium,” *Advanced Materials Technologies*, vol. 9, no. 8, p. 2301732, 2024.
- [41] J. A. Espina, M. H. Cordeiro, M. Milivojevic, I. Pajić-Lijaković, and E. H. Barriga, “Response of cells and tissues to shear stress,” *Journal of Cell Science*, vol. 136, no. 18, p. jcs260985, 2023.
- [42] H. Kim, J. S. Kim, and C. K. Chung, “Identification of cerebral cortices processing acceleration, velocity, and position during directional reaching movement with deep neural network and explainable ai,” *NeuroImage*, vol. 266, p. 119783, 2023.
- [43] M. Boyer, L. Bouyer, J.-S. Roy, and A. Campeau-Lecours, “Reducing noise, artifacts and interference in single-channel emg signals: A review,” *Sensors*, vol. 23, no. 6, p. 2927, 2023.
- [44] A. Feduccia, G. Agin-Liebes, C. M. Price, N. Grinsell, S. Paradise, and D. M. Rabin, “The need for establishing best practices and gold standards in psychedelic medicine,” *Journal of Affective Disorders*, vol. 332, pp. 47–54, 2023.
- [45] R. Cuocolo, M. Caruso, T. Perillo, L. Ugga, and M. Petretta, “Machine learning in oncology: a clinical appraisal,” *Cancer letters*, vol. 481, pp. 55–62, 2020.
- [46] F. Khan, Y. Gulzar, S. Ayoub, M. Majid, M. S. Mir, and A. B. Soomro, “Least square-support vector machine based brain tumor classification system with multi model texture features,” *Frontiers in Applied Mathematics and Statistics*, vol. 9, p. 1324054, 2023.
- [47] Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, and X. Liang, “An improved random forest based on the classification accuracy and correlation measurement of decision trees,” *Expert Systems with Applications*, vol. 237, p. 121549, 2024.
- [48] M. A. Assaad, M. I. Saleh, and R. Mahrousseh, “A novel framework for accurate brain tumor detection in mri scans using cnn, mlp, and knn techniques,” 2025.
- [49] H. Yu, L. T. Yang, Q. Zhang, D. Armstrong, and M. J. Deen, “Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives,” *Neurocomputing*, vol. 444, pp. 92–110, 2021.

- [50] X. Wang, “Application of multimodal fusion in automatic image classification: Combining cnn and rnn model,” *Intelligent Decision Technologies*, vol. 19, no. 2, pp. 928–942, 2025.
- [51] K. Faust, Q. Xie, D. Han, K. Goyle, Z. Volynskaya, U. Djuric, and P. Diamandis, “Visualizing histopathologic deep learning classification and anomaly detection using nonlinear feature space dimensionality reduction,” *BMC bioinformatics*, vol. 19, no. 1, p. 173, 2018.
- [52] M. RV, Y. G. N, M. HM, B. KN, and M. R, “Brain tumor classification on mri images using fine-tuned deep learning models,” *Machine Learning for Computational Science and Engineering*, vol. 1, no. 1, p. 22, 2025.
- [53] S. K. Zhou, H. N. Le, K. Luu, H. V. Nguyen, and N. Ayache, “Deep reinforcement learning in medical imaging: A literature review,” *Medical image analysis*, vol. 73, p. 102193, 2021.
- [54] K. Borys, Y. A. Schmitt, M. Nauta, C. Seifert, N. Krämer, C. M. Friedrich, and F. Nensa, “Explainable ai in medical imaging: An overview for clinical practitioners—saliency-based xai approaches,” *European journal of radiology*, vol. 162, p. 110787, 2023.
- [55] E.-S. A. El-Dahshan, H. M. Mohsen, K. Revett, and A.-B. M. Salem, “Computer-aided diagnosis of human brain tumor through mri: A survey and a new algorithm,” *Expert Systems with Applications*, vol. 41, no. 11, pp. 5526–5545, 2014.
- [56] L. Debnath and F. A. Shah, *Wavelet transforms and their applications*, vol. 434. Springer, 2015.
- [57] K. C. Vidyasagar, K. R. Kumar, G. A. Sai, M. Ruchita, and M. J. Saikia, “Signal to image conversion and convolutional neural networks for physiological signal processing: A review,” *Ieee Access*, vol. 12, pp. 66726–66764, 2024.
- [58] K. Fujinoki and K. Ashizawa, “Directional lifting wavelet transform for image edge analysis,” *Signal Processing*, vol. 216, p. 109188, 2024.
- [59] C. M. Akujuobi, *Wavelets and wavelet transform systems and their applications*. Springer, 2022.
- [60] M. Moradi, “Wavelet transform approach for denoising and decomposition of satellite-derived ocean color time-series: Selection of optimal mother wavelet,” *Advances in Space Research*, vol. 69, no. 7, pp. 2724–2744, 2022.
- [61] I. H. Latif, S. H. Abdulredha, and S. K. A. Hassan, “Discrete wavelet transform-based image processing: A review,” *Al-Nahrain Journal of Science*, vol. 27, no. 3, pp. 109–125, 2024.
- [62] D. V. Puri, P. H. Kachare, I. Al-Shourbaji, A. Jabbari, R. Kirner, and A. Alameen, “Design of efficient bi-orthogonal wavelets for eeg-based detection of schizophrenia,” *Engineering Science and Technology, an International Journal*, vol. 68, p. 102090, 2025.

