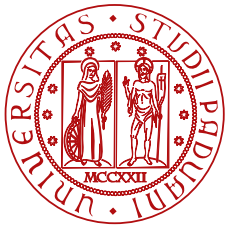


Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

Master Thesis in Data Science

Convolutional Neural Networks for Classification of Neurodevelopmental Disease Using fMRI Data

Master Candidate

Bita Darabi

Student ID 2013005

Supervisor

Prof. Daniele Marinazzo

University of Ghent

Co-supervisor

Prof. Tomaso Erseghe

University of Padova

Academic Year
2023/2024

I sincerely thank those who have made this thesis possible. First and foremost, my heartfelt appreciation goes to my family. To my parents for their unwavering love and support throughout my academic journey, whose encouragement and belief in my abilities have been a constant source of strength.

I am profoundly grateful to Professor Tomaso Erseghe at the University of Padova for his guidance and invaluable insights, which significantly enriched my research experience. His mentorship has been instrumental in my academic development. Special thanks are due to my esteemed supervisor, Professor Daniele Marinazzo, whose expertise and attention to detail have greatly enhanced the quality of my work. His patience and willingness to educate me have been genuinely inspiring even when I took missteps.

I also express my deep appreciation for my friends Mohammad, Mahsa and Sepideh, whose support, camaraderie, and unwavering faith in me have been of immense comfort during this challenging but rewarding journey. To all, I express my sincere appreciation and gratitude for accompanying me on this scholarly voyage. Your support has been a beacon of light in moments of doubt and has consistently guided me toward realizing my academic goals.

Abstract

This thesis presents an approach to analyzing functional Functional Magnetic Resonance Imaging (fMRI) data by introducing a novel Spatiotemporal Model that seamlessly integrates U-NET and 3D Convolutional Neural Network (3DCNN) with long short-term memory (LSTM) networks. This integration marks a significant advancement in neuroimaging analysis, addressing the complex challenge of capturing brain activity's complex spatial and temporal dynamics.

Traditional fMRI data analysis methods are limited by their inability to simultaneously map the brain's spatial structures and temporal patterns. These restrictions limit the depth of insights into neural functions and cognitive processes. To overcome these limitations, our Spatio-Temporal Model uses the U-NET and 3DCNN which allows for precisely identifying brain regions. Additionally, LSTM networks are integrated to model the temporal dependencies within the fMRI time series. This enhances the understanding of how spatial patterns of brain activity evolve over time.

The model provides new perspectives for exploring cognitive processes and neural functions. The results highlight the potential of this advanced analytical tool in understanding neurodevelopmental diseases and setting new approaches for future research in neuroimaging.

This research addresses a critical gap in fMRI data analysis and opens the door for significant contributions to neuroscience, potentially impacting diagnostic and therapeutic strategies for neurodevelopmental disorders. The thesis represents a step forward in neuroimaging analysis, demonstrating the potential of combining deep learning technologies to uncover the complex mechanisms of the brain.

Contents

List of Figures

List of Tables

List of Acronyms

CNN Convolutional Neural Networks

CNP Consortium for Neuropsychiatric Phenomics

CN cognitively normal

SMC significant memory concern

DALYs disability-adjusted life years

NDDs Neurodevelopmental disorders

EMCI early mild cognitive impairment

MCI mild cognitive impairment

LMCI late mild cognitive impairment

AD Alzheimer's disease

ICA independent component analysis

KPCA kernel principal component analysis

MsRNN multi-scale recurrent neural network

tSNE t-Distributed Stochastic Neighbor Embedding

SVM Support Vector Machine

FEP first-episode psychosis

AUC Area Under the Curve

HC healthy controls

SCH schizophrenia

BD bipolar disorder

ADHD attention-deficit/hyperactivity disorder

MRI Magnetic Resonance Imaging

sMRI structural Magnetic Resonance Imaging

ODD oppositional defiant disorder

MLP multilayer perceptron

ANN artificial neural network

NLP natural language processing

ECG Electrocardiogram

HARDI High Angular Resolution Diffusion Imaging

RS-fMRI Resting-State fMRI

T-fMRI task-based fMRI

NIH National Institutes of Health

NIFTI Neuroimaging Informatics Technology Initiative

LSTM long short-term memory

RelU Rectified linear unit

3DCNN 3D Convolutional Neural Network

fMRI Functional Magnetic Resonance Imaging

RNN Recurrent Neural Network

FSL FMRIB Software Library

BOLD Blood Oxygen Level Dependent

1 Introduction

1.1 Brain function and Brain disease

This section provides an overview of brain functions, various disorders, and current diagnostic tools. It then delves deeper into neurodevelopmental disorders, which are the primary focus of this thesis. Specifically, it discusses schizophrenia, bipolar disorder, and ADHD, offering brief explanations of each. This segment aims to lay the foundation for understanding the complex interplay between brain function and disorders, setting the stage for exploring the use of machine learning in classifying these conditions based on neuroimaging data.

1.1.1 Brain function

In everyday life, outside the intricate field of neuroanatomy, our perception of the brain tends to be quite abstract. We rarely, if ever, think of it in terms of its physical structure. Instead, when we imagine the brain, we might picture something like a sizable gray walnut. This simplified view underscores the distance between our daily experiences and the complex reality of brain anatomy [1].

Our brain is a vital organ that underpins numerous critical human survival functions. It is composed of several parts that collaboratively oversee cognition, motion, sensation, and emotions. This organ is categorized into three primary sections: the forebrain, midbrain, and hindbrain. The forebrain, particularly the cerebrum, facilitates higher-order functions such as thinking, remembering,

language skills, and sensory interpretation. The midbrain is critical for reflexive actions and controlling eye movements. Meanwhile, the hindbrain is essential for managing core life-sustaining activities like breathing and heart rates, besides aiding in movement coordination. Located within the hindbrain, the cerebellum is instrumental in refining motor skills and ensuring equilibrium. The brainstem, forming a bridge to the spinal cord, is crucial for transmitting signals across the body and to the cerebral cortex. As the epicenter of intellect, the brain orchestrates behaviors, deciphers sensory inputs, and prompts physical actions, underscoring its indispensable role in human life and functionality[9], [12]

1.1.2 Brain disease

Brain diseases can significantly impair a person's cognitive abilities, behavior, and daily functioning. These conditions span several categories, including autoimmune disorders like Multiple Sclerosis (MS), which attacks the nervous system's insulation; epilepsy, characterized by seizures; and various infections that can lead to symptoms like confusion and headaches.

Mental health disorders are another group of brain diseases. According to reports, mental health disorders affect one in five adults and diminish the quality of life and functional capacity, with conditions ranging from anxiety and depression to schizophrenia. Treatment usually includes both medication and therapy.

Neurodegenerative diseases are a group of disorders, such as Alzheimer's and Parkinson's, that result from abnormal protein accumulation, leading to memory, thought, and movement issues. They're mainly identified by the gradual loss of structure or function of neurons, which is called neurodegeneration [37].

Neurodevelopmental disorders are a category of brain disorders, such as ADHD and autism, which impact brain growth and development and are often managed by pediatric specialists.

Other brain diseases, including Strokes and traumatic brain injuries, arise from blood supply issues or physical trauma and can cause severe brain damage, affecting speech, movement, and cognitive functions.

Another group of brain diseases includes brain tumors, whether spreading from elsewhere or originating in the brain, which vary in severity based on their growth and invasiveness.

The prevalence of brain diseases varies, with conditions like Alzheimer's and autism being more common, while others like meningitis have become rarer due to vaccinations. Brain diseases pose significant challenges to individuals' health and well-being despite their varied incidence rates [10].

1.1.3 Current Means of Diagnosis

The burden of brain diseases in Europe is immense, as mentioned in the research by Olesen and Leonardi, which provides a detailed evaluation [40]. Their study utilizes disability-adjusted life years (DALYs) to gauge the overall impact, finding that such diseases constitute about 35% of total DALYs in Europe. This study underscores the massive portion of health-related impairments attributed to brain diseases and advocates for a proportional investment in medical education, research funding, and healthcare to address the complexities of brain conditions.

Brain disease diagnosis typically relies on clinical symptom scores and doctors' experience and subjectivity, and inefficiencies can sometimes lead to misdiagnoses [54].

There is a range of brain disorders, including neurodevelopmental diseases, movement disorders, epilepsy, strokes, and neuro-oncology, causing diverse clinical evaluations to be tailored to diagnose and treat these diseases. Early detection and accurate diagnosis for effective treatment are of great importance[21].

Diagnostic tools for brain disorders include neurological exams and brain imaging techniques like CT, MRI, and PET scans, and the analysis of fluid from the brain and spinal cord can detect abnormalities such as bleeding or infections [11].

Among all means of diagnosis, clinical evaluation of patients with brain diseases is very important and effective in understanding and diagnosing these conditions. Diseases are defined by symptoms and signs that affect an individual's normal functioning, and clinical assessment is a structured process undertaken by healthcare providers to collect symptoms and signs, determine a syndrome, and hypothesize about the underlying disease[21].

In [16], Jeste and Geschwind bring attention to the necessity of patient stratification in neurodevelopmental disorders due to their heterogeneity, suggesting that more refined subgroups might be needed for effective treatment. They also point to the potential for the timing of interventions to impact effectiveness, not-

ing that treating adults may not address neurobiological consequences already established during development.

In recent years, there has been some research on using neuroimaging in diagnosing mood and other brain disorders. While neuroimaging techniques, particularly MRI and fMRI, are on the brink of transitioning from research to clinical practice, they face significant limitations in widespread clinical application. These tools, despite their demonstrated ability to identify structural and functional changes related to mood disorders, primarily cater to neurological disorders. Their application within psychiatric diagnostics is yet to be fully realized due to challenges in generalizability and the subtlety of neuroanatomical variations in psychiatric conditions. This underscores the urgent need for further research and innovation. Integrating machine learning algorithms with neuroimaging data should also be considered to create individualized assessments, potentially bridging the gap between research findings and patient-specific care. An "ideal" neuroimaging tool would perform whole-brain analyses, accommodate disease heterogeneity, and offer robust validation to inform diagnostic and prognostic assessments in a personalized medicine framework[51].

Despite the promise shown, significant work remains in developing neuroimaging biomarkers that can reliably inform clinical decisions for mood and brain disorders; this involves overcoming challenges such as ensuring the specificity and sensitivity of biomarkers amidst the heterogeneity of psychiatric disorders and the necessity for more rigorous and extensive validation processes. The quest for reliable neuroimaging biomarkers is essential for enhancing the diagnosis, treatment, and management of mood disorders [50].

1.1.4 Neurodevelopmental Diseases

Neurodevelopmental disorders (NDDs) encompass a spectrum of conditions impacting brain functional development. These disorders can vary widely in severity, from mild forms that permit relatively normal living to profound disorders necessitating continuous care throughout life. Conditions classified under neurodevelopmental disorders cover a range of problems, such as behavioral conduct disorders, Attention Deficit Hyperactivity Disorder (ADHD), cerebral palsy, and speech and language difficulties [38].

Researchers are learning more about how genetics play a crucial role in neurodevelopmental disorders (NDDs), which include conditions like autism,

attention-deficit/hyperactivity disorder (ADHD), and intellectual disabilities. They have found that specific gene changes can significantly impact the development and behavior of individuals with these disorders. It is important to mention that many of these conditions might not be completely separate. Instead, they could be part of a larger spectrum that also includes conditions like schizophrenia and bipolar disorder. This suggests that these different disorders might share some common genetic factors. Thanks to advancements in genetic testing, doctors are getting better at diagnosing these disorders early on; this is important because the sooner a condition is diagnosed, the sooner treatment can start, which can lead to better outcomes for the person affected [36].

People did not pay much attention to neurodevelopmental disorders until about 50 years ago. Now, there is a shift towards not just surviving illnesses but improving life quality, leading to a better understanding of NDDs. The latest mental health guide has even added a new category for NDDs, showing that these issues start early in life and are related to how the brain develops, sometimes affected by problems during pregnancy or birth. ADHD and learning disorders are getting more research and attention today, even though they were not consistently recognized in the past. There is concern about overdiagnosing ADHD, the importance of recognizing development stages in diagnosis, and the need for a combined approach to treatment and understanding NDDs. NDDs can be a significant burden, affecting health, society, and the economy, and pose issues because the healthcare system is more set up for acute (short-term) health problems rather than chronic (long-term) ones. Other health conditions alongside NDDs can provide insights into how the brain works and stress the importance of looking at patient care more holistically. Future research is needed to understand better the differences and long-term effects of NDDs, calling for a more integrated healthcare approach to tackle these issues effectively [13].

In this paper we are going to classify between 3 neurodevelopmental diseases and healthy controls, these three diseases are: Schezophrenia, Bipolar and ADHD. Here, we briefly discuss each of these conditions.

Schizophrenia is a serious psychiatric condition with an unknown cause. Presently, there is discussion regarding a combination of genetic and environmental factors contributing to the outset of this disorder. This disorder is intricate, and patients may display various symptoms, varying between individuals or even within the same individual over time [19]. Studies using animal models have indicated that developmental hippocampal lesions can lead to disconnec-

tions within the prefrontal cortex. Further insights from magnetic resonance imaging and postmortem examinations have highlighted deficits within the temporoprefrontal neuronal circuit [52].

Bipolar disorder, is known as one of the most challenging mood disorders, it can cause considerable distress, disability, and burden not just for patients but also for their families and caregivers [48]. Differentiating between unipolar and bipolar depression during the early stages of the illness could aid in providing targeted and effective treatment strategies [24]. This disorder can impact individuals from all backgrounds and genders, affecting approximately 1 in every 100 people at some stage in their lives. While the exact cause remains uncertain, it's thought to involve a combination of genetic predisposition, biological variations in brain function, and environmental factors like intense stress or significant life events [8].

