ADVANCEMENTS IN MRI-BASED TECHNIQUES FOR NEUROLOGICAL DISORDER DIAGNOSIS: A REVIEW OF MACHINE LEARNING APPROACHES

A Preliminary Exam Report
Submitted to
The Temple University Graduate Board

---

in Partial Fulfillment
of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY

---

by

Zahra Sadeghi Adl
April 2024

Examining Committee Members:

Dr. Feroze Mohamed, Department of Radiology
Dr. Iyad Obeid, Department of Electrical and Computer Engineering
Dr. Joseph Picone, Department of Electrical and Computer Engineering

ABSTRACT

Advancements in MRI-Based Techniques for Neurological Disorder Diagnosis: A Review of Machine Learning Approaches

by

Zahra Sadeghi Adl

Neurological disorders are a major global burden on healthcare systems, requiring timely and precise diagnosis. Magnetic Resonance Imaging (MRI) has emerged as a cornerstone in the assessment of neurological disorders, offering non-invasive visualization of anatomical structures and pathological changes. This report provides a overview of current advancements in MRI-based techniques for diagnosis and prognosis of neurological disorders, focusing on machine learning approaches. We investigate the following recently published papers: "A deep learning model for detection of cervical spinal cord compression in MRI scans," "Machine learning-based classification of chronic traumatic brain injury using hybrid diffusion imaging," along with "Quantitative Analysis in Cervical Spinal Cord Injury Patients Using Diffusion Tensor Imaging and Tractography." Every publication is thoroughly reviewed, explaining the authors' approaches, conclusions, and contributions to the field. To further facilitate future research, the report identifies important research limitations and challenges.

TABLE OF CONTENTS

CHAPTER 1

Introduction: Deciphering Hidden Insights in MRI

Neurological disorders represent a wide range of medical conditions that have a substantial impact on the nervous system. These conditions frequently cause major impairments in everyday functioning and quality of life. Early detection and accurate diagnosis of these disorders are pivotal for effective treatment and management [Khalil et al. (2018)]. Magnetic Resonance Imaging (MRI), in contrast to other imaging modalities including computed tomography (CT) scan, provides detailed non-invasive imaging of the brain and spinal cord with greater soft tissue contrast, making it an important instrument in the diagnosis of neurological disorders [Hu et al. (2022)].

While MRI has been instrumental in providing anatomical insights into neurological disorders, traditional MRI techniques primarily offer qualitative assessments of tissue characteristics, relying on visual interpretation by radiologists. These techniques are inherently subjective and may lack sensitivity to subtle changes in tissue properties. However, recent advancements in quantitative analysis have revolutionized the field, enabling objective measurements of tissue-specific parameters. This quantitative approach has the potential to provide deeper insights into tissue microstructure, function, and pathology, enhancing diagnostic accuracy and treatment decisions [Davatzikos (2019)].

In recent years, the potential of quantitative analysis of MRI has been further enhanced by machine learning (ML) methods [S. Wang & Summers (2012)]. These advanced techniques can extract hidden information from MRI images, enabling a more comprehensive understanding of neurological disorders. By analyzing these

imaging data, subtle patterns and associations that may not be apparent to the human eye can be identified. This capability has the potential to improve diagnosis and prognosis, as well as treatment strategies for neurological disorders [Chan & Siegel (2018)].

The objectives of this report are threefold. First, to provide a brief review of neurological disorders, MRI technology, and fundamentals of ML. Second, to analyze the methodologies, findings, and implications of three key research papers in this area, namely "A deep learning model for detection of cervical spinal cord compression in MRI scans" [Merali et al. (2021)], "Machine learning-based classification of chronic traumatic brain injury using hybrid diffusion imaging" [Muller et al. (2023)], and "Quantitative Analysis in Cervical Spinal Cord Injury Patients Using Diffusion Tensor Imaging and Tractography" [Park et al. (2022)]. Third, to discuss potential future directions for research and clinical applications of MRI-based techniques for neurological disorder detection, based on the gaps and limitations identified in the existing literature. By achieving these objectives, this report aims to contribute to the ongoing advancement of MRI-based techniques for neurological disorder detection and diagnosis.

CHAPTER 2

Overview: Neurological Disorders

Neurological disorders encompass a wide range of conditions affecting the central and peripheral nervous systems, leading to significant morbidity and mortality globally. The impact of neurological disorders on individuals and society as a whole is profound, often leading to significant disability, diminished quality of life, and increased healthcare expenditure. Diagnosis and management of neurological disorders pose considerable challenges due to the complexity of the nervous system and the diverse nature of these conditions [Khalil et al. (2018)].

In this chapter, we provide a brief overview of the human nervous system and common neurological disorders. By understanding the fundamentals of nervous system and common neurological disorders, we can better understand challenges and complexities of assessment in neurological patients, including MRI-based methods.

2.1 Neurological Disorder Categories

Neurological disorders are categorized based on their primary location affected, the primary type of dysfunction involved, and the primary type of cause. This classification helps in understanding the underlying pathology and in devising appropriate treatment strategies.

The causes of neurological disorders are multifactorial and can include trauma, infections, genetic factors, and environmental influences. For instance, traumatic brain injuries, are a leading cause of disability and death worldwide, with long-term consequences on cognitive and motor functions [L. G. Smith et al. (2019)]. Infections

like meningitis and encephalitis can result in severe neurological complications, highlighting the importance of early diagnosis and treatment. Moreover, genetic factors play a significant role in conditions such as Alzheimer's disease and autism spectrum disorder, emphasizing the need for further research in genomics to understand these complex disorders [Mai & Paxinos (2011)].

While numerous neurological disorders exist, we focus on three specific conditions that have been the subject of the papers that will be explored in this report: chronic Traumatic Brain Injury (cTBI), Spinal Cord Injury (SCI), and Degenerative Cervical Myelopathy (DCM). Each of these disorders poses unique clinical and scientific questions, driving investigation into their underlying mechanisms, diagnostic modalities, and therapeutic interventions.

2.1.1 Chronic Traumatic brain injury

Traumatic Brain Injury (TBI) is a multifaceted and often life-altering condition with numerous potential acute and chronic neurological consequences, contributing to approximately 1 million deaths in the United States over the last two decades [Daugherty (2021)]. This condition can result from various causes, including falls, motor vehicle accidents, sports injuries, and assaults. The effects of TBI can be diverse and may include physical, cognitive, emotional, and behavioral changes. These effects can vary depending on the severity and location of the injury, as well as the individual's age, health, and other factors [L. G. Smith et al. (2019)]. The neuropathology of chronic traumatic brain injury (cTBI) consists of a primary injury that is a direct consequence of traumatic insult and a secondary injury that results from a cascade of molecular and cellular events, including cell death, axonal injury, and inflammation [Jordan (2000)].

The Trail Making Test (TMT) is a valuable tool for assessing cognitive impairments in individuals with TBI. TMT-A primarily evaluates visual attention and pro-

cessing speed, while TMT-B assesses higher-order cognitive functions such as cognitive flexibility and task switching. In TBI assessment, longer completion times and increased error rates on both parts of the test indicate greater cognitive dysfunction, reflecting challenges in attention, processing speed, and executive functions. TMT results provide clinicians with objective measures to diagnose TBI-related cognitive impairments, tailor interventions, monitor treatment progress, and evaluate rehabilitation outcomes.

2.1.2 Spinal cord injury

Spinal Cord Injury (SCI) is a traumatic event with profound and life-altering consequences. It occurs when the spinal cord sustains damage, typically from a sudden blow or impact to the spine. This damage can result in a loss of sensation, movement, and function below the level of the injury. Beyond this, there are social, emotional, and economic consequences for patients, their families, and society at large (eg, employment, relationships, community access, isolation) [Kirshblum et al. (2007)].

The American Spinal Injury Association (ASIA) Impairment Scale (AIS) and injury level are crucial elements in the assessment and classification of SCI patients [Aarabi et al. (2017)]. The ASIA score provides a standardized method for evaluating the severity of SCI based on motor and sensory function, while the injury level indicates the specific vertebral level affected by the injury. The ASIA score categorizes SCI into five grades:

1. AIS A: Complete loss of motor and sensory function below the level of injury.

2. AIS B: Sensory function preserved, but no motor function below the neurological level, including the sacral segments S4-S5.

3. AIS C: Motor function preserved below the neurological level, and more than

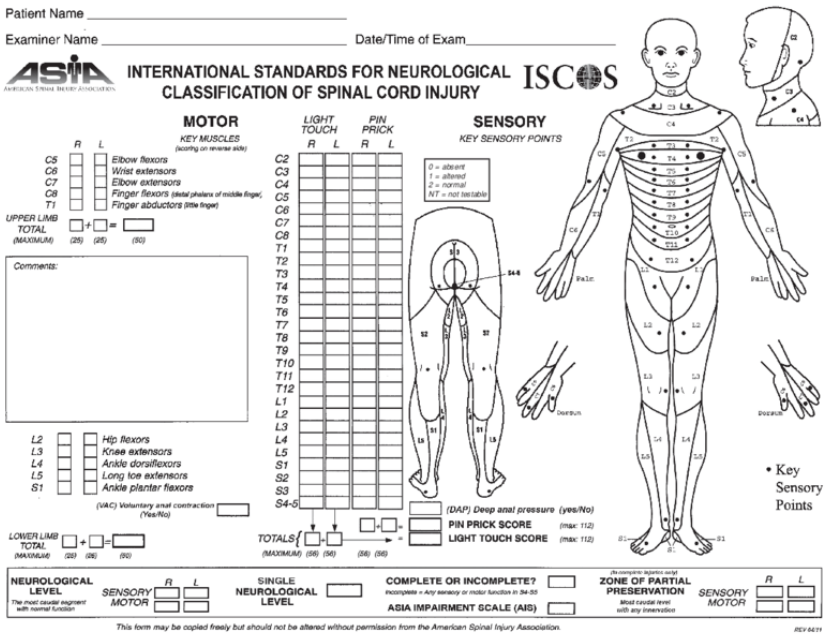

Figure 1: The ASIA examination to evaluate motor and sensory function in a patient with spinal cord injury. The ASIA test helps classify the severity of injury and guides treatment decisions for optimal patient care [Aarabi et al. (2017)].

half of the key muscles below the neurological level have a muscle grade lower than median score.

4. AIS D: Motor function preserved below the neurological level, and at least half of the key muscles below the neurological level have a muscle grade of higher than median score.

5. AIS E: Normal motor and sensory function.

The injury level indicates the specific vertebral level affected by the spinal cord injury. It is determined based on the neurological level of injury (NLI), which corresponds to the most caudal segment of the spinal cord with normal sensory and motor function on both sides of the body. The NLI is determined through sensory and motor examination following the ISNCSCI (International Standards for Neurological Classification of Spinal Cord Injury) guidelines (Figure 1). The ASIA score

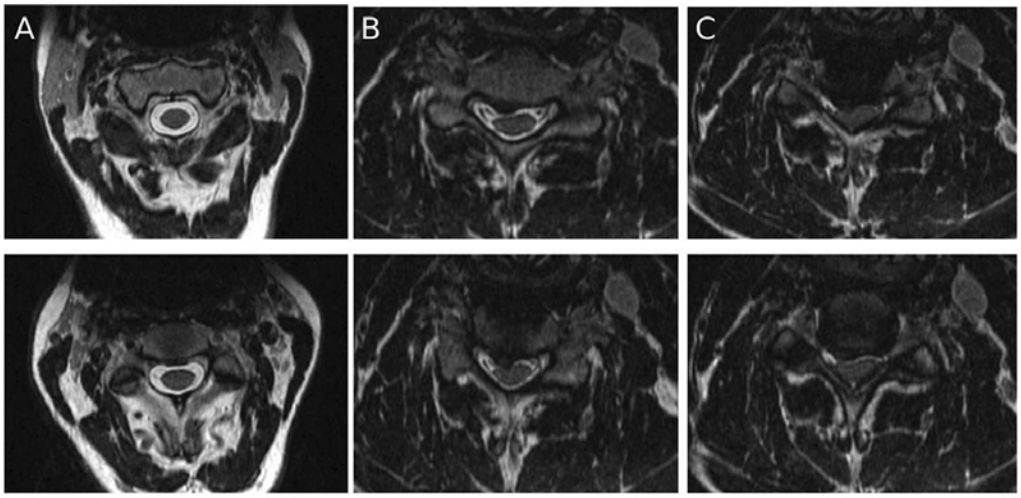

Figure 2: MRI images showing (A) normal spinal cord, (B) partial spinal cord compression, and (C) circumferential spinal cord compression [Merali et al. (2021)].

and injury level provide valuable information for clinicians to establish prognosis, guide treatment decisions, monitor patient progress, and evaluate the effectiveness of interventions.

