ROYAUME DU MAROC
Université Mohammed V - RABAT -
جامعة محمد الخامس – الرباط –
Faculté des sciences
كلية العلوم

CENTRE D'ETUDES DOCTORALES - SCIENCES ET TECHNOLOGIES

N° d'ordre 3873

THESE

En vue de l'obtention du : ***DOCTORAT***

Structure de Recherche : Laboratoire de Recherche en Informatique et Télécommunications
Discipline : Sciences de l'ingénieur
Spécialité : Informatique et Télécommunications

Présentée et soutenue le 08/12/2023 par :

Fatima Zahra BENABDALLAH

Machine and Deep Learning based Approaches for Autism Spectrum Disorder Identification

Membres du jury

Mohammed Rziza	PES, Faculté des Sciences, Université Mohammed V de Rabat	Président
Aicha Majda	PES, Faculté des Sciences Juridiques, Economiques et Sociales, Université Moulay Ismail, Meknès	Rapporteur / Examinateur
Abdellatif Kobbane	PES, École Nationale Supérieure d'Informatique et d'Analyse des Systèmes, Université Mohammed V de Rabat	Rapporteur / Examinateur
Yahya Benkaouz	PH, Faculté des Sciences, Université Mohammed V de Rabat	Rapporteur / Examinateur
Dounia Lotfi	PH, Faculté des Sciences, Université Mohammed V de Rabat	Examinateur
Brahim Alibouch	PA, Faculté des sciences, Université Ibn zohr, Agadir	Invité
Ahmed Drissi el Maliani	PH, Faculté des Sciences, Université Mohammed V de Rabat	Co-directeur de Thèse
Mohammed El Hassouni	PES, Faculté des Lettres et Sciences Humaines, Université Mohammed V de Rabat	Directeur de Thèse

Année Universitaire : 2023-2024

✉ Faculté des Sciences, avenue Ibn Battouta, BP. 1014 RP, Rabat –Maroc
☎ 00212 (05) 37 77 18 76 🖷 00212(05) 37 77 42 61; http://www. fsr.um5.ac.ma

To my mother

To my family and friends

As well as anyone else who has been a part of my academic journey

Please accept my sincere appreciation.
Your guidance, support and encouragement have been instrumental in the successful completion of this thesis.

Acknowledgements

First of all, I would like to thank the director of the Laboratory of Research in Information technology and Telecommunication, **Prof. Dr. Khalid El Minaoui** for his support and cooperation throughout my doctoral studies.

My heartfelt appreciation is to my supervisors, **Prof. Dr. Mohammed El Hassouni** and **Prof. Dr. Ahmed Drissi El Maliani**, for their exceptional support and invaluable guidance throughout my doctoral journey. Their unwavering commitment to my success and their wealth of knowledge have been instrumental in shaping my academic and professional growth. Working alongside such esteemed mentors has been an enriching experience, and I am truly fortunate to have had the opportunity to learn from them. Their mentorship has not only enhanced my research capabilities but has also inspired me to strive for excellence in all aspects of my work.

I would also like to express my gratitude to **Prof. Dr. Mohammed Rziza** for accepting to preside the jury committee and agreeing to review my thesis manuscript. Thank you for being an integral part of me being able to make it to the end.

Additionally, I want to thank all the rapporteurs/examiners, **Profs. Dr. Aicha Majda, Dr. Abdellatif Kobbane, and Dr. Yahya Benkaouz**. Thank you for your time and effort, reading and examining my thesis and for providing valuable feedback and thought-provoking questions. Your insights have greatly enriched the quality of my work.

Special thanks go to my examiner **Prof. Dr. Dounia Lotfi** for her immense help on several aspects of my work, for her support and kindness and for generously sharing her thoughts and experience.

I would also like to thank, **Prof. Dr. Brahim Alibouch**, for accepting to be part of the jury committee, for his time and his insightful questions and suggestions.

Abstract

The task of accurately identifying autism spectrum disorder presents a significant challenge within the realm of neurodevelopmental disorders. This study aimed to delve into autism theories, specifically, the notions of overconnectivity and underconnectivity within the autistic brain, employing a range of computational methods and deep learning techniques. The research began by investigating different preprocessing strategies, emphasizing their impact on the accurate classification of the autistic subjects. Subsequently, novel approaches, including a mathematical elimination approach, an enhancement approach as well as an RGB image mimicking technique based on autism tailored methods were introduced to investigate and validate the existing theories of overconnectivity and underconnectivity with an overall goal of improving autism identification. The results demonstrated the complexity and context-dependent nature of neural connectivity patterns in ASD. Interestingly, the study led to achieving a high autism identification accuracy of 96% and revealed that no single computation method or combination proved to be universally effective, as the results consistently changed with different approaches. This led to the conclusion that even methods initially considered incompatible could yield remarkable results when thoroughly investigated and tested in diverse contexts. These findings highlight the importance of exploring multiple strategies and theories in future endeavors towards early autism diagnosis. By incorporating a range of tools and methodologies, and continually refining the detection system, the goal of accurate and early autism identification can be significantly advanced.

Keywords: Autism spectrum disorder; fMRI-rs; Connectivity; Machine Learning; Deep learning; Mathematical elimination; enhancement; RGB image mimicking.

Résumé

Dans le domaine des troubles neurodéveloppementaux, identifier avec précision les troubles du spectre autistique présente un grand défi. Cette étude vise à explorer les théories de l'autisme, en particulier les notions de surconnectivité et de sous-connectivité, par l'emploi d'une gamme de méthodes informatiques et de techniques d'apprentissage en profondeur. On commence par étudier différentes stratégies de prétraitement, en mettant l'accent sur leur impact sur la classification précise des sujets autistes. Par la suite, de nouvelles approches, notamment une approche d'élimination mathématique, une approche de renforcement ainsi qu'une technique d'imitation d'images RVB basée sur des méthodes adaptées à l'autisme, ont été introduites pour étudier et valider les théories existantes de surconnectivité et de sous-connectivité dans le but global d'améliorer l'identification de l'autisme. Les résultats ont démontré la complexité et la nature contextuelle des schémas de connectivité neuronale dans les TSA. Il est intéressant de noter que l'étude a permis d'atteindre une précision d'identification de l'autisme de 96% et a révélé qu'aucune méthode ou combinaison de calcul ne s'est avérée universellement efficace, les résultats changeant constamment avec les différentes approches. On peut donc en conclure que même des méthodes initialement considérées comme incompatibles peuvent donner des résultats remarquables lorsqu'elles sont étudiées et testées en profondeur dans divers contextes. Ces résultats soulignent l'importance d'explorer de multiples stratégies et théories dans les efforts futurs en vue d'atteindre un diagnostic précoce de l'autisme. En intégrant une gamme d'outils et de méthodologies et en affinant continuellement le système de détection, l'objectif d'une identification précise et précoce de l'autisme peut être considérablement avancé.

Mots-clefs: Troubles du spectre autistique; IRMf-rs; Connectivité; Apprentissage automatique; Apprentissage en profondeur; Elimination mathématique; Renforcement; Imitation d'images RVB.

Résumé étendu

L'autisme, ou le trouble du spectre autistique est un trouble neurodéveloppemental caractérisé par un large éventail de symptômes, de degrés de sévérité et de particularités individuelles. Son diagnostic est un diagnostic comportemental reposant sur un processus complexe impliquant plusieurs évaluations. Le diagnostic rapide de ce trouble permet un accès rapide à des interventions et des soutiens adaptés, améliorant ainsi les perspectives à long terme pour les personnes affectées. Mais pour atteindre un diagnostic précoce, il faut trouver une nouvelle méthode alternative de diagnostic, car le diagnostic comportemental même quand il est mené tôt il n'est jugé fiable qu'après l'âge de 2 ans.

Par conséquent, dans cette thèse, nous nous sommes tournés vers l'étude des images cérébrales pour étudier la possibilité de créer un nouveau système de diagnostic précoce. Les images cérébrales peuvent être acquises par plusieurs techniques. Pour étudier l'autisme la technique que nous avons jugé la plus adaptée est l'IRM fonctionnel.

En intégrant les théories de la connectivité dans le cerveau autiste nous avons alors commencé notre quête de l'identification de l'autisme.

Dans le premier chapitre, nous introduisons le contexte et la problématique ainsi que les objectives poursuivis.

Le deuxième chapitre offre un aperçu général sur l'autisme, son historique, sa prévalence, diagnostic et traitement. Il cite également les différentes techniques d'acquisition d'images cérébrales et explique le choix de l'utilisation des images d'IRM fonctionnel. Ce chapitre également comprend la présentation de la base de données utilisée ainsi que les importantes méthodes utilisées dans les différents tests et analyses menés.

Le troisième chapitre représente la première contribution de ce travail, qui est une approche d'élimination basée sur l'apprentissage automatique. Ici, nous avons intégré les théories de l'autisme dans notre investigation pour mettre en évidence leur importance. Nous avons également vérifié l'impact de plusieurs éléments, comme le prétraitement et les caractéristiques des personnes affectées par l'autisme, sur l'identification de ce dernier.

Le quatrième chapitre qui est la deuxième contribution de ce travail, intègre l'apprentissage en profondeur dans la quête de l'identification de l'autisme. Il explore comment l'apprentissage en profondeur, en particulier les réseaux de neurones convolutifs, peuvent améliorer la détection de l'autisme.

Dans le cinquième chapitre, nous allons plus loin dans l'utilisation de l'apprentissage en profondeur en introduisant une méthode de renforcement d'information novatrice. Cette méthode tire profit des théories de l'autisme vérifiées dans la première contribution. Le meilleur résultat de précision atteint avec cette approche est de 96%. Ce résultat démontre le potentiel remarquable des techniques avancées de l'apprentissage en profondeur, et de l'intégration des théories de l'autisme dans l'investigation de ce dernier.

Finalement, dans le dernier chapitre, nous récapitulons les principales conclusions de notre thèse. Nous soulignons comment nos méthodes ont permis d'améliorer la détection de l'autisme, en mettant en évidence l'importance de l'intégration des théories de connectivité du cerveau autiste dans le processus de détection.