The third class of neurodevelopmental disease that we are going to classify in this paper is ADHD. ADHD is one of the most prevalent neurobehavioral disorders seen in children seeking treatment [23], and can also be diagnosed in adolescents and adults [4]. It often coexists with other psychiatric issues such as oppositional defiant disorder (ODD), conduct disorder, mood, and anxiety disorders, as well as cigarette and substance use disorders [7]. Throughout life, untreated ADHD carries significant social and societal costs, including academic and work-related underperformance, involvement in delinquent behavior, impaired motor vehicle safety, and challenges in interpersonal relationships [7] [5].

1.1.5 fMRI

fMRI is a relatively recent technology with the capacity to characterize and classify brain disorders [19] It is a category of imaging techniques created to illustrate regional, time-dependent alterations in brain metabolism [3]. These metabolic shifts can arise from cognitive state changes induced by tasks or spontaneous processes in the resting brain [22]. Since its introduction in 1990, this technique has been extensively utilized in numerous studies of cognition for clinical purposes, including surgical planning, monitoring treatment effectiveness, and serving as a biomarker in pharmacological and training interventions [22]; also, there's a lot of hope that fMRI data could be used to understand and group brain disorders like Alzheimer's disease, schizophrenia, mild brain injuries, addiction, or bipolar disorder based on biological measurements [19].

1.2 Methodology

The methodology section delves into the computational strategies employed to process brain imaging data. It begins with the basics of multilayer neural networks, followed by the principles of convolutional neural networks, which are adapted to image data analysis. The discussion progresses to 3DCNNs, suitable for examining volumetric brain scans, and then to Long Short-Term Memory networks, which handle data sequences and temporal patterns. The U-Net architecture is outlined for its application in medical image segmentation. Finally, the integration of LSTM layers into U-Net is discussed, showcasing a hybrid approach to analyze both spatial and temporal brain data features.

1.2.1 Multilayer Neural Network

The multilayer neural network, often referred to as a MLP, is a fundamental type of artificial neural network (ANN). It is tailored to approximate complex, nonlinear relationships from the input data to the corresponding outputs. The general structure of an MLP can be described by the following expressions:

Let $y = \mathcal{E}_{mlp}(x; \theta_{mlp})$ denote the MLP transformation, where $\mathcal{E}_{mlp}$ signifies the transformation function that maps an input x to an output y using a set of parameters θ_{mlp}. This transformation function can be decomposed into:

$$h = W^T \cdot x + b, \tag{1.1}$$

$$z = V^T \cdot h + u, \tag{1.2}$$

$$y = \sigma(z), \tag{1.3}$$

with W^T and V^T representing the weight matrices, b and u the bias vectors. Here, h and z embody the hidden layer representations. The function $\sigma(\cdot)$ is a nonlinear activation function applied to z to obtain the final output y. For a visual depiction, refer to Figure 1.1.

As depicted in Figure 1.1, the MLP initially transforms the input data into a hidden layer representation, applying a non-linear activation function to capture complex patterns within the data. The ultimate output is obtained following

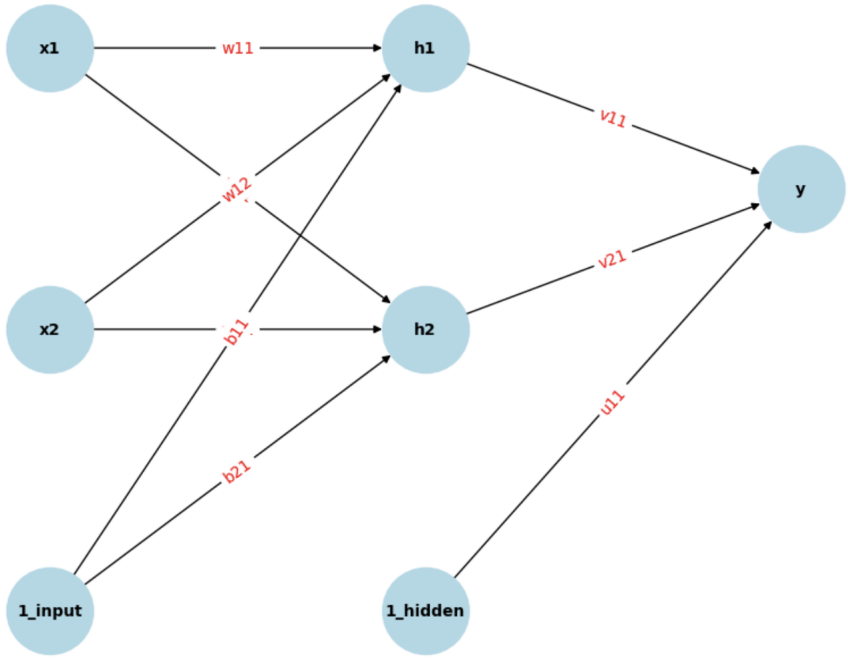


Figure 1.1: An MLP consisting of one input layer, one hidden layer, and one output layer.

the hidden layer processing. A common activation function employed in such networks is the Rectified Linear Unit Rectified linear unit (RelU), defined by:

$$y = \max(0, x). \tag{1.4}$$

The network is trained to emulate the intricate nonlinear mapping between inputs and outputs by iterative adjustment of the network parameters. The weights W and V, along with biases b and u, are optimized during the training process using the backpropagation algorithm [28].

1.2.2 Convolutional Neural Network

Convolutional Neural Networks (CNN) represent a class of ANNs that are particularly adept at identifying patterns in visual data. Their architecture is straightforward but effective, enabling the autonomous detection of pertinent features within images, eliminating the need for manual feature extraction [39]. The term "convolutional" in CNNs is derived from a mathematical operation known as convolution, a specialized linear operation applied to matrices. CNNs are composed of several distinct layers: these include the convolutional layer,

the non-linearity (activation) layer, the pooling (subsampling) layer, and the fully-connected (dense) layer. Parameters are inherent to the convolutional and fully-connected layers, while the pooling and non-linearity layers operate without adjustable parameters. CNNs have demonstrated remarkable efficacy in various machine learning applications, particularly in tasks involving image data. They have achieved significant success in large-scale image classification datasets like ImageNet, fields of computer vision, and even in natural language processing (NLP), where their results have been exceptionally impressive [2].

In Convolutional Neural Networks, the convolution operation is fundamental. This process involves a function integrating point-wise multiplication and shifting across the input data to produce a transformed output. The convolution is mathematically defined as the integral of the product of two functions after one is reversed and shifted, denoted as $f * g$ for functions f and g.

For discrete data, the convolution is represented as a summation:

$$(f * g)(x) = \sum_{\alpha=-\infty}^{\infty} f(\alpha)g(x - \alpha). \tag{1.5}$$

Here, a specific kernel $g(x)$ is utilized to extract relevant patterns from the original data $f(x)$. The outcomes, or patterns, are then held in the resulting function $f * g$. In a deep neural network, the kernel is initialized randomly and refined throughout the training process to discern patterns in the input data automatically.

The practical inputs for CNN can vary from real-world images to videos and audio signals, as represented by $f(x)$ in the equations. During the network training phase, the kernel or filter is updated via the backpropagation algorithm. A 2D convolution example is depicted in Figure 1.2, where the left matrix represents the input, the middle matrix is the kernel, and the right matrix displays the filtered output, showing the detected patterns.

The convolutional layers of a CNN are useful at adapting to 1D, 2D, and 3D data, capturing static features in images and dynamic patterns in temporal data. Convolutional Neural Networks in the one-dimensional form are primarily applied in signal processing applications, including the classification of Electrocardiogram (ECG) signals [15] and speech recognition [61]. Two-dimensional Convolutional Neural Networks are utilized for tasks associated with image processing, such as classifying images with multiple labels [59] and categorizing medical imagery [29]. Three-dimensional Convolutional Neural

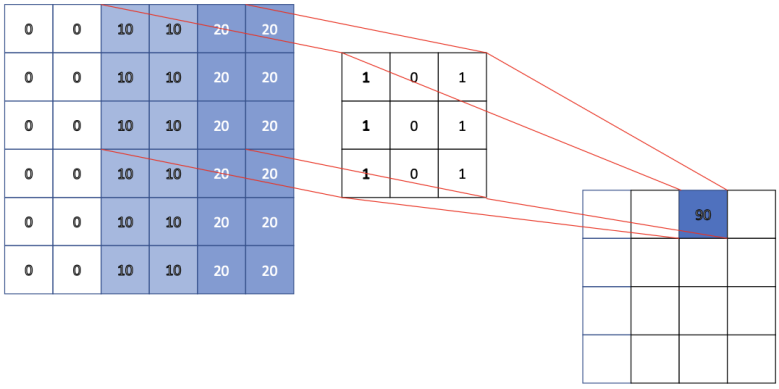


Figure 1.2: 2D convolution operation example. The left matrix represents the input data, the middle matrix is a 3x3 convolutional kernel, and the right matrix is the output data matrix, highlighting the extracted patterns.

Networks are specifically designed to analyze volumetric data. They have shown notable success in areas such as segmenting brain lesions [27] and identifying human actions [26].

1.2.3 3D Convolutional Neural Networks

A 3DCNN is a neural network specifically designed to process and analyze three-dimensional data. While a 2D CNN, as depicted in Figure 1.3, is suitable for handling matrix-like data (such as images) and can not handle 3D temporal datasets. On the other hand, 3D CNN can interpret volumetric data where depth, height, and width are all crucial components, as depicted in Figure 1.4. This unique feature makes 3D CNN highly proficient at analyzing data with three-dimensional spatial relationships, allowing them to capture patterns and features that unfold across all three axes. These networks use 3D convolutions to effectively learn from data represented in three dimensions, which enables them to extract spatial hierarchies and context in a way that flat 2D filters cannot achieve. An example of a 3D CNN architecture is C3D [55], widely used in video analysis tasks.

3D CNNs can process various three-dimensional datasets such as medical imaging (fMRI and CT scans), video sequences, and any data where understanding the context within volume or time is crucial. Medical imaging benefits from 3D CNNs as the depth of the scan provides vital insights into anatomy or pathology that are not apparent in 2D slices alone. Additionally, 3D CNNs are suitable for video analysis as the added dimension captures the temporal progression.

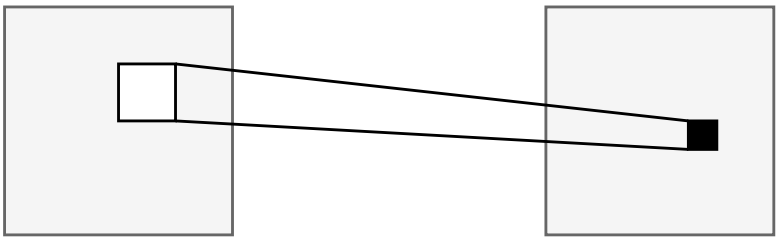


Figure 1.3: A sample of 2D CNN

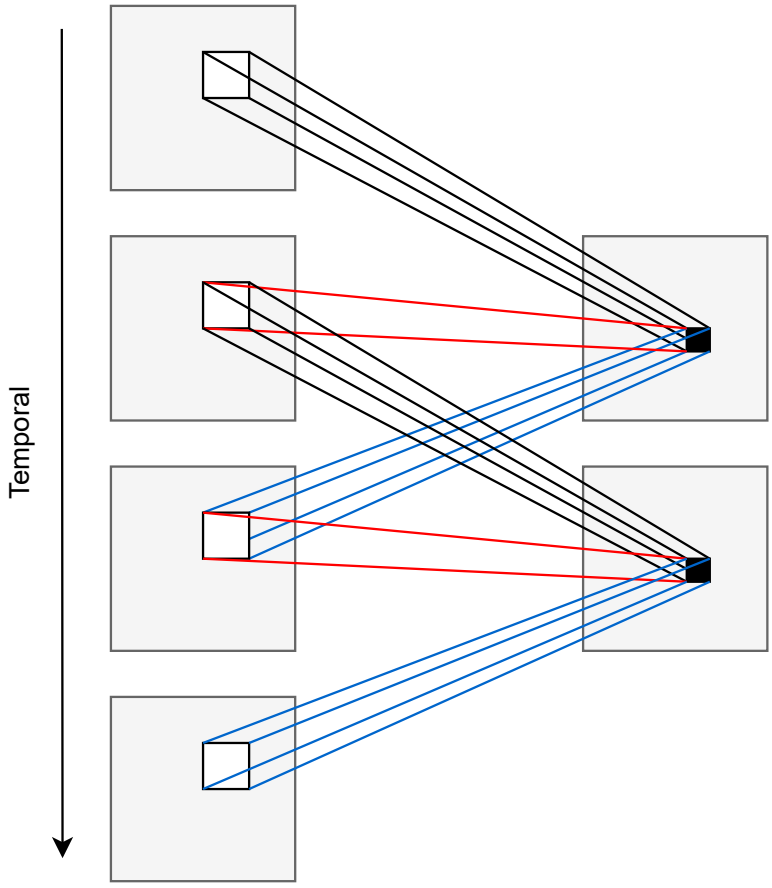


Figure 1.4: A sample of 3D CNN depicts weight values of three dimensions, where connections of the same color represent the shared weight.

It is ideal for applications like action recognition, where comprehending the sequence of movements is crucial for accurate classification or analysis. Similarly, the temporal dimension is essential in fMRI datasets, and 3D CNNs can be used to process them effectively.

The applications of 3D CNNs are significant, particularly in the context of fMRI and brain research, showcasing the depth of analysis and insight they provide. In neuroimaging, 3D CNNs are crucial for analyzing fMRI data, iden-

tifying and classifying brain activity patterns, assessing neural connectivity, and understanding brain function in various health and disease states. They are instrumental in advancing research in mental health, where they contribute to studying neurological disorders, brain development, and aging. Additionally, 3D CNN facilitates interpreting complex brain imaging data, supporting advancements in neurosurgical planning and the development of new therapeutic strategies.

The primary difference between 3D CNNs and their 2D counterparts lies in their ability to process an additional dimension. This depth allows 3D CNNs to capture patterns over time or through a volume, which is particularly beneficial for datasets where the third dimension conveys crucial information. However, this additional dimension also means that 3D CNNs are typically more computationally intensive and may require more data to train effectively. They also require careful parameter tuning and can be sensitive to overfitting, representing a trade-off between complexity and performance.