2.1.3 Degenerative Cervical Myelopathy

Degenerative Cervical Myelopathy (DCM), also referred to as cervical spondylotic myelopathy, is a prevalent and progressive neurological condition. It occurs when the spinal cord in the neck becomes compressed due to degenerative changes in the cervical spine. This compression can result from a variety of factors, including disc herniation, bone spurs, and thickening of ligaments [Karadimas et al. (2013)]. Figure 2 shows normal and compressed spinal cord due to DCM.

2.2 Summary

Neurological disorders encompass a diverse range of conditions affecting both the central and peripheral nervous systems, imposing significant burdens globally. In this chapter, we explored the complexities of the human nervous system and common neu-

rological disorders, focusing on their classification, causes, and clinical presentations.

The human nervous system, consisting of the central nervous system (CNS) and peripheral nervous system (PNS), regulates vital physiological functions. Understanding these systems' organization is crucial for grasping neurological disorder pathophysiology.

Neurological disorders are classified based on location, dysfunction type, and causes, requiring tailored assessment and treatment. Traumatic brain injury (TBI), spinal cord injury (SCI), and degenerative cervical myelopathy (DCM) are notable examples due to their prevalence and profound impacts. TBI, resulting from various traumas, leads to diverse physical, cognitive, and emotional changes, necessitating comprehensive management. SCI, often caused by spinal trauma, results in sensory and motor loss, requiring specialized care. DCM, marked by spinal cord compression, poses diagnostic and therapeutic challenges, underscoring the importance of early intervention.

CHAPTER 3

Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that has revolutionized the field of neurology, providing clinicians with unparalleled insights into the structure and function of the brain and spinal cord [Hu et al. (2022)]. In this chapter, we go through fundamentals of MRI, exploring its diverse sequences, clinical applications, challenges encountered, and its profound impact on modern healthcare.

3.1 Principles of MRI

MRI works by harnessing the natural magnetic properties of atoms within the body, particularly hydrogen atoms, which are abundant in water and fat molecules. At its core, MRI relies on the interaction between radio waves and the magnetic field created by a large magnet. When a patient is placed within the MRI machine, their body's hydrogen atoms align with this magnetic field. Radiofrequency pulses are then applied, causing these hydrogen atoms to temporarily deviate from their aligned position.

Once the radiofrequency pulse is turned off, the hydrogen atoms gradually return to their original alignment with the magnetic field. As they do so, they emit signals that are picked up by the MRI machine's receiver coils. These signals are then processed by a computer to generate detailed images of the body [Hu et al. (2022)].

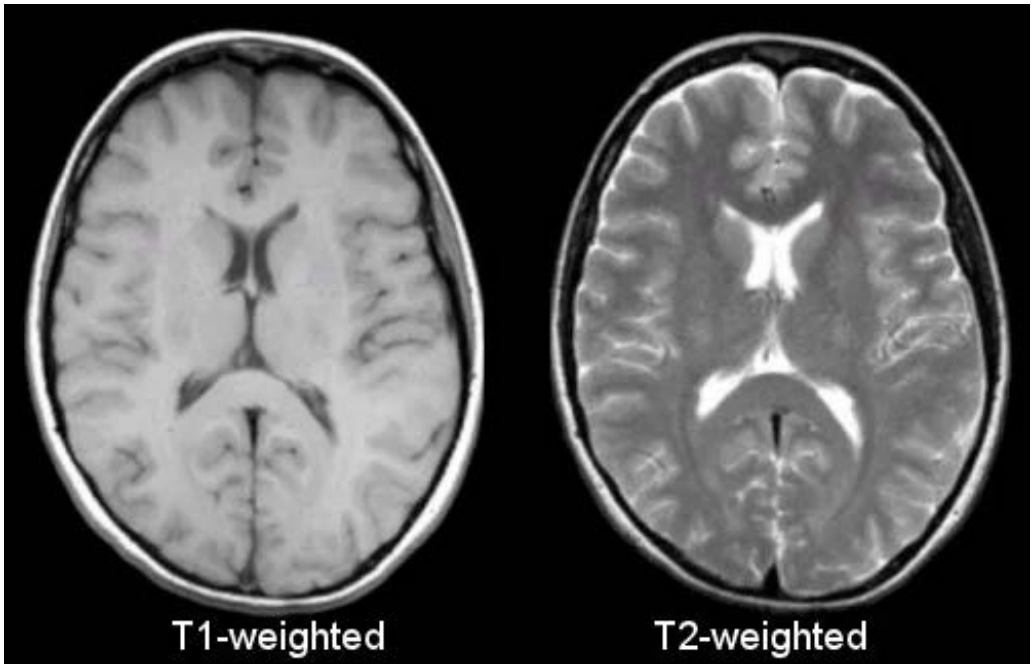

Figure 3: Spin echo sequences of brain [Trifan et al. (2017)].

3.2 MRI Sequences

MRI sequences are diverse imaging techniques designed to highlight specific tissue characteristics and physiological processes. These sequences are essential for generating high-quality images used in the diagnosis and management of various medical conditions. Understanding the principles and applications of MRI sequences is crucial for optimizing image acquisition protocols and interpreting results accurately.

At the core of MRI sequences lies the manipulation of magnetic field gradients and radiofrequency pulses to generate image contrast. By altering these parameters, MRI sequences can selectively highlight different tissue properties, such as relaxation times, diffusion rates, and perfusion characteristics [Hu et al. (2022)].

3.2.1 Spin Echo Sequences

One of the most commonly used categories of MRI sequences is spin echo sequences. These include T1-weighted and T2-weighted imaging techniques, each pro-

viding unique contrast properties (Figure 3).

T1-weighted imaging is a MRI technique that offers detailed anatomical information about the brain and spinal cord. The term "T1-weighted" indicates that the image is weighted to emphasize differences in the T1 relaxation times of various tissues. T1 relaxation time, also known as spin-lattice relaxation time, is the time it takes for the protons in a tissue to realign with the magnetic field after being perturbed by radiofrequency (RF) pulses during an MRI scan. It reflects the rate at which energy is transferred from excited protons back to the surrounding lattice or "spin lattice". In T1-weighted images, cerebrospinal fluid (CSF) typically appears dark, while gray matter and white matter appear bright. This characteristic makes T1-weighted images particularly useful for identifying abnormalities such as tumors, hemorrhages, and lesions. These abnormalities can cause discernible alterations in the appearance of gray and white matter in the image, aiding in diagnostic interpretation and treatment planning.

T2-weighted imaging is another MRI technique used to visualize anatomical structures within the brain and spinal cord. Similar to T1-weighted imaging, T2-weighted images are named as such because they are weighted to highlight differences in the T2 relaxation times of various tissues. T2 relaxation time, also known as spin-spin relaxation time, is the time it takes for the transverse magnetization (or "spin-spin" interactions) to decay after RF pulses are applied during an MRI scan. It reflects the rate at which energy is lost through interactions between neighboring protons. In T2-weighted images, cerebrospinal fluid (CSF) appears bright, while gray matter and white matter exhibit varying shades of gray. This characteristic makes T2-weighted imaging particularly useful for detecting abnormalities such as edema, inflammation, and demyelination. These pathological changes often manifest as alterations in the signal intensity of gray and white matter, enabling clinicians to identify and evaluate conditions like multiple sclerosis, infections, and vascular lesions

[Hu et al. (2022)].

3.2.2 Diffusion Sequences

Diffusion-weighted imaging (DWI) is a MRI technique that captures the motion of water molecules within biological tissues. By measuring the random Brownian motion of water molecules, DWI provides insights into tissue microstructure and integrity. In healthy tissue, water molecules exhibit relatively unrestricted movement, whereas in regions with barriers such as cell membranes or fiber tracts, water diffusion is constrained. DWI detects these variations in water diffusion by quantifying the magnitude and direction of diffusion within each voxel of the image. The degree of diffusion restriction is measured using a parameter known as the apparent diffusion coefficient (ADC), which reflects tissue microstructural properties. Areas with high cellularity, such as tumors, typically exhibit lower ADC values due to increased cellular density and restricted water diffusion. Conversely, regions of free water diffusion, such as cerebrospinal fluid-filled spaces, display higher ADC values. By analyzing these diffusion characteristics, DWI provides clinicians with valuable information about tissue integrity and pathology, facilitating the early detection and characterization of various neurological disorders, including strokes, tumors, and demyelination in conditions such as multiple sclerosis .

Diffusion Tensor Imaging (DTI) is an extension of DWI that specifically focuses on mapping the orientation and direction of water diffusion within tissues. DTI essentially acts as a compass, discerning not only how water moves but also its preferred directionality within biological structures (4). This capability allows for the visualization and mapping of white matter tracts in the brain, crucial for understanding brain connectivity. DTI achieves this by applying multiple diffusion-sensitizing gradients in different directions, which enable the estimation of a diffusion tensor at each voxel in the image. The diffusion tensor represents the magnitude and

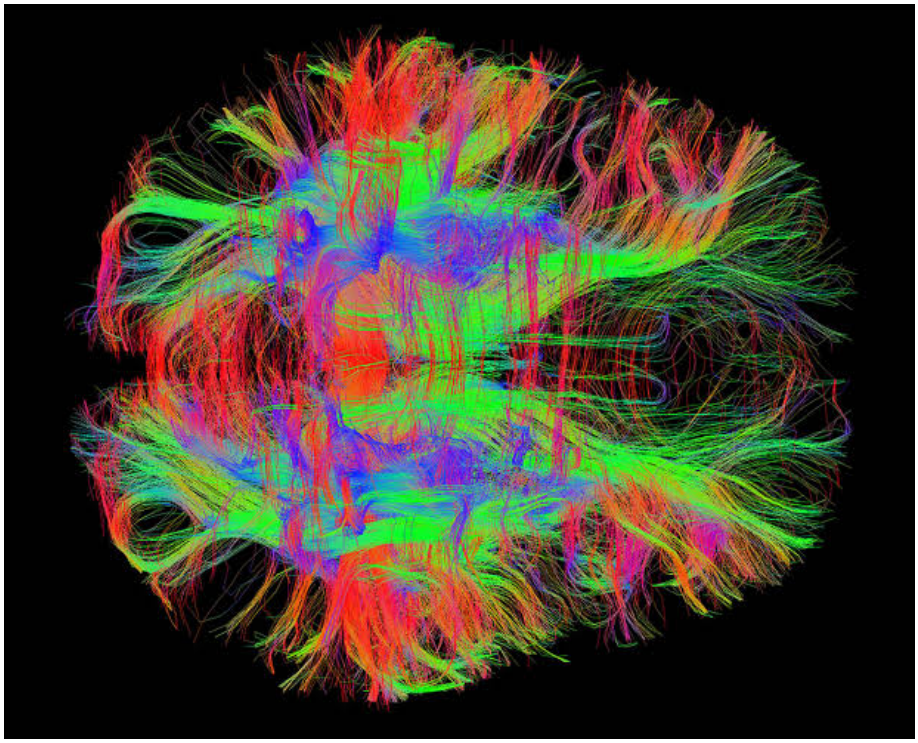

Figure 4: White matter fiber tracts in the adult human brain [Saygin et al. (2017)].

directionality of water diffusion, providing information about the underlying tissue microstructure and organization. By analyzing the principal eigenvector of the diffusion tensor, DTI can determine the primary direction of water diffusion within each voxel, thus facilitating the reconstruction of white matter fiber tracts. This ability to map white matter pathways is particularly valuable in studying various neurological conditions, including traumatic brain injury and neurodegenerative diseases, where alterations in white matter integrity are evident [Hu et al. (2022)].

Within DTI, metrics like Fractional Anisotropy (FA), Mean Diffusivity (MD), Radial Diffusivity (RD), and Axial Diffusivity (AD) are commonly analyzed. FA quantifies the degree of anisotropy of water diffusion, particularly in white matter. Anisotropy refers to the directional dependence of diffusion, indicating the extent to which water molecules preferentially diffuse along specific orientations rather than equally in all directions. It ranges from 0 to 1, where 0 represents isotropic diffusion (equal diffusion in all directions) and 1 represents perfectly anisotropic diffusion (dif-

fusion along a single direction). MD reflects the average diffusion magnitude in all directions within a voxel. It is sensitive to changes in tissue microstructure, such as alterations in cell density, edema, or tissue damage. Increased MD values are often associated with tissue pathology, inflammation, or neuronal loss. RD measures diffusion perpendicular to fiber bundles and is indicative of myelin integrity. Elevated RD values may suggest demyelination or damage to the myelin sheath surrounding axonal fibers. AD measures diffusion along fiber bundles, reflecting axonal integrity. Changes in AD values can indicate axonal degeneration, regeneration, or alterations in axonal density. Decreased AD values may signify axonal loss, while increased AD values may suggest axonal regeneration or compensatory changes in response to injury or disease. Together, FA, MD, RD, and AD provide comprehensive insights into white matter microstructure, crucial for understanding neurological conditions and assessing treatment efficacy [Hu et al. (2022)].