List of Figures

List of Tables

Abbreviations

ASD	**A**utism **S**pectrum **D**isorder
TDC	**T**ypically **D**eveloping **C**ontrols
ADHD	**A**ttention **D**eficit **H**yperactivity **D**isorder
BOLD	**B**lood **O**xygen **L**evel **D**ependent
NICHCY	**N**ational **I**ssues **C**enter for **H**andicapped **C**hildren and **Y**outh
DSM	**D**iagnostic and **S**tatistical **M**annual
PDDNOS	**P**ervasive **D**evelopmental **D**isorder **N**ot **O**therwise **S**pecified
PDD	**P**ervasive **D**evelopmental **D**isorders
ICD	**I**nternational **C**lassification of **D**iseases
ADI-R	**A**utism **D**iagnostic **I**nterview - **R**evised
ADOS	**A**utism **D**iagnostic **O**bservation **S**chedule
DISCO	**D**iagnostic **I**nterview for **S**ocial and **CO**mmunication Disorders
3Di	**D**evelopmental, **D**imensional, and **D**iagnostic **I**nterview
CARS-2	**C**hildhood **A**utism **R**ating **S**cale - Second Edition
CT	**C**omputed **T**omography
CAT	**C**omputerized **A**xial **T**omography
MRI	**M**agnetic **R**esonance **I**maging
RF	**R**adio **F**requency
fMRI	**F**unctional **M**agnetic **R**esonance **I**maging
SPECT	**S**ingle-**P**hoton **E**mission **C**omputed **T**omography
PET	**P**ositron **E**mission **T**omography
MEG	**M**agneto**e**ncephalo**g**raphy
EPI	**E**cho **P**lanar **I**maging
ATP	**A**denosine **T**riphosphate
ABIDE	**A**utism **B**rain **I**maging **D**ata **E**xchange
RS-fMRI	**R**esting **S**tate **f**unctional **M**agnetic **R**esonance **I**maging
CCS	**C**onnectome **C**omputation **S**ystem
CPAC	**C**onfigurable **P**ipeline for the **A**nalysis of **C**onnectomes
DPARSF	**D**ata **P**rocessing **A**ssistant for **R**esting-**S**tate **F**MRI
ROI	**R**egions **o**f **I**nterest
AAL	**A**utomated **A**natomical **L**abeling
EZ	**E**ickhoff-**Z**illes
HO	**H**arvard-**O**xford

TT	**T**alairach-**T**ournoux
MST	**M**inimum **S**panning **T**ree
MaxST	**Max**imum **S**panning **T**ree
HC	**H**ierarchical **C**lustering
TP	**T**rue **P**ositive
FP	**F**alse **P**ositive
TN	**T**rue **N**egative
FN	**F**alse **N**egative
SVM	**S**upport **V**ector **M**achine
IM	**I**nitial **M**atrices
HC-LM	**H**ierarchical **C**lustering- **L**ocal connectivity **M**atrices
HCLM-MaxST	**H**ierarchical **C**lustering- **L**ocal connectivity **M**atrices minus **MaxST**
M-MST	**M**atrix minus the **M**inimum **S**panning **T**ree
HC-LRM	**H**ierarchical **C**lustering- **L**ong **R**ange connectivity **M**atrices
HCM-MST	**H**ierarchical **C**lustering **Minus MST**
ANN	**A**rtificial **N**eural **N**etworks
CNN	**C**onvolutional **N**eural **N**etworks
RNN	**R**ecurrent **N**eural **N**etworks
CM	**C**onnectivity **M**atrices
VGG	**V**isual **G**eometry **G**roup
ReLU	**R**ectified **L**inear **U**nit
RN	**R**esidual **N**etwork

Contents

Chapter 1

Introduction and background

1.1 Context and motivation

In the last decades, the humanity has been confronted with the emergence of growing numbers of psychological disorders that challenged all knowledge and required expert attention all around the globe.

Autism spectrum disorder (ASD) is part of these pandemic psychological disorders. ASD is a neurodevelopmental disorder that exhibits a range of genetic, pathophysiological, and environmental conditions. Its symptoms include qualitative impairments in social interaction as problems in using nonverbal behaviors, facial expression and body postures, failure to develop peer relationships, lack of spontaneous seeking of others and lack of social/emotional reciprocity. They also include the manifestation of qualitative impairments in communication as a delay or lack of the development of spoken language with no attempt to compensate through alternative modes of communication, inability to initiate or sustain a conversation with others, stereotyped and repetitive use of language or idiosyncratic language. Autistic children also manifest restricted, repetitive, and stereotyped patterns of behavior, interests, and activities as giving abnormal intensity or focus to a preoccupation with one or more stereotyped and restricted patterns, inflexibility when adhering to a certain routine, stereotyped and repetitive motor mannerisms that may include one part or the whole body and persistent preoccupation with parts of objects [1].

The variation in the symptoms, their severity and combinations lead to a spectrum of autism manifestations. These latter lead to difficulty to diagnose and digging out the origin of this disorder. It also renders finding the exact deficit location in the brain challenging, since every symptom can be related to a different brain zone.

During the early stages of childhood, the brain undergoes a crucial process of neural circuit formation, where non-functional and unnecessary neurons are selectively removed to enhance the efficiency and power of the functioning ones. This natural pruning process results in a decrease in the total number of neurons, while the remaining neurons establish intricate connectivity as each day progresses. However, in children with ASD, this vital pruning process is impaired, leading to the retention of a larger number of neurons compared to typically developing brains.

According to [2], brain development involves two significant processes that entail substantial loss of neural elements. The first process is naturally occurring cell death, where more than 50% of neurons

are eliminated. The second process, known as synaptic exuberance and pruning, involves the generation of an abundance of connections followed by the systematic elimination of up to 50% of those connections. In contrast, [3] highlights that males with autism exhibit an abnormal excess of prefrontal neurons, with a mean of approximately 67% more neurons than observed in the control group. This excess of neurons in the prefrontal region adds complexity to the understanding of brain development in individuals with autism and underscores the intricate nature of the disorder's neurobiological basis.
Several studies also investigated early brain development in infants at high risk for ASD or those who later receive an ASD diagnosis. These studies have observed differences in brain growth trajectories and structural characteristics. One notable finding is that children who later develop ASD tend to show accelerated brain growth during the first year of life. This accelerated growth is often followed by a plateau or deceleration in brain growth in the second year. These patterns have been observed in various brain regions, including the frontal cortex, temporal cortex, and amygdala [4], [5], [6].
During the adolescent period, individuals with autism spectrum disorder may face an increased risk of developing seizures and experiencing behavior problems and psychiatric symptoms. Several studies have highlighted these challenges during this developmental stage [7], [8], [9].
Further more, children and adolescents with ASD may also develop maladaptive behaviors including irritability, hyperactivity, aggressivity, depression, and anxiety, among others [10]. However, It is important to note that these maladaptive behaviors can vary widely among individuals with ASD, and their underlying causes and manifestations may differ.
In adulthood, in some cases, the core symptoms of autism spectrum disorder may become milder compared to early development. This phenomenon is often referred to as "improvement" or "amelioration" of symptoms. While the severity of symptoms can vary widely among individuals with ASD, research suggests that there may be a reduction in certain core symptoms as individuals transition into adulthood. Several studies have explored this topic and reported findings indicating a decrease in the severity of core ASD symptoms in adulthood [11], [12]. These age behavioral improvements, may be a result of the maturation and the stabilization of the disease processes. This may also be because the brain returns to a normal size and the growth to a normal pace in comparison to the overgrowth the brain goes through in childhood.
In general, autism characteristics can manifest before the age of 3 years old. Many abnormalities in the social interaction, language as used in social communication, or symbolic or imaginative play can be remarked. However, due to some factors as lack of parental perception and unavailability of diagnostic centers, the average age of autism detection has reached 5 years old [13].
Moreover, the DSM-5, that specify the criteria to autism diagnosis, has been criticized for its potential limitations in capturing the full heterogeneity of symptoms and co-occurring conditions that can accompany ASD. The DSM-5 criteria primarily focus on core symptoms related to social communication deficits, restricted and repetitive behaviors, and sensory sensitivities. Some individuals with ASD may present with additional symptoms and co-occurring conditions that are not adequately addressed by the DSM-5 criteria. These may include intellectual disabilities, language impairments, attention-deficit/hyperactivity disorder (ADHD), anxiety, depression, and other psychiatric or medical conditions. Furthermore, the most cognitively able cases and those with less classic autism presentations can also present a challange, diagnosis wise. All these contribute in some cases to excluding affected individuals from the correctly diagnosed subjects [14], [15].

When we exclude the aforementioned diagnosis problems, early behavioral characteristics of ASD can be remarked at the first year of life with a rate of diagnosis stability that reaches 90%[16] and can be confirmed around the age of 2 through a bunch of behavioral tests. But this process is very costly and cannot be provided to all children.

According to experts, early and accurate detection of autism spectrum disorder (ASD) through the implementation of tailored intervention programs can significantly enhance the life of the affected child as they progress into adulthood [17]. Early identification enables a comprehensive understanding of the child's potential and the severity of the disorder, facilitating the provision of appropriate support and targeted interventions. Such early interventions and services are vital in helping children from a very young age to develop essential skills, including speech, mobility, and social interaction. By offering these crucial therapies and support, the child's quality of life improves, increasing their potential to smoothly integrate into society as they grow.

Given this objective, ASD research strives not only to find a cure but also to expedite the detection process using non-harmful methods, especially considering that the affected individuals are children. This necessitates careful consideration when selecting the means and methods of diagnosis.

Brain imaging techniques offer promise as a potential means of autism diagnosis. The reason is that the age is not a key factor. Hence an early diagnosis may be possible. A wide array of anatomical and functional neuroimaging techniques is at our disposal, including X-ray, MRI, MEG, fMRI, PET, and SPECT [18], [19], [20]. Anatomical techniques enable the visualization of the brain's structure and its various components, providing valuable insights into its morphology. In contrast, functional neuroimaging serves a distinct purpose, aiming to unravel the cognitive organization of the brain and its dynamic processes, unraveling the brain communication processes.