1.2.4 Long Short-Term Memory(LSTM)

LSTM networks, a type of Recurrent Neural Network (RNN), are specifically designed to learn from sequences, making them an ideal choice for time-series data or any dataset where context over time is important. In the area of fMRI and other neuroimaging data, LSTMs provide distinct advantages in understanding the temporal dynamics and dependencies within brain activity patterns.

LSTM networks have practical applications in the field of fMRI research. They are useful for analyzing time-series data from brain scans, which allows researchers to study the dynamic changes and connectivity patterns that occur over time. For example, they can be used to track the progression of neuronal activity across different brain regions during cognitive tasks, providing valuable insights into how information flows and evolves through neural circuits. This type of analysis is crucial for understanding the temporal complexity of brain functions and can reveal how different brain regions interact over time during various mental processes.

fMRI studies require understanding changes in brain activity over extended periods. LSTMs are particularly advantageous because they can remember and utilize past information. For instance, LSTMs can analyze subtle changes in brain activity patterns over months or years and identify predictive biomarkers

of disease progression in neurodevelopmental disorders such as Alzheimer's. LSTMs are being used in mental health research to analyze fMRI data and detect patterns or anomalies related to psychiatric conditions [14]. These networks examine the temporal sequences of brain activity, helping to identify characteristic patterns of neural dynamics associated with specific disorders [14]. This information can potentially aid in diagnosing and monitoring mental health conditions. For example, LSTMs can detect subtle changes in brain activity that correspond to fluctuations in mood or anxiety levels, providing a valuable tool for understanding and treating mental health disorders [45].

An LSTM unit has three important gates, as depicted in Figure 1.6: the input, output, and forget gates. These gates are responsible for deciding whether new input should be allowed to enter the unit (input gate), discarding information that is no longer needed (forget gate), or affecting the output of the unit at the current time step (output gate). To summarize the mathematical operations within an LSTM cell, we can say that:

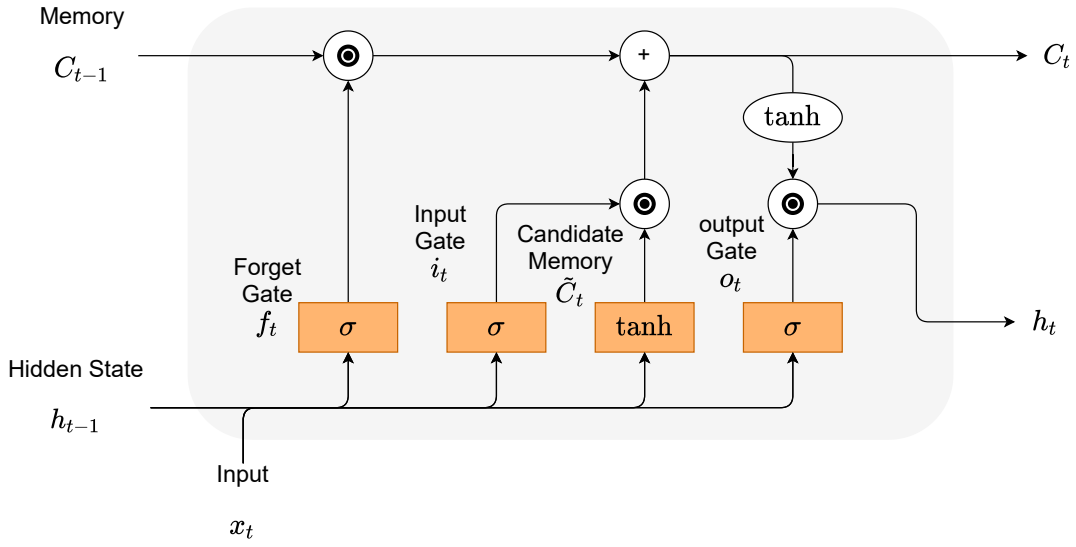


Figure 1.5: Architecture of a LSTM cell

- Forget Gate (f_t): This gate decides what information is discarded from the cell state. It looks at the previous output h_{t-1} and the current input x_t and applies a sigmoid function.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1.6}$$

  Here, σ is the sigmoid function, W_f is the weight matrix for the forget gate, and b_f is the bias.

- Input Gate (i_t) and Candidate Cell State ($\tilde{C}_t$): The input gate decides which values will be updated, and a tanh layer creates a vector of new candidate values that could be added to the state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{1.7}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{1.8}$$

- Update Cell State (C_t): The old cell state C_{t-1} is updated to the new cell state C_t. The update is done by multiplying the old state by f_t and adding the product of the input gate i_t and the candidate cell state $\tilde{C}_t$.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{1.9}$$

- Output Gate (o_t) and Output (h_t): The output gate decides the next hidden state. The hidden state contains information on previous inputs. A sigmoid layer filters the state, which is then output after being processed by a tanh layer.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{1.10}$$

$$h_t = o_t * \tanh(C_t) \tag{1.11}$$

The LSTM's ability to update, forget, or output information allows it to capture long-term dependencies in sequential data, making it particularly useful in contexts where the temporal dimension is crucial, such as analyzing fMRI time series data where the temporal evolution of brain activity is key for understanding neural processes.

1.2.5 U-Net

The U-Net architecture was initially designed for biomedical image segmentation [47] and has shown exceptional performance in this field. Its unique U-shaped structure consists of a contracting path to capture context and an expanding path for precise localization, as shown in Figure 1.6. This makes it particularly effective in dealing with the high variability of medical images, which is why it is widely used for segmenting anatomical structures or regions of interest in various imaging modalities, including fMRI.

The U-Net architecture has a unique design that consists of two primary paths: a contracting path and an expansive one. The contracting path, also called the encoder, captures the context within the data using a series of convolutional and max pooling layers. These layers reduce the spatial dimensions while increasing the feature depth. In contrast, the expansive path, also known

as the decoder, aims to enable precise localization by recovering the spatial resolution through up-convolution and concatenation operations with corresponding feature maps from the contracting path. The concatenation process is essential as it combines high-level feature information with spatial detail, facilitating precise segmentation. Furthermore, at the final layer, a 1x1 convolution is applied to map the feature-rich output to the desired number of classes for classification.

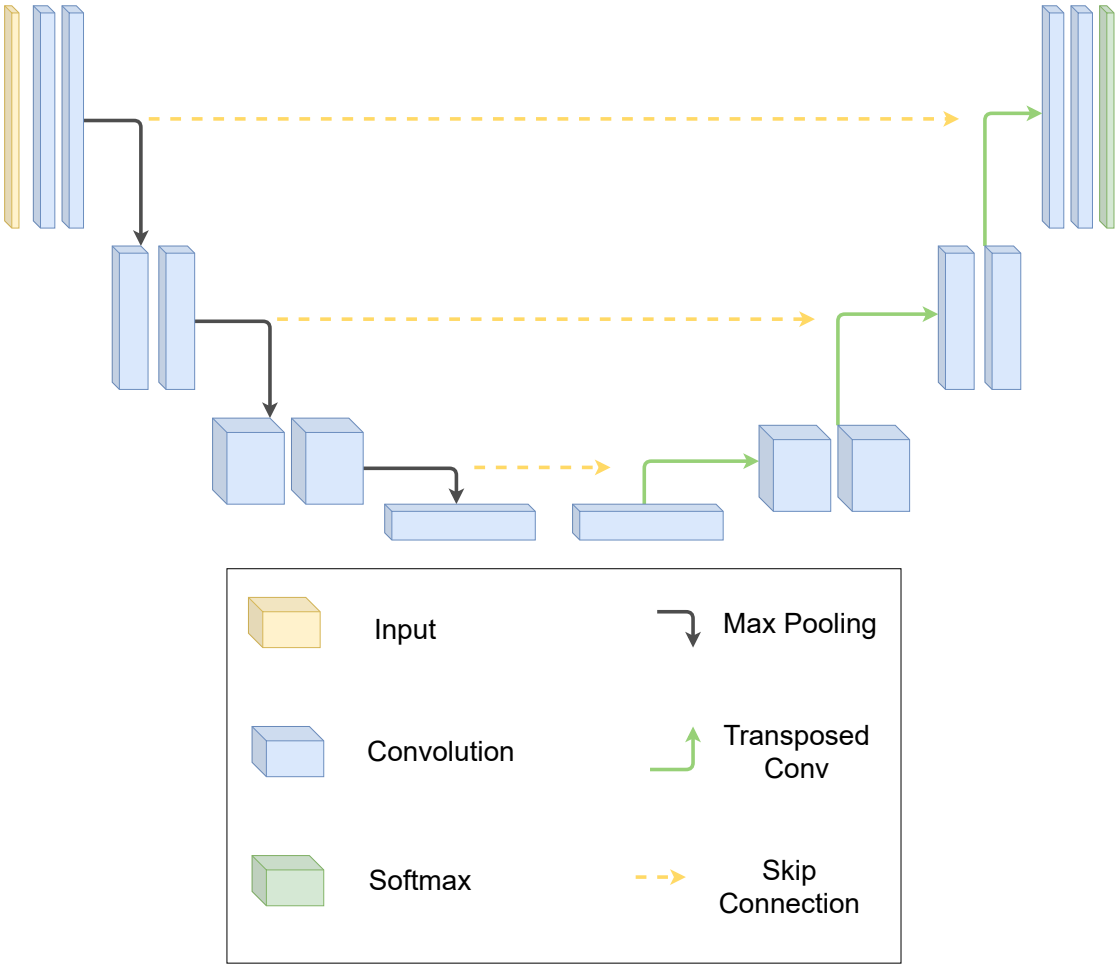


Figure 1.6: Architecture of U-Net model

U-Net can be utilized in the analysis of fMRI data to perform tasks such as segmenting brain regions and aiding in identifying and analyzing functional brain networks or areas that are activated under specific conditions [18]. It can be applied to isolate particular brain regions engaged during cognitive tasks, emotional responses, or resting states. This precise segmentation can facilitate the study of brain function, connectivity, and the effects of various

neurological conditions on brain structure and activity, making it an invaluable tool for researchers.

Furthermore, U-Net's ability to handle three-dimensional data makes it well-suited for analyzing volumetric fMRI scans, where capturing spatial context is crucial. Its architecture integrates local information from downsampling paths with global information through upsampling paths, which is critical to its success in producing high-resolution segmentations. This is particularly beneficial in fMRI studies, where distinguishing between closely adjacent brain regions or detecting subtle changes in brain activity is essential.

U-Net is a powerful tool in clinical research that can aid in developing diagnostic tools based on fMRI. Segmenting regions affected by stroke or identifying patterns associated with neurodevelopmental diseases such as Alzheimer's or Parkinson's can provide valuable insights into disease progression and outcomes. Automating the segmentation process with U-Net can also significantly speed up the analysis of large fMRI datasets, giving researchers more time to focus on higher-level interpretation and hypothesis testing.

1.2.6 Spatio-Temporal Model

Analyzing functional fMRI data is critical for understanding the complex dynamics of the brain's activity. Traditional methods often struggle to capture the whole essence of the brain's spatial structures and temporal patterns simultaneously. The spatio-Temporal model that merges U-NET 3DCNN with LSTM networks, as depicted in Figure. 1.7 aims to elevate the analysis of fMRI data by concurrently mapping the brain's spatial and temporal dimensions.

This advanced model is designed to navigate the brain's complex spatial configurations and the dynamic nature of neural signals over time. The U-NET 3DCNN component, proven for its proficiency in medical image segmentation, captures essential spatial details across multiple resolutions. This capability is crucial for accurately identifying specific brain regions in fMRI scans. Complementarily, integrating LSTM networks allows for adept modeling of temporal dependencies within fMRI time series, offering a deeper understanding of how spatial patterns of brain activity evolve.

Fusing U-NET's spatial precision with LSTM's temporal insight creates a powerful tool for comprehensive fMRI data analysis. This model uncovers the dynamic interactions between different brain regions, offering new pathways

for exploring cognitive processes and neural functions and potentially revolutionizing the understanding of neurodevelopmental diseases. By leveraging the model's detailed analysis, researchers can gain invaluable insights into the brain's operational mechanisms embedded in the complex spatiotemporal neural activity.

These limitations informed this model's development of existing analytical methods to capture the dual complexities of fMRI data. Integrating U-NET with LSTM layers addresses these challenges head-on, significantly contributing to neuroimaging analysis. It opens doors to novel explorations and understandings of the brain's functionality, setting a new standard for fMRI data analysis.

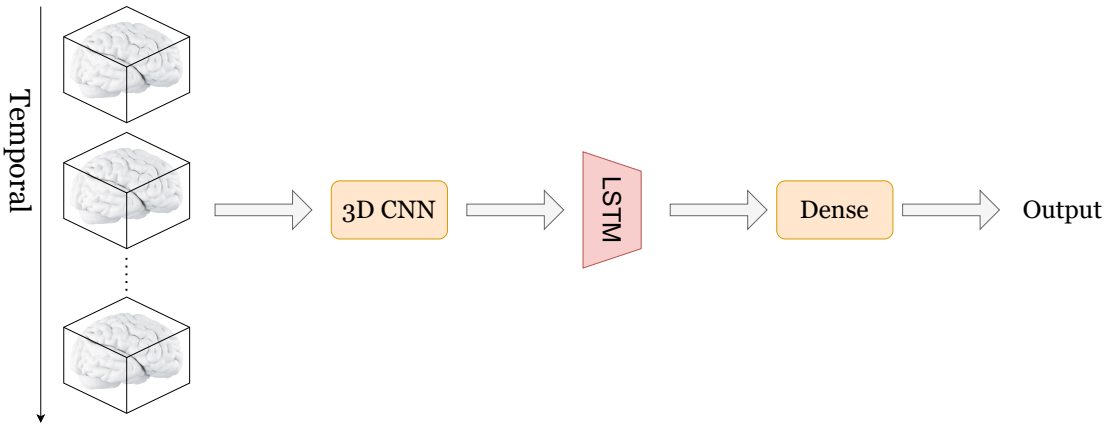


Figure 1.7: Spatio-temporal Model Workflow, showing the process from sequential input through a 3D CNN and LSTM to the dense layer and output

1.2.7 Integrating LSTM Layers into U-Net

The integration of Long-Short-Term Memory (LSTM) layers into the U-Net architecture represents a pivotal enhancement for analyzing fMRI data. This enables the model to capture temporal dependencies alongside spatial features effectively. This hybrid model is particularly adept at addressing dynamic processes in the brain, such as monitoring the progression of brain tumors over time.