Neurite Orientation Dispersion and Density Imaging (NODDI) is a sophisticated MRI technique that goes beyond DTI by providing detailed information about tissue microstructure. NODDI distinguishes between different types of barriers that restrict water movement, such as cell membranes and myelin, and quantifies the density and orientation of neurites (axons and dendrites) within the brain. By characterizing the complexity of tissue microstructure, NODDI contributes to our understanding of neurological disorders such as Alzheimer's disease, offering insights into alterations in neuronal architecture beyond what DWI and DTI can provide.

NODDI introduces metrics like Orientation Dispersion Index (ODI) and Volume Fraction of Cerebrospinal Fluid (Vic), which further enhance our ability to assess tissue composition and understand neuroplasticity and structural changes in neurological conditions. ODI quantifies the dispersion of neurite orientations within a voxel. Higher ODI values suggest greater complexity or dispersion of neurite orientations, potentially indicating increased neuroplasticity or structural alterations in

neurological disorders. ODI is particularly useful for assessing white matter integrity, as changes in neurite orientation dispersion may reflect alterations in axonal density or myelination. Vic represents the volume fraction of cerebrospinal fluid within a voxel. Vic shows how much of the MRI voxel is filled with isotropic tissue components, like CSF or regions with freely moving water molecules, versus tissue with organized structures like neurites. Higher Vic values indicate a larger proportion of isotropic tissue, suggesting areas with low cellular density or high extracellular space. These NODDI metrics offer valuable insights into the underlying microstructural changes associated with neurological disorders, aiding in diagnosis, prognosis, and treatment evaluation [Zhang et al. (2012)].

3.3 MRI in Clinical Applications

MRI is indispensable in clinical practice for diagnosing and monitoring neurological disorders [Stewart et al. (1992)]. Various MRI sequences are being used for specific clinical scenarios, each offering unique advantages for evaluating different neurological conditions. For example, T1-weighted images are ideal for assessing anatomy and detecting structural abnormalities, while T2-weighted and FLAIR (Fluid-attenuated inversion recovery) sequences are sensitive to pathology such as edema, inflammation, and demyelination.

The interpretation of MRI findings extends beyond visual inspection of images; it involves integrating imaging data with clinical information and ancillary tests to formulate accurate diagnoses and prognoses. Quantitative analysis techniques, such as volumetry and texture analysis, provide objective measures of tissue characteristics and disease burden, facilitating disease monitoring and treatment response assessment. For example, in patients with multiple sclerosis (MS), volumetric analysis of brain lesions can quantify the extent of disease progression and assess treatment efficacy. Additionally, texture analysis of MRI images can differentiate between different

tissue types and provide insights into tissue heterogeneity, aiding in the characterization of tumors and predicting treatment outcomes. Furthermore, machine learning algorithms trained on large datasets can aid in automated lesion segmentation and classification, accelerating image analysis and reducing inter-observer variability. For instance, in the context of brain tumor diagnosis, machine learning models trained on annotated MRI images can automatically segment tumors and classify different tumor subtypes based on their imaging features. This not only streamlines the diagnostic process but also enhances consistency and accuracy in lesion identification, leading to more reliable treatment planning and patient management [Davatzikos (2019)].

Despite technological advancements, MRI remains susceptible to various artifacts that can confound image interpretation. Understanding the underlying causes of artifacts and employing appropriate mitigation strategies are essential for obtaining diagnostically accurate images. For instance, motion artifacts, which are commonly encountered in pediatric or claustrophobic patients, can be minimized through proper patient preparation and sedation techniques. For example, pediatric patients may require sedation to reduce movement during scanning, while claustrophobic patients may benefit from pre-scan counseling and the use of open MRI systems to alleviate anxiety and minimize motion. Moreover, advanced motion correction techniques integrated into MRI scanners, such as prospective motion correction and real-time tracking, can further mitigate motion artifacts by adjusting imaging parameters during data acquisition to compensate for patient movement. Susceptibility artifacts are another common issue, often arising from metallic implants or air-tissue interfaces. For instance, in patients with dental implants or orthopedic hardware, susceptibility artifacts can cause signal distortion, making it challenging to assess nearby anatomical structures. In such cases, careful sequence selection and optimization are necessary to minimize artifacts. Techniques like using specialized sequences such as susceptibility-weighted imaging (SWI) or adjusting imaging parameters to reduce susceptibility

effects can help mitigate signal distortion and improve image quality. Additionally, employing alternative imaging modalities like CT or ultrasound may be considered for certain patients with metallic implants to avoid MRI-related artifacts altogether. Furthermore, the use of advanced post-processing algorithms, such as artifact correction algorithms or image reconstruction techniques, can aid in reducing susceptibility artifacts and enhancing image quality, particularly in challenging cases [T. B. Smith (2010)].

3.4 Summary

MRI is a versatile imaging modality widely used in clinical practice for evaluating neurological disorders. Understanding the principles of MRI sequences and their clinical applications is crucial for interpreting imaging findings accurately. Spin Echo sequences provide detailed anatomical information, while diffusion sequences offer insights into tissue microstructure and pathology. MRI plays a central role in diagnosing and monitoring neurological conditions, providing valuable information for treatment planning and patient management.

CHAPTER 4

Machine Learning Fundamentals

Machine learning (ML) is a subfield of artificial intelligence (AI) that focuses on the development of algorithms and models capable of learning from data to make predictions or decisions without being explicitly programmed. ML techniques have become increasingly prevalent across various industries due to their ability to extract insights from large datasets and automate complex tasks. In this chapter, we provide a brief overview of machine learning fundamentals, covering both traditional machine learning techniques and deep learning methods [P. Wang et al. (2021)]. For each method, we will explore specific examples of machine learning algorithms utilized in the papers summarized in this report in the following chapter.

4.1 Traditional Machine Learning

Traditional machine learning methods are algorithms that learn patterns and make predictions based on input data. These methods often rely on statistical techniques to identify relationships and patterns in the data. Here, we discuss some fundamental traditional ML algorithms, including Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Linear Regression.

4.1.1 Support Vector Machines

Support Vector Machines (SVM) is a supervised learning algorithm used for classification tasks. In SVM, the algorithm aims to find the hyperplane that best separates the data points into different classes in feature space. The hyperplane is chosen to

maximize the margin, which is the distance between the hyperplane and the nearest data points from each class, also known as support vectors [Chandra & Bedi (2021)].

For linearly separable data, the decision boundary is represented by the equation of a hyperplane:

$$f(x) = w^T x + b, \tag{4.1-1}$$

Where:$f(x)$ is the decision function, w is the weight vector, x is the input feature vector, and b is the bias term.

The optimization objective of SVM is to maximize the margin while minimizing the classification error. This optimization problem can be formulated as a convex quadratic programming problem.

4.1.2 k-Nearest Neighbors

k-Nearest Neighbors (KNN) is a simple and intuitive algorithm used for both classification and regression tasks. In KNN, the prediction for a new data point is based on the majority class (for classification) or the average value (for regression) of its k nearest neighbors in feature space [Guo et al. (2003)].

For classification, the predicted class label of a new data point is determined by a majority vote among its k nearest neighbors. The distance metric used to measure the similarity between data points is typically Euclidean distance, although other distance metrics such as Manhattan or Minkowski distance can also be used.

The value of k is a hyperparameter that needs to be tuned during model training. A smaller value of k results in a more flexible model with higher variance, while a larger value of k leads to a smoother decision boundary with lower variance.

4.1.3 Linear Regression

Linear Regression is a supervised learning algorithm used for predicting continuous target variables based on one or more input features [Su et al. (2012)]. In linear

regression, the relationship between the input features X and the target variable y is modeled using a linear function:

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n, \tag{4.1-2}$$

Where y is the target variable, $x_1, x_2, ..., x_n$ are the input features, and $\beta_0, \beta_1, \beta_2, ..., \beta_n$ are the coefficients (parameters) of the model.

The goal of linear regression is to learn the values of the coefficients that minimize the residual sum of squares between the observed and predicted values of the target variable. Linear regression assumes a linear relationship between the input features and the target variable, which may not always hold true in practice. However, it remains a powerful and widely used algorithm due to its simplicity and interpretability.

4.1.4 Decision Trees

Decision Trees are a versatile and interpretable supervised learning algorithm used for both classification and regression tasks. In decision trees, the data is split recursively into subsets based on the values of input features, with the goal of maximizing the homogeneity (purity) of the resulting subsets [Suthaharan (2016)].

At each node of the tree, the algorithm selects the feature and the split point that best separates the data into different classes (for classification) or reduces the variance of the target variable (for regression). This process is repeated recursively until a stopping criterion is met, such as reaching a maximum tree depth or minimum number of data points in a leaf node.

The decision tree algorithm creates a tree-like structure where each internal node represents a decision based on a feature, each branch represents the outcome of the decision, and each leaf node represents the predicted class label or value of the target variable.

4.1.5 Random Forests

Random Forests are an ensemble learning technique that combines multiple decision trees to improve the predictive performance and robustness of the model [Rigatti (2017)]. In a Random Forest, each decision tree is trained on a random subset of the training data and a random subset of the input features.

During training, each decision tree is built independently, and the final prediction is obtained by aggregating the predictions of all trees (e.g., taking the majority vote for classification or averaging the predictions for regression).

4.2 Deep Learning

Deep learning is a subset of machine learning that uses artificial neural networks with multiple layers to learn complex patterns in large amounts of data. Deep learning algorithms have shown remarkable performance in various tasks such as image recognition, natural language processing, and speech recognition. Here, we discuss two fundamental deep learning architectures: Convolutional Neural Networks (CNN) and Residual Networks (ResNet) [P. Wang et al. (2021)].

4.2.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a class of deep neural networks particularly well-suited for processing and analyzing structured grid data, such as images [O'shea & Nash (2015)]. CNNs are composed of multiple layers, including convolutional layers, pooling layers, and fully connected layers.

Convolutional layers are the building blocks of CNNs and consist of filters (also known as kernels) that slide over the input image, performing convolution operations to extract local features. Each filter detects specific patterns or features in the input image, such as edges, textures, or shapes.

Considering a convolutional layer with L filters and denoting the input of the convolutional layer by $\mathbf{C}$, we can express the lth convolutional map, $\mathbf{O}^{(l)}$, corresponding to the lth filter as

$$\mathbf{O}^{(l)} = \sigma(\mathbf{C} * \mathbf{f}^{(l)}), \tag{4.2-3}$$

where '$*$' denotes 2-D convolution, σ is the activation function, and $\mathbf{f}^{(l)}$ is the lth 2-D convolutional filter.

Pooling layers are used to reduce the spatial dimensions of the feature maps while preserving the most important information. Max pooling and average pooling are common pooling operations used in CNNs. Fully connected layers are used to make predictions based on the high-level features extracted by the convolutional and pooling layers. These layers connect every neuron in one layer to every neuron in the next layer.

4.2.2 Residual Networks

Residual Networks (ResNet) are a type of deep neural network architecture introduced to address the vanishing gradient problem in very deep neural networks. The key innovation of ResNet is the use of skip connections, also known as identity mappings, to enable the flow of information through the network more effectively [Koonce (2021)].

In a standard feedforward neural network, each layer learns a mapping from its input to its output. However, as the network becomes deeper, it becomes increasingly difficult to optimize the training objective due to the vanishing gradient problem. Vanishing gradient is a phenomenon that occurs during the training of deep neural networks, where the gradients of the loss function with respect to the network parameters become very small as they propagate backward through the network layers. This can significantly slow down or even halt the training process, as small gradients lead to negligible updates to the network weights, making it difficult for the model

to learn meaningful representations from the data.

Backpropagation is the primary algorithm used to train neural networks by iteratively updating the network parameters (weights) based on the gradients of the loss function with respect to those parameters. During backpropagation, gradients are computed by recursively applying the chain rule of calculus to propagate the error from the output layer back through the network layers.

The vanishing gradient problem arises when the gradients diminish as they propagate backward through many layers of the network. This is often caused by the use of activation functions with limited range, such as sigmoid or tanh functions, which saturate for large or small inputs, leading to gradients close to zero. As a result, the updates to the weights in earlier layers become negligible, hindering the training process.