In the first years, children can vary in their rate of development, therefore, appearances can be deceiving [21]. Using anatomical imaging to detect autism may then be untrustworthy. The alternative is studying the functionality of the brain. In other words, studying the brain network communication. This may give more insight about the differences between autistics and controls. In this light, FMRI, Functional MRI, present an important advantage compared to its peers, as it stands out as the most widely used neuroimaging technique, primarily due to its availability and safety, particularly for use with children. This non-invasive method identifies various neural connections by measuring the temporal correlation between fluctuations in the Blood Oxygen Level Dependent (BOLD) signal of distinct anatomical regions [22]. The resultant output consists of four dimensional images, offering the visualization of the brain's structure and functional activities.

fMRI can be employed in two different settings: resting state and task-based paradigms. Although task-based fMRI may be better suited to investigate certain theories related to autism, such as the weak coherence theory [23], it may pose challenges for many individuals with ASD, especially those who are considered low-functioning or young children [24]. In contrast, resting state fMRI does not necessitate a constrained experimental setup or the active and focused participation of subjects. It has also demonstrated the ability to capture interactions between brain regions, potentially leading to the identification of diagnostic biomarkers for neuropathological conditions [25]. Hence, rs-fMRI can be concluded to be very fitted to study ASD.

1.2 Goals

When talking about autism we refer to a vast spectrum of symptoms that vary in severity and occurrence. Some autistic subjects can present all the symptoms while others may have less severe and more subtle behavioral features that may only manifest after a big change in environment. Moreover, the symptoms manifestation and severity can change throughout development. All these factors render the diagnosis difficult and may lead to delaying it.
Another setback that faces the early detection of autism is the vagueness that encloses autism causes. As there are theories of hereditary autism [26], [27] but also theories about environment induced autism [28], [29]. Lead author John Lewis, a researcher at the Montreal Neurological Institute and Hospital of McGill University and the Ludmer Centre for Bioinformatics and Mental Health, found a correlation between network inefficiencies and autism symptom severity. He concluded that identifying the earliest signs of autism is important because it may allow for diagnosis before behavioral changes appear, leading to earlier intervention and better prospects for a positive outcome [30].
By identifying the brain regions with anomalies and pinpointing the neural dysfunctions associated with autism, researchers can narrow down the genetic factors and mechanisms potentially responsible for the development of autism. This, in turn, could pave the way for the development of preventive measures aimed at curtailing the progression of the disorder.
Therefore, in this thesis we study the possibility of early diagnosis based on brain images with the objective of eradicating the drawbacks that the actual used methods of diagnosis present and paving the way for a future detection system that can restrain autism before the appearance of its symptoms. To achieve this goal, we need to adapt the means of diagnosis to young children and verify their safety of use. Then apply the methods and tools that will help extract any useful information related to the disorder. Cluster this information and analyze it to build a system that can permit to classify the tested subjects into their true category with the highest possible accuracy.
Through this work we also pursued other minor goals as:
Verifying the impact of certain characteristics as the age and gender on the autism detection.
Levering the effects, the choice of the preprocessing strategy can have on the final result of the fMRI images analysis.
Finding methods and tools that can encapsulate some of the most recurrent autism theories, with a primary aim to verify their authenticity and, in tandem, enhance the detection system—ultimately giving rise to a tailor-made autism detection framework.
Underscoring the potential of utilizing multiple combinations in offering fresh insights into alternative outcomes by employing a mix-and-match approach across various methods and datasets. This includes evaluating the effectiveness of machine learning and deep learning techniques in classifying subjects as either autistic or non-autistic.

1.3 Contributions

Through this thesis, by conducting multiple analysis with different batches of component and strategies, we were able to reach a variety of conclusions concerning the autism detection that go from the preprocessing phase to the classification process.
The contributions of this work can be summarized in the following:

Firstly, by testing different preprocessing strategies, we were able to highlight their impact on the results of the following analysis. Additionally, the mix and match strategy showed that there is no perfect preprocessing strategy. As every strategy can achieve good results when coupled with its matched analysis method. Hence, there is no bad or perfect preprocessing pipeline.
Another point worth mentioning concerns the impact of age. In the first tests conducted, it was visible that the subjects with age over 18 had the lowest accuracy whilst those younger were more easily separable. This result corroborates the theories regarding the decreasing effect that age have on the severity of the disorder, especially on the subjects receiving medical help.
Continuing on the same path of theories investigation, another important contribution stems from searching for methods that can embody some of the autism theories to explore this disorder in depth. This approach was crowned by finding evidence of a deficit of under-connectivity between the regions of the autistic brain.
This also opened the path to using some new approaches to utilizing brain images while reducing the calculation consumption by exploring the brain connectivity instead of the raw four-dimensional fMRI images. In this same track, we developed new methods that permit to retrieve the maximum information possible by mimicking, condensing and mathematically modifying methods usually used with other types of data to fit our investigation and keep the cost of analysis to a minimum.
Furthermore, this work embraced the heterogeneity of the data to deduce the most effective results since autism spectrum disorder, as its name suggests has a spectrum of characteristics and a spectrum of manifestations. Hence, the need of a strategy permitting to detect autism without any concern about other factors. Since when studying groups of subjects sharing the same characteristics, the results become biased to those traits and when changing the tested subjects to others with a slight difference, the results become void. However, by keeping the heterogeneity of the data and reducing the elimination of images to the strict minimum by only excluding damaged material, the study can then be generalized.
Last but not least, by introducing mathematical theories into the autism detection combined with the autism tailored methods the accuracy of detection was greatly impacted in a good way and permitted to reach new highs in the autism detection. In other words, by pooling all the information extracted from the above-mentioned conclusions, we were able to create a final detection system that permitted to classify autistic subjects with an accuracy that reached the highest value of 96%.

1.4 Thesis outline

This thesis is organized as follows:

- Chapter 1: Introduction to the thesis topic, background information and theory.
- Chapter 2: Presentation of the autism spectrum disorder, the imaging techniques, the database, preprocessing and ASD-adapted methodology.
- Chapter 3: Investigation of ASD theories using a novel approach of elimination.
- Chapter 4: Features extraction for autism detection using a combination of machine and deep learning.

- Chapter 5: Details of the experiment of connections enhancement for autism detection using deep learning.
- Chapter 6: Conclusion and future directions.

Chapter 2

Data and methods

Finding an alternative mean of diagnosis is one of the big poles of autism research. Using brain images is one of the leads researchers have been pursuing to achieve this goal. However, a major obstacle faced in analyzing brain images is the limited availability of data. Early studies on ASD using fMRI images encountered the issue of small dataset sizes, which hampers the generalization of results.
To surmount this critical concern, the Autism Brain Imaging Data Exchange (ABIDE) database emerged as a groundbreaking solution. This repository constitutes an expansive trove of fMRI images gleaned from a substantial array of subjects spanning diverse international locations. The inclusion encompasses a tapestry of variables, including age demographics, IQ levels, genders, and other pertinent characteristics. The ABIDE initiative is delineated into two distinct parts, namely ABIDE I and ABIDE II with a preprocessed version for ABIDE I.
The influence of the preprocessing on the eventual outcomes of any analytical endeavor cannot be overstated. The selection of these pipelines is a paramount determinant of the conclusions drawn from the data [31] and [32]. The advent of a shared resource such as ABIDE casts its brilliance in facilitating meaningful juxtapositions with other research revelations.
Henceforth, our work takes the preprocessed ABIDE database as the cornerstone for our investigative and exploratory pursuits into Autism. Prior to delving into the tests and analyses, this chapter begins by unfolding the tapestry of Autism disorder, from tracing its historical roots to examining its diagnostic and treatment paradigms. The chapter continues by giving a brief overview of the different imaging techniques with an explanation regarding the reason behind the choice of using fMRI as the main data source.
Later, in this same chapter, we introduce more in detail the ABIDE database and its preprocessed version. Finally, we present the focal methods of this work, namely, the autism tailored methods.

2.1 Autism Spectrum Disorder

2.1.1 History

The term autism was first used around 1911 by the Swiss psychiatrist Eugen Bleuler [33]. He used autism to describe a group of symptoms related to schizophrenia. According to the perspective of Bleuler, autistic thinking is distinguished by the presence of infantile wishes aimed at escaping from

unsatisfying realities. Individuals exhibiting autistic thinking tend to replace these unsatisfying aspects with fantasies and hallucinations. In this view, the individual retreats into a world of their own imagination as a coping mechanism to distance themselves from distressing or dissatisfying aspects of reality. The term "autism" finds its origin in the Greek word "autos" which means "self". It describes conditions in which a person experiences a disconnection from the social world, leading to the phenomenon of an "isolated self".

In the 1940s, Leo Kanner, a renowned psychiatrist and the founding director of the child psychiatry program at the Johns Hopkins University School of Medicine, introduced the term 'autism' into psychiatric classification [34]. His work centered around observing and describing the behavior of several children who displayed withdrawn and socially disconnected behaviors. Kanner's groundbreaking research culminated in the first systematic description of what he termed "Autistic Disturbance of Affective Contact," which served as the initial comprehensive characterization of this distinct neurodevelopmental disorder.

Following the publication of Kanner's paper in 1943, there was considerable debate about the classification of the described children. There was even a belief that those characteristics were an early manifestation of schizophrenia. This lack of clear understanding surrounding autism's nature resulted in its classification in DSM-II as an infantile psychosis, grouped within the broader diagnostic category of childhood schizophrenia [35].

Later, due to the resemblance in phenotypic and clinical features, Autism and schizophrenia were linked in many researchers' minds until the 1960s. This overlap between schizophrenia and autism was also related to the fact that some of the individuals diagnosed with ASD subsequently develop schizophrenia symptoms. It has even been demonstrated later that the two disorders co-occur frequently [36].

It was only in 1971 that medical professionals began to have a separate understanding of autism, and the conceptual separation of the autistic disorder from schizophrenia was ultimately achieved by carefully delineating symptomatic differences (such as age of onset), examining family histories, and comparing treatment responses in individuals with suspected adult schizophrenia and those with infantile autism. These thorough investigations allowed for a clear distinction between the two conditions, paving the way for their separate classification and facilitating a more accurate understanding of their unique characteristics and treatment approaches. [37].

While Kanner's initial description of 'infantile autism' was indeed replicated in the scientific literature, it took considerable time for his diagnostic term to be officially included in the American Psychiatric Association's Diagnostic and Statistical Manual (DSM). Specifically, it was not until 1980, approximately 37 years later, that 'infantile autism' was formally incorporated into the DSM-III [38].