Encoder and LSTM Layers: Temporal Feature Analysis

The architecture retains the original U-Net's encoder section for spatial information compression, extracting feature maps from the input image. Positioned strategically after the encoder and before the decoder, the LSTM layers introduce a novel processing stage. They interpret the depth dimension of the encoder's output feature maps as a temporal sequence. This reinterpretation allows the LSTM to analyze the progression of spatial features across time, akin to observing changes in a series of frames. By doing so, the LSTM layers can identify and memorize patterns of change in the anatomical structures or pathology captured in successive fMRI scans, effectively bridging the spatial and temporal dimensions of the data.

Decoder: Enhanced Spatial-Temporal Reconstruction

The temporal feature maps, enriched with dynamic insights from the LSTM layers, are then conveyed to the decoder. The decoder, leveraging this spatial and temporal information composite, embarks on reconstructing the segmented output with heightened accuracy. This process ensures that the segmentation not only delineates the morphology and location of brain structures but also encapsulates their temporal dynamics. Such an enriched output is invaluable, offering nuanced insights into the evolution of pathological or anatomical changes over time, thereby significantly enhancing the utility of fMRI data in clinical and research settings.

2

State of the Art

In this section, we provide a selective review of recent studies on machine learning/deep learning algorithms for fMRI data and disease classification.

Research using fMRI spans various fields; for some notable areas of study, we can mention natural image reconstruction from fMRI data using deep learning, as in the Roman Beliy paper [6]. Research on fMRI data and its neurophysiological implications is explored in [32], while [20] delves into the study of the human visual cortex using fMRI. The anatomy of language is comprehensively reviewed in [44]. The current landscape of fMRI research in the neuroscience domain is highlighted by [43], and [42] projects the future trajectory of fMRI in cognitive science studies, suggesting a rich field of ongoing and future explorations.

In this work, we focus on building a classification model to classify three neurodevelopmental disorders based on fMRI data. Many research studies have been proposed to classify neurodevelopmental disorders based on fMRI data. In recent years, numerous studies leveraging machine learning and deep learning have emerged for this aim, showing a marked preference for these techniques due to their effectiveness in classifying such disorders.

For this kind of research, we can mention studies such as [49]. This research presents a deep learning-based methodology for recognizing Alzheimer's disease using fMRI data. It utilizes a CNN with the LeNet-5 architecture for classifying fMRI images of Alzheimer's patients against controls. The study demonstrates a high accuracy rate of 96.85% in classifying the Alzheimer's condition. It suggests that deep learning, particularly CNNs, is a powerful tool for distinguishing clinical data from healthy data in fMRI and holds promise

for future diagnostic tools. The research indicates a potential expansion of this methodology to predict different stages of Alzheimer's across various age groups.

Deep learning models have also been applied to the study of early detection and staging of Alzheimer's, especially crucial for older populations. These models excel in identifying and categorizing the stages of Alzheimer's disease, from mild cognitive impairments to more advanced stages; for example, [46] discusses the development of a deep learning model for diagnosing Alzheimer's disease through multi-class classification of its various stages, utilizing resting-state fMRI data. The study employs a ResNet-18 architecture to distinguish between different stages of Alzheimer's disease, including cognitively normal (CN), significant memory concern (SMC), early mild cognitive impairment (EMCI), mild cognitive impairment (MCI), late mild cognitive impairment (LMCI), and Alzheimer's disease (AD). The model is trained from scratch and uses transfer learning to improve its performance. The research demonstrates that this approach can achieve high accuracy, suggesting its potential for assisting early diagnosis and clinical decision-making in Alzheimer's disease.

In schizophrenia diagnosis, there have been advancements through the FMRI and deep learning methods. [62] improves classification effects of magnetic resonance data by applying a CNN algorithm, specifically the VGG16 net, enhanced with transfer learning. The study demonstrates an 84.3% classification accuracy, suggesting that deep learning can significantly aid in early schizophrenia diagnosis and address small sample and high-dimensional data classification challenges, thereby enhancing the generalization capability of deep learning models in clinical settings.

In a different research, Luca Steardo Jr. et al.2020 [53] provide a comprehensive examination of how Support Vector Machine (SVM) techniques are employed in the diagnosis of schizophrenia using fMRI data. They highlight the effectiveness of SVM in distinguishing between patients with schizophrenia and healthy controls. Their review underscores the potential of SVM as a non-invasive, cost-effective tool for early-stage identification of schizophrenia, emphasizing its superiority in accuracy compared to traditional diagnostic methods. This approach not only aids in early detection but also contributes to the refinement of clinical diagnostic processes, potentially saving crucial time in patient care management.

In another study [63], where they achieved a high accuracy in schizophre-

nia classification, they used FMRI data and applied independent component analysis (ICA) and kernel principal component analysis (KPCA).

In [60], a multi-scale recurrent neural network (MsRNN) model is developed for accurate multi-class classification of schizophrenia, psychotic bipolar disorder, schizoaffective disorder, and healthy controls. The framework also employs t-Distributed Stochastic Neighbor Embedding (tSNE) clustering to visualize relationships between disorders and uses a leave-one-feature-out method for biomarker identification. It achieves significant classification accuracy and provides insights into the discriminative brain regions associated with these disorders.

Another research that shows the potential of deep learning approaches in enhancing diagnostic accuracy for psychiatric disorders is [30], which is focused on first-episode psychosis (FEP), bipolar disorder (BD), and healthy controls (HC) using structural Magnetic Resonance Imaging (sMRI) data. They introduce two novel end-to-end CNN architectures. These CNN models outperformed traditional classifiers in distinguishing between FEP, BD, and HC based on gray matter volume images, demonstrating high precision, recall, F1-score, accuracy, and Area Under the Curve (AUC) in both binary and three-way classification tasks.

For the research that has been conducted for ADHD disorder, we can mention [31]. This research assesses the diagnostic capability of analyzing resting-state fMRI data for ADHD, employing functional connectivity and machine learning techniques. The study meticulously identified significant differences in brain regions, such as the anterior cingulate cortex and cerebellum, between ADHD patients and healthy controls. Utilizing a linear discriminant analysis classifier, it achieved an average classification accuracy of 80.08%, demonstrating the effectiveness of resting-state fMRI combined with machine learning in diagnosing ADHD.

In another study by Mao et al. (2019) [35], they applied a spatio-temporal CNN without reliance on hand-crafted features. They use granular computing for analyzing fMRI scans, resulting in a method that surpasses traditional classification accuracies, reaching a significant 71.3 percent on the ADHD-200 dataset. This research shows the potential of deep learning models to capture complex spatio-temporal patterns in brain activity.

In [25], a classification framework is designed to address the challenges posed by the heterogeneity of multi-site medical datasets, specifically applied

to the ADHD-200 collection. The study leverages decision trees within a hierarchical classification scheme to improve the interpretability and diagnostic support of computational models dealing with ADHD. This approach aims at enhancing the diagnosis and understanding of ADHD by integrating diverse data sources, thus reflecting on the potential of hierarchical and interpretable models in clinical decision support systems.

3

Datasets and models

3.1 Dataset

This segment discusses the data sources, detailing the preprocessing steps undertaken for the fMRI data used in this thesis. It provides an overview of the data, including how it was accessed and its general characteristics. The section further elaborates on the data processing techniques employed to adapt it for the spatio-temporal models explored in the research. It outlines the data structure requirements of our models and discusses how the data was processed, considering the limitations of resources available for training. This explanation ensures a comprehensive understanding of the data's journey from collection to its final form.

3.1.1 Data Description

The data used in this thesis was obtained from the OpenfMRI database, and its accession number is ds000030 [56]. It is titled 'UCLA Consortium for Neuropsychiatric Phenomics LA5c Study'. This study is initiated by Consortium for Neuropsychiatric Phenomics (CNP) and is done by Bilder and colleagues [57]. CNP is funded by the National Institutes of Health (NIH) Roadmap Initiative and aims to uncover the genetic and environmental foundations behind variations in psychological and neural system phenotypes. It seeks to understand how the human genome influences complex psychological syndromes and aims to revolutionize the development of new treatments for neuropsychiatric

disorders.

This study has enrolled a diverse group of individuals, including 119 HC from the community and others diagnosed with schizophrenia (SCH) (50), BD (49), and ADHD (40), aged 21-50. These participants were selected based on specific criteria from the Los Angeles area and underwent extensive neuropsychological testing and fMRI scanning. Inclusion criteria were based on NIH racial and ethnic designations and educational attainment, with exclusions applied to minimize confounding in genetic research. Language for testing was determined for bilingual individuals through a verbal fluency test, and all participants were screened for various health conditions, substance use, and psychiatric history. Drug use was checked via urine analysis [56]. During the course of the investigation, various methods were employed to gather data, including conducting interviews, utilizing rating scales, implementing self-report assessments, and conducting neurocognitive evaluations through traditional paper-and-pencil tests as well as digital formats. Specifically, the collection of neuroimaging data comprised techniques such as structural Magnetic Resonance ImagingsMRI, High Angular Resolution Diffusion Imaging (HARDI), Resting-State fMRI (RS-fMRI), and five distinct types of task-based fMRI (T-fMRI).

The fMRI data used in this study was preprocessed according to established protocols on the OpenFMRI platform [17]. The preprocessing pipeline included data conversion, anonymization, motion correction, and brain extraction using FMRIB Software Library (FSL) tools, along with cortical reconstruction utilizing Freesurfer. Multi-level fMRI analysis was performed with FSL, incorporating nuisance regressors for motion. Additionally, the study incorporated manual quality control checks for brain extraction quality and the accuracy of functional and anatomical registrations. Details and parameters for each processing step are accessible via the OpenFMRI GitHub repository [41].

3.1.2 Selection and pre-processing

In the previous section, the distribution of samples among the groups was detailed: 119 for HC, 50 for SCH, 49 for BD, and 40 for ADHD. While preparing the data for the computational model, it was noted that fMRI data is considerably large, and the aggregate volume of the dataset was significant, leading to size limitations for uploads on the university's data management system. To address this, 19 samples were randomly excluded from the HC group, which had the

largest initial sample size. This exclusion was done to maintain data integrity and comply with the system's constraints. The updated frequency of each group post-exclusion is depicted in Figure 3.1.

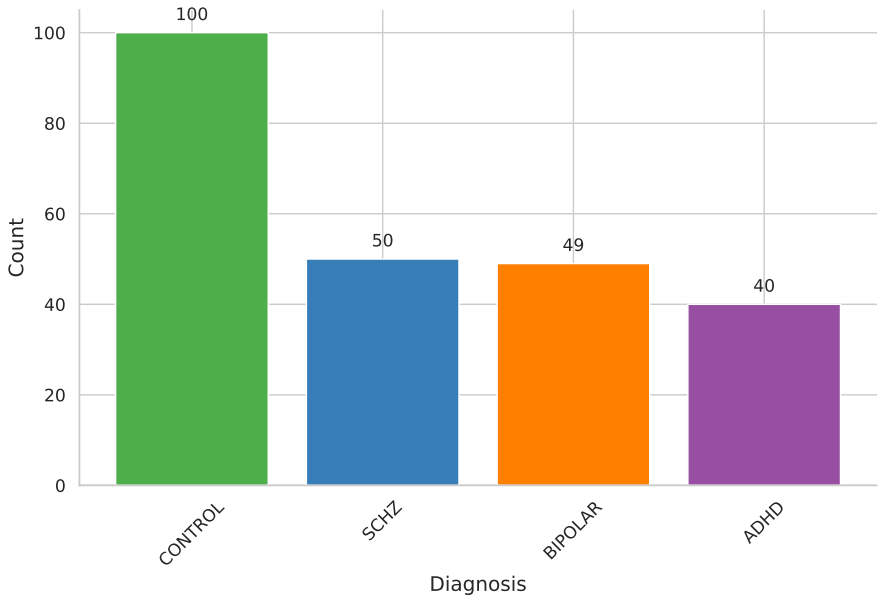


Figure 3.1: Number of Individuals in Each Diagnosis Group.

In this study, each participant's original fMRI data comprised four-dimensional images shaped $(65, 77, 49, 152)$, in Neuroimaging Informatics Technology Initiative (NIFTI) format. These dimensions correspond to the spatial (height, width, depth) and temporal (time points) axes of brain activity.

To align the data with the input requirements of a combined CNN and LSTM architecture, a transposition was performed to place the time points as the first axis so the shape of each sample would be $(152, 65, 77, 49)$.

Due to computational limits, a subset of 40 time points was extracted [3.3] from the full 152 after discarding the initial 10 to avoid signal instability at the start of the sequence.

The final dataset maintained the spatial resolution $(65 \times 77 \times 49)$ voxels and included 40 time points, balancing the need to analyze spatial and temporal features for the machine learning process.

In the CNN models utilized in this thesis, the original spatial shape of fMRI data was preserved. However, for U-Net CNN architectures employed in this study, the preprocessing required adjustments due to the model's structure. The original dimensions of the fMRI scans were $65 \times 77 \times 49$. The network archi-

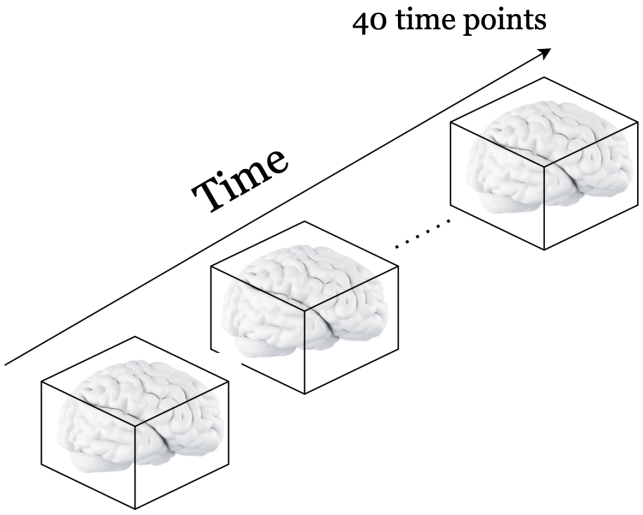


Figure 3.2: Sequential Extraction of Brain Activity over 40 Time Points

tecture incorporates sequences of down-sampling and up-sampling operations that halve and then double the input volume size, respectively.