Residual Networks (ResNet) address the vanishing gradient problem by introducing skip connections, also known as identity mappings, that bypass one or more layers and directly feed the input of a layer to its output. This allows the network to learn residual mappings, capturing the difference between the desired output and the input to the skipped layers. By enabling the flow of information through the network more effectively, ResNet architectures can effectively train very deep networks without suffering from the vanishing gradient problem.

Mathematically, the output of a residual block in a ResNet can be expressed as follows:

$$y = F(x, W) + x, \tag{4.2-4}$$

Where x is the input to the residual block, $F(x, W)$ is the residual mapping learned by the block, and W are the parameters (weights) of the block.

ResNet architectures typically consist of several residual blocks stacked on top of each other, with shortcut connections (skip connections) between blocks. This allows ResNet to effectively train very deep networks, with hundreds or even thousands of

layers, without suffering from the vanishing gradient problem.

4.3 Summary

In this chapter, we have explained the fundamentals of machine learning (ML), a vital subset of artificial intelligence (AI) revolutionizing various industries. We explored traditional ML techniques such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Linear Regression, Decision Trees, and Random Forests, each offering unique strengths in pattern recognition and prediction tasks. Additionally, we introduced deep learning methodologies, focusing on Convolutional Neural Networks (CNNs) and Residual Networks (ResNet), which excel in processing complex data structures like images and sequences.

CHAPTER 5

Machine Learning and MRI Analysis

As discussed in previous chapters, MRI scans serve as a key indicator for clinical decision-making about neurological disorders. Traditional methods for diagnosing neurological disorders rely on clinical assessment, manual interpretation of radiological imaging, and neurological examination. For instance, consider a patient presenting with symptoms suggestive of Degenerative Disc Disease (DDD). Clinicians may conduct a physical examination and order an MRI scan to assess the severity of spinal cord compression. However, manually analyzing MRI images to differentiate between benign age-related changes and pathological compression can be challenging. Moreover, variations in interpretation among radiologists can further complicate diagnosis [S. Wang & Summers (2012)].

Quantitative analysis of MRI represents a powerful approach for extracting objective measurements of tissue properties and physiological parameters from MRI data [Tofts (2005)]. There is growing recognition that computer models may assist in the initial interpretation of medical imaging studies and rapidly flag studies with pathologic findings. Machine learning (ML) methods have been demonstrated to be effective for various medical purposes [Chong et al. (2015), Mohamed et al. (2022), VergaraVictor et al. (2017)]. These techniques offer a promising approach to enhance the detection and characterization of neurological disorders from MRI scans. Machine learning models can automatically extract relevant features from MRI images and be helpful for diagnosis. These models can also aid in quantifying the degree of severity of disorders, providing valuable information for treatment planning and monitoring

disease progression [Lundervold & Lundervold (2019).

In this chapter, we explain the principles, techniques, and clinical applications of quantitative analysis in MRI, highlighting its significance in advancing medical diagnostics and research.

5.1 Structural MRI

As discussed in Chapter 3, structural MRI images are widely used in clinical practice for diagnosing and monitoring neurological disorders due to their high resolution and ability to provide detailed anatomical information. In recent years, there has been a growing interest in using these images for developing models to better understand the structure of the nervous system and predict neurological disorders. These models leverage machine learning algorithms to analyze large datasets of structural images and identify patterns associated with specific conditions. In this section, we will provide an overview of related works by going through a recently published paper, "A deep learning model for detection of cervical spinal cord compression in MRI scans" [Merali et al. (2021)] and analyze the method, findings, and implications. At the end of this section, we will explore the gaps in structural image-based models.

5.1.1 Background

Machine learning techniques have garnered significant attention in recent years for their potential to enhance the analysis of structural MRI (sMRI) images in various neurological and psychiatric disorders [Lundervold & Lundervold (2019)].

Brain disorders, such as Alzheimer's disease and Parkinson's disease, are among the most common and extensively studied neurological conditions. Structural MRI plays a crucial role in diagnosing and monitoring these disorders, as it can detect changes in brain structure and volume that are characteristic of these conditions.

Spinal cord disorders, on the other hand, are less studied and often overlooked in

neuroimaging research. This is partly due to the technical challenges associated with imaging the spinal cord, which is smaller and more susceptible to motion artifacts compared to the brain. However, recent advancements in MRI technology, such as high-resolution imaging and motion correction techniques, have made it possible to obtain high-quality images of the spinal cord. As a result, there is a growing interest in using structural MRI to study spinal cord disorders.

Deep convolutional neural networks have shown promise in this area and have been tested in a variety of pathology categories. Researchers have made use of convolutional neural networks (CNNs) for automated segmentation of spinal imaging [Michopoulou et al. (2009)]. In addition, previous studies have utilized computer vision methods, including CNNs, to extract quantitative parameters from cervical spinal cord MRI scans [Jin et al. (2016)]. The majority of studies have focused on classifying MRI images of the lumbar spine [Castro-Mateos et al. (2014)].

5.1.2 Case-Study 1: DCM Classification

As explained in chapter 2, degenerative cervical myelopathy (DCM) is a prevalent condition characterized by progressive compression of the cervical spinal cord, resulting in substantial functional impairment and diminished quality of life for affected individuals. With the increasing volume of medical imaging data, there is a growing recognition of the potential for deep learning models to assist in interpreting MRI scans, particularly in primary-care settings. In this section, we will explain a recently published paper, "A deep learning model for detection of cervical spinal cord compression in MRI scans" [Merali et al. (2021)]. This study aimed to develop and validate a deep learning model specifically for detecting cervical spinal cord compression in MRI scans, with the potential to improve the efficiency and objectivity of interpretation.

Data Acquisition

T2-weighted MRI data were obtained retrospectively from 605 patients enrolled in the AO Spine CSM North America (CSM-NA) or AO Spine CSM International (CSM-I) clinical studies. These studies had prior institutional research ethics board approval and were conducted in accordance with relevant guidelines. Eligible patients, meeting specific criteria including age, evidence of cervical spinal cord compression, and symptomatic DCM, were included (289 patients). MRI studies were collected in DICOM format, accompanied by baseline clinical data such as demographic information and modified Japanese Orthopedic Association (mJOA) scores, a standardized measure used to assess the severity of symptoms in patients with DCM.

Data Pre-processing and Labeling

A comprehensive data pre-processing protocol was implemented to prepare the MRI data for analysis. This included anonymization of MRI studies, conversion of axial T2-weighted sequences into JPEG images, and normalization of pixel values. The anonymization process removed patient identifiers using specialized software, ensuring data privacy. The acquired MRI images were then converted into JPEG format and down-sampled to a uniform size of 299×299 pixels. Subsequently, pixel values were normalized between 0 and 1 to facilitate consistent analysis. Importantly, the heterogeneity in the number of MRI slices was preserved to maintain dataset diversity.

Following pre-processing, two senior neurosurgical residents independently examined each image to label spinal cord compression as "compressed" or "non-compressed." Inter-rater reliability was assessed using Cohen's kappa metric, and discrepancies were resolved through consensus discussion. This rigorous labeling process ensured the accuracy and reliability of the labeled dataset, serving as the ground-truth reference for model development and validation.

Dataset	Training/Validation (n = 201)	Holdout (n = 88)	p-value
Age (median)	55	56	0.65
Gender (male)	63%	66%	0.53
Baseline mJOA (median)	13	13	0.72
MRI Scanner Manufacturer	GE Medical Systems (n = 98) Siemens (n = 12)	GE Medical Systems (n = 92) Siemens (n = 66) Philips Medical Systems (n = 21)	0.12
Slice thickness median (range)	3 (2–5) mm	3 (2–5) mm	0.82
Voxel size median (range)	0.3516 (0.2539–0.7813) mm	0.3516 (0.2539–0.7813) mm	0.65

Table 1: Demographics of patients and scanner parameters in the training/validation and holdout datasets.

Two distinct patient cohorts were defined: a training/validation dataset and a holdout dataset. The training/validation dataset comprised 75% of patients from each site (201 patients), chosen randomly, while the remaining 25% constituted the holdout dataset (88 patients). Demographic information from each dataset is seen in Table 1. Patient age, gender, baseline mJOA, MRI scanner manufacturer, and MRI image parameters, did not significantly differ between the Training/Validation Dataset and the Holdout Dataset.

Model Architecture and Training

The ResNet-50 convolutional neural network architecture was utilized in this study. Transfer learning was employed, where the initial weights for the network were adapted from those developed during the ImageNet competition Russakovsky et al. (2014). Transfer learning involves initializing model weights from a pre-trained model, facilitating faster tuning of the model for specific tasks. In this case, the weights from a model trained on a different larger dataset will be used to start the training, instead of radom initialization, leveraging the learned representations of simple image features like edges.

The fully connected layers of the ResNet-50 network were replaced with a set of fully connected layers with randomly initialized weights. The number of neurons in these fully connected layers was varied to explore different network configurations.

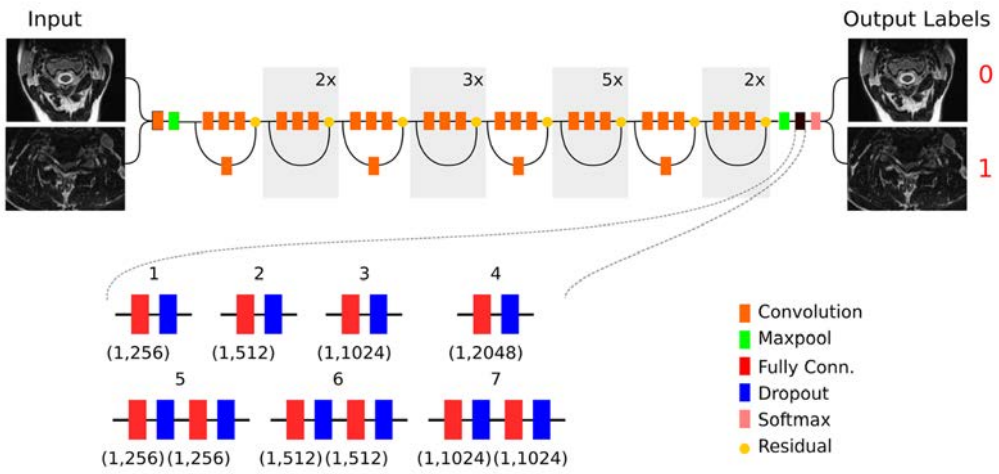

Figure 5: Overview of the convolutional neural network model architecture. The convolutional layers (orange) were derived from the Resnet-50 model, while the fully connected layers were modified for the classification task. Seven separate model configurations were tested as shown.

Dropout layers were also incorporated to mitigate overfitting during training. Figure 5 shows the model.

Several network configurations were tested, each varying in the number of fully connected layers and neurons, as well as the dropout rate. These configurations were evaluated using the training/validation dataset, with performance metrics such as binary cross-entropy loss and accuracy monitored during training. The Adam optimizer was used during model training, with hyperparameters such as learning rate and momentum tuned using a random search strategy.

After training, the best performing model architecture was selected based on its performance on the validation set. This model was then tested on a holdout dataset, and various performance metrics including area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and f1 score were calculated.

Additionally, class activation maps (CAMs) were generated to visualize the regions of the image that contributed most to the model's classification decisions. CAMs help to interpret the model's behavior and identify relevant image features associated with each class.

Model Architecture	Epochs	Validation Accuracy	Validation Loss	Training Accuracy	Training Loss
Model 1	50	89.88%	0.4223	97.12%	0.0892
Model 2	50	90.95%	0.3093	99.13%	0.0219
Model 3	50	91.09%	0.3979	98.99%	0.0292
Model 4	54	91.53%	0.3562	99.10%	0.0260
Model 5	62	90.78%	0.2932	99.12%	0.0238
Model 6	77	92.41%	0.2569	99.03%	0.0284
Model 7	69	92.23%	0.2593	99.23%	0.0234

Table 2: Comparison of model performance during training.

Results

The Training/Validation Dataset, consisting of 201 patients, was utilized for model training. Seven neural network configurations were trained, with Model 6 achieving the highest accuracy and lowest binary cross-entropy loss on the validation set (Table 2). Model 6, based on the ResNet-50 CNN architecture, included two fully connected layers with 512 neurons each and two dropout layers with 30% dropout 5. Model 6 was tested on the holdout dataset, achieving an AUC of 0.94, sensitivity of 0.88, specificity of 0.89, and f1-score of 0.82. For each patient in the holdout dataset, the model's classification output was compared to human labelers' assignments, with a median AUC of 0.94 across patients (Table 3).