- [63] W. Wang, L. Li, Z. Qu, and X. Yang, “Enhanced damage segmentation in rc components using pyramid haar wavelet downsampling and attention u-net,” *Automation in Construction*, vol. 168, p. 105746, 2024.
- [64] J. GhoshDastider, D. Pal, and P. K. Mishra, “Evidence of kolmogorov like scalings and multifractality in the rainfall events,” *arXiv preprint arXiv:2405.03463*, 2024.
- [65] T. H. Ali *et al.*, “Choosing an appropriate wavelet for varx time series model analysis,” *Journal of Economics and Administrative Sciences*, vol. 31, no. 146, pp. 174–196, 2025.
- [66] J. Touthang, “Wavelets and their recent applications,” *Mathematics and Computer Science for Real-World Applications*, pp. 505–538, 2025.
- [67] A. D. Le, S. Jin, S. Seo, Y.-S. Bae, and T. Q. Nguyen, “Biorthogonal lattice tunable wavelet units and their implementation in convolutional neural networks for computer vision problems,” *IEEE Open Journal of Signal Processing*, 2025.
- [68] J. Paillard, J. F. Hipp, and D. A. Engemann, “Green: A lightweight architecture using learnable wavelets and riemannian geometry for biomarker exploration with eeg signals,” *Patterns*, vol. 6, no. 3, 2025.
- [69] D. V. Puri, S. L. Nalbalwar, and P. P. Ingle, “Eeg-based systematic explainable alzheimer’s disease and mild cognitive impairment identification using novel rational dyadic biorthogonal wavelet filter banks,” *Circuits, Systems, and Signal Processing*, vol. 43, no. 3, pp. 1792–1822, 2024.
- [70] M. Sinith, G. A, and C. KR, “Biorthogonal wavelets,” in *Wavelets and Multi-Resolution Analysis: A Signal Processing Approach*, pp. 105–123, Springer, 2025.
- [71] T. Guo, T. Zhang, E. Lim, M. Lopez-Benitez, F. Ma, and L. Yu, “A review of wavelet analysis and its applications: Challenges and opportunities,” *IEEE Access*, vol. 10, pp. 58869–58903, 2022.
- [72] Z.-M. Cao, Y.-Y. Chen, J. Jiang, Z.-L. Cao, and B. Tu, “Extraction of moiré fringes’ phase information based on the morlet wavelet,” *Applied Optics*, vol. 63, no. 8, pp. 1971–1981, 2024.
- [73] T. H. Chowdhury, K. N. Poudel, and Y. Hu, “Time-frequency analysis, denoising, compression, segmentation, and classification of pcg signals,” *Ieee Access*, vol. 8, pp. 160882–160890, 2020.
- [74] M. T. H. Chowdhury, K. N. Poudel, and Y. Hu, “Automatic phonocardiography analysis using discrete wavelet transform,” in *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing*, pp. 1–8, 2019.
- [75] T. M. Lai *et al.*, “Wavelet transform based analysis of high impedance faults,” *Turkish Journal of Electrical Power and Energy Systems*, 2024.
- [76] S. Gaikwad, M. B. Mali, S. O. Rajankar, and S. A. Shirsat, “Enhancing medical imaging: Leveraging discrete wavelet transform and its variants for advanced image fusion,” *Journal of Integrated Science and Technology*, vol. 13, no. 5, pp. 1109–1109, 2025.