In the 1980s, according to the National Dissemination Center for Children with Disabilities (NICHCY) [39], a new term Pervasive Developmental Disorders emerged and was first used to describe a class of disorders with many similarities to autism. However, it was not introduced in the DSM-III as a new term. The specific diagnostic criteria including the differentiation between different subtypes and the inclusion of specific disorders under the PDD umbrella, were only defined in 1987 and refined in subsequent editions of the Diagnostic and Statistical Manual of Mental Disorders, starting with the DSM-III-R [40]. This class of disorders share common characteristics that can be categorized into the following groups: difficulties in social interaction, challenges in imaginative activities, impaired

verbal and nonverbal communication skills, and a tendency to engage in a limited number of repetitive interests and activities.
Later in 1994, the manual used by professional to diagnose psychiatric disorders the Diagnostic and Statistical Manual of Mental Disorders published a revised version DSM-IV [41] that listed five disorders under the category of Pervasive Developmental Disorders, namely, Autistic Disorder, Rett's Disorder, Childhood Disintegrative Disorder, Asperger's Disorder, and the Pervasive Developmental Disorder Not Otherwise Specified, or PDDNOS.
In 2013, the American Psychiatric Association introduced the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders, DSM-5 [42], which replaced DSM-IV and brought significant changes to the diagnostic definition of autism. In the DSM-5, the individual disorders of autistic disorder, Asperger's disorder, and PDD-NOS were amalgamated under the unified term "autism spectrum disorder" (ASD). This change aimed to recognize the continuum of symptoms and severity observed in individuals with autism, providing a more comprehensive and inclusive framework for diagnosing and understanding the condition.

2.1.2 Characteristics

As autism history has gone through many changes since the first use of the term "autism", the characteristics definition of this disorder has also seen many changes. From Kanner to DSM, every update added a new triad of characteristics broadening the spectrum of autism.
In his 1943 publication, Kanner highlighted two fundamental features of what he termed "inborn autistic disturbances of affective contact" [34]. Firstly, he emphasized the presence of severe difficulties in social interaction and connectedness right from the beginning of life. Secondly, he identified a resistance to change or a strong insistence on sameness. This term "sameness" was employed to describe stereotyped movements observed in autistic children, such as body rocking and hand flapping. Kanner believed that these repetitive movements were ways for the child to maintain a sense of consistency and stability in their world.
On the other hand, Asperger's study revealed characteristics that resembled personality disorders, and he also reported similar issues in the fathers of the cases he examined. Asperger's work sparked a continuing debate on the boundaries of autism, leading to the exploration of the "broader autism phenotype" and discussions on the concept of neurodiversity [35].
In DSM-II, the only available category to describe individuals with early childhood onset of severe developmental disturbances, as described by Kanner, was "childhood schizophrenic reaction." This highlights the historical challenges in accurately defining and diagnosing autism and the evolving understanding of the spectrum of neurodevelopmental disorders.
However, this anomaly was a distinctive concept in its own right and not the earliest manifestation of schizophrenia. Thus, the emergence of a new definition of autism as proposed by Rutter in 1978 which included delayed and deviant social and language abilities with restricted interests and repetitive behaviors that appears beyond the general developmental level.
The American National Society for Autistic Children [43] put forward a definition of autism that encompassed not only unusual rates and sequences of development (similar to Rutter's concept) but also highlighted hypo- and hyper-sensitivities to the environment. This definition underscored the

importance of considering sensory sensitivities as a core aspect of autism, acknowledging that individuals with autism may exhibit atypical responses to sensory stimuli. These sensory sensitivities can manifest as either hypo-sensitivity, where there is decreased sensitivity to certain sensory inputs, or hyper-sensitivity, where there is heightened sensitivity and over-responsiveness to sensory stimuli. The inclusion of these sensory aspects in the definition expanded the understanding of autism beyond developmental patterns and contributed to a more comprehensive characterization of the disorder. Leading in addition to other lines of research to the establishment of the validity of autism as a diagnostic concept in DSM-III [40].

By conducting a comprehensive comparison of the clinical phenomenology of autism, including age and family history of schizophrenia and childhood schizophrenia, researchers found distinct differences between these concepts [44], [45]. Additionally, the treatment patterns used to approach these disorders further highlighted their dissimilarity. Structured teaching approaches appeared to be more effective for children with autism, in contrast to the unstructured psychotherapy used in schizophrenia treatment [46].

Moreover, evidence of the strong genetic basis of autism emerged, with higher rates of concordance observed in monozygotic twin pairs compared to same-sex dizygotic twin pairs [47]. This finding discredited the disproven "refrigerator mother" theory of autism [48] and lent support to the notion that autism has a biological origin.

Based on these substantial considerations, autism, previously referred to as "infantile autism," was officially included as a diagnostic category for the first time in the third edition of the Diagnostic and Statistical Manual [40]. It was placed within a new class of conditions known as Pervasive Developmental Disorders (PDDs), encompassing disorders characterized by delays in the development of multiple basic functions, including socialization and communication. However, it became evident that individuals with autism exhibit changes over time that deviate from the classic presentation of the disorder in infancy. As a result, the term "residual infantile autism" was introduced to accommodate cases where individuals once met the criteria for infantile autism but no longer do so at a later stage in their development. This acknowledgment allowed for a more nuanced understanding of the diverse presentations of autism and emphasized the importance of considering developmental changes when diagnosing and supporting individuals with this condition.

With the introduction of the residual form of autism in DSM-III, concerns emerged regarding the rigidity and lack of a developmental orientation in the diagnosis. As a response to these concerns, significant changes were made in the 1987 revision of DSM-III, known as DSM-III-R. One of the most important changes was the transition from "infantile autism" to "autistic disorder" as the name for the condition.

In DSM-III-R, a new polythetic set of 16 detailed criteria was provided, organized into the standard three major domains of dysfunction observed in autism, namely, the qualitative impairments in reciprocal social interaction, the impairments in communication, and the restricted interests/resistance to change and repetitive movements. The diagnosis of autistic disorder was considered valid when at least eight out of the total criteria were met.

Meanwhile, the World Health Organization's International Classification of Diseases, 10th edition (ICD-10) [49], adopted a different approach with two diagnostic guides, one for clinical work and one for research. However, this resulted in a divergence between the DSM and ICD-10 views, potentially

complicating research comparisons and international collaborations, especially in fields like genetics and epidemiology. In response to these challenges, the fourth edition of DSM (DSM-IV) [41] underwent major revisions.
In DSM-IV, the traditional three-category model was retained, but the final set of criteria underwent changes, becoming more detailed. The minimum number of criteria required for a diagnosis of autism was reduced to include at least six of the characteristic features.
Later, building upon the foundations laid by DSM-IV and decades of research, DSM-5 [14] was introduced, representing a significant shift in the conceptualization of autism. The DSM-5 merged the Pervasive Developmental Disorders into the autism spectrum disorder (ASD) category, transforming autism diagnosis from a multi-categorical system to a single diagnosis based on multiple dimensions. This unification better reflected the diverse presentations and variations within the autism spectrum, providing a more comprehensive and flexible framework for understanding and diagnosing autism.

2.1.3 Prevalence

Autism prevalence has also gone through a significant change throughout history since it was first introduced as a psychiatric disorder. In the beginning it was considered a rare disorder that amounted to nearly 5 per 10,000 emergence of autism among children known to have special needs in the former London Borough of Camberwell. This study was conducted by Lorna Wing and Judith in 1979 on children with IQ under 70 [50].
During that year, Wing and Gould made a significant discovery by identifying a subgroup of children, comprising approximately 15 per 10,000, who exhibited challenges in social interaction, communication, and imagination, along with repetitive and stereotyped behaviors. Despite not entirely fitting the classic description of autism presented by Kanner, these children were classified within the broader category of the autism spectrum.
As a result of their research in the Camberwell study, the overall prevalence rate for all children with special needs falling within the autism spectrum was estimated to be around 20 in every 10,000 children [50].
During the 1990s, Stephan Ehlers and Christopher Gillberg conducted a study aimed at exploring the prevalence of Asperger syndrome and other autism-related profiles among children with an IQ of 70 or above, who attended mainstream schools [51]. Asperger syndrome was initially described by Hans Asperger in 1944, focusing on children who shared many similarities with those diagnosed with Kanner autism. However, the key distinction was that individuals with Asperger syndrome exhibited average to superior grammatical language abilities.
Inclusion of Asperger syndrome in this study was warranted due to its alignment with the core characteristics of autism, such as the triad of impairments involving social interaction, communication, and imagination, as well as the presence of repetitive patterns of activities, akin to those observed in Kanner autism.
Based on the number of children identified in their research, they computed a prevalence rate of 36 per 10,000 for children who were confirmed to have Asperger syndrome and an additional 35 per 10,000 for those experiencing social difficulties. Some of these children might have met the criteria for Asperger syndrome if more comprehensive information had been available, but they were undoubtedly within the autism spectrum. Teachers of these children had previously noticed social and/or educational

differences, yet they had not been able to pinpoint the exact cause for these disparities.
Later on, after a three-decade-long study focusing on children of average or high intelligence who struggled with social interaction, Sula Wolff referenced Ehlers and Gillberg's study. She regarded the total rate of 71 per 10,000 mentioned in their paper as encompassing the children she examined and categorized as belonging to the most 'subtle' end of the autism spectrum [52].
In the beginning of the twenty first century the autistic disorder began to get more attention and more surveys of its prevalence began seeing the light. In 2005, the Office of National Statistics conducted a survey in Great Britain to assess the mental health of children and young people, revealing a prevalence rate of 90 in 10,000 for individuals categorized as having autism [53]. However, this figure did not distinguish between autism, Asperger syndrome, or other profiles within the autism spectrum.
A year later, Gillian Baird and her colleagues published a report based on a survey of children aged between 9 and 10 years old in the South Thames region. The study included children who had already received a diagnosis of autism or had difficulties with social and communication skills. Additionally, children considered at risk of being undetected cases were included if they had a statement of special educational needs. The diagnosis conducted followed the criteria outlined in the ICD-10.
The results of this study indicated a prevalence rate of 38.9 in 10,000 for 'childhood autism', and a rate of 77.2 in 10,000 for other conditions within the autism spectrum, resulting in an overall prevalence figure of 116 in 10,000 [54].
In 2008, the Autism and Developmental Disabilities Monitoring Network in the USA investigated eight-year-old children in 14 states. They found an overall rate of 1 in 88 of autism, with around five times as many boys as girls diagnosed.
Later, from telephone surveys undertaken in 2011 and 2012 of parents of children aged between 6 and 17 years old, the National Center for Health Statistics in the USA published findings. The report showed a prevalence rate for autism of 1 in 50, [55].
In the same period, a higher prevalence rate of 2.64% emerged in a study done in South Korea, where it was reported that two thirds of the people on the autism spectrum were in the mainstream school population, and had never been diagnosed before. [56].
Hence, taking into consideration studies from different parts of the world, researchers put forward an estimation of 62 autistic children in every 10,000. According to them the increase in estimation autism showed in the last decades was related to the broadening diagnostic criteria, service availability and increasing awareness of autism among professionals and the public, [57].
From then on, the prevalence rate of autism kept on increasing as published on the center of disease control and prevention reports reaching in 2023, an estimation of 1 in 36 children diagnosed with autism spectrum disorder (ASD) in the U.S, according to data of 2020. With 1 in every 23 boys identified with autism and 1 in every 87 girls identified with autism. Making the rate of boys being identified with autism four times more likely than girls. Autism was also found to affect all ethnic and socioeconomic groups [58].