Class activation maps (CAMs) were generated to visualize the features contributing to the model's classifications. True positive examples showed activation over clinically relevant areas such as the spinal cord and cerebrospinal fluid (CSF) spaces, while false negatives sometimes relied on features outside the spinal canal or focused on relevant areas but resulted in incorrect classifications (Figure 6). The successful development and validation of the deep learning model represent a significant advancement in automated diagnostic tools for spinal disorders. By leveraging a large dataset of cervical spine MRI scans, the study has demonstrated the feasibility of training an existing convolutional neural network for this novel medical imaging classification task. The model's high performance metrics underscore its potential to enhance di-

	Area Under the Curve (SD)	p-value
Entire Holdout Dataset (n = 88)	0.94 (0.08)	
Age (years)		
< 40 (n = 9)	0.88 (0.14)	0.12
40–65 (n = 63)	0.95 (0.06)	0.78
> 65 (n = 16)	0.92 (0.09)	0.45
mJOA		
18 (n = 2)	1.00 (0)	0.94
15–17 (n = 22)	0.96 (0.04)	0.67
12–14 (n = 39)	0.92 (0.09)	0.62
< 12 (n = 25)	0.95 (0.07)	0.77
MRI Scanner Manufacturer		
GE Medical Systems (n = 52)	0.94 (0.07)	0.82
Siemens (n = 25)	0.93 (0.06)	0.71
Philips Medical Systems (n = 11)	0.95 (0.08)	0.74
Location		
North America (n = 33)	0.95 (0.07)	0.68
South America (n = 16)	0.93 (0.09)	0.67
Europe (n = 21)	0.91 (0.08)	0.78
Asia Pacific (n = 18)	0.93 (0.07)	0.81

Table 3: Model performance on the holdout dataset stratified by patient characteristics and MRI scanner manufacturer.

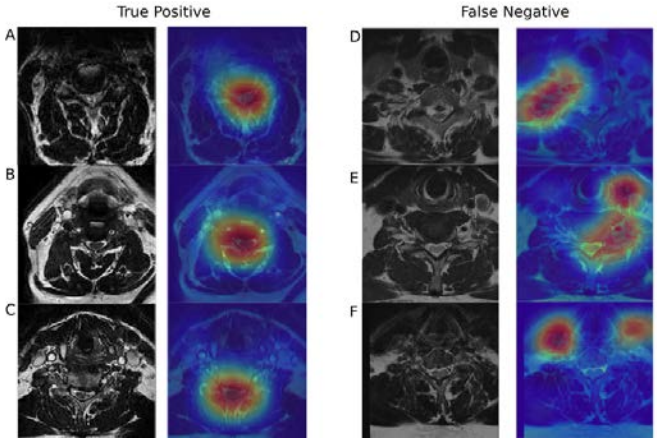

Figure 6: Class activation maps were generated for both correctly classified (true positives) and incorrectly classified (false negatives) example images. Blue indicates no activation, while red indicates maximal activation. In true positive cases, maximal activation typically occurred over the spinal canal and spinal cord. However, false negatives sometimes showed activation over irrelevant areas like paraspinal muscles or vascular structures. In some false negatives, activation occurred over the spinal cord and canal, but also over seemingly unrelated areas of the image.

agnostic workflows in clinical settings, providing rapid and accurate interpretation of MRI scans to aid clinical decision-making.

Limitations

The study's reliance on a dataset primarily composed of patients with confirmed Degenerative Cervical Myelopathy (DCM) who underwent surgical intervention could limit the generalizability of the model. This patient population may not fully represent the spectrum of DCM severity and may exclude asymptomatic or mildly symptomatic individuals who opt for conservative management. Additionally, the exclusion of patients who did not undergo surgical treatment may introduce selection bias and hinder the model's applicability to a broader patient cohort.

Furthermore, the study's binary classification system, which categorizes spinal cord segments as either compressed or non-compressed, may oversimplify the complexity of spinal cord compression severity. This dichotomous approach fails to account for variations in compression degree and the nuanced impact on clinical outcomes. Incorporating more nuanced classification schemes or severity scoring systems could provide a more comprehensive assessment of spinal cord compression severity and improve the model's predictive accuracy.

The study's reliance on subjective interpretations by radiologists for labeling may introduce variability and potential biases into the dataset. Objective measures or advanced imaging techniques, such as quantitative analysis algorithms or diffusion tensor imaging, could provide more standardized and reliable assessments of spinal cord abnormalities. Additionally, the use of healthy segments from patients' cords as normal labels may not fully capture the potential impact of spinal cord injuries on adjacent regions. Future studies should consider including healthy controls to establish a more robust baseline for comparison and enhance diagnostic accuracy.

5.2 Diffusion MRI

Traditional neuroimaging techniques such as structural MRI provide valuable anatomical information but often lack the sensitivity to detect subtle microstructural changes in the brain. Diffusion imaging techniques, including diffusion tensor imaging (DTI) and neurite orientation dispersion and density imaging (NODDI), offer unique insights into tissue microarchitecture and white matter integrity. By leveraging machine learning (ML) algorithms, researchers aim to harness the rich information encoded in diffusion images to improve the understanding and management of neurological disorders.

In this section, we will provide an overview of related works in the field,and then we will go through two recently published papers, "Machine learning-based classification of chronic traumatic brain injury using hybrid diffusion imaging" [Muller et al. (2023)] and "Quantitative Analysis in Cervical Spinal Cord Injury Patients Using Diffusion Tensor Imaging and Tractography" [Park et al. (2022)] and analyze the method, findings, and implications. At the end of this section, we will explore the limitations of these studies.

5.2.1 Background

Previous studies have demonstrated the potential of biomarkers derived from diffusion data, such as NODDI and DTI, in combination with machine learning techniques, for detecting neurodegenerative diseases like Alzheimer's abnormalities [Prasuhn et al. (2020)]. However, there is a notable gap in research regarding the application of machine learning to infer cognitive deficits in patients with chronic traumatic brain injury (cTBI). Current algorithms often lack the capability to make predictions without age-matched controls or neglect to explore the diagnostic potential of higher-order diffusion models, such as NODDI [Qu et al. (2021)]. Thus, there exists a need for further investigation into the development of machine learning models that can effec-

tively analyze diffusion data to infer cognitive impairments in cTBI patients.

5.2.2 Case-Study 2: TBI Classification

As explianed in chapter 2, traumatic brain injury (TBI) is a significant public health concern, often leading to chronic impairments and disabilities. Conventional neuroimaging methods, such as T1-weighted imaging, may not adequately capture the underlying neuropathology of chronic TBI. Therefore, there is a growing interest in leveraging advanced imaging techniques, such as diffusion tensor imaging (DTI) and neurite orientation dispersion imaging (NODDI), to develop biomarkers for detecting and monitoring chronic TBI. In this study, machine learning (ML) algorithms were employed to analyze hybrid diffusion imaging (HYDI) data and develop biomarkers for inferring symptom severity in chronic TBI patients.

Data Acquisition

A total of 59 subjects were recruited for this study, comprising 17 men and 42 women, with an average age of 47 $\pm$ 15 years, all experiencing chronic symptoms resulting from concussion-induced mild traumatic brain injury (mTBI) (Malec et al., 2007). Among the 59 cTBI subjects, 22 had sustained a single concussion. All subjects met criteria for mild traumatic brain injury, including no significant amnesia, and no structural brain injury evident on MRI, such as hematoma, contusion, or brain stem injury. Symptoms emerging after TBI could include headache, hypersensitivity to auditory or visual stimuli, balance problems, cognitive problems, or emotional problems (e.g., depression or anxiety).

Clinical assessment of TBI subjects included a battery of self-reported measures and cognitive tests, administered on the same day as the imaging study. Subjects were classified based on individual scores compared to the mean value of the entire cohort for each of the 21 tested neuropsychological outcomes, categorized as favor-

able or unfavorable outcomes. Only two of the tests were utilized in this paper for classification.

Data Pre-processing and Labeling

The initial preprocessing of raw DICOM data involved correcting susceptibility-induced distortion. The output was then used to align all volumes to the b0 image for eddy current correction. B0, or the "b-zero" image, refers to a specific type of image acquired during diffusion magnetic resonance imaging (dMRI). In dMRI, different images are obtained using varying magnetic field gradients, which provide information about the diffusion properties of water molecules in tissues. The b0 image is acquired with no diffusion weighting, meaning that it does not sensitively measure water diffusion and serves as a reference or baseline image. Prior to fitting the DTI and NODDI models, denoising was performed, which utilizes dMRI noise level estimation and denoising based on random matrix theory.

DTI parameter maps, including fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), radial diffusivity (RD), and mean diffusivity. Higher-order diffusion metrics, such as axonal density (intra-cellular volume fraction, Vic) and orientation dispersion index (ODI), were computed from the NODDI component of the analysis.

Diffusion and T1-weighted images were aligned to a common template in Montreal Neurological Institute (MNI152) standard space. Region-based metrics were then calculated for each subject by averaging the diffusion and T1 metrics within each of the 20 regions from the Johns Hopkins University white matter tractography atlas. As a result, for each metric (FA, MD, AD, RD, ODI, Vic, T1), 20 feature was generated representing that feature in corresponding region of the brain.

Subjects' outcomes were classified as favorable or unfavorable based on their performance in neuropsychological tests, specifically Trail making A (sec) and Trail mak-

ing B (sec). In Trail Making A, participants are instructed to connect a sequence of numbers as quickly as possible, while in Trail Making B, they are required to alternate between numbers and letters in ascending order. For both tests, shorter completion times indicate better cognitive function. Each subject's score was compared to the mean value of the entire cohort, with scores lower than the mean considered favorable and scores higher than the mean considered unfavorable. Participants with missing values or scores were excluded from the final analysis.

Model Architecture and Training

First, a feature selection was performed using DT, to select the most important regions for the classification. Feature extraction and ranking preceded the training of machine learning models, a crucial step aimed at enhancing inference accuracy while mitigating overfitting risks. The proposed classification pipeline consisted of a feature ranking module followed by a classification model. Notably, the decision tree (DT) module ranked brain regions based on feature occurrence frequency, with the top six features selected for subsequent classification. K-nearest neighbors (K-NN) was chosen as the classifier due to its equitable treatment of features based on sample distance, aligning with the objective of the study.

Evaluation of machine learning models was conducted through a K-fold cross-validation (CV) approach, a standard practice in the assessment of model performance, particularly in scenarios with limited datasets [Anguita et al. (2012)]. This approach involved splitting the dataset into training and testing sets, with K subsets generated for iterative testing. The models were trained on the training set and evaluated using the testing set, ensuring consistent evaluation conditions across iterations (Figure 7). By employing this rigorous evaluation methodology, the study aimed to ascertain the robustness and generalizability of the machine learning-based classification model in predicting TBI severity based on imaging metrics. The accu-

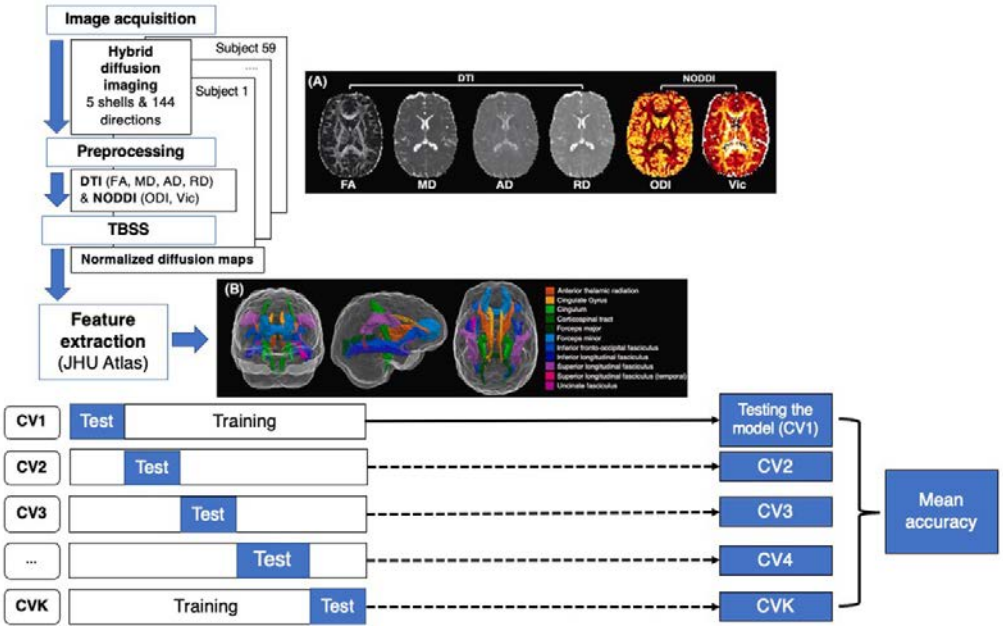

Figure 7: Flow chart representing ML approach for cTBI classification, including image acquisition, preprocessing, image normalization and skeletonization using TBSS, and atlas registration. The dataset was divided into training and test datasets for K-fold CV, calculating the mean accuracy of each model.

racy of classification models based on DTI, NODDI, and T1-weighted imaging was compared.