- [77] M. Talbi, R. Baazaoui, and B. Nasraoui, “A novel method of image denoising based on 2d dual-tree dwt and swt,” *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 22, no. 4, pp. 1–20, 2024.
- [78] M. A. Belkadi and A. Daamouche, “A robust qrs detection approach using stationary wavelet transform,” *Multimedia Tools and Applications*, vol. 80, no. 15, pp. 22843–22864, 2021.
- [79] K. Liu, Y. Guo, and B. Su, “Image denoising network based on subband information sharing using dual-tree complex wavelet,” *Neural processing letters*, vol. 55, no. 8, pp. 10975–10991, 2023.
- [80] H. NAIMI, A. B. H. ADAMOU-MITICHE, and L. MITICHE, “Hybrid thresholding lifting dual tree complex wavelet transform with wiener filter for quality assurance of medical image,” in *2021 12th International Conference on Information and Communication Systems (ICICS)*, pp. 228–233, IEEE, 2021.
- [81] C. Jiang, W. Yang, Y. Wang, and D. Zhao, “Gnss signal processing based on improved lifting wavelet transform with prior constraint,” *Scientific Reports*, vol. 15, no. 1, p. 20019, 2025.
- [82] B. Latreche, A. Merrad, B. Ali, H. Naimi, and S. Slami, “A robust dual-layer medical image watermarking scheme based on matrix factorization in the lwt domain for e-healthcare applications,” *Multimedia Tools and Applications*, vol. 84, no. 25, pp. 29883–29913, 2024.
- [83] W. K. Fadhil, Z. M. Ahmeed, B. F. Jameel, O. A. Azawi, A. M. Saeed, and M. R. Zaki, “Biomedical signal processing: Advanced techniques, applications, and emerging trends,” *European Journal of Applied Science, Engineering and Technology*, vol. 3, no. 4, pp. 49–54, 2025.
- [84] Q. Wei, S. Zeng, Z. Jiang, F. Xu, S. Shi, W. Wen, Z. Qin, Z. Lou, and K. Li, “Topological representation based on wavelet transform as a novel imaging biomarker for tumor diagnosis in ultrasound images: A comprehensive study,” *Computer Methods and Programs in Biomedicine*, p. 108859, 2025.
- [85] B. Zhang, J. M. Fadili, and J.-L. Starck, “Wavelets, ridgelets, and curvelets for poisson noise removal,” *IEEE Transactions on image processing*, vol. 17, no. 7, pp. 1093–1108, 2008.
- [86] C. Tian, M. Zheng, W. Zuo, B. Zhang, Y. Zhang, and D. Zhang, “Multi-stage image denoising with the wavelet transform,” *Pattern Recognition*, vol. 134, p. 109050, 2023.
- [87] H. Naimi, A. B. H. Adamou-Mitiche, and L. Mitiche, “Medical image denoising using dual tree complex thresholding wavelet transform and wiener filter,” *Journal of King Saud University-Computer and Information Sciences*, vol. 27, no. 1, pp. 40–45, 2015.
- [88] H. Naimi, “Performance and quality assurance of medical image using hybrid thresholding wavelet transform with wiener filter,” *Australian Journal of Electrical and Electronics Engineering*, vol. 19, no. 3, pp. 294–299, 2022.

- [89] K. Hackmack, F. Paul, M. Weygandt, C. Allefeld, J.-D. Haynes, A. D. N. Initiative, *et al.*, “Multi-scale classification of disease using structural mri and wavelet transform,” *Neuroimage*, vol. 62, no. 1, pp. 48–58, 2012.
- [90] B. Kaushik, “Efficient compression techniques for medical image storage and transmission: A comprehensive review,” *International Journal of Image and Graphics*, p. 2650024, 2024.
- [91] J. Wu, J. Li, J. Yang, and S. Mei, “Wavelet-integrated deep neural networks: A systematic review of applications and synergistic architectures,” *Neurocomputing*, p. 131648, 2025.
- [92] C. Fan, M. Yun, L. Chen, H. Zhao, and B. Peng, “Dwmfusion: A lightweight medical image fusion network based on discrete wavelet convolution and selective state space models,” *Expert Systems with Applications*, p. 128040, 2025.
- [93] J. Mudegaonkar and D. Yadav, “Enhanced brain mri disease classification via wavelet decomposition-infused cnn architecture featuring residual blocks,” *Journal of Integrated Science and Technology*, vol. 13, no. 1, pp. 1005–1005, 2025.
- [94] S. Shivadekar, *Cognitive Artificial Intelligence for Health and Climate: Deep Models, Interpretability, and Decision Support*. Deep Science Publishing, 2025.
- [95] B. Chinnaiyan, S. Balasubaramanian, M. Jeyabalu, and G. S. Warriar, “Ai applications—computer vision and natural language processing,” *Model Optimization Methods for Efficient and Edge AI: Federated Learning Architectures, Frameworks and Applications*, pp. 25–41, 2025.
- [96] S. Cohen, “The evolution of machine learning: Past, present, and future,” in *Artificial Intelligence in Pathology*, pp. 3–14, Elsevier, 2025.
- [97] M. K. Sumanth Kamath, “9 intelligent solutions for future smart industries,” *Industry 5.0: A New Revolution Through Human-Centric Solution*, vol. 17, p. 129, 2025.
- [98] A. S. Batista *et al.*, “The transforming clinical practice: The role of ai-powered medical assistants in enhancing healthcare efficiency and decision-making,” *Journal of Next-Generation Research 5.0*, 2025.
- [99] A. Alharthi, A. Alqurashi, T. Alharbi, M. Alammar, N. Aldosari, H. Bouchekara, Y. Shaaban, M. S. Shahriar, and A. Al Ayidh, “Explainable ai for sensor signal interpretation to revolutionize human health monitoring: A review,” *IEEE Access*, 2025.
- [100] A. Nazir, A. Hussain, M. Singh, and A. Assad, “Deep learning in medicine: advancing healthcare with intelligent solutions and the future of holography imaging in early diagnosis,” *Multimedia Tools and Applications*, vol. 84, no. 17, pp. 17677–17740, 2025.
- [101] J. H. Korteling, G. C. van de Boer-Visschedijk, R. A. Blankendaal, R. C. Boonekamp, and A. R. Eikelboom, “Human-versus artificial intelligence,” *Frontiers in artificial intelligence*, vol. 4, p. 622364, 2021.