2.1.4 Diagnosis

The fact that Autism is one of the most challenging disorders is partially due to the difficulty of diagnosis. To this day, autism diagnosis is mainly based on behavioral tests and/or relatives and

entourage reviews. Most of these assessments are generally developed for early detection as autism is a disorder that shows early. However, no final conclusion is decided before the child reaches the age of two years old.

From the diagnostic assessments used to measure ASD, some are semi-structural while other are computerized. From the semi-structural tests, we can cite ADI-R, ADOS and DISCO [59].

The Autism Diagnostic Interview – Revised (ADI-R) relies on information provided by parents or caregivers for assessment. This diagnostic tool is suitable for children and adults aged 18 months and older. It employs a diagnostic algorithm that aligns with both DSM-5 and ICD-11 criteria. The evaluation focuses on reciprocal social interaction, communication, and restricted and repetitive behaviors to aid in diagnosing autism.

On the other hand, the Autism Diagnostic Observation Schedule (ADOS) evaluates social communication, interaction, and play habits through direct observation. It consists of various modules tailored to the individual's verbal abilities, whether they are children or adults. Each module comprises specific activities that enable the examiner to identify behaviors consistent with an Autism Spectrum Disorder (ASD) diagnosis.

As for DISCO, The Diagnostic Interview for Social and Communication Disorders, it is used to assess a suspected autistic patient over a lapse of time. In this test it is important to have an assessor who is well acquainted with the subject to monitor the subject over a course of time that can be measured in years. Its objective is to delineate behavioral patterns that have developed throughout the growing process of the subject.

The computerized assessment used for ASD detection is The Developmental, Dimensional, and Diagnostic Interview (3Di) [60] which is dedicated to parents of autistic subjects with no restriction of age. It measures the intensity of symptoms across the autistic spectrum as well as offering information about potential comorbidities.

Another test worth mentioning is the Childhood Autism Rating Scale – Second Edition The (CARS-2) [61] that aims to differentiate between children with moderate or severe Autism Spectrum Disorder (ASD) and those with other cognitive disorders, including developmental delays. This assessment tool is commonly utilized in research studies and complements other diagnostic evaluations to aid in clinical diagnoses. It can be administered either by the subject's caregivers or by clinicians.

All the above-mentioned tests rely solely on observations, which cannot be completely trusted especially in the first years of childhood. Hence, research turned toward finding a new mean of diagnosis that can be conducted early and be more trustworthy.

Hence, risk factors including genetics, prenatal conditions, and sociodemographic backgrounds were investigated.

According to researchers, genetics are involved in the vast majority of autistic cases [62]. Children born to older parents have been found to be at a higher risk for having autism. Parents who have a child with ASD have a 2 to 18 percent chance of having a second child who is also affected. Among identical twins, if one child has autism, the other will be affected about 36 to 95 percent of the time. In non-identical twins, if one child has autism, then the other is affected about 31 percent of the time. Looking into the gender differences in ASD manifestation, the official accepted ratio is of 4:1 [63], males to females, with several studies reporting ratios varying from 2:1 to 7:1, males to females. This imbalance engendered different explanations. A social one based on the male bias suggesting that

females are being marginalized and under-recognized. Alternatively, some proponents argued that females might possess a protective effect, making them less susceptible to developing autism.
Biologically, two theories, the extreme male brain theory and the fetal testosterone theory, shed light on the relationship between autism and testosterone levels in the amniotic fluid during pregnancy. According to these theories, higher levels of testosterone in the amniotic fluid are associated with improved abilities in analyzing and understanding complex systems and patterns. However, at the same time, they may lead to a reduction in empathetic traits. This theory also suggests that certain features may become exaggerated, such as larger brain volume, over-connectivity within hemispheres, enlarged amygdala, and reduced inter-hemispheric connectivity due to a diminished corpus callosum. Supporting the testosterone theory further, elevated testosterone levels in both males and females have been connected to some autistic behavioral traits, like avoiding eye contact [64].
Investigations into the prenatal environment have revealed significant factors that contribute to the development of autism in children. Notably, maternal drug intake, illness, immune response, stress, and exposure to pollution have been identified as potential influencers. Maternal substance abuse, alcohol or drug use for mental illness treatments like selective serotonin uptake inhibitors and sodium valproate, have all been associated with an increased risk of ASD [65]. Additionally, the presence of certain genetic susceptibilities related to drug intake, whether for mental illness or abused, increases the likelihood of psychiatric disorders in offspring. Moreover, illnesses like hypertension, obesity, asthma, diabetes, and autoimmune disorders have also been linked to possible autistic symptoms in children. Maternal immune activation was also found to affect a child's neurodevelopmental outcome. According to [66] Maternal infection during critical stages of pregnancy leads to acute fetal inflammation. The consequences of acute fetal neuroinflammation on early neurodevelopmental consequences may facilitate the development of psychopathological and neuropathological phenotypes shared between schizophrenia and autism.
On another hand, Air pollution not only presents a threat to pregnant woman resulting in a higher risk of having autistic children, but also increase the chance of postnatal infants to develop autism throughout life [67].
Even with the advance in autism investigation, most children are still being diagnosed after the age of four.
Since autism is a neuro developmental disorder, research also turned to studying the brain to locate the dysfunction origin, with the objective of early diagnosis of this disorder.
However, the huge complexity of the brain renders identification of the said dysfunction very difficult. As all brain zones are not only physically related but also functionally related. On the bright side, huge advance has been made.

2.1.5 Treatment

Even though, there is no known medication to completely treat autism, there are some available medications that help the children with some symptoms as autism-associated agitation and irritability.
However, medication works best when paired with therapy to develop socialization and other life skills. These therapies vary depending on the specific needs of each individual.
A therapist with expertise in ASD treatments can conduct sessions aimed at fostering various skills,

such as language and social abilities. One-on-one support within the classroom or enrollment in specialized education programs can also be beneficial.
Behavioral interventions play a crucial role in reducing negative behaviors while encouraging positive ones. Additionally, family counseling can provide valuable support and guidance for families of individuals with ASD.
Moreover, since autism usually co-occur with other health problems, certain medications may help. According to [68] more than half the children with autism have one or more chronic sleep problems. Children with autism also suffer of the Deficient Hyperactivity Disorder (ADHD) with an estimation of 30 to 70 percent [69]. Additionally, almost a third of the autistic population have epilepsy [70] and 11 to 40 percent of the autistic children can also display anxiety disorders [71]. Depression is also one of the ailments that touches autistic children [72] with a prevalence that goes from 7 percent in childhood to 26 percent in adulthood. Schizophrenia is also one of the disorders which prevalence increases with age to reach 35 percent. Moreover, children with autism are nearly eight times more likely to suffer from one or more chronic gastrointestinal disorders than other children. Autistic children are also more affected by obesity than the non-autistic.
Hence, early detection can be life changing as children affected by autism can benefit from many interventions such as the aforementioned treatments as well as speech and occupational therapy. The early intervention can improve learning, communication and social skills, as well as underlying brain development. It can also permit to reduce the risk of developmental regression, or loss of skills, that affects around 1 in 5 children with autism and typically occurs between ages 1 and 3.
Farther more, it provides the best opportunity to support healthy development and deliver benefits and minimize the disorder impact across the lifespan.

2.2 Imaging techniques

One of the axes of research for autism detection is based on visualizing and studying the brain. The neuroimaging presents a key to uncovering the secrets of the brain. The different imaging techniques available provides a vast choice of procedures that can be categorized into structural and/or functional. The structural techniques permit the visualization of the morphology of the brain and its different components. While the goal of functional neuroimaging is to understand the cognitive organization of the brain and its dynamics, in typical condition and when affected by a disease or a mental disorder.

2.2.1 Structural imaging techniques

The structural imaging techniques are non-invasive procedures that use different types of technology to provide a representation of a structure without having to remove the skin or bone that protects that structure. They are not only used to visualize the brain but also the other parts of the human body. However, in the neurological field they have been of big help in tracing many diseases. The techniques that have received big acknowledgement and been vastly used are mostly X-ray, CT scan and MRI.

2.2.1.1 X-ray

X-rays are a type of high-energy electromagnetic radiation characterized by a short wavelength that has the ability to penetrate solids and ionize gases [73]. Widely employed in the field of medicine, X-rays are generated by an X-ray machine and directed towards a specially treated metallic plate positioned behind the patient's body. As the radiation beam interacts with the plate, it causes darkening. Soft tissues moderately impede X-rays, resulting in a gray appearance on the plate, while hard tissues like bones largely obstruct the rays, creating a lighter "shadow." Consequently, X-rays are particularly effective for visualizing rigid anatomical structures such as teeth and bones.

2.2.1.2 Computed Tomography

Tomography is a method of obtaining sectional images of the body. Computed tomography (CT), also known as computerized tomography, is an imaging technique that utilizes computer analysis of multiple cross-sectional X-rays to provide detailed information about structures within the body. This technique was developed in the 1970s and is based on the principle that X-rays are absorbed or deflected at various levels as they pass through the body. During a CT scan, the patient lies on a motorized platform while a computerized axial tomography (CAT) scanner revolves 360 degrees around them, capturing X-ray images from multiple angles [74].