Results

First, five ML algorithms were employed, including support vector machine (SVM), k-nearest neighbors (K-NN), logistic regression (LR), random forest (RF), and decision tree (DT), to classify outcomes of Trail making A (sec) and Trail making B (sec) as favorable or unfavorable based on the DTI, NODDI, and T1 extracted features in all 20 regions of the brain. The scikit-learn library in Python was utilized for model development. The results are shown in Figure 8. This shows how DTI, and NODDI features are better metrics for predicting outcomes in TBI.

Then the proposed method was tested using the extracted metrics. ML-based models utilizing DTI and NODDI metrics demonstrated significantly higher accuracy

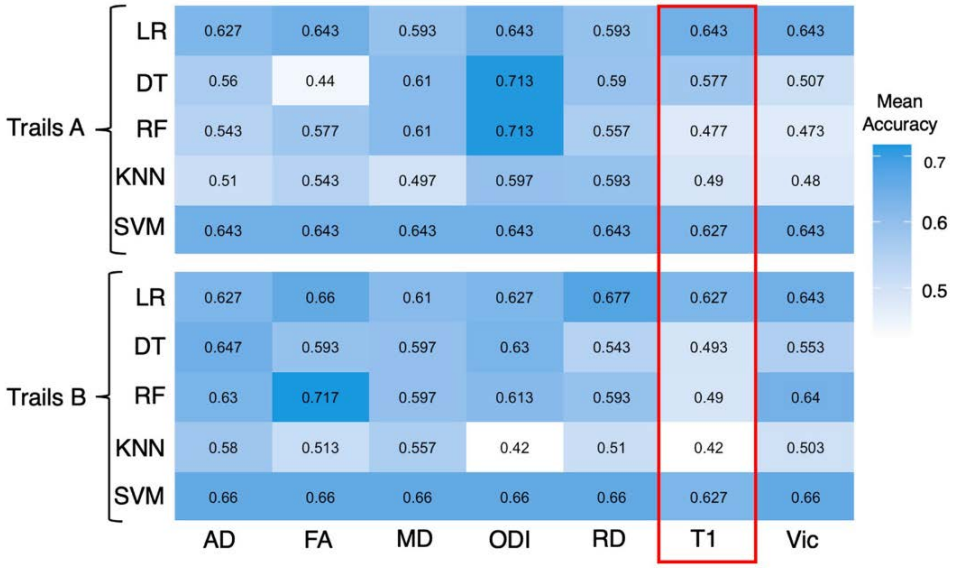

Figure 8: Heatmap showing mean accuracy performance of different ML algorithms for a trail making A and B. All 20 regions are used as features for the above figure. Mean accuracy results based on T1 inferences are highlighted in red and are expressed in percentages.

Metric (% accuracy range)	Mean accuracy (range) (average) (range)	*P*-value*
T1 (51.7–56.8%)	55.1% (51.7–56.8%)	–
DTI (58.7–73.0%)		
FA	61.0% (44.0–72.7%)	0.030
AD	61.5% (51.0–73.0%)	0.009
MD	61.0% (49.7–68.3%)	0.005
RD	61.0% (51.0–71.3%)	0.004
NODDI (64.0–72.3%)		
Vic	59.0% (47.3–66.0%)	0.036
ODI	63.7% (42.0–72.3%)	0.001

Table 4: Performance across all ML algorithms, including LR, DT, RF, K-NN, SVM, and combined feature selection with K-NN.

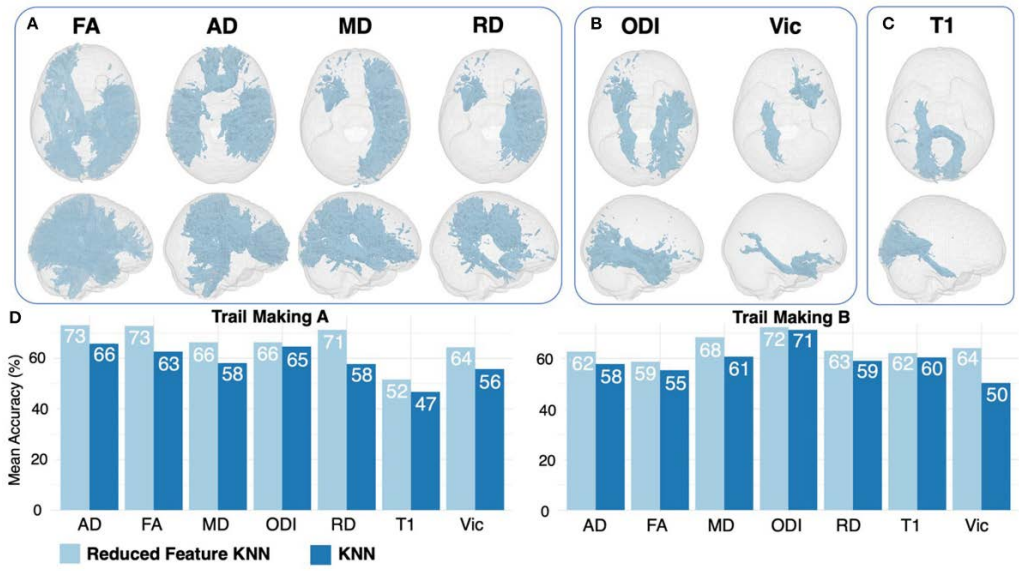

Figure 9: Feature ranking results for DTI (A), NODDI (B), and T1 (C) regions. Features are displayed if they were ranked as significant for both trail making A and B. Results of 6-feature KNN are displayed in light blue, compared with 20-feature KNN results in dark blue (D).

in classifying chronic TBI compared to conventional T1-weighted imaging. Across ML algorithms, the mean accuracy ranged from 58.7% to 73.0% for DTI-based models and 64.0% to 72.3% for NODDI-based models (Table 4). Feature selection using DT methods improved classification accuracy by approximately 11% (Figure 9).

This study highlights the potential of ML algorithms to utilize advanced diffusion imaging techniques for inferring symptom severity in chronic TBI patients. DTI and NODDI metrics outperformed conventional T1-weighted imaging in predicting cognitive impairment, suggesting their utility as biomarkers for chronic TBI. Feature selection methods, such as DT, offer a promising approach to improve model performance and identify relevant brain regions associated with clinical outcomes.

Limitations

One of the primary limitations of many studies in this field, including the present study, is the reliance on predefined metrics rather than directly utilizing the images

themselves. Incorporating image-based models alongside traditional metrics could enhance the understanding and accuracy of outcomes assessment.

While ML algorithms offer powerful tools for analyzing complex imaging data, the interpretability of model predictions remains a challenge. Understanding the underlying features and mechanisms driving classification decisions is essential for gaining insights into the pathophysiology of chronic TBI and translating research findings into actionable clinical interventions. Future studies should prioritize the development of interpretable ML models that provide clinically relevant insights.

Current studies often focus solely on analyzing imaging features, overlooking the potential synergies between clinical and imaging features. Integrating multimodal data, including neuropsychological assessments, clinical symptoms, and imaging biomarkers, could enhance the predictive accuracy and clinical utility of ML models.

Furthermore, using mean values as the sole metric for defining favorable and unfavorable outcomes may introduce sensitivity to the data, particularly in small sample sizes. Alternative approaches, such as robust statistical methods or machine learning techniques for outcome classification, may offer more reliable metrics for outcome assessment in future studies.

5.2.3 Case-Study 3: Cervical SCI Outcome Prediction

Spinal cord injury (SCI) is a debilitating condition with profound implications for patients' neurological and functional outcomes. Conventional MRI has limitations in assessing the microstructural changes within the spinal cord tissue, necessitating the use of advanced imaging techniques such as DTI and tractography. This section critically analyzes a study investigating the clinical utility of DTI and tractography in predicting outcomes after traumatic cervical SCI. The study also assesses whether the predictability differs before and after surgery.

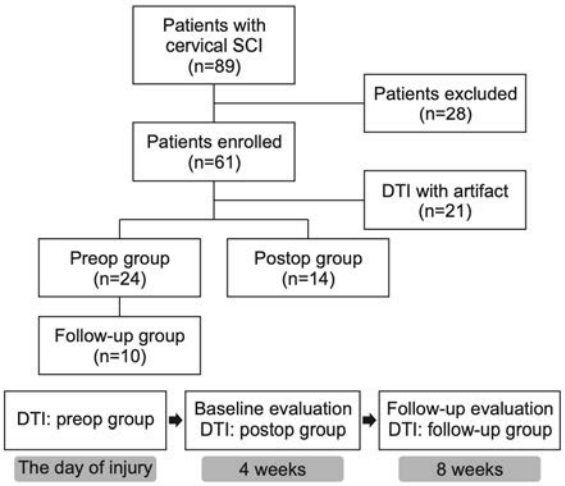

Figure 10: Flowchart of a schematic representation of the clinical study design.

Data Acquisition

A prospective randomized clinical trial involving acute-stage cervical spinal cord injury (SCI) patients with American Spinal Injury Association impairment scale (AIS) of B, C, and D was conducted. Inclusion criteria included acute traumatic injuries, AIS B, C, and D, neurological levels of injury (NLI) corresponding to C7 or higher, and receipt of surgical management. Exclusion criteria comprised combined or preexisting brain lesions, combined peripheral nerve injuries, and patients with artifact-laden images. Of the 89 initially enrolled patients, 28 were excluded, resulting in a total of 61 subjects. Patients were randomly assigned to preoperative or postoperative groups, with those undergoing DTI before surgery assigned to the former and those examined four weeks after surgery to the latter. Patients who underwent DTI before surgery and repeated the examination eight weeks post-injury were included in the follow-up group. The final analysis included data from 38 patients with cervical SCI (preop, n=24; postop, n=14; follow-up, n=10). The flowchart of patient enrollment and the clinical trial design is depicted in Figure 10. Neurological assessment was performed according to the International Standards for the Neurological Classification of Spinal Cord Injury (ISNCSCI) developed by the American Spinal Injury Association

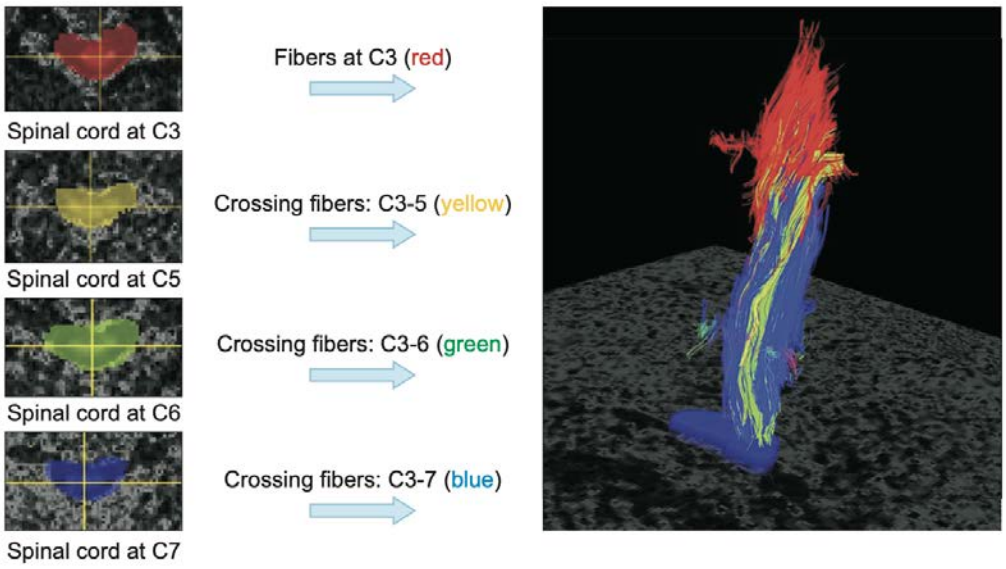

Figure 11: Example of tractographic analysis.

(ASIA). Baseline evaluation occurred at four weeks post-injury, followed by a follow-up assessment eight weeks later. AIS and NLI were determined for all patients, and upper extremity motor scores (UEMSs) and lower extremity motor scores (LEMSs) were evaluated. Sensory scores for light touch (LT) and pinprick (PP) were recorded separately for upper and lower extremity dermatomes.

Activities of daily living (ADL) performance was evaluated using the Korean version of the Modified Barthel Index (K-MBI) and Functional Independence Measure (FIM) at four and eight weeks post-injury. The K-MBI consists of 10 items, and the FIM assesses function in six areas, generating total, motor, and cognitive scores.