- [102] I. H. Sarker, “Machine learning: Algorithms, real-world applications and research directions,” *SN computer science*, vol. 2, no. 3, p. 160, 2021.
- [103] C. Mishra and D. Gupta, “Deep machine learning and neural networks: An overview,” *IAES international journal of artificial intelligence*, vol. 6, no. 2, p. 66, 2017.
- [104] S. Jain *et al.*, “A comprehensive survey of deep learning: Architectures, applications, and challenges,” *Journal of Artificial Intelligence*, 2024.
- [105] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electronic markets*, vol. 31, no. 3, pp. 685–695, 2021.
- [106] M. O. K. Mendonça, S. L. Netto, P. S. R. Diniz, and S. Theodoridis, “Machine learning: Review and trends,” in *Signal Processing and Machine Learning Theory* (P. S. R. Diniz, ed.), ch. 13, pp. 869–959, Academic Press, 2023.
- [107] S. B. Kotsiantis, I. Zaharakis, P. Pintelas, *et al.*, “Supervised machine learning: A review of classification techniques,” *Emerging artificial intelligence applications in computer engineering*, vol. 160, no. 1, pp. 3–24, 2007.
- [108] A. Géron, “Hands-on machine learning with scikit-learn, keras, and tensorflow: concepts,” *Aurélien Géron-Google Kitaplar*, yy <https://books.google.com.tr/books>, 2019.
- [109] K. Sun, A. Roy, and J. M. Tobin, “Artificial intelligence and machine learning: Definition of terms and current concepts in critical care research,” *Journal of Critical Care*, vol. 82, p. 154792, 2024.
- [110] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [111] S. A. Fayaz, S. Jahangeer Sidiq, M. Zaman, and M. A. Butt, “Machine learning: An introduction to reinforcement learning,” *Machine Learning and Data Science: Fundamentals and Applications*, pp. 1–22, 2022.
- [112] A. B. Musa, “Comparative study on classification performance between support vector machine and logistic regression,” *International Journal of Machine Learning and Cybernetics*, vol. 4, pp. 13–24, 2013.
- [113] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, “An overview on the advancements of support vector machine models in healthcare applications: A review,” *Information*, vol. 15, no. 4, p. 235, 2024.
- [114] M. Suyal and P. Goyal, “A review on analysis of k-nearest neighbor classification machine learning algorithms based on supervised learning,” *International Journal of Engineering Trends and Technology*, vol. 70, no. 7, pp. 43–48, 2022.
- [115] M. Sonawane and C. Dhawale, “Evaluation and analysis of few parametric and nonparametric classification methods,” in *2016 Second International Conference on Computational Intelligence & Communication Technology (CICT)*, pp. 14–21, IEEE, 2016.