2.2.1.3 Magnetic resonance imaging

Magnetic resonance imaging (MRI) is a medical imaging method that harnesses a phenomenon in nuclear physics discovered in the 1930s [75]. Unlike other techniques, MRI does not expose patients to radiation and relies on magnetic fields and radio waves. The procedure involves the use of an MRI scanner, essentially a large tube housing a powerful circular magnet. This magnet generates a strong magnetic field that aligns the hydrogen atoms' protons within the body.
By applying a radio frequency (RF) magnetic pulse at the appropriate frequency, the hydrogen nuclei absorb energy and subsequently emit a brief, weak signal known as the MR signal. This signal is detected by the RF coils in the MRI system, allowing the creation of an image.
Early MRI scanners were initially rudimentary, but advancements in digital computing and electronics propelled their refinement, surpassing other techniques in precise imaging. Presently, MRI has become a standard tool in Radiology due to its ability to produce high-resolution images with excellent tissue contrast.

2.2.2 Functional imaging techniques

In ASD investigation, structural imaging techniques have focused on detecting the brain structures with physical differences that might be related to this disorder. In [76] thicker frontal cortex and in [77] a thinner temporal cortex were reported in patients with ASD. These areas are notable because the frontal cortex is linked to communication and language abilities and the temporal cortex is linked to auditory processing (ie. language input), both of which are issues that many with ASD struggle

with.
However, the drawback of using structural imaging in autism detection is the fact that the rate of development of children in the first years vary tremendously from child to child. Hence, the physical attributes of the brain that can pop up as showing abnormality can only be due to the growing pace and not be related to any problem.
Hence, a need for a method that does not rely on brain structure but goes deeper into the functioning of the brain. This is where functional imaging techniques come in handy. Functional imaging techniques include but is not restricted to functional MRI (fMRI), MEG, MRS, PET, and SPECT.

2.2.2.1 Single-photon emission computed tomography

Single-photon emission computed tomography (SPECT) is a nuclear imaging technique frequently employed in diagnostic medicine [78]. SPECT generates a three-dimensional image displaying the distribution of a radioactive tracer that is injected into the bloodstream and subsequently absorbed by specific tissues.
The SPECT procedure consists of several stages. In cerebral studies, the patient is initially seated in a calm, dimly lit room and instructed to refrain from reading or speaking for a minimum of 10 minutes. Subsequently, the patient receives an injection of a radio-labeled tracer compound, carefully selected by a nuclear pharmacist. If sedation is required, it is preferably administered after the tracer injection. Following the injection, a variable waiting period is observed to allow the tracer to circulate and be absorbed by the target tissues. The duration of this waiting period may vary based on the specific study, tracer selection, and tracer dosage, and it can range up to 90 minutes.
After the waiting period has elapsed, the patient is positioned within the detector apparatus. In certain studies, vasodilatory medications may be used to assess tissue perfusion. Once the patient is properly positioned, and any necessary medications have been administered, the detector rotates around the patient, capturing planar scans at intervals of 3 to 6 degrees. These scans are subsequently combined to generate the final three-dimensional image. While SPECT provides detailed information about the tissues, it does have limitations. Small metabolic abnormalities detected by SPECT can often be challenging to precisely locate without corresponding anatomical imaging.

2.2.2.2 Positron emission tomography

Positron emission tomography (PET) [79] is a functional imaging technique that utilizes radiotracers to visualize and measure metabolic and physiological processes within the body. PET relies on the detection of small amounts of radioactive substances known as positron emitters. Commonly used positron emitters in PET imaging include carbon-11, oxygen-15, nitrogen-13, and fluorine-18, each with its own specific imaging purposes based on the target process in the body.
The radioactive isotopes used in PET imaging have short half-lives, such as 2 minutes for oxygen-15 and 109 minutes for fluorine-18. Regional tracer concentration is imaged using the properties of positron decay and annihilation. After emission from the parent nucleus, the emitted positron travels a short distance within the tissue until it is slowed down and combined with a free electron, forming a positronium. This positronium then decays through annihilation, producing a pair of gamma rays.

These gamma rays, with an energy of 511 kiloelectron volts (keV) each, travel in nearly opposite directions. PET detectors, arranged in pairs aligned collinearly, capture and detect these coincident gamma rays. This "electronic collimation" technique is what allows PET to be much more sensitive (over 100 times) than conventional nuclear medical techniques like single-photon emission tomography (SPECT), which uses gamma cameras and lead collimators [80].
However, despite the high sensitivity of 3D PET, there is a significant amount of scattered radiation that can interfere with the quality of the final 3D acquisition. Scattered radiation refers to gamma rays that deviate from their original path due to interactions within the tissue. These scattered gamma rays can be detected and measured, affecting the overall quality and accuracy of the PET images obtained.

2.2.2.3 Magnetoencephalography

Magnetoencephalography (MEG) [81] is a non-invasive functional neuroimaging technique used to map brain activity by detecting and recording the magnetic fields generated by the electrical currents naturally occurring in the brain. MEG employs highly sensitive magnetometers to measure the magnetic fields, which typically range from femto-tesla to pico-tesla in strength. It offers exceptional temporal resolution, accurately capturing the timing of neuronal activity, and also provides good spatial resolution.
During MEG recording, the magnetic fields are generated by the electric currents produced by activated neurons in the brain. According to the right-hand rule, an electric current is always associated with a magnetic field that is perpendicular to its direction. Since the magnetic permeability of biological tissues is similar to that of empty space, the magnetic field is not significantly distorted by the scalp or skull.
The source of the magnetic fields recorded by MEG is the dendritic currents of pyramidal neurons that fire synchronously and in parallel. Axonal and synaptic currents, along with their magnetic fields, tend to cancel each other out. Consequently, the synchronous activation of pyramidal neurons produces detectable magnetic fields outside the head, allowing MEG to capture the neural activity.
The amplitude of the magnetic fields generated by the brain is extremely small, typically in the range of a few hundred femto-tesla ($10^{-15}T$). By comparison, the Earth's magnetic field ranges from 10^{-4} to $10^{-5}T$. Therefore, the challenge lies in detecting and recording these small magnetic fields while effectively shielding out the relatively stronger Earth's magnetic fields. This challenge is addressed by employing superconducting quantum interference detectors (SQUIDs), which act as highly sensitive magnetic field detectors. These SQUIDs are housed within magnetically shielded rooms, which provide a controlled environment free from external magnetic interference.

2.2.2.4 Functional Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging (fMRI) [82] revolutionized the landscape of human systems-level brain research. It is a class of imaging methods developed in order to demonstrate regional, time-varying changes in brain metabolism. This technique brought many key advantages for the brain research community and have been used in an exceptionally large number of studies in the

cognitive neurosciences, clinical psychiatry/psychology, and presurgical planning.
The popularity of fMRI derives from its widespread availability. It can be carried out on clinical scanners, which enabled rapid growth using the clinical scanner infrastructure. Moreover, its price is more affordable when compared to other technologies. Therefore, it can be found in hospitals worldwide. FMRI is also a non-invasive technique that does not require any injection of radioisotope or other pharmacologic agent. This makes it safe to use on all population categories. Especially, on young children.
The spatial resolution of fMRI images is also considered a great advantage when compared to other functional imaging techniques. It can show functionally informative voxels as low as 0.5 mm x 0.5 mm x 0.5 mm. FMRI is also characterized by speed. The acquiring time of the sequential imaging volumes can be as low as 200ms for every sequence. Moreover, fMRI has sufficient sensitivity to create functional maps on individuals in minutes.
Functional MRI involves the use of an MRI scanner equipped with hardware and software to scan an entire volume of data using a pulse sequence called echo planar imaging (EPI) [83]. It captures the activity of the brain by mapping the metabolic changes that can be consequent to task-induced cognitive state changes or the result of unregulated processes in the resting brain.
FMRI measures the small changes in blood flow that occur with brain activity. All the actions and reactions of the human body induces a reaction in the brain that may include the formation and propagation of action potentials, binding of vesicles to the pre-synaptic junction, the release of neurotransmitters across the synaptic gap, their reception and regeneration of action potentials in the postsynaptic structures, scavenging of excess neurotransmitters, etc. All these require energy in the form of adenosine triphosphate (ATP) [84]. This nucleotide is produced principally by the mitochondria from glycolytic oxygenation of glucose, and its production results in carbon dioxide as a byproduct. When a region of the brain activates, the additional neural firing and other increased signaling processes result in a locally increased energy requirement, in turn resulting in up-regulated cerebral metabolic rate of oxygen in the affected brain region.
FMRI measurements can be accomplished with different techniques. Bolus tracking measures the increase of regional cerebral blood volume. Spin tagging assesses the regional cerebral blood flow. However, since its discovery in 1990, the blood-oxygen-level dependent (BOLD) contrast have been the most used making it the standard technique. This technique takes advantage of the different magnetic properties of oxygenated and deoxygenated hemoglobin to generate an image contrast [85].

2.3 Data

2.3.1 ABIDE

Data availability can play a significant role in the advance of research. Hence, with the goal of facilitating discovery and comparisons in the Autism research, the Autism Brain Imaging Data Exchange (ABIDE) initiative has aggregated functional and structural brain imaging data collected from laboratories around the world in a base named ABIDE [86].
ABIDE includes two large-scale collections: ABIDE I and ABIDE II. Each collection was created through the aggregation of datasets independently collected across more than 24 international brain imaging laboratories and made available to investigators throughout the world, consistent with open

science principles, such as those at the core of the International Neuroimaging Data-sharing Initiative. The Autism Brain Imaging Data Exchange I (ABIDE I) is the first ABIDE initiative. It was launched in August 2012 and involved 17 international sites. ABIDE I aggregates collected resting state functional magnetic resonance imaging (RS-fMRI), anatomical and phenotypic datasets made available for data sharing on multiple platforms. This database yields 1112 dataset, including 539 from individuals with ASD and 573 from typical controls with different range of age comprised between 7 and 64 years. Since its establishment, ABIDE I has been used in multi research works across the globe. Through these works, it was demonstrated that brain imaging can be a mean to capture whole brain and/or regional properties of the brain connectome related to autism spectrum disorder. Hence, the feasibility of aggregating resting state fMRI and structural MRI data across sites was proved very beneficial for autism detection. In accordance with HIPAA guidelines and 1000 Functional Connectomes Project / INDI protocols, all datasets have been anonymized, with no protected health information included.
Later, ABIDE II was established to further promote discovery science of the brain connectome in ASD. It saw the light in June 2016 and provided the researchers with new data. ABIDE II aggregates more than 1000 additional datasets with greater phenotypic characterization, particularly in regard to measures of core ASD and associated symptoms. In addition, two collections including longitudinal samples of data collected from 38 individuals at two time points with intervals between 1 and 4 years were added. To date, ABIDE II involves 19 sites, 10 charter institutions and 7 new members. The number of samples comprises 1114 datasets from 521 individuals with ASD and 593 controls with age range between 5 and 64 years. In accordance with HIPAA guidelines and 1000 Functional Connectomes Project / INDI protocols, in ABIDE II also, all datasets are anonymous, with no protected health information included.