To accommodate these operations without dimensional inconsistencies or data cropping, the dimensions of input images were modified to be evenly divisible by 2^n, where n represents the number of down-sampling steps in the network and is a power of 2.

Aiming to conserve as much information as possible, image dimensions were resized to powers of 2, closely matching the largest original dimension (77). Consequently, spatial dimensions were adjusted to $80 \times 80 \times 80$, incorporating padding to the width, height, and depth axes to maintain uniform volumes processable by the U-Net architecture. This preparation is crucial for applying CNNs to fMRI data, enabling the network to learn and extract spatial features from the scans effectively.

3.2 Models

This section outlines the technical implementation of neural network models developed for analyzing fMRI data, focusing on their structural components and the rationale behind their design. It starts with a simple 3DCNN model without considering the temporal aspect of fMRI data, then there is an overview of the basic 3DCNN-LSTM model, emphasizing its dual capability to dissect spatial

features through 3DCNN layers and temporal dynamics using LSTM layers.

Furthermore, the section transitions to elaborate on a more sophisticated U-NET 3DCNN and LSTM model. This architecture is specifically engineered to tackle the multifaceted spatial-temporal patterns present in fMRI data. It adopts a U-Net structure known for its efficiency in medical image segmentation, combining it with LSTM to analyze temporal dependencies.

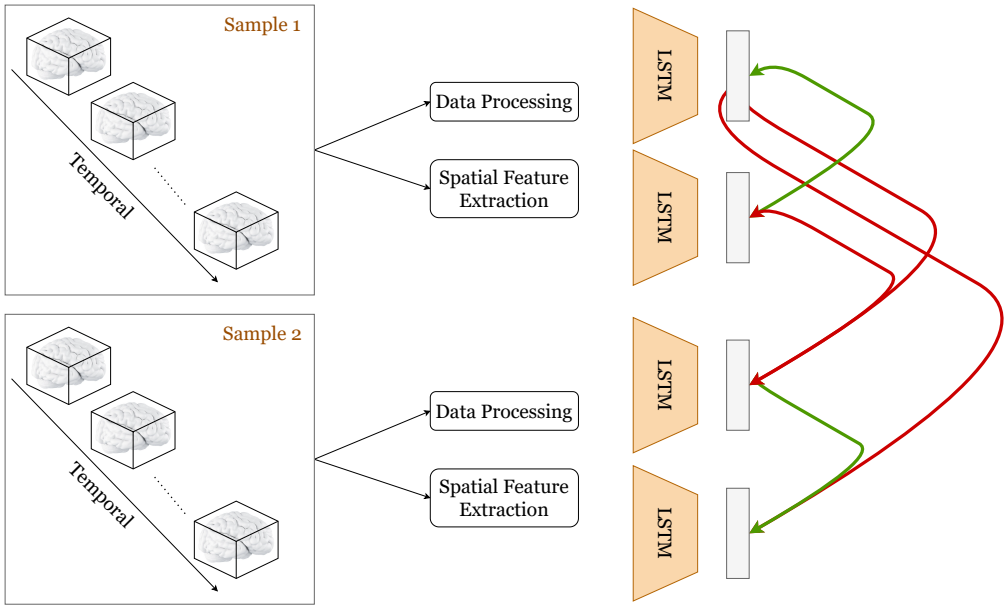


Figure 3.3: Simple overview of the proposed model

3.2.1 3DCNN Model

In this study, a 3D Convolutional Neural Network (3DCNN) is utilized to analyze functional magnetic resonance imaging (fMRI) data. This method excels in extracting spatial information from brain images but does not account for temporal changes or the sequence of data points over time. The decision to implement this model stems from an interest in evaluating the classification outcomes when temporal information is intentionally disregarded and each time point is considered separately.

Each time point in the fMRI data is treated as independent, with no inherent sequence among them. This involves shuffling the time points to ensure that consecutive slices from the same category (e.g., images pertaining to the same patient or condition) are not adjacent within the dataset. The dataset comprises 100 HC, 50 SCH, 49 BD, and 40 ADHD samples. From each sample, 40 time

points are extracted, yielding 4000 slices for HC, 2000 for SCHZ, 1960 for BD, and 1600 for ADHD. These slices are then shuffled and divided into training, testing, and validation sets to be processed by the 3DCNN model.

The model starts with two convolutional layers. These layers are key to the model's ability to pick up on patterns. The first convolutional layer has 16 filters, while the second one has 32 filters. Each filter uses a small, three-dimensional kernel of size (3,3,3), which moves through the input volume to produce a feature map. This process transforms the spatial input data into a form the model can use to learn patterns. Following each convolutional layer, there's a MaxPooling layer with a pool size of (2,2,2). MaxPooling is a downsampling technique that reduces the dimensionality of the feature maps, making the model more efficient and helping to prevent overfitting by extracting the most dominant features. After each MaxPooling layer, a Dropout layer is applied, with a dropout rate of 25%. After processing through the convolutional and pooling layers, the model employs a Flattening layer to convert the 3D feature maps into a 1D vector. This is essential for transitioning to the next phase, which involves two fully connected (dense) layers. The first dense layer has 64 units and uses the ReLU activation function to introduce non-linearity, enabling the model to learn complex patterns. The final layer adjusts its number of units to match the number of classes using for classification, and utilizes a softmax activation function. This function is crucial for multi-class classification, as it calculates the probability distribution across the classes, with the model predicting the class having the highest probability.

Despite the relatively simple structure of the 3DCNN model, the considerable volume of data and the intricacy of handling 3D spatial dimensions present a significant challenge for training.

3.2.2 Basic 3DCNN-LSTM model

As mentioned in previous part, 3DCNNS are used to extract spatial information from fMRI data. To analyze temporal sequences, LSTM layers are employed. A basic model, consisting of a set of 3DCNN layers, is used as a starting point for analysis. Despite its simplicity, this model has a large number of parameters due to the convolution operations that take place across the three spatial dimensions of the data. This complexity highlights the challenges in capturing the intricate spatial-temporal dynamics that are inherent in fMRI datasets.

This 3DCNN is followed by LSTM layers to harness the temporal dimension of fMRI data. The model architecture begins with TimeDistributed Conv3D layers, employing a (3,3,3) kernel size with 'RelU' activation and 'same' padding to maintain dimensionality. Sequential layers include MaxPooling3D with pool-size of shape (2,2,2) for dimensionality reduction and dropout with rate of 0.25 to prevent overfitting. Three concolutional layesr with same Maxpooling and dropout has been used bu with increasing filter size in each convolution, the first convolution has 16 filters and the second and third convolutions 32 and 64 filters respectively. After convolutional layers, a Flatten layer feeds into LSTM unit with 64 hidden nodes. After LSTM a dense layer is added with 128 nodes and 'RelU' activation function, fallowed by 0.5 dropout rating. Finally, a dense layer with n nodes, where n is the number of our classes, comes with a 'softmax' activation function.

After this layer, the models are compiled. In this part, different learning rates were tested. An adaptive learning rate strategy was implemented utilizing the ReduceLROnPlateau callback from TensorFlow/Keras. This method adjusts the learning rate by reducing it when the validation loss does not improve over specified epochs. It aims for finer model tuning by minimizing steps in the parameter space upon encountering minimal loss improvements with a given reduction factor. The learning rate continues to decrease to a minimum threshold, facilitating a dynamic adjustment to the training process for improved efficiency and performance.

3.2.3 U-NET 3DCNN AND LSTM MODEL

The architecture of the U-Net + LSTM model is designed to process the complex spatial-temporal patterns present in four-dimensional fMRI data. The input layer is tailored to accommodate data shaped $(40, 80, 80, 80, 1)$, where the dimensionality represents time points followed by the spatial volume of the brain scans, all with a single channel.

At the outset, the first convolutional block introduces two TimeDistributed Conv3D layers, each employing 18 filters of (3,3,3) dimensionality. These layers utilize ReLU activation functions and 'same' padding, ensuring the spatial dimensions are retained after the operation. This block is complemented by a TimeDistributed MaxPooling3D layer with a (2,2,2) pool size, which serves to reduce the spatial dimensions by half. A subsequent TimeDistributed Dropout

layer with a dropout rate of 0.25 is incorporated to mitigate the risk of overfitting.

As the data proceeds through the model, it encounters successive convolutional blocks that mirror the initial structure, doubling the number of filters in each consecutive block. Following the first block, the subsequent TimeDistributed Conv3D layers in these blocks employ 36, 72, and finally 144 filters, respectively. Each block is interspersed with TimeDistributed MaxPooling3D to continue the dimensionality reduction and TimeDistributed Dropout, with the latter's rate fixed at 0.5, deepening the network's robustness.

The network reaches its bottleneck, the most abstracted feature representation, with a pair of TimeDistributed Conv3D layers that expand the feature set to 288 filters. This central juncture captures the highest-level features before the model transitions into the expansive path.

The expansive path mirrors the encoding path in a symmetrical fashion but in reverse, utilizing TimeDistributed Conv3DTranspose layers to restore the spatial dimensions progressively. Each layer in this path employs strides of (2,2,2), effectively doubling the volume with each step. Concurrently, these upsampled outputs are concatenated with their corresponding feature maps from the downsampling path, weaving in fine-grained details for precise localization.

Flattening the multidimensional data, a TimeDistributed Flatten layer transitions the network from convolutional processing to sequential, allowing the following LSTM layer with 128 units to interpret the temporal aspects of the data. This LSTM layer captures the dynamics over time, a critical feature for fMRI analysis.

Culminating the network's layers, a Dense layer with 512 neurons applies ReLU activation to learn complex representations from the LSTM output. Finally, the output layer, employing softmax activation, articulates the probability distribution over the target classes, providing the classification output for each fMRI sequence.

4 Analysis

4.1 Performance Analysis

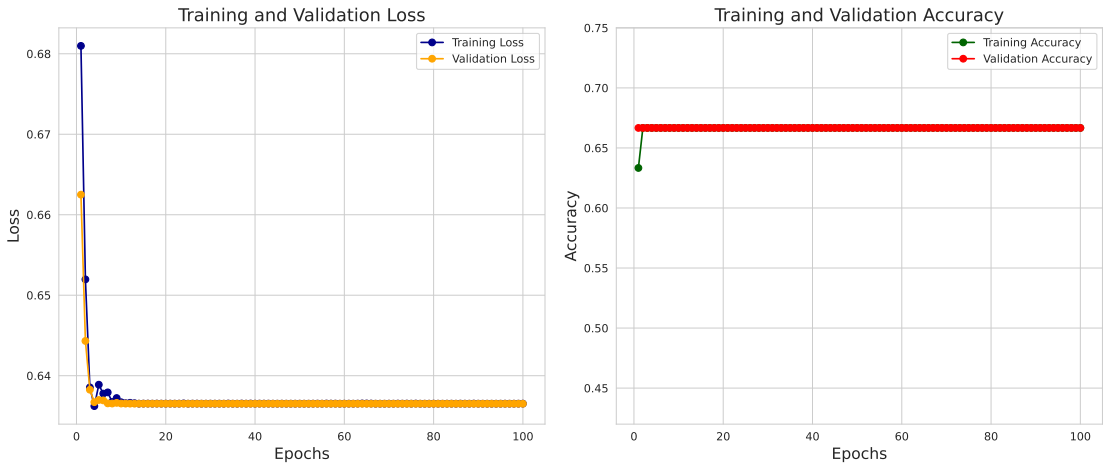
The base model's results described in part 1.2.3 are shown in table 4.1 and, the experimental results of U-NET 3DCNN and LSTM, described in 1.2.5 is shown in table 4.2.

Table 4.1: Training performance of different group comparisons with basic 3DCNN-LSTM model

Groups	Training loss	Validation loss	Training accuracy	Validation accuracy
HC-SCH	0.6365	0.6366	0.6667	0.6667
HC-ADHD	0.5990	0.5990	0.7149	0.7150
HC-BD	0.6313	0.6367	0.6741	0.6666
SCH-ADHD	0.6872	0.6882	0.556	0.556
SCH-BD	0.7000	0.7000	0.5100	0.5000
BD-ADHD	0.6877	0.6870	0.5472	0.5556

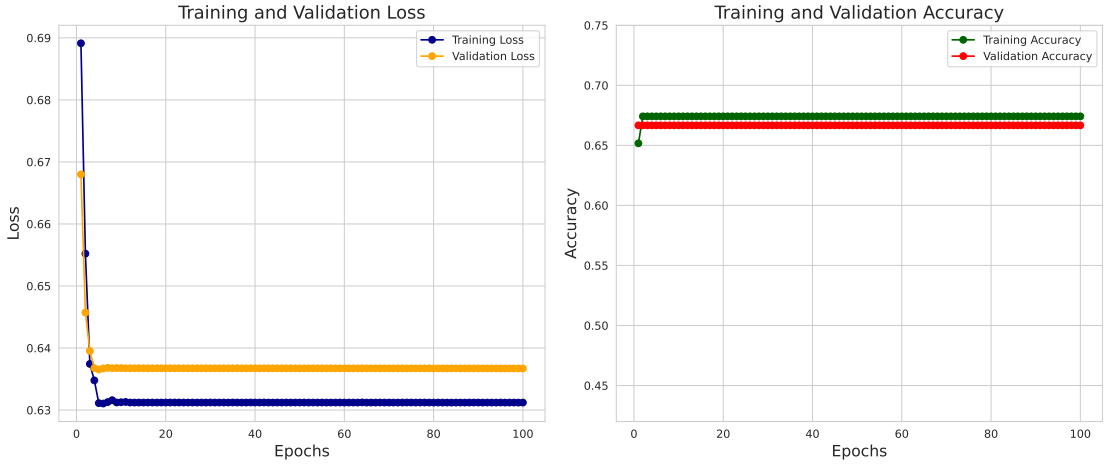
For both the base model and U-NET-LSTM models, more or less the same performance is monitored. The classification results in both model architectures highlight a better accuracy between the control and disease-specific groups, achieving a relatively higher accuracy compared to classification between diseased groups.