Multiplanar MRI of the cervical spine was acquired on a 1.5T MR scanner using a T2-weighted fast spin-echo sequence. DTI was obtained via diffusion-weighted single-shot echo-planar imaging (EPI) between spinal levels C2 and T1.

Data Pre-processing and Labeling

Rotationally invariant parameters such as fractional anisotropy (FA) and apparent diffusion coefficient (ADC) values were calculated on a voxel-by-voxel basis. Trac-

tography reconstructed fiber projections throughout the white matter, producing a voxelwise map of fiber orientation (Figure 11).

Statistical Analysis

Statistical analyses were conducted using PASW Statistics(IBM Corp, Armonk, NY, USA). The Mann-Whitney U test was employed to compare numerical data between two groups. The Pearson chi-square test was utilized to analyze baseline categorical data such as sex, cause, initial NLI and AIS, and the failure rate of DTI analysis between the preoperative (preop) and postoperative (postop) groups. Changes in neurologic scores, functional scores, and quantitative DTI parameters between initial and follow-up statuses of patients in the follow-up group were compared using the Wilcoxon signed-rank test.

Spearman correlation analysis was performed to delineate the relationships between neurologic and functional assessments and quantitative DTI parameters (FA, ADC, fiber numbers, and connection rate) at each cervical level. A p-value less than 0.05 was considered statistically significant.

Results

In this study, an analysis was conducted to investigate changes in neurologic, functional, and diffusion tensor imaging (DTI) parameters following surgery compared to preoperative statuses in cervical spinal cord injury (SCI) patients. Between the preoperative (preop) and postoperative (postop) groups, no significant differences were found in demographic characteristics or surgical approaches.

Both groups demonstrated significant improvements in upper extremity motor scores (UEMS), Korean version of the Modified Barthel Index (K-MBI), and total Functional Independence Measure (FIM) scores after 4 weeks. Additionally, significant improvements were observed in sensory scores and total FIM scores in the preop

	Preop group (n=24)	Postop group (n=14)	p-value
FA			
C3	0.772±0.078	0.606±0.112	0.000*
C4	0.700±0.102	0.568±0.136	0.006*
C5	0.664±0.102	0.612±0.120	0.066
C6	0.609±0.134	0.604±0.153	0.873
C7	0.676±0.125	0.573±0.173	0.085
Cinj	0.621±0.110	0.607±0.171	0.575
ADC			
C3	0.845±0.133	1.165±0.396	0.001*
C4	0.877±0.179	1.202±0.394	0.003*
C5	0.893±0.184	1.133±0.381	0.016*
C6	0.926±0.251	1.084±0.347	0.060
C7	0.777±0.346	1.134±0.372	0.009*
Cinj	0.902±.0241	1.178±0.432	0.015*
Fiber No.			
C3	1245.78±279.24	874.36±415.10	0.009*
C4	1267.48±294.06	736.36±477.08	0.001*
C5	1262.09±313.31	618.57±440.89	0.000*
C6	1157.04±293.19	432.79±373.51	0.000*
C7	813.65±430.23	288.71±299.86	0.000*
Crossing fiber No.			
C3-5	348.48±300.92	259.43±275.32	0.363
C3-6	235.70±275.26	143.93±236.59	0.137
C3-7	49.61±128.45	63.64±155.04	0.484

Table 5: DTI parameters between the preop and postop groups.

	Preop Group																	
	FA					ADC					Fiber No.					Crossing fiber No.		
	C3	C4	C5	C6	C7	C3	C4	C5	C6	C7	C3	C4	C5	C6	C7	C3-5	C3-6	C3-7
UE_FU																		
Motor	-0.021	0.079	-0.138	0.021	-0.172	0.004	0.075	0.207	0.011	0.338	0.296	0.101	0.219	0.042	0.387	0.221	0.130	0.165
Sensory																		
Light touch	0.161	0.571*	0.181	-0.142	-0.248	-0.142	-0.622*	-0.071	0.055	0.406	0.082	0.071	0.121	0.170	0.334	0.089	0.007	0.091
Pinprick	0.073	0.213	-0.003	-0.285	0.046	-0.121	-0.355	0.051	0.086	0.038	0.394	0.317	0.216	0.211	0.375	0.201	0.007	0.041
Total	0.084	0.466	0.206	-0.118	-0.108	-0.198	-0.561*	-0.051	0.006	0.025	0.191	0.148	0.127	0.125	0.345	0.071	-0.060	0.-007
LE_FU																		
Motor	-0.044	0.019	-0.258	-0.224	-0.441	0.192	0.207	0.504*	0.284	0.582*	0.124	0.143	0.116	-0.056	0.168	-0.004	0.007	0.016
Sensory																		
Light touch	0.113	-0.035	0.019	-0.146	-0.171	0.020	0.061	0.362	0.228	0.317	0.254	0.212	0.196	0.126	0.226	-0.038	-0.027	-0.135
Pinprick	0.078	0.017	0.106	0.045	-0.101	-0.101	-0.026	0.205	-0.014	0.213	0.364	0.378	0.257	0.077	0.288	0.072	-0.064	-0.135
Total	0.095	-0.032	-0.007	-0.075	-0.170	-0.018	0.007	0.324	0.105	0.025	0.364	0.319	0.215	0.097	0.200	-0.004	-0.010	-0.072
K-MBI_FU	0.280	0.128	0.149	0.060	0.035	-0.255	0.161	0.215	0.007	0.156	0.333	0.353	0.358	0.420	0.457*	0.233	0.211	0.152
FIM_FU																		
Self-care	0.190	0.053	0.142	0.126	0.089	-0.193	0.262	0.157	0.012	0.101	0.323	0.299	0.361	0.480*	0.461*	0.317	0.325	0.312
Sphincter control	0.002	0.037	0.081	-0.117	0.039	0.030	0.233	0.247	0.262	0.249	0.083	0.188	0.167	0.146	0.256	-0.080	-0.048	-0.076
Transfer	0.228	-0.006	-0.111	-0.210	0.048	-0.210	0.219	0.381	0.165	0.099	0.312	0.359	0.307	0.308	0.381	0.115	0.084	0.041
Total	0.331	0.106	0.110	-0.040	0.078	-0.278	0.172	0.275	0.065	0.093	0.310	0.342	0.329	0.454*	0.482*	0.195	0.213	0.186

Table 6: Correlation analysis between follow-up (FU) neurologic and functional findings and DTI parameters in the preop group.

	Postop Group																	
	FA					ADC					Fiber No.					Crossing fiber No.		
	C3	C4	C5	C6	C7	C3	C4	C5	C6	C7	C3	C4	C5	C6	C7	C3-5	C3-6	C3-7
UE																		
Motor	0.736*	0.138	0.331	0.284	0.204	-0.366	0.325	0.096	0.185	-0.143	0.150	0.148	0.355	0.273	0.152	0.239	0.314	0.341
Sensory																		
Light touch	0.035	-0.334	-0.366	-0.174	0.016	-0.211	0.407	0.344	0.265	-0.006	0.095	0.116	0.198	0.044	0.054	0.142	0.059	0.123
Pinprick	-0.006	-0.391	-0.469	-0.253	-0.012	-0.126	0.295	0.228	0.156	0.000	-0.027	-0.027	0.059	-0.095	0.024	0.020	-0.032	0.071
Total	-0.006	-0.391	-0.469	-0.253	-0.012	-0.126	0.295	0.228	0.156	0.000	-0.027	-0.027	0.059	-0.095	0.024	0.020	-0.032	0.071
LE																		
Motor	0.007	-0.074	0.104	-0.026	-0.163	0.264	0.617*	0.576*	0.535	0.171	0.548*	0.496	0.690*	0.619*	-0.372	0.534*	0.491	0.396
Sensory																		
Light touch	-0.342	0.064	-0.150	0.257	0.321	0.150	-0.214	0.000	-0.235	-0.150	0.504	0.339	0.305	0.192	0.287	0.380	0.358	0.271
Pinprick	-0.342	0.064	-0.150	0.257	0.321	0.150	-0.214	0.000	-0.235	-0.150	0.504	0.339	0.305	0.192	0.287	0.380	0.358	0.271
Total	-0.342	0.064	-0.150	0.257	0.321	0.150	-0.214	0.000	-0.235	-0.150	0.504	0.339	0.305	0.192	0.287	0.380	0.358	0.271
K-MBI	0.366	0.149	0.292	0.022	-0.333	-0.033	0.311	0.096	0.245	0.426	0.460	0.535*	0.690*	0.563*	-0.099	0.599*	0.539*	0.522
FIM																		
Self-care	0.579*	0.096	0.202	0.094	-0.161	-0.348	0.305	0.136	0.296	0.283	0.246	0.336	0.434	0.442	0.232	0.444	0.419	0.550*
Sphincter control	0.150	0.252	0.402	0.122	-0.287	0.113	0.153	0.023	0.107	0.408	0.621*	0.694*	0.804*	0.632*	-0.119	0.676*	0.628*	0.501
Transfer	0.348	0.084	0.262	-0.042	-0.435	-0.095	0.318	0.117	0.295	0.510	0.413	0.528	0.662*	0.566*	-0.123	0.576*	0.516	0.577*
Total	0.338	0.132	0.264	-0.006	-0.360	-0.105	0.261	0.085	0.261	0.476	0.451	0.570*	0.697*	0.607	-0.011	0.608*	0.548*	0.564*

Table 7: Correlation analysis between baseline neurologic and functional findings and DTI parameters in the postop group

group. However, no significant differences were noted between the preop and postop groups in these assessments.

Comparing quantitative DTI parameters between the two groups revealed significant differences. The preop group exhibited higher fractional anisotropy (FA) values at C3 and C4 levels and lower apparent diffusion coefficient (ADC) values at C3, C4, C5, and C7 levels compared to the postop group. Moreover, the preop group showed higher fiber numbers at all levels and lower failure rates in DTI analysis. Spearman correlation analysis indicated significant associations between preoperative DTI parameters and baseline neurologic and functional scores, with FA and ADC values correlating with motor and sensory scores, respectively (Table 5).

Postoperative DTI parameters were significantly correlated with baseline clinical statuses and functional outcomes. Lower FA values and higher ADC values postoperatively suggested axonal membrane damage and disorganization within fiber tracts. Fiber numbers and crossing fiber numbers significantly decreased postoperatively compared to preoperative values. Correlation analyses revealed associations between postoperative DTI parameters and clinical scores, with significant correla-

tions observed particularly between postoperative fiber numbers and functional scores (Table 6 and 7).

While the study had limitations including sample size and varied surgical approaches, it highlights the significance of preoperative and postoperative DTI in predicting neurological and functional outcomes in cervical SCI patients. The findings emphasize the importance of interpreting DTI separately before and after surgery for a comprehensive understanding of patient prognosis.

Limitations

The study's sample size was relatively small, and the follow-up period was relatively short. Larger cohort studies with longer follow-up durations are needed to validate the findings and better understand the long-term prognostic value of DTI and tractography in cervical SCI.

The diverse array of surgical approaches utilized in treating cervical SCI patients introduces variability in postoperative DTI findings. Future studies could explore the impact of different surgical techniques on DTI parameters and establish standardized protocols for surgical interventions to minimize confounding factors in DTI analysis.

CHAPTER 6

Discussion

6.1 Explored Questions

Through the analysis of these three papers, it becomes evident that machine learning (ML) holds significant promise in enhancing both structural and diffusion tensor imaging (DTI) classifications. Studies focusing on structural MRI showed the capability of ML algorithms to distinguish between different neurological disorders with high accuracy, enabling early and precise diagnoses. Moreover, the application of ML in DTI analysis showcases its potential in predicting functional outcomes in conditions like traumatic brain injury (TBI), where traditional imaging methods may fall short. Interestingly, while ML has been extensively utilized in various medical imaging domains, its application for spinal cord classifications remains relatively understudied, presenting a significant area for exploration and potential improvement. Using ML techniques for spinal cord classifications could offer valuable insights into neurological disorders affecting the spine, paving the way for more accurate diagnoses and tailored treatment strategies.

The choice of imaging sequences plays a crucial role in the accuracy and efficacy of diagnostic models. DTI has emerged as a particularly valuable sequence, offering insights into tissue microarchitecture and white matter integrity. Studies examining TBI highlight the enhanced accuracy achieved by incorporating DTI metrics such as fractional anisotropy and mean diffusivity into predictive models, enabling more precise assessments of neurological recovery and guiding personalized rehabilitation strategies. Additionally, the utility of T2-weighted images in degenerative cervical

myelopathy (DCM) classification underscores the importance of leveraging diverse imaging sequences to capture comprehensive anatomical and pathological information. Integrating multiple imaging modalities, including DTI and T2-weighted imaging, can enhance diagnostic accuracy and provide a more holistic understanding of neurological disorders.