2.3.2 ABIDE Preprocessed

Preprocessing in the medical field is very important, especially in the case of heterogeneous imaging data. Because, the different types of imaging machines used to produce the brain images have different rendering with different characteristics [87]. Preprocessing helps diminishing those differences by giving the data a same format and correcting all abnormalities.
The preprocessing techniques applicable on functional images are vast and different, it goes from timing correction to giving the same format to data from different acquisitions among other corrections. The preprocessing on imaging data can be applied through different pipelines of methods and strategies. Its focal point is to reduce the noise of the data, since a small amount of noise can lead to results that can be very far from reality.
Combining the multiple preprocessing methods can lead to different preprocessing pipelines. However, the choice of the preprocessing pipeline has been proven to impact the results of any analysis [31]. A same analysis done to the same data with different preprocessing pipelines can give different results. This showcases the importance of the choice of the preprocessing pipeline.
To cope with the dilemma of the choice of the preprocessing pipeline, an open sharing of preprocessed neuroimaging data from the Autism Brain Imaging Data Exchange was released. The database concerned was ABIDE I. This latter was preprocessed by five different teams using their preferred tools. The functional preprocessing was performed using the Connectome Computation System (CCS), the Configurable Pipeline for the Analysis of Connectomes (CPAC), the Data Processing Assistant

for Resting-State fMRI (DPARSF) and the NeuroImaging Analysis Kit.
Due to the controversies surrounding bandpass filtering and global signal regression, four different preprocessing strategies were performed with each pipeline. These strategies were made to add filtering with and without global signal correction in addition to the initial pipeline with no added filtering also with and without global signal correction.
These strategies were named as follows:

- TrueTrue: This strategy includes band-pass filtering and global signal regression.
- TrueFalse: In this strategy only the band-pass filtering is included.
- FalseFalse: Neither the band-pass filtering nor the global signal regression is included.
- FalseTrue: Only the global signal regression is included.

Moreover, to limit the variance between outputs, statistical derivatives for each pipeline and strategy were calculated by the CPAC software.
In the state of art most research papers use the Configurable Pipeline for the Analysis of Connectomes. This provides a comparison base for autism research. Additionally, ABIDE preprocessed have an advantage of processing time and space gain.

2.3.3 Brain parcellation

In the context of fMRI data, each image encapsulates a vast amount of information. In our research, we have chosen to concentrate on studying the connectivity of the brain. However, it's important to acknowledge that the brain consists of billions of interconnected synapses, resulting in colossal connectivity matrices. This complexity poses significant challenges in interpretation.
To address this issue, atlases offer a solution by reducing the complexity. Atlases group the synapses of the brain into specific regions known as regions of interest (ROIs). Essentially, an atlas establishes a relationship between a set of coordinates and an anatomical label, facilitating a simpler understanding of brain interactions [88].

2.3.3.1 AAL

The AAL atlas, referenced in [89], provides a framework for labeling brain activation. It is constructed based on an anatomical parcellation of the Montreal Neurological Institute (MNI) single-subject brain. The AAL atlas is bundled with the AAL Toolbox. To achieve functional resolution (3x3x3 mm3), the AAL atlas is subdivided using nearest-neighbor interpolation, resulting in the division of the brain into 116 regions.

2.3.3.2 Dosenbach

The Dosenbach 160 atlas, which is distributed with DPARSF/DPABI and mentioned in [90], consists of 160 spheres with a radius of 4.5 mm, positioned at specific coordinates. These regions were identified based on meta-analyses of task-related fMRI studies. The Dosenbach atlas is known for

its reliability and is particularly suitable for constructing functional networks, as opposed to purely anatomical atlases. Its design takes into account the functional characteristics of the brain, making it a valuable tool for investigating brain connectivity and network analysis.

2.3.3.3 CC200

The CC200 atlas, described in [91], divides the brain into 200 regions based on functional parcellation. To achieve this, a two-stage spatially-constrained functional procedure was applied to preprocessed and unfiltered resting-state data from an independent dataset comprising 41 individuals aged between 18 and 55 years old.
For each individual, a grey matter mask was created by averaging individual-level grey matter masks obtained through automated segmentation. Connectivity graphs were then constructed by considering each voxel within the grey matter mask as a node and the temporal correlations with neighboring voxels on a 3D level as edges. From these connectivity graphs, association matrices were generated, setting the connectivity between voxels to 1 if they belonged to the same region of interest (ROI) and 0 otherwise.
Later, a group-level correspondence matrix was constructed by averaging the individual-level association matrices and subsequently partitioned into 200 regions using normalized cut clustering. The resulting analysis at the group level was fractionated into functional resolution using nearest-neighbor interpolation.
The labels for each ROI of the CC200 atlas were assigned based on their overlap with the Eickhoff-Zilles (EZ) atlas [92], Harvard-Oxford (HO) atlas [93], Talaraich and Tournoux (TT) atlas [94], and the AAL atlas.
The EZ atlas is a functional atlas derived from the max-propagation atlas found in the SPM Anatomy Toolbox. To make it compatible with template space, the EZ atlas was transformed and aligned accordingly. Additionally, the atlas was fractionated into functional resolution, allowing for more detailed analysis of brain regions and their functional properties.
The HO atlas, which is distributed with FSL, underwent a series of steps to create both cortical and subcortical probabilistic atlases. Initially, the atlas was subjected to a threshold of 25%, retaining regions with a probability above this threshold. The subcortical regions, representing white matter, gray matter, cerebrospinal fluid, and the brainstem, were then excluded from the subcortical atlas. The cortical and subcortical regions were combined to form a comprehensive atlas, and like the EZ atlas, it was also fractionated into functional resolution to enable detailed functional analysis.
Finally, the TT atlas, distributed with AFNI, went through a registration process to align it with a common template space. After coregistration, the TT atlas was further warped to ensure accurate alignment with the template space. Similar to the other atlases, it was also fractionated into functional resolution using nearest-neighbor interpolation, providing a functional parcellation that can be used for detailed analysis of brain regions and their functional characteristics.

2.3.4 Time series

A time series refers to a sequence of data points or observations collected and recorded over time. It is a collection of data points ordered chronologically, with each data point associated with a specific

time or time interval. Time series data is often used to study how variables change over time and to identify patterns, trends, and relationships in the data [95].
In time series, the time dimension is a fundamental component, and the data points are typically recorded at regular intervals (e.g., hourly, daily, monthly) or irregular intervals depending on the nature of the data collection process. Time series data can be univariate, involving a single variable observed over time, or multivariate, involving multiple variables observed simultaneously over time.
Time series data can originate from various sources and fields, such as economics, finance, meteorology, healthcare, engineering, and many others [96]. It can be taken on any variable that changes over time.
Time series analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for. In particular, a time series allows one to see what factors influence certain variables from period to period.
In the human body every action and reaction engage fluctuations in the brain activity. Going further in the specification, looking, hearing, talking, feeling, etc, induce different patterns of fluctuations and involve different brain zones. These fluctuations are captured in the functional brain images. As the brain can be segmented into a group of specialized regions. Each one of these regions is responsible for one or more activities. Even in rest with closed eyes, the brain is still highly active, and the patterns of activity in this resting state are thought to reveal particular networks of areas that often act together.
The time series of brain functional imaging contain rhythmic activity that reflects neural oscillations [97], these oscillations represent the fluctuations in the excitability of populations of neurons that we refer to as regions of interest. These oscillations can be slow or fast they can last over a long period of time or be transient. Thus, time series permit to capture the interactions of the brain regions into a one-dimensional signal.
Time series analysis techniques applied in brain research include spectral analysis (e.g., Fourier analysis, wavelet analysis), time-frequency analysis (e.g., spectrograms, wavelet transforms), connectivity analysis (e.g., coherence, phase synchronization), and advanced machine learning algorithms to extract temporal patterns or predict brain states [98], [99].
By leveraging time series analysis in brain research, scientists gain insights into brain function, connectivity, cognitive processes, and neurological disorders. It aids in understanding the brain's temporal dynamics, identifying biomarkers, and developing diagnostic tools and therapeutic interventions for neurological and psychiatric conditions.

2.3.5 Connectivity matrices

The primary objective of connectomes, as discussed in [100], is to create comprehensive maps of connections between neural elements that are distributed anatomically. These elements can include individual neurons, specific groups of neurons, or large-scale brain regions. The connectivity between any two nodes in a network is typically represented using a two-dimensional matrix.
In a network with N nodes, the potential number of connections can be on the order of N^2. Consequently, a connectivity matrix is constructed, where the rows and columns correspond to the different nodes in the network. The matrix element positioned at the intersection of each row and column encodes information about the connection between the nodes represented in that row and column.

Given that the brain consists of billions of synapses that are interconnected, the resulting connectivity matrices are enormous, leading to a highly complex system that is challenging to interpret. However, the use of atlases helps alleviate this complexity by approaching the brain at a regional level rather than focusing on individual synapses.
Furthermore, employing atlases allows for more comprehensible connectivity matrices as they provide a condensed description of the pairwise connections between regions that serve a common function or are involved in a shared activity.
To construct a connectivity matrix, denoted as CM, for a brain network consisting of N nodes, a square matrix is created. This matrix has N rows and N columns, with each row and corresponding column representing a unique node in the network. The connectivity values estimated for each pair of nodes are then populated in the elements of this matrix, resulting in an NxN connectivity matrix. The subscripts of CM are used to index each element, where the first subscript, i, represents the row index, and the second subscript, j, represents the column index. This matrix representation is highly versatile and can be utilized to encode various structural properties of a network.