In figures 4.1 and 4.2, you can see the training curves and how the accuracy and loss are changing over epochs.



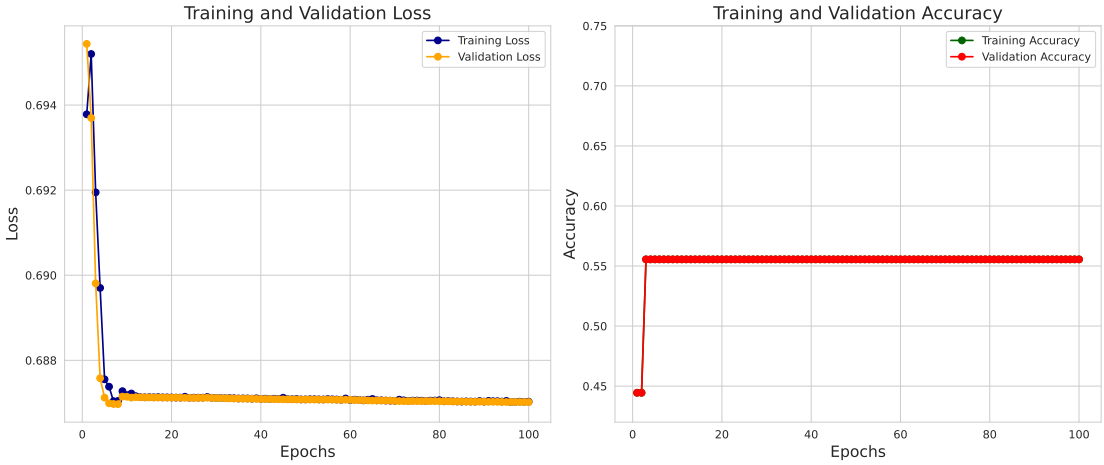
(a) SCH vs. HC

(b) ADHD vs. HC

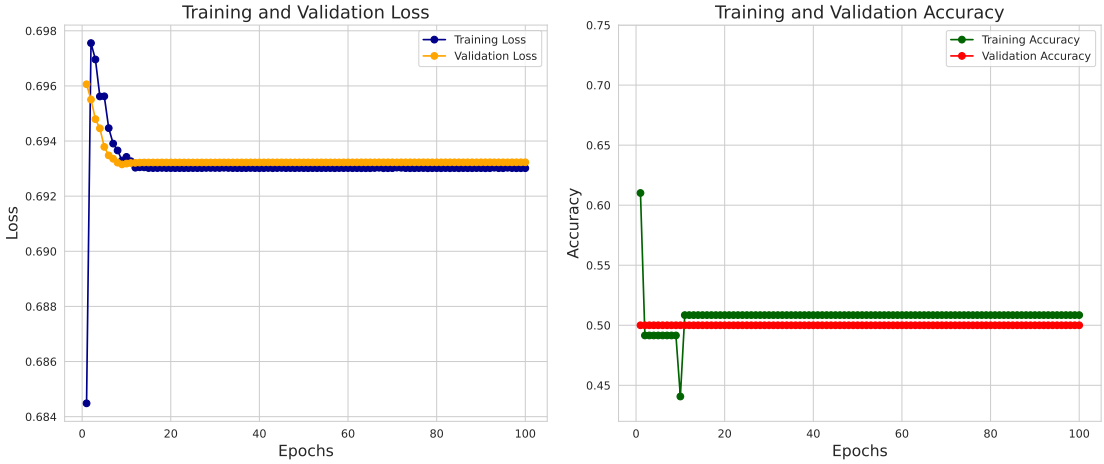


(c) BD vs. HC

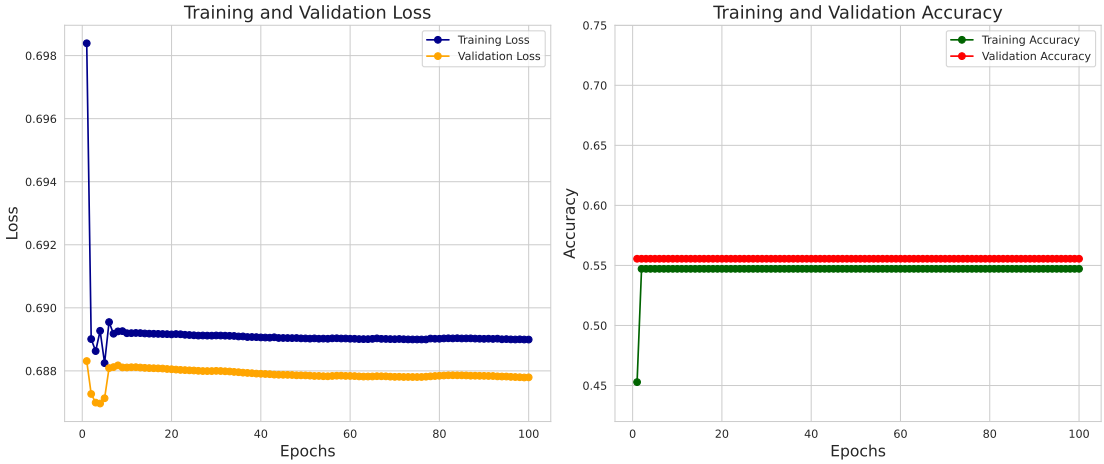
Figure 4.1: Performance Analysis of Classification Model in Distinguishing Healthy Controls from Disorder Groups.



(a) SCH vs. ADHD



(b) SCH vs. BD



(c) ADHD vs. BD

Figure 4.2: Performance Analysis of Classification Model in Distinguishing Disorder Groups

Table 4.2: Training performance of different group comparisons with U-NET-LSTM model

Groups	Training loss	Validation loss	Training accuracy	Validation accuracy
HC-SCH	0.6665	0.665	0.6667	0.6667
HC-ADHD	0.5983	0.5983	0.7143	0.7143
HC-BD	0.6312	0.6367	0.6742	0.6767
SCH-ADHD	0.6870	0.6870	0.5556	0.5556
SCH-BD	0.6930	0.6932	0.5010	0.5000
BD-ADHD	0.6890	0.6877	0.5470	0.5555

Fine-tuning the models involved implementing a dynamic learning rate approach. This process entailed reducing the learning rate incrementally whenever there was no decrease in loss after three epochs. The rationale behind this strategy is to enhance the optimization process of the models, which could potentially lead to more accurate predictions. In conjunction with this, a range of optimizers—Adam, SGD, and RMSprop—were evaluated for their impact on the convergence of the models. However, it was observed that the models quickly reached peak performance and demonstrated limited improvement in subsequent epochs.

Experiments were also conducted with varying filter sizes in the convolutional layers. The base model underwent training sessions with two distinct sets of filter configurations: one progressing from 8, 16, to 32, and another from 16, 32, to 64. The purpose of these experiments was to determine whether a greater number of filters, allowing for more intricate feature extraction, would result in improved performance of the model.

The distinctive characteristic of the U-NET architecture, where the number of filters expands with each convolutional layer, underwent rigorous testing as well. The number of filters was doubled at each layer, starting from a baseline of 8 in some cases and 16 in others. This method of exponentially increasing filters was chosen to assess if a more complex hierarchy of features would enhance the discriminative capabilities of the model.

For the LSTM layers, unit counts of 32 and 64 were trialed, investigating the models' performance with varying degrees of complexity. Additionally, to pursue a more in-depth analysis within the computational constraints, models incorporating two consecutive LSTM layers were explored, aiming to assess the potential benefits of a deeper sequential learning architecture.

For a valid evaluation of the model's performance, an example of a confusion matrix for one of the classifications is presented here. In figure 4.3, you can see the confusion matrix for classification between Healthy control and Schizophrenia in the U-NET-LSTM model, and in table 4.3, you can see the metrics related to this matrix. The accuracy is approximately 66.67%; while this may initially suggest a moderate level of predictive ability, a deeper analysis leveraging additional performance metrics reveals a more nuanced picture. Precision stands at 44.44%. This relatively low precision indicates that, of all the instances predicted as one group (here as HC), less than half were correct. Such a figure highlights the models' tendency to misclassify diseased instances as healthy.

The recall is recorded at 66.67%. This metric suggests that the model is reasonably capable of detecting healthy instances when they are present; however, as evidenced by the precision rate, this sensitivity comes at an increased cost in the form of false alarms.

The F1 Score, which combines precision and recall, is 53.33%, reflecting the imbalance between precision and recall. This score underscores the models' limited efficacy in balancing the trade-off between incorrectly classifying diseased instances and missing healthy instances.

It is important to consider that when one group is relatively larger than the other group in classification, the accuracy metric might be inflated due to this imbalance. This phenomenon, known as the accuracy paradox, can mislead the evaluation of model performance if accuracy is considered in isolation.

In scenarios where healthy cases are more prevalent, a model can achieve high accuracy simply by predicting every instance as healthy. However, such a model would be practically deficient, as its ability to distinguish between healthy and disease is compromised.

In conclusion, while the models achieve an accuracy rate of over 50%, this statistic is misleading. The low precision rate confirms that the models are prone to false positives.

The results of the 3DCNN model explained in part 3.2.1 is shown in table 4.4.

The accuracy for 3DCNN model shows slightly better results, though the temporal information is not considered here; it has higher accuracies compared to both the basic 3DCNN-LSTM model and the U-NET 3DCNN-LSTM model.

While this accuracy is higher, it shows an overfitting behavior. Comparing training accuracies with test accuracies shows that even though the results are

Metric	Value
Precision	0.445
Recall	0.667
F1 Score	0.533
Accuracy	0.667

Table 4.3: Performance metrics for the classification between healthy control and schizophrenia disorder in U-NET-LSTM model

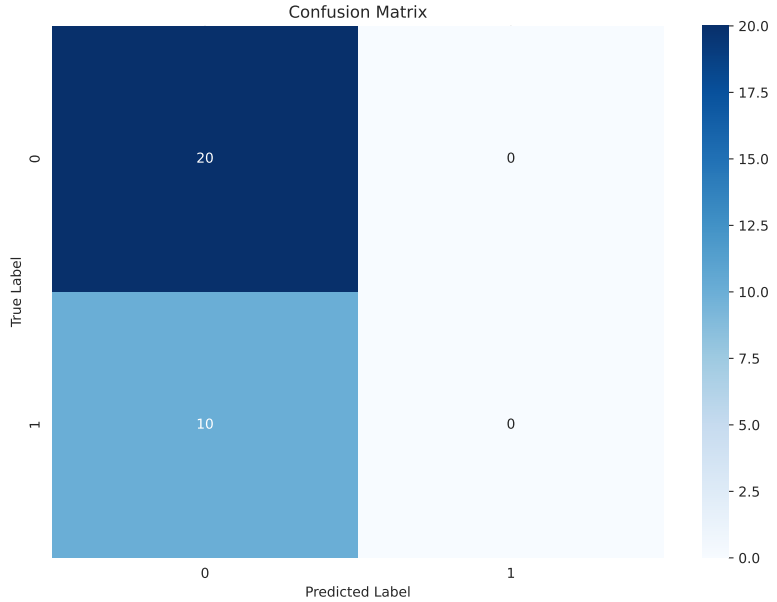


Figure 4.3: Confusion matrix for HC-SCH classification in U-NET-LSTM model

higher, the model does not generalize well to new data. To mitigate this problem, a dropout with a 0.5 rate was added to the last layer before the output, but it didn't affect the performance much.

In figure 4.4 and 4.5, you can see the training curves of this model.

Table 4.4: Training performance of different group comparisons with 3DCNN model

Groups	Training loss	Validation loss	Training accuracy	Validation accuracy
HC-SCH	0.0044	3.4582	0.9975	0.6291
HC-ADHD	0.0012	5.8201	0.9994	0.7642
HC-BD	0.0018	4.1745	0.9985	0.6383
SCH-ADHD	0.0030	2.9564	0.9990	0.7500
SCH-BD	0.0063	4.9458	0.9966	0.5037
BD-ADHD	0.0043	5.7112	0.9995	0.5847

(a) SCH vs. HC

(b) ADHD vs. HC

(c) BD vs. HC

Figure 4.4: Performance Analysis of Classification Model in Distinguishing Healthy Controls from Disorder Groups.

(a) SCH vs. ADHD

(b) SCH vs. BD

(c) ADHD vs. BD

Figure 4.5: Performance Analysis of Classification Model in Distinguishing Disorder Groups.

5

Discussion and Future Works

5.1 Discussion

In exploring the classification of neurodevelopmental disorders using fMRI data, it is essential to consider the intrinsic complexities of the data derived from Blood Oxygen Level Dependent (BOLD) signals. The BOLD signal, which underpins fMRI imaging, is not a plain measurement of neural activity but rather an indirect marker based on changes in blood flow and oxygenation levels. This dependence introduces a layer of complexity, as BOLD signals are susceptible to various physiological factors beyond neural activity, such as fluctuations in heart rate, respiration, and even subtle movements by the subject [33] and the link between the fMRI signal and the underlying neural activity is unclear [34].

These physiological factors can introduce noise into the fMRI data, potentially confounding the signals associated with the neurodevelopmental conditions under investigation. The complexity of disentangling these intertwined signals presents a significant challenge for any computational model attempting to classify disorders based on fMRI data. While advanced machine learning models, such as the 3D CNN+LSTM approach adopted in this study, hold promise for extracting meaningful patterns from complex datasets, the inherent variability and noise within BOLD signals may limit the model's ability to achieve high classification accuracy.