The timing of imaging acquisitions, particularly in relation to surgical interventions, significantly influences the interpretation and predictive value of imaging data. Studies focusing on traumatic cervical spinal cord injury (SCI) underscore the importance of interpreting preoperative and postoperative DTI and tractography data separately. Preoperative DTI and tractography demonstrate superior fractional anisotropy (FA) and apparent diffusion coefficient (ADC) values, with lower interpretation failure rates compared to postoperative data. However, postoperative DTI and tractography data more accurately reflect the patient's clinical state at the time of evaluation. This highlights the critical role of timing in imaging acquisitions and suggests that DTI and tractography should be interpreted in the context of the patient's surgical history to predict clinical outcomes effectively.

6.2 Unexplored Questions and Future Works

Despite significant advancements in the application of machine learning (ML) and neuroimaging techniques, several questions remain unexplored, offering avenues for future research. Understanding the impact of demographic factors, such as age, on imaging biomarkers and predictive models is essential for tailoring treatment strategies to individual patients. Moreover, addressing challenges related to reducing artifact effects in imaging data and identifying optimal ML methods for different imaging modalities can further enhance the accuracy and reliability of diagnostic models. Exploring these uncharted territories holds promise for refining existing techniques and developing novel approaches to neuroimaging analysis.

The influence of age on neuroimaging biomarkers and predictive models is a critical area that warrants further investigation. Age-related changes in brain structure and function may impact the interpretation of imaging data and the efficacy of diagnostic models. Understanding how age influences the relationship between imaging metrics and clinical outcomes can inform the development of age-specific diagnostic algorithms and treatment strategies. Future studies should explore the role of age as a potential confounding factor in neuroimaging analysis and incorporate age-related considerations into predictive models for neurological disorders.

Artifacts in neuroimaging data can significantly compromise the accuracy and reliability of diagnostic models. Addressing artifact effects requires comprehensive strategies for data preprocessing and quality control to minimize their impact on imaging analysis. Techniques such as motion correction, spatial filtering, and artifact detection algorithms can help mitigate artifact effects and improve the robustness of neuroimaging analyses. Future research should focus on developing advanced artifact reduction methods tailored to specific imaging modalities and optimizing preprocessing pipelines to ensure high-quality imaging data for accurate diagnosis and prediction.

Choosing the most appropriate machine learning (ML) method for neuroimaging analysis depends on various factors, including the nature of the data, the complexity of the task, and the desired clinical outcomes. Different ML algorithms, such as support vector machines, neural networks, decision trees, and ensemble methods, offer unique advantages and may be more suitable for specific applications. Comparative studies evaluating the performance of different ML methods on neuroimaging datasets can help identify the most effective approaches for different tasks. Additionally, incorporating domain knowledge and expert input into ML model development can enhance interpretability and clinical relevance. Future research should explore the strengths and limitations of various ML methods in neuroimaging analysis and

establish guidelines for selecting the optimal approach based on specific diagnostic and prognostic objectives.

CHAPTER 7

CONCLUSION

The integration of magnetic resonance imaging (MRI) with advanced quantitative analysis techniques and machine learning (ML) methods represents a significant advancement in the field of neurological disorder detection and diagnosis. MRI has become a cornerstone in the imaging evaluation of neurological disorders, offering detailed anatomical insights into the brain and spinal cord.

Traditionally, MRI has relied on qualitative assessments by radiologists, which may be subjective and lack sensitivity to subtle changes in tissue properties. However, recent advancements in quantitative analysis have transformed MRI into a powerful tool for objective measurement of tissue-specific parameters. This quantitative approach has the potential to provide deeper insights into tissue microstructure, function, and pathology, thereby enhancing diagnostic accuracy and informing treatment decisions.

These ML-based methods further enhances the capabilities of MRI-based techniques for neurological disorder detection. ML algorithms can extract hidden patterns and associations from MRI data, enabling a more comprehensive understanding of neurological disorders. By leveraging ML, MRI has the potential to improve diagnosis, prognosis, and treatment strategies for neurological disorders, ultimately benefiting patients and healthcare providers alike.

Through the analysis of key research papers and discussions on the current state of the art in MRI-based techniques for neurological disorder detection, this report explored the potential and challenges in this rapidly evolving field. While significant

progress has been made, there are still unexplored questions and opportunities for future research.

Moving forward, addressing questions such as the impact of age on imaging biomarkers, strategies for reducing artifact effects, and the selection of optimal ML methods will be crucial for further advancing MRI-based techniques for neurological disorder detection. Collaborative efforts between researchers, clinicians, and industry partners will be essential for translating these advancements into tangible benefits for patients and healthcare providers.

BIBLIOGRAPHY

Aarabi, B., Sansur, C. A., Ibrahimi, D. M., Simard, J. M., Hersh, D. S., Le, E., ... Akhtar-Danesh, N. (2017). Intramedullary lesion length on postoperative magnetic resonance imaging is a strong predictor of asia impairment scale grade conversion following decompressive surgery in cervical spinal cord injury. *Neurosurgery*, *80*(4), 610–620.

Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., Ridella, S., et al. (2012). The'k'in k-fold cross validation. In *Esann* (Vol. 102, pp. 441–446).

Castro-Mateos, I., Pozo, J. M., Lazary, A., & Frangi, A. F. (2014). 2d segmentation of intervertebral discs and its degree of degeneration from t2-weighted magnetic resonance images. In *Medical imaging 2014: Computer-aided diagnosis* (Vol. 9035, pp. 310–320).

Chan, S., & Siegel, E. L. (2018, 11). Will machine learning end the viability of radiology as a thriving medical specialty? *British Journal of Radiology*, *92*(1094), 20180416.

Chandra, M. A., & Bedi, S. (2021). Survey on svm and their application in image classification. *International Journal of Information Technology*, *13*(5), 1–11.

Chong, S.-L., Liu, N., Barbier, S., & Ong, M. E. H. (2015). Predictive modeling in pediatric traumatic brain injury using machine learning. *BMC medical research methodology*, *15*, 1–9.

Daugherty, J. (2021). Differences in state traumatic brain injury–related deaths, by principal mechanism of injury, intent, and percentage of population living in rural areas—united states, 2016–2018. *MMWR. Morbidity and mortality weekly report*, *70*.

Davatzikos, C. (2019). Machine learning in neuroimaging: Progress and challenges. *NeuroImage*, *197*, 652-656.

Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). Knn model-based approach in classification. In *On the move to meaningful internet systems 2003: Coopis, doa, and odbase: Otm confederated international conferences, coopis, doa, and odbase 2003, catania, sicily, italy, november 3-7, 2003. proceedings* (pp. 986–996).

Hu, L., Yang, S., Jin, B., & Wang, C. (2022). Advanced neuroimaging role in traumatic brain injury: A narrative review. *Frontiers in Neuroscience*, *16*. doi: 10.3389/fnins.2022.872609

Jin, R., Luk, K. D., Cheung, J., & Hu, Y. (2016). A machine learning based prognostic prediction of cervical myelopathy using diffusion tensor imaging. In *2016 ieee international conference on computational intelligence and virtual environments for measurement systems and applications (civemsa)* (p. 1-4). doi: 10.1109/CIVEMSA .2016.7524318

Jordan, B. D. (2000). Chronic traumatic brain injury associated with boxing. In *Seminars in neurology* (Vol. 20, pp. 179–186).

Karadimas, S. K., Erwin, W. M., Ely, C. G., Dettori, J. R., & Fehlings, M. G. (2013). Pathophysiology and natural history of cervical spondylotic myelopathy. *Spine*, *38*(22S), S21–S36.

Khalil, M., Teunissen, C. E., Otto, M., Piehl, F., Sormani, M. P., Gattringer, T., ... others (2018). Neurofilaments as biomarkers in neurological disorders. *Nature Reviews Neurology*, *14*(10), 577–589.

Kirshblum, S. C., Priebe, M. M., Ho, C. H., Scelza, W. M., Chiodo, A. E., & Wuermser, L.-A. (2007). Spinal cord injury medicine. 3. rehabilitation phase after acute spinal cord injury. *Archives of physical medicine and rehabilitation*, *88*(3), S62–S70.

Koonce, B. (2021). Resnet 50. *Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization*, 63–72.

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, *29*(2), 102–127.

Mai, J. K., & Paxinos, G. (2011). *The human nervous system*. Academic press.

Merali, Z., Wang, J. Z., Badhiwala, J. H., Witiw, C. D., Wilson, J. R., & Fehlings, M. G. (2021). A deep learning model for detection of cervical spinal cord compression in mri scans. *Scientific reports*, *11*(1), 10473.

Michopoulou, S. K., Costaridou, L., Panagiotopoulos, E., Speller, R., Panayiotakis, G., & Todd-Pokropek, A. (2009). Atlas-based segmentation of degenerated lumbar intervertebral discs from mr images of the spine. *IEEE Transactions on Biomedical Engineering*, *56*(9), 2225-2231.

Mohamed, M., Alamri, A., Mohamed, M., Khalid, N., O'Halloran, P., Staartjes, V., & Uff, C. (2022). Prognosticating outcome using magnetic resonance imaging in patients with moderate to severe traumatic brain injury: a machine learning approach. *Brain Injury*, *36*(3), 353–358.

Muller, J. J., Wang, R., Milddleton, D., Alizadeh, M., Kang, K. C., Hryczyk, R., ... others (2023). Machine learning-based classification of chronic traumatic brain injury using hybrid diffusion imaging. *Frontiers in Neuroscience*, *17*, 1182509.

O'shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

Park, G. S., Kim, T. U., Lee, S. J., Hyun, J. K., & Kim, S. Y. (2022). Quantitative analysis in cervical spinal cord injury patients using diffusion tensor imaging and tractography. *Annals of Rehabilitation Medicine*, *46*(4), 172.

Prasuhn, J., Heldmann, M., Münte, T. F., & Brüggemann, N. (2020). A machine learning-based classification approach on parkinson's disease diffusion tensor imaging datasets. *Neurological research and practice*, *2*, 1–5.

Qu, Y., Wang, P., Liu, B., Song, C., Wang, D., Yang, H., ... others (2021). Ai4ad: Artificial intelligence analysis for alzheimer's disease classification based on a multisite dti database. *Brain Disorders*, *1*, 100005.

Rigatti, S. J. (2017). Random forest. *Journal of Insurance Medicine*, *47*(1), 31–39.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... Fei-Fei, L. (2014). Imagenet large scale visual recognition challenge. *CoRR*, *abs/1409.0575*. Retrieved from `http://dblp.uni-trier.de/db/journals/corr/corr1409.html#RussakovskyDSKSMHKKBBF14`

Saygin, Z. M., Kliemann, D., Iglesias, J. E., van der Kouwe, A. J., Boyd, E., Reuter, M., ... others (2017). High-resolution magnetic resonance imaging reveals nuclei of the human amygdala: manual segmentation to automatic atlas. *Neuroimage*, *155*, 370–382.

Smith, L. G., Milliron, E., Ho, M.-L., Hu, H. H., Rusin, J., Leonard, J., & Sribnick, E. A. (2019). Advanced neuroimaging in traumatic brain injury: an overview. *Neurosurgical Focus*, *47*(6), E17.

Smith, T. B. (2010). Mri artifacts and correction strategies. *Imaging in Medicine*, *2*(4), 445.

Stewart, W. A., Parent, J. M., Towner, R. A., & Dobson, H. (1992). The use of magnetic resonance imaging in the diagnosis of neurological disease. *The Canadian Veterinary Journal*, *33*(9), 585.

Su, X., Yan, X., & Tsai, C.-L. (2012). Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(3), 275–294.

Suthaharan, S. (2016). Decision tree learning. *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*, 237–269.

Tofts, P. (2005). *Quantitative mri of the brain: measuring changes caused by disease*. John Wiley & Sons.

Trifan, G., Gattu, R., Haacke, E. M., Kou, Z., & Benson, R. R. (2017). Mr imaging findings in mild traumatic brain injury with persistent neurological impairment. *Magnetic resonance imaging*, *37*, 243–251.

VergaraVictor, M., MayerAndrew, R., KiehlKent, A., et al. (2017). Detection of mild traumatic brain injury by machine learning classification using resting state functional network connectivity and fractional anisotropy. *Journal of neurotrauma*.

Wang, P., Fan, E., & Wang, P. (2021). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognition Letters*, *141*, 61–67.

Wang, S., & Summers, R. M. (2012). Machine learning and radiology. *Medical image analysis*, *16*(5), 933–951.

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). Noddi: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, *61*(4), 1000–1016.