- [116] K. P. Sinaga and M.-S. Yang, "Unsupervised k-means clustering algorithm," *IEEE access*, vol. 8, pp. 80716–80727, 2020.
- [117] H. A. Salman, A. Kalakech, and A. Steiti, "Random forest algorithm overview," *Babylonian Journal of Machine Learning*, vol. 2024, pp. 69–79, 2024.
- [118] R. Kumar, B. Goswami, S. M. Mhatre, and S. Agrawal, "Naive bayes in focus: a thorough examination of its algorithmic foundations and use cases," *Int. J. Innov. Sci. Res. Technol*, vol. 9, no. 5, pp. 2078–2081, 2024.
- [119] J.-z. Feng, Y. Wang, J. Peng, M.-w. Sun, J. Zeng, and H. Jiang, "Comparison between logistic regression and machine learning algorithms on survival prediction of traumatic brain injuries," *Journal of critical care*, vol. 54, pp. 110–116, 2019.
- [120] X. Ran, Y. Xi, Y. Lu, X. Wang, and Z. Lu, "Comprehensive survey on hierarchical clustering algorithms and the recent developments," *Artificial Intelligence Review*, vol. 56, no. 8, pp. 8219–8264, 2023.
- [121] V. Vashishth, A. Chhabra, and D. K. Sharma, "Gmmr: A gaussian mixture model based unsupervised machine learning approach for optimal routing in opportunistic iot networks," *Computer Communications*, vol. 134, pp. 138–148, 2019.
- [122] C. C. Aggarwal, *Neural Networks and Deep Learning*. Cham: Springer, 2018.
- [123] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [124] I. Goodfellow, Y. Bengio, and A. Courville, "Deep feedforward networks," *Deep learning*, vol. 1, pp. 161–217, 2016.
- [125] P. Kvello and N. Gericke, "Identifying knowledge important to teach about the nervous system in the context of secondary biology and science education—a delphi study," *PloS one*, vol. 16, no. 12, p. e0260752, 2021.
- [126] J.-W. Lin, "Artificial neural network related to biological neuron network: a review," *Advanced Studies in Medical Sciences*, vol. 5, no. 1, pp. 55–62, 2017.
- [127] A. Halužan Vasle and M. Moškon, "Synthetic biological neural networks: From current implementations to future perspectives," *Biosystems*, vol. 237, p. 105164, 2024.
- [128] A. Dongare, R. Kharde, A. D. Kachare, *et al.*, "Introduction to artificial neural network," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 2, no. 1, pp. 189–194, 2012.
- [129] Y. A. Christobel and R. J. Suji, "A comprehensive review on neural network architectures," *Journal of Computational Analysis and Applications*, vol. 33, no. 7, pp. 1443–1448, 2024.
- [130] T. Szandala, "Review and comparison of commonly used activation functions for deep neural networks," in *Bio-inspired neurocomputing*, pp. 203–224, Springer, 2020.
- [131] M. Kumar, U. Mehta, and G. Cirrincione, "Enhancing neural network classification using fractional-order activation functions," *AI Open*, vol. 5, pp. 10–22, 2024.

- [132] A. Dureja and P. Pahwa, “Analysis of non-linear activation functions for classification tasks using convolutional neural networks,” *Recent Patents on Computer Science*, vol. 12, no. 3, pp. 156–161, 2019.
- [133] S. K. Roy, S. Manna, S. R. Dubey, and B. B. Chaudhuri, “Lisht: Non-parametric linearly scaled hyperbolic tangent activation function for neural networks,” in *International Conference on Computer Vision and Image Processing*, pp. 462–476, Springer, 2022.
- [134] S. Maharjan, A. Alsadoon, P. Prasad, T. Al-Dalain, and O. H. Alsadoon, “A novel enhanced softmax loss function for brain tumour detection using deep learning,” *Journal of neuroscience methods*, vol. 330, p. 108520, 2020.
- [135] Y. Song, A. Schwing, R. Urtasun, *et al.*, “Training deep neural networks via direct loss minimization,” in *International conference on machine learning*, pp. 2169–2177, PMLR, 2016.
- [136] M. G. Abdolrasol, S. S. Hussain, T. S. Ustun, M. R. Sarker, M. A. Hannan, R. Mohamed, J. A. Ali, S. Mekhilef, and A. Milad, “Artificial neural networks based optimization techniques: A review,” *Electronics*, vol. 10, no. 21, p. 2689, 2021.
- [137] S. H. Haji and A. M. Abdulazeez, “Comparison of optimization techniques based on gradient descent algorithm: A review,” *PalArch's Journal of Archaeology of Egypt/Egyptology*, vol. 18, no. 4, pp. 2715–2743, 2021.
- [138] X. Zhou, “Understanding the convolutional neural networks with gradient descent and backpropagation,” in *Journal of Physics: Conference Series*, vol. 1004, p. 012028, IOP Publishing, 2018.
- [139] N. Ketkar, J. Moolayil, N. Ketkar, and J. Moolayil, “Convolutional neural networks,” *Deep learning with Python: learn best practices of deep learning models with PyTorch*, pp. 197–242, 2021.
- [140] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, “A review of convolutional neural networks in computer vision,” *Artificial Intelligence Review*, vol. 57, no. 4, p. 99, 2024.
- [141] M. A. Hossain and M. S. A. Sajib, “Classification of image using convolutional neural network (cnn),” *Global Journal of Computer Science and Technology*, vol. 19, no. 2, pp. 13–14, 2019.
- [142] H. Gholamalinezhad and H. Khosravi, “Pooling methods in deep neural networks, a review,” *arXiv preprint arXiv:2009.07485*, 2020.
- [143] L. Zhao and Z. Zhang, “A improved pooling method for convolutional neural networks,” *Scientific Reports*, vol. 14, p. 1589, 2024.
- [144] S. S. Basha, S. R. Dubey, V. Pulabaigari, and S. Mukherjee, “Impact of fully connected layers on performance of convolutional neural networks for image classification,” *Neurocomputing*, vol. 378, pp. 112–119, 2020.