According to the analysis of the models and performance metrics, both the base model and the U-NET-LSTM model show similar performance and are not

capable of accurately classifying the data into subsequent groups. The 3DCNN model, which lacks temporal information but has more statistics, performs relatively better in some cases.

In both main models, accuracy improves to a certain point but does not improve further, indicating that the model falls into a local minimum and cannot escape from it. One reason for this is the limited amount of data available for this classification. It can be suggested that with more data and statistics, there is the possibility to overcome this issue. This is also demonstrated by the use of the 3DCNN model, which, even without considering temporal information, performed slightly better because the model had access to more data, as depicted in figure 5.1. The F1 score in this model is relatively better than that in the U-NET-LSTM model, as shown in table 5.2.

Table 5.1: Performance Metrics Across Different Groups in 3CNN model without temporal information

Groups	Precision	Recall	F1 Score	Accuracy
HC-SCH	0.6528	0.6816	0.5537	0.6330
HC-ADHD	0.4921	0.6330	0.6361	0.6816
HC-BD	0.6499	0.6700	0.6539	0.6700
SCH-ADHD	0.4840	0.4930	0.4860	0.4930
SCH-BD	0.4731	0.4737	0.4706	0.4737
BD-ADHD	0.2894	0.4472	0.3470	0.4472

Table 5.2: Performance Metrics Across Different Groups in U-NET-LSTM model

Groups	Precision	Recall	F1 Score	Accuracy
HC-SCH	0.6528	0.6816	0.5537	0.6330
HC-ADHD	0.5102	0.7142	0.5952	0.7142
HC-BD	0.4444	0.4444	0.5333	0.6666
SCH-ADHD	0.3086	0.5555	0.3968	0.5555
SCH-BD	0.2500	0.5000	0.3333	0.5000
BD-ADHD	0.3086	0.5555	0.3968	0.5555

Another important aspect to mention is that this data is inherently complex for classification since different groups are highly correlated with each other, as shown in [58] and depicted in figure 5.1.

Given this context, three primary challenges hinder the effective classification of this dataset using the proposed models. Firstly, the limited volume of

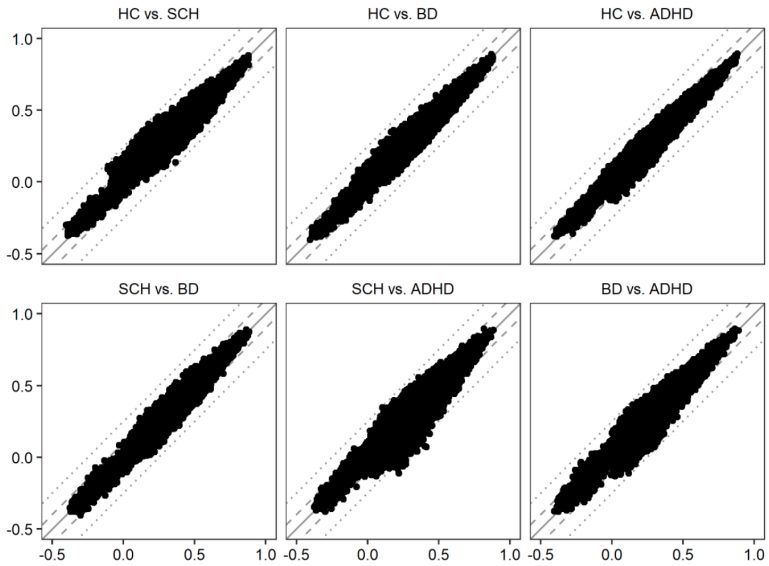


Figure 5.1: Comparison of mean Pearson coefficients between participant categories [58]

data and, by extension, the restricted statistical base available for training these models pose a significant obstacle. This scarcity of data undermines the models' ability to navigate and interpret the dataset's complexities effectively.

Secondly, the dataset may not provide sufficient discriminative information for classification tasks. This could stem from the data's nature or the manner in which it is processed and presented to the models.

Lastly, the potential inadequacy of the models themselves in addressing this specific classification challenge should be considered. This might be due to limitations in deep learning models when the data is limited or when the data is highly intertwined among different groups.

Together, these issues underscore the challenges of classifying highly correlated data groups and highlight the need for enhanced data volume, improved data quality, or more sophisticated modeling approaches to achieve better performance.

5.2 Future work

The analysis of brain disorder classification through neuroimaging data in this thesis unveils significant opportunities for enhancing our understanding and methodologies in this field. The challenge of limited dataset size stands out

among the critical insights garnered. Achieving higher model performance is contingent upon the expansion of these datasets. By developing collaborations with multiple research centers, there's potential to aggregate a richer, more varied set of fMRI data. Such an approach could reveal patterns previously obscured by the constraints of smaller datasets, thereby enriching our analytical depth.

Increasing the temporal dimension of fMRI data can also help achieve better results. The dynamic nature of brain activity, captured over time, holds untapped potential for improving classification accuracy. Delving into sophisticated temporal analysis techniques could significantly enhance our ability to discern and classify neurological conditions more precisely.

Current models, while effective to a degree, often struggle with the complexity inherent in highly correlated data groups found in neuroimaging. There's a burgeoning need to explore, develop, or adapt algorithms specifically designed for neuroimaging data. Such advancements could incorporate elements of unsupervised or semi-supervised learning, offering a more nuanced approach to navigating the intricate patterns of brain activity.

This study also highlights the complex challenges in researching brain disorders with fMRI technology. It underscores the critical need for ongoing research into more sophisticated data preprocessing techniques and model architectures capable of navigating the complexities of fMRI data. Future work may benefit from incorporating strategies specifically designed to mitigate the impact of physiological noise and enhance the signal-to-noise ratio, potentially improving the model's capacity to classify neurodevelopmental disorders accurately.

The ultimate goal of this research trajectory is clinical translation. The path forward involves not just technical advancements but also the rigorous validation of these neuroimaging-based models in clinical settings. Assessing their practicality, sensitivity, specificity, and predictive value in real-world environments is critical for understanding their implications for patient care and the broader medical community.

Finally, ethical and cost considerations must be considered for fMRI data collection. As we venture further into neuroimaging-based diagnostics, finding cost-effective and ethically sound approaches to implement these technologies in healthcare systems is critical. Balancing economic feasibility with ethical standards is essential for the widespread adoption and beneficial impact of these advancements.

In pursuing these interconnected paths, the field can progress towards fully harnessing the potential of neuroimaging and machine learning in diagnosing and understanding brain disorders. This concerted effort promises to pave the way for more personalized and effective patient care, marking significant strides in our quest to address the complexities of brain health.

6

Conclusion

Given the advancements in deep learning models and their increasing use in various fields, one might assume they would yield promising results in classifying conditions like healthy communication versus neurodevelopmental disorders using neuroimaging data. It's crucial to note that the results vary between using MRI datasets and fMRI datasets, with our focus here being on fMRI.

Several studies have demonstrated relatively good outcomes in this area. For example, research by Zheng et al.[62] reported accuracy of 71%, Liang et al.[31] achieved an 80% accuracy rate in distinguishing ADHD, and Mao et al. [35] also reported a 71% in classification accuracy using spatio-temporal models.

The source and collection site of the fMRI data also plays a significant role in the variability of results, with each site potentially showing slightly different outcomes.

Regarding the data utilized in this study, it is crucial to acknowledge the complexity of its nature. Furthermore, there exists similarity in data across different groups, which complicates classification. One potential solution to this issue is to augment the dataset with additional instances, although this approach is controversial due to the high cost associated with collecting fMRI data. However, examining the results presented in Table 5.1 provides hope that increasing the dataset size may yield marginally improved results, even without accounting for the critical temporal information in fMRI. Therefore, it can be argued that augmenting the fMRI dataset with temporal information may lead to more robust results.

References

[1] Sandra Ackerman. "Major Structures and Functions of the Brain". In: *Discovering the Brain*. National Academies Press (US), 1992. URL: https://www.ncbi.nlm.nih.gov/books/NBK234157/ (visited on 04/02/2024).

[2] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a Convolutional Neural Network". In: *2017 International Conference on Engineering and Technology (ICET)*. 2017 International Conference on Engineering and Technology (ICET). Aug. 2017, pp. 1–6. DOI: 10.1109/ICEngTechnol.2017.8308186. (Visited on 02/29/2024).

[3] P. A. Bandettini et al. "Time Course EPI of Human Brain Function during Task Activation". In: *Magnetic Resonance in Medicine* 25.2 (June 1992), pp. 390–397. ISSN: 0740-3194. DOI: 10.1002/mrm.1910250220. pmid: 1614324.

[4] R. A. Barkley and J. Biederman. "Toward a Broader Definition of the Age-of-Onset Criterion for Attention-Deficit Hyperactivity Disorder". In: *Journal of the American Academy of Child and Adolescent Psychiatry* 36.9 (Sept. 1997), pp. 1204–1210. ISSN: 0890-8567. DOI: 10.1097/00004583-199709000-00012. pmid: 9291721.

[5] Russell A. Barkley. "Major Life Activity and Health Outcomes Associated with Attention-Deficit/Hyperactivity Disorder". In: *The Journal of Clinical Psychiatry* 63 Suppl 12 (2002), pp. 10–15. ISSN: 0160-6689. pmid: 12562056.

[6] Roman Beliy et al. "From Voxels to Pixels and Back: Self-supervision in Natural-Image Reconstruction from fMRI". In: *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates, Inc., 2019. URL: %5Curl%7Bhttps://proceedings.neurips.cc/paper_files/paper/2019/hash/7d2be41b1bde6ff8fe45150c37488ebb-Abstract.html%7D (visited on 03/01/2024).

[7] Joseph Biederman et al. "Young Adult Outcome of Attention Deficit Hyperactivity Disorder: A Controlled 10-Year Follow-up Study". In: *Psychological Medicine* 36.2 (Feb. 2006), pp. 167–179. ISSN: 0033-2917. DOI: 10.1017/S0033291705006410. pmid: 16420713.

[8] *Bipolar Disorder - National Institute of Mental Health (NIMH)*. URL: https://www.nimh.nih.gov/health/publications/bipolar-disorder (visited on 02/26/2024).

[9] *Brain Basics: Know Your Brain | National Institute of Neurological Disorders and Stroke*. URL: https://www.ninds.nih.gov/health-information/public-education/brain-basics/brain-basics-know-your-brain (visited on 04/02/2024).

[10] *Brain Diseases: Definition & Types*. Cleveland Clinic. URL: https://my.clevelandclinic.org/health/diseases/22934-brain-diseases (visited on 04/02/2024).

[11] *Brain Disorders: Causes, Symptoms, and Diagnosis*. Healthline. July 16, 2012. URL: https://www.healthline.com/health/brain-disorders (visited on 04/03/2024).

[12] *Brain: How It Works, Function, Parts & Conditions*. Cleveland Clinic. URL: https://my.clevelandclinic.org/health/body/22638-brain (visited on 04/02/2024).

[13] Elisa Cainelli and Patrizia Bisiacchi. "Neurodevelopmental Disorders: Past, Present, and Future". In: *Children* 10.1 (Dec. 24, 2022), p. 31. ISSN: 2227-9067. DOI: 10.3390/children10010031. pmid: 36670582. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9856894/ (visited on 04/08/2024).

[14] Zhe Sage Chen et al. "Modern Views of Machine Learning for Precision Psychiatry". In: *Patterns* 3.11 (Nov. 11, 2022), p. 100602. ISSN: 2666-3899. DOI: 10.1016/j.patter.2022.100602. pmid: 36419447. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9676543/ (visited on 03/19/2024).

[15] *Classification of ECG Signals Based on 1D Convolution Neural Network | IEEE Conference Publication | IEEE Xplore*. URL: https://ieeexplore.ieee.org/document/8210784 (visited on 02/29/2024).

[16] *Clinical Trials for Neurodevelopmental Disorders: At a Therapeutic Frontier*. DOI: 10.1126/scitranslmed.aad9874. URL: https://www.science.org/doi/10.1126/scitranslmed.aad9874 (visited on 04/04/2024).

[17] *Data Processing Stream*. URL: https://legacy.openfmri.org/data-processing-stream/ (visited on 03/11/2024).

[18] Josepheen De Asis-Cruz et al. "FetalGAN: Automated Segmentation of Fetal Functional Brain MRI Using Deep Generative Adversarial Learning and Multi-Scale 3D U-Net". In: *Frontiers in Neuroscience* 16 (June 7, 2022). ISSN: 1662-453X. DOI: 10.3389/fnins.2022.887634. URL: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2022.887634/full (visited on 03/19/2024).

[19] Oguz Demirci et al. "A Review of Challenges in the Use of fMRI for Disease Classification / Characterization and A Projection Pursuit Application from A Multi-site fMRI Schizophrenia Study". In: *Brain Imaging and Behavior* 2.3 (Sept. 1, 2008), pp. 207–226. ISSN: 1931-7565. DOI: 10.1007/s11682-008-9028-1. URL: https://doi.org/10.1007/s11682-008-9028-1 (visited on 02/25/2024).

[20] Stephen A. Engel et al. "fMRI of Human Visual Cortex". In: *Nature* 369.6481 (June 1994), pp. 525–525. ISSN: 1476-4687. DOI: 10.1038/369525a0. URL: https://www.nature.com/articles/369525a0 (visited on 03/18/2024).

[21] Stéphane Epelbaum and Federica Cacciamani. "Clinical Assessment of Brain Disorders". In: *Machine Learning for Brain Disorders*. Ed. by Olivier Colliot. New York, NY: Humana, 2023. ISBN: 978-1-07-163194-2 978-1-07-163195-9. pmid: 37988502. URL: http://www.ncbi.nlm.nih.gov/books/NBK597487/ (visited on 04/04/2024).

[22] Gary H. Glover. "Overview of Functional Magnetic Resonance Imaging". In: *Neurosurgery clinics of North America* 22.2 (Apr. 2011), pp. 133–139. ISSN: 1042-3680. DOI: 10.1016/j.nec.2010.11.001. pmid: 21435566. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3073717/ (visited on 02/23/2024).