- [145] A. Lindberg and L. Larsson, “Evaluating the performance of extended convolutional networks: A comparative study between vgg-16 and vgg-23 for image classification,” 2024.
- [146] M. S. Alam *et al.*, “Optimizing vgg16 deep learning model with enhanced hunger games search for logo classification,” *Scientific Reports*, 2024.
- [147] M. Tan and Q. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.
- [148] M. Ali *et al.*, “Enhancing brain tumour multi-classification using efficient-net b0-based intelligent diagnosis for internet of medical things (iomt) applications,” *Information*, vol. 15, no. 8, p. 489, 2024.
- [149] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241, Springer, 2015.
- [150] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. Haddadi Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli, and D. Merhof, “Medical image segmentation review: The success of u-net,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 10076–10095, 2024.
- [151] L. Borawar and R. Kaur, “Resnet: Solving vanishing gradient in deep networks,” in *Proceedings of International Conference on Recent Trends in Computing: ICRTC 2022*, pp. 235–247, Springer, 2023.
- [152] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2024.
- [153] D. M. Ahmed, M. M. Hassan, and R. J. Mstafa, “A review on deep sequential models for forecasting time series data,” *Applied Computational Intelligence and Soft Computing*, vol. 2022, no. 1, p. 6596397, 2022.
- [154] Y. Yu, X. Si, C. Hu, and J. Zhang, “A review of recurrent neural networks: Lstm cells and network architectures,” *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.
- [155] S. M. Al-Selwi, M. F. Hassan, S. J. Abdulkadir, A. Muneer, E. H. Sumiea, A. Alqushaibi, and M. G. Ragab, “Rnn-lstm: From applications to modeling techniques and beyond—systematic review,” *Journal of King Saud University Computer and Information Sciences*, vol. 36, no. 5, p. 102068, 2024.
- [156] M. A. Muzammil, S. Javid, A. K. Afridi, R. Siddineni, M. Shahabi, M. Haseeb, F. Fariha, S. Kumar, S. Zaveri, and A. J. Nashwan, “Artificial intelligence-enhanced electrocardiography for accurate diagnosis and management of cardiovascular diseases,” *Journal of Electrocardiology*, vol. 83, pp. 30–40, 2024.

- [157] Y. Yang, F. Li, X. Qin, H. Wen, X. Lin, and D. Huang, "Feature separation and adversarial training for the patient-independent detection of epileptic seizures," *Frontiers in Computational Neuroscience*, vol. 17, p. 1195334, 2023.
- [158] S. Karthik, M. Santhosh, M. S. Kavitha, and A. C. Paul, "Automated deep learning based cardiovascular disease diagnosis using ecg signals.," *Computer Systems Science & Engineering*, vol. 42, no. 1, 2022.
- [159] S. Bhatlawande, S. Shilaskar, S. Pramanik, and S. Sole, "Multimodal emotion recognition based on the fusion of vision, eeg, ecg, and emg signals," *International journal of electrical and computer engineering systems*, vol. 15, no. 1, pp. 41–58, 2024.
- [160] E. Abdelfattah, S. Joshi, and S. Tiwari, "Machine and deep learning models for stress detection using multimodal physiological data," *IEEE Access*, 2025.
- [161] H. Laçi, K. Sevrani, and S. Iqbal, "Deep learning approaches for classification tasks in medical x-ray, mri, and ultrasound images: a scoping review," *BMC Medical Imaging*, vol. 25, no. 1, p. 156, 2025.
- [162] M. U. Khan, G. Chetty, R. Goecke, and R. Fernandez-Rojas, "A systematic review of multimodal signal fusion for acute pain assessment systems," *ACM Computing Surveys*, 2025.
- [163] H. Uyanik, A. Sengur, M. Salvi, R.-S. Tan, J. H. Tan, and U. R. Acharya, "Automated detection of neurological and mental health disorders using eeg signals and artificial intelligence: A systematic review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 15, no. 1, p. e70002, 2025.
- [164] A. R. Freeda, R. Kanthavel, and A. Anju, "Scalability issues in ai computing in large-scale networks," in *AI for Large Scale Communication Networks*, pp. 395–414, IGI Global, 2025.
- [165] D. An, M. Lim, and S. Lee, "Challenges for data quality in the clinical data life cycle: systematic review," *Journal of Medical Internet Research*, vol. 27, p. e60709, 2025.
- [166] N. Bernard and K. Balog, "A systematic review of fairness, accountability, transparency, and ethics in information retrieval," *ACM Computing Surveys*, vol. 57, no. 6, pp. 1–29, 2025.
- [167] R. Nanmaran, S. Srimathi, G. Yamuna, S. Thanigaivel, A. Vickram, A. Priya, A. Karthick, J. Karpagam, V. Mohanavel, and M. Muhibbullah, "Investigating the role of image fusion in brain tumor classification models based on machine learning algorithm for personalized medicine," *Computational and Mathematical Methods in Medicine*, vol. 2022, no. 1, p. 7137524, 2022.
- [168] S. Das and R. S. Goswami, "Review, limitations, and future prospects of neural network approaches for brain tumor classification," *Multimedia Tools and Applications*, vol. 83, no. 15, pp. 45799–45841, 2024.
- [169] M. N. Yeasmin, M. Al Amin, T. J. Joti, Z. Aung, and M. A. Azim, "Advances of ai in image-based computer-aided diagnosis: A review," *Array*, p. 100357, 2024.

- [170] M. J. Leming, E. E. Bron, R. Bruffaerts, Y. Ou, J. E. Iglesias, R. L. Gollub, and H. Im, "Challenges of implementing computer-aided diagnostic models for neuroimages in a clinical setting," *NPJ Digital Medicine*, vol. 6, no. 1, p. 129, 2023.
- [171] N. Heidari and A. Iosifidis, "Geometric deep learning for computer-aided design: A survey," *arXiv preprint arXiv:2402.17695*, 2024.
- [172] T. Habuza, A. N. Navaz, F. Hashim, F. Alnajjar, N. Zaki, M. A. Serhani, and Y. Statsenko, "Ai applications in robotics, precision medicine, and medical image analysis: an overview and future trends," *Inform. Med. Unlocked*, vol. 24, no. 100596, pp. 10–1016, 2021.
- [173] R. Alvarado, "Should we replace radiologists with deep learning? pigeons, error and trust in medical ai," *Bioethics*, vol. 36, no. 2, pp. 121–133, 2022.
- [174] O. Abda and H. Naimi, "Advanced brain tumor mr image classification using a combination undecimated wavelet transform, efficientnet-b0 and pca via multi-svm analysis," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 49, no. 2, pp. 601–613, 2025.
- [175] M. Nickparvar, "Brain tumor mri dataset." https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset?select=Training&fbclid=IwAR10g86NhVHRrsqjksPTXD4c9BHPHm5Dw5FdEfdFvJV_qdeJHppB10w32Kw, 2021. Accessed on December 28, 2023.
- [176] J. Cheng, "Brain tumor dataset." https://figshare.com/articles/dataset/brain_tumor_dataset/1512427, 2024. Accessed on April 9, 2024.
- [177] S. Bhuvaji, "Sartaj: Brain tumor classification (mri)." <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>, 2024. Accessed on April 9, 2024.
- [178] O. Abda and H. NAIMI, "Enhanced brain tumor mri classification using stationary wavelet transform, resnet50v2, and lstm networks," *ITEGAM-JETIA*, vol. 11, no. 51, pp. 127–133, 2025.
- [179] M. Nickparvar, "Brain tumor mri dataset." <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>, 2020. Accessed: 2025-06-15.
- [180] S. Kumar and D. P. Mankame, "Optimization driven deep convolution neural network for brain tumor classification," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 3, pp. 1190–1204, 2020.
- [181] R. Hashemzahi, S. J. S. Mahdavi, M. Kheirabadi, and S. R. Kamel, "Detection of brain tumors from mri images base on deep learning using hybrid model cnn and nade," *biocybernetics and biomedical engineering*, vol. 40, no. 3, pp. 1225–1232, 2020.
- [182] C. Narmatha, S. M. Eljack, A. A. R. M. Tuka, S. Manimurugan, and M. Mustafa, "A hybrid fuzzy brain-storm optimization algorithm for the classification of brain tumor mri images," *Journal of ambient intelligence and humanized computing*, pp. 1–9, 2020.