[23] Laurence L. Greenhill et al. "Practice Parameter for the Use of Stimulant Medications in the Treatment of Children, Adolescents, and Adults". In: *Journal of the American Academy of Child and Adolescent Psychiatry* 41 (2 Suppl

Feb. 2002), 26S–49S. ISSN: 0890-8567. DOI: 10.1097/00004583-200202001-00003. pmid: 11833633.

[24] Dominik Grotegerd et al. "Discriminating Unipolar and Bipolar Depression by Means of fMRI and Pattern Classification: A Pilot Study". In: *European Archives of Psychiatry and Clinical Neuroscience* 263.2 (Mar. 1, 2013), pp. 119–131. ISSN: 1433-8491. DOI: 10.1007/s00406-012-0329-4. URL: https://doi.org/10.1007/s00406-012-0329-4 (visited on 02/26/2024).

[25] Sarah Itani, Fabian Lecron, and Philippe Fortemps. "A Multi-Level Classification Framework for Multi-Site Medical Data: Application to the ADHD-200 Collection". In: *Expert Systems with Applications* 91 (Jan. 1, 2018), pp. 36–45. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2017.08.044. URL: https://www.sciencedirect.com/science/article/pii/S0957417417305924 (visited on 03/05/2024).

[26] Shuiwang Ji et al. "3D Convolutional Neural Networks for Human Action Recognition". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.1 (Jan. 2013), pp. 221–231. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2012.59. URL: https://ieeexplore.ieee.org/document/6165309 (visited on 02/29/2024).

[27] Konstantinos Kamnitsas et al. "Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation". In: *Medical Image Analysis* 36 (Feb. 2017), pp. 61–78. ISSN: 1361-8423. DOI: 10.1016/j.media.2016.10.004. pmid: 27865153.

[28] Y. LeCun et al. "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4 (Dec. 1, 1989), pp. 541–551. ISSN: 0899-7667. DOI: 10.1162/neco.1989.1.4.541. URL: https://doi.org/10.1162/neco.1989.1.4.541 (visited on 02/27/2024).

[29] Qing Li et al. "Medical Image Classification with Convolutional Neural Network". In: *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*. 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV). Dec. 2014, pp. 844–848. DOI: 10.1109/ICARCV.2014.7064414. URL: https://ieeexplore.ieee.org/document/7064414 (visited on 02/29/2024).

[30] Zhuangzhuang Li et al. "Deep Learning Based Automatic Diagnosis of First-Episode Psychosis, Bipolar Disorder and Healthy Controls". In: *Computerized Medical Imaging and Graphics* 89 (Apr. 1, 2021), p. 101882. ISSN: 0895-6111. DOI: 10.1016/j.compmedimag.2021.101882. URL: https://www.sciencedirect.com/science/article/pii/S0895611121000306 (visited on 03/05/2024).

[31] Sheng-Fu Liang et al. "Differentiation between Resting-State fMRI Data from ADHD and Normal Subjects: Based on Functional Connectivity and Machine Learning". In: *2012 International Conference on Fuzzy Theory and Its Applications (iFUZZY2012)*. 2012 International Conference on Fuzzy Theory and Its Applications (iFUZZY2012). Nov. 2012, pp. 294–298. DOI: 10.1109/iFUZZY.2012.6409719. URL: https://ieeexplore.ieee.org/abstract/document/6409719 (visited on 03/05/2024).

[32] N. K. Logothetis et al. "Neurophysiological Investigation of the Basis of the fMRI Signal". In: *Nature* 412.6843 (July 12, 2001), pp. 150–157. ISSN: 0028-0836. DOI: 10.1038/35084005. pmid: 11449264.

[33] Nikos K. Logothetis and Josef Pfeuffer. "On the Nature of the BOLD fMRI Contrast Mechanism". In: *Magnetic Resonance Imaging* 22.10 (Dec. 2004), pp. 1517–1531. ISSN: 0730-725X. DOI: 10.1016/j.mri.2004.10.018. pmid: 15707801.

[34] Nikos K. Logothetis et al. "Neurophysiological Investigation of the Basis of the fMRI Signal". In: *Nature* 412.6843 (July 2001), pp. 150–157. ISSN: 1476-4687. DOI: 10.1038/35084005. URL: https://www.nature.com/articles/35084005 (visited on 04/09/2024).

[35] Zhenyu Mao et al. "Spatio-Temporal Deep Learning Method for ADHD fMRI Classification". In: *Information Sciences* 499 (Oct. 1, 2019), pp. 1–11. ISSN: 0020-0255. DOI: 10.1016/j.ins.2019.05.043. URL: https://www.sciencedirect.com/science/article/pii/S0020025519304475 (visited on 03/05/2024).

[36] Deborah J. Morris-Rosendahl and Marc-Antoine Crocq. "Neurodevelopmental Disorders—the History and Future of a Diagnostic Concept". In: *Dialogues in Clinical Neuroscience* 22.1 (Mar. 2020), pp. 65–72. ISSN: 1294-8322. DOI: 10.31887/DCNS.2020.22.1/macrocq. pmid: 32699506. URL:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7365295/ (visited on 04/07/2024).

[37] *Neurodegenerative Diseases*. National Institute of Environmental Health Sciences. URL: https://www.niehs.nih.gov/research/supported/health/neurodegenerative (visited on 02/20/2024).

[38] *Neurodevelopmental Disorders*. East and North Hertfordshire NHS Trust. URL: %5Curl%7Bhttps://www.enherts-tr.nhs.uk/services/community-paediatrics/neurodevelopmental-disorders/%7D (visited on 04/08/2024).

[39] Keiron O'Shea and Ryan Nash. *An Introduction to Convolutional Neural Networks*. Dec. 2, 2015. DOI: 10.48550/arXiv.1511.08458. arXiv: 1511.08458 [cs]. URL: http://arxiv.org/abs/1511.08458 (visited on 02/28/2024). preprint.

[40] J. Olesen and M. Leonardi. "The Burden of Brain Diseases in Europe". In: *European Journal of Neurology* 10.5 (Sept. 2003), pp. 471–477. ISSN: 1351-5101. DOI: 10.1046/j.1468-1331.2003.00682.x. pmid: 12940825.

[41] Russ Poldrack. *Poldrack/Openfmri*. Oct. 5, 2023. URL: https://github.com/poldrack/openfmri (visited on 03/11/2024).

[42] Russell A. Poldrack. "The Future of fMRI in Cognitive Neuroscience". In: *NeuroImage* 62.2 (Aug. 15, 2012), pp. 1216–1220. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2011.08.007. pmid: 21856431. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4131441/ (visited on 03/18/2024).

[43] Russell A. Poldrack. "The Role of fMRI in Cognitive Neuroscience: Where Do We Stand?" In: *Current Opinion in Neurobiology* 18.2 (Apr. 2008), pp. 223–227. ISSN: 0959-4388. DOI: 10.1016/j.conb.2008.07.006. pmid: 18678252.

[44] Cathy J. Price. "The Anatomy of Language: A Review of 100 fMRI Studies Published in 2009". In: *Annals of the New York Academy of Sciences* 1191 (Mar. 2010), pp. 62–88. ISSN: 1749-6632. DOI: 10.1111/j.1749-6632.2010.05444.x. pmid: 20392276.

[45] Mirjam Quaak et al. "Deep Learning Applications for the Classification of Psychiatric Disorders Using Neuroimaging Data: Systematic Review and Meta-Analysis". In: *NeuroImage : Clinical* 30 (Feb. 10, 2021), p. 102584. ISSN: 2213-1582. DOI: 10.1016/j.nicl.2021.102584. pmid: 33677240. URL:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8209481/ (visited on 03/19/2024).

[46] Farheen Ramzan et al. "A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer's Disease Stages Using Resting-State fMRI and Residual Neural Networks". In: *Journal of Medical Systems* 44.2 (Dec. 18, 2019), p. 37. ISSN: 1573-689X. DOI: 10.1007/s10916-019-1475-2. URL: https://doi.org/10.1007/s10916-019-1475-2 (visited on 03/03/2024).

[47] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4. DOI: 10.1007/978-3-319-24574-4_28.

[48] Kelly A. Sagar et al. "Altered Affective Processing in Bipolar Disorder: An fMRI Study". In: *Journal of Affective Disorders* 150.3 (Sept. 25, 2013), pp. 1192–1196. ISSN: 0165-0327. DOI: 10.1016/j.jad.2013.05.019. URL: %5Curl%7Bhttps://www.sciencedirect.com/science/article/pii/S0165032713003777%7D (visited on 02/26/2024).

[49] Saman Sarraf and Ghassem Tofighi. "Deep Learning-Based Pipeline to Recognize Alzheimer's Disease Using fMRI Data". In: *2016 Future Technologies Conference (FTC)*. 2016 Future Technologies Conference (FTC). San Francisco, CA, USA: IEEE, Dec. 2016, pp. 816–820. ISBN: 978-1-5090-4171-8. DOI: 10.1109/FTC.2016.7821697. URL: http://ieeexplore.ieee.org/document/7821697/ (visited on 03/01/2024).

[50] J. B. Savitz, S. L. Rauch, and W. C. Drevets. "Clinical Application of Brain Imaging for the Diagnosis of Mood Disorders: The Current State of Play". In: *Molecular Psychiatry* 18.5 (May 2013), pp. 528–539. ISSN: 1476-5578. DOI: 10.1038/mp.2013.25. URL: https://www.nature.com/articles/mp201325 (visited on 04/04/2024).

[51] C. Scarpazza et al. "Translating Research Findings into Clinical Practice: A Systematic and Critical Review of Neuroimaging-Based Clinical Tools for Brain Disorders". In: *Translational Psychiatry* 10.1 (Apr. 20, 2020), pp. 1–16. ISSN: 2158-3188. DOI: 10.1038/s41398-020-0798-6. URL: https://www.nature.com/articles/s41398-020-0798-6 (visited on 04/04/2024).

[52] Andrea Schmitt et al. "Schizophrenia as a Disorder of Disconnectivity". In: *European Archives of Psychiatry and Clinical Neuroscience* 261.2 (Aug. 25, 2011), p. 150. ISSN: 1433-8491. DOI: 10.1007/s00406-011-0242-2. URL: https://doi.org/10.1007/s00406-011-0242-2 (visited on 02/23/2024).

[53] Luca Steardo et al. "Application of Support Vector Machine on fMRI Data as Biomarkers in Schizophrenia Diagnosis: A Systematic Review". In: *Frontiers in Psychiatry* 11 (2020). ISSN: 1664-0640. URL: https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2020.00588 (visited on 03/04/2024).

[54] Jing Teng et al. "Brain Disease Research Based on Functional Magnetic Resonance Imaging Data and Machine Learning: A Review". In: *Frontiers in Neuroscience* 17 (Aug. 17, 2023), p. 1227491. ISSN: 1662-4548. DOI: 10.3389/fnins.2023.1227491. pmid: 37662098. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10469689/ (visited on 04/03/2024).

[55] Du Tran et al. "Learning Spatiotemporal Features With 3D Convolutional Networks". In: Proceedings of the IEEE International Conference on Computer Vision. 2015, pp. 4489–4497. URL: https://openaccess.thecvf.com/content_iccv_2015/html/Tran_Learning_Spatiotemporal_Features_ICCV_2015_paper.html (visited on 03/19/2024).

[56] *UCLA Consortium for Neuropsychiatric Phenomics LA5c Study*. URL: https://legacy.openfmri.org/dataset/ds000030/ (visited on 03/11/2024).

[57] © Stanford University, Stanford, and California 94305. *UCLA Consortium for Neuropsychiatric Phenomics LA5c Study*. Research from Stanford University - Spotlight at Stanford. URL: https://exhibits.stanford.edu/data/catalog/mg599hw5271 (visited on 03/11/2024).

[58] Wout Van Echelpoel. *Theme Modeling for the Stratification of Neurological Patiënten. /*. 2023.

[59] Yunchao Wei et al. "HCP: A Flexible CNN Framework for Multi-Label Image Classification". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38.9 (Sept. 2016), pp. 1901–1907. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2015.2491929. URL: https://ieeexplore.ieee.org/document/7305792 (visited on 02/29/2024).

[60] Weizheng Yan et al. "Mapping Relationships among Schizophrenia, Bipolar and Schizoaffective Disorders: A Deep Classification and Clustering Framework Using fMRI Time Series". In: *Schizophrenia Research*. Computational Approaches to Understanding Psychosis 245 (July 1, 2022), pp. 141–150. ISSN: 0920-9964. DOI: 10.1016/j.schres.2021.02.007. URL: https://www.sciencedirect.com/science/article/pii/S0920996421000773 (visited on 03/05/2024).

[61] Jianfeng Zhao, Xia Mao, and Lijiang Chen. "Speech Emotion Recognition Using Deep 1D & 2D CNN LSTM Networks". In: *Biomedical Signal Processing and Control* 47 (Jan. 1, 2019), pp. 312–323. ISSN: 1746-8094. DOI: 10.1016/j.bspc.2018.08.035. URL: https://www.sciencedirect.com/science/article/pii/S1746809418302337 (visited on 02/29/2024).

[62] JinChi Zheng et al. "Diagnosis of Schizophrenia Based on Deep Learning Using fMRI". In: *Computational and Mathematical Methods in Medicine* 2021 (Nov. 9, 2021), e8437260. ISSN: 1748-670X. DOI: 10.1155/2021/8437260. URL: https://www.hindawi.com/journals/cmmm/2021/8437260/ (visited on 03/03/2024).

[63] Yuan Zhou et al. "Functional Disintegration in Paranoid Schizophrenia Using Resting-State fMRI". In: *Schizophrenia Research* 97.1 (Dec. 1, 2007), pp. 194–205. ISSN: 0920-9964. DOI: 10.1016/j.schres.2007.05.029. URL: https://www.sciencedirect.com/science/article/pii/S0920996407002393 (visited on 03/04/2024).