- [183] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, “Multi-class brain tumor classification using residual network and global average pooling,” *Multimedia Tools and Applications*, vol. 80, no. 9, pp. 13429–13438, 2021.
- [184] S. A. Nawaz, D. M. Khan, and S. Qadri, “Brain tumor classification based on hybrid optimized multi-features analysis using magnetic resonance imaging dataset,” *Applied Artificial Intelligence*, vol. 36, no. 1, p. 2031824, 2022.
- [185] D. O. Patil and S. T. Hamde, “A new enhanced local optimal-oriented pattern descriptor using steerable wavelet transform for mri brain tumor classification,” *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 47, no. 1, pp. 369–384, 2023.
- [186] N. Remzan, Y. E. Hachimi, K. Tahiry, and A. Farchi, “Ensemble learning based-features extraction for brain mr images classification with machine learning classifiers,” *Multimedia Tools and Applications*, vol. 83, no. 19, pp. 57661–57684, 2024.
- [187] B. Babu Vimala, S. Srinivasan, S. K. Mathivanan, Mahalakshmi, P. Jayagopal, and G. T. Dalu, “Detection and classification of brain tumor using hybrid deep learning models,” *Scientific reports*, vol. 13, no. 1, p. 23029, 2023.
- [188] F. Zulfiqar, U. I. Bajwa, and Y. Mehmood, “Multi-class classification of brain tumor types from mr images using efficientnets,” *Biomedical Signal Processing and Control*, vol. 84, p. 104777, 2023.
- [189] A. A. Asiri, A. Shaf, T. Ali, M. A. Pasha, A. Khan, M. Irfan, S. Alqahtani, A. Alghamdi, A. H. Alghamdi, A. F. A. Alshamrani, *et al.*, “Advancing brain tumor detection: harnessing the swin transformer’s power for accurate classification and performance analysis,” *PeerJ Computer Science*, vol. 10, p. e1867, 2024.
- [190] A. Alshuhail, A. Thakur, R. Chandramma, T. Mahesh, A. Almusharraf, V. Vinoth Kumar, and S. B. Khan, “Refining neural network algorithms for accurate brain tumor classification in mri imagery,” *BMC Medical Imaging*, vol. 24, no. 1, p. 118, 2024.
- [191] M. Amin, K. M. Nahar, H. Gharaibeh, A. Nasayreh, N. Alsalmanc, A. Alomar, M. Malkawi, N. Alqasem, A. Smerat, R. A. Zitar, *et al.*, “Diet transformer model with pca-ade integration for advanced multi-class brain tumor classification,” *Intelligence-Based Medicine*, vol. 11, p. 100192, 2025.
- [192] A. B. Subba and A. K. Sunaniya, “Computationally optimized brain tumor classification using attention based googlenet-style cnn,” *Expert Systems with Applications*, vol. 260, p. 125443, 2025.

Author's Publications

International Publications

1. ABDA Oussama, NAIMI Hilal. Advanced brain tumor MR image classification using a combination undecimated wavelet transform, EfficientNet-B0 and PCA via Multi-SVM analysis. **Iranian Journal of Science and Technology, Transactions of Electrical Engineering**, Volume 49, pages 601–613, (2025).
2. ABDA Oussama, NAIMI Hilal. Enhanced brain tumor MRI classification using stationary wavelet transform, ResNet50V2, and LSTM networks. **Journal of Engineering and Technology for Industrial Applications (JETIA)**, 2025, v.11, n.51, p. 127–133.

International Communications

1. ABDA Oussama, NAIMI Hilal. Hybrid Brain Tumor Classification Using DenseNet201 and MobileNetV2 Features with Multi-Layer Perceptron. **5th International Conference on Frontiers in Academic Research (ICFAR)**, September 25–26, 2025, Konya, Turkey. Oral presentation.
2. ABDA Oussama, NAIMI Hilal. Brain Tumor MRI Classification Using ResNet50 and Xception Feature Fusion with Random Forest. **5th International Conference on Frontiers in Academic Research (ICFAR)**, September 25–26, 2025, Konya, Turkey. Oral presentation.
3. ABDA Oussama, NAIMI Hilal. Deep Learning for MRI-Based Medical Image Classification. **3rd International Conference on Pioneer and Innovative Studies (ICPIS 2025)**, October 06–07, 2025, Konya, Turkey. Oral presentation.
4. ABDA Oussama, NAIMI Hilal. Deep Learning for Efficient Medical Image Classification Across Multiple Modalities. **5th International Conference on Trends in Advanced Research (ICTAR 2025)**, October 16–17, 2025, Konya, Turkey. Oral presentation.