2.3.5.1 Covariance matrices

In mathematics and statistics, covariance serves as a measure of the relationship between two random variables. It quantifies the degree of variation shared by the two variables [101]. Essentially, covariance provides insight into how two variables change together.
Covariance is expressed in units, which are determined by multiplying the units of the two variables being analyzed. The resulting covariance value can be positive or negative. A positive covariance suggests that the two variables tend to move in the same direction, meaning they exhibit similar patterns of variation. Conversely, a negative covariance indicates that the two variables tend to move in opposite directions, demonstrating an inverse relationship.
To compute the covariance between two variables, X and Y, the calculation involves measuring the differences between the X-component and the mean of X, as well as the differences between the Y-component and the mean of Y. These differences are multiplied together, and the products are summed. Finally, the sum is divided by the number of samples to obtain the covariance value.
Mathematically this can be formulated as:

$$Cov(X,Y) = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{n} \tag{2.1}$$

Where:
X_i – the values of the X-variable
Y_i – the values of the Y-variable
$\bar{X}$ – the mean (average) of the X-variable
$\bar{Y}$ – the mean (average) of the Y-variable
n – the number of data points

2.3.5.2 Correlation matrices

While covariance measures the direction of a relationship between two variables, correlation [102] goes a step further and quantifies the strength of that relationship. Positive values for both covariance and correlation indicate that the variables move in the same direction, while negative values indicate movement in opposite directions.
When it comes to correlation, a coefficient close to +1 indicates a strong positive correlation, meaning the variables have a strong linear relationship in the same direction. A coefficient close to -1 represents a strong negative correlation, indicating a strong linear relationship in the opposite direction. A coefficient close to zero suggests a weak or no linear relationship between the variables.
Covariance, on the other hand, measures the overall variability of two random variables from their expected values. While it provides information about the direction of the relationship, it doesn't provide insight into the strength of the relationship or the degree of dependency between the variables.
Correlation, being a scaled measure of covariance, not only captures the direction of the relationship but also quantifies its strength. One crucial distinction between covariance and correlation is that correlation coefficients are dimensionless. Unlike covariance, correlation coefficients have no associated units and are purely numerical values, making them more interpretable and comparable across different studies or datasets.
The correlation between two random variables (X, Y) is obtained by computing the variables mean and standard deviation, then summing all the squares of the difference between every sample and the mean of the variable. Divide by the number of samples and do the square root of this fraction. By using the covariance, the correlation can be computed using the formula:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \tag{2.2}$$

Where:
$\rho(X,Y)$ – the correlation between the variables X and Y
$Cov(X,Y)$ – the covariance between the variables X and Y
σ_X – the standard deviation of the X-variable
σ_Y – the standard deviation of the Y-variable

2.3.5.3 Tangent space matrices

When working with brain images, the most used connectivity matrices are the correlation matrices. However, new matrices have emerged that are computed in the tangent space [103].
Tangent space connectivity matrices are mathematical constructs used in the field of differential geometry to describe the local geometric properties of a manifold. In particular, they provide a way to measure how tangent vectors on a manifold are related to each other.
A manifold is a geometric space that locally resembles Euclidean space. In simple terms, a manifold is a geometric object that appears flat or Euclidean when viewed up close but may have more complicated global properties.
At each point on the manifold, there is a tangent space, which is a vector space that approximates the manifold's geometry at that point. Tangent vectors in the tangent space represent directions and

velocities that can be taken from the given point.
The tangent space connectivity matrix describes the relationship between these tangent vectors at different points on the manifold. The size of the matrix depends on the dimension of the manifold and the number of tangent vectors being considered.
Each entry of the matrix corresponds to the inner product between two tangent vectors at different points. The inner product measures the extent to which the vectors are aligned with each other. A high inner product value indicates that the vectors are similar in direction, while a low value indicates that they are orthogonal or pointing in different directions.
By examining the tangent space connectivity matrix, one can gain insight into the connectivity and curvature properties of the manifold. The matrix can reveal information about how the manifold is locally connected, whether it has areas of high curvature or flat regions, and how tangent vectors change as one moves along the manifold.
In the medical field and precisely in autism detection, the tangent space connectivity matrices have been tested in [104] and gave good results. Because these matrices permit to go one step further at a group level, they couple the information from all interactions in a unique group connectome using a geometrical framework. Hence, allowing to measure interactions in the common space called the tangent space [105].

2.3.6 Brain connectivity graphs

Connectivity matrices are a two-dimensional representation of the activity of the brain. To visualize this information into a three-dimensional space, the graphs are the best tools to be used.
A brain connectivity graph, also known as a connectome or brain network, is a graphical representation of the connections between different regions or nodes in the brain. It describes the structural or functional interactions among various brain areas and provides insights into the organization and communication patterns within the brain.
A brain connectivity graph $g = (V, E)$ typically consists of a finite set V of vertices or nodes that represent specific brain regions or anatomical locations, and a finite set of edges $E \subseteq VV$, that represent the connections or interactions between these regions. The nodes can be defined based on anatomical landmarks, functional parcellation, or a combination of both. The edges can represent structural connections (e.g., white matter tracts) or functional connections (e.g., statistical dependencies between brain activity).
Brain connectivity graphs provide a valuable framework for studying the complex interactions and organization of the brain, advancing our understanding of brain function and dysfunction. There are two primary types of brain connectivity graphs:

- Structural Connectivity Graph: This type of graph represents the anatomical connections between brain regions. It is derived from techniques such as diffusion MRI (dMRI) and tractography, which allow the estimation of white matter fiber tracts connecting different regions. In a structural connectivity graph, each node represents a specific brain region, and the edges represent the physical connections or fiber tracts between them.

- Functional Connectivity Graph: This graph represents the statistical dependencies or correlations between the activities of different brain regions. It is derived from functional neuroimaging data, such as fMRI, EEG, or MEG. Functional connectivity is based on the temporal synchrony or coherence of neural activity between different brain regions. In a functional connectivity graph, each node represents a brain region, and the edges indicate the strength or similarity of functional activity between them.

Overall, brain connectivity graphs provide valuable insights into the complex network architecture of the brain and its underlying dynamics. They allow researchers to investigate how different brain regions interact and cooperate to support various cognitive functions, behaviors, and neurological processes. Connectivity graphs can reveal important information about information flow, integration, segregation, network hubs, and communication pathways within the brain. By applying graph theory and network analysis techniques, properties of brain networks can be quantified. These latter can be used to identify network-level features, connectivity patterns, and alterations associated with specific brain disorders, cognitive processes, or states.
Brain connectivity graphs have diverse applications in neuroscience, cognitive science, clinical research, and computational modeling. They contribute to our understanding of brain function, provide insights into neurological and psychiatric conditions, aid in the development of biomarkers, and facilitate the advancement of personalized medicine and brain-based interventions.

2.4 Autism theories adapted tools

As aforementioned, the growth process of the autistic brain can lead to abnormalities that primarily affect neural patterning and wiring. These abnormalities often result in an excess of local and short-distance cortical interactions, which can hinder the proper functioning of long-distance interactions between different regions of the brain [106].
One theory that provides an explanation for these abnormalities is the weak central coherence [107]. This latter stipulates that cognitive differences observed in individuals with autism can be attributed to a reduced ability to integrate contextual information, coupled with a tendency towards local rather than global processing. Essentially, individuals with autism tend to exhibit a processing style that is focused on details, emphasizing specific features of objects while potentially overlooking the broader configurations and contextual meanings.
Functional neuroimaging studies consistently demonstrate that performing even simple cognitive tasks involves activation in multiple brain regions, with successful performance relying on interactions among these regions. However, in concordance with the aforementioned theories the abnormally reduced level of neuronal interactions between cortical areas, leads to alterations in region-specific activation. This disruption in long-range neural connectivity is thought to contribute to the core features of ASD.
Furthermore, research findings indicate that individuals with ASD exhibit poor coordination of brain activity within cortical networks, particularly in frontal and parietal cortices, compared to typically developing individuals [108]. More recent studies utilizing group-wise analysis have identified significant deficiencies in long-range connectivity in terms of both intra- and inter-connectivity weights

[109]. These findings further support the concept of long-range neural disconnectivity as a characteristic feature of ASD. On the other hand, there is also short-range intra-regional connectivity, and on this issue, there is less empirical evidence and theoretical agreement [110].
To identify autism successfully, integrating these theories in the investigation process can be a decisive factor. Therefore, verifying these theories by shaping them into a testable form, through tools that permit to extract information that can be said to be autism tailored, can turn autism identification into a new path that might lead to breakthrough in this disorder future diagnosis.

2.4.1 The spanning trees

In the mathematical field of graph theory, a spanning tree is a concept that refers to a subgraph of a connected graph that includes all the vertices of the original graph while maintaining the connectivity and having no cycles.
A graph consists of a set of vertices or nodes and a set of edges that connect pairs of vertices. A spanning tree is a subset of the edges that retains the connectivity of the graph but eliminates any cycles. As a result, a spanning tree is always acyclic and connected [111].
In other words, a spanning tree includes all the vertices of the original graph and ensures that every vertex is reachable from any other vertex in the tree. This property ensures that the spanning tree captures the entire connectivity of the graph.
Furthermore, with no cycles, there are no closed paths or loops within the tree. Hence, removing any edge from the spanning tree would result in a disconnected graph.
In the research field, the spanning tree is considered as an unbiased graph theory tool that simplifies high-order network structure while preserving its core framework [112], [113].
A graph can have many spanning trees, however, when a condition is met, specific trees can be extracted as the widely known minimum spanning tree (MST). Hence, using this graph tools, we can model some of the brain connectivity deficits based on the chosen condition.

2.4.1.1 Minimum spanning tree

As mentionned above, a spanning tree T of an undirected weighted graph G is a subgraph that includes all vertices of G, and connects them together without any cycles. A graph may have several spanning trees where the most known is the minimum spanning tree.
The minimum spanning tree connects all the vertices of a connected, weighted graph while minimizing the sum of the weights of its edges. It seeks to find a spanning tree with the lowest possible sum of edge weights.
These trees can help identify the most cost-efficient or shortest paths between nodes in a graph, ensuring connectivity while minimizing overall costs or distances.
However, for the sake of autism detection, by extracting the brain connections that have the weakest connections, the minimum spanning tree can permit to investigate the underconnectivity deficit of the autistic brain, since weak connections can be related to this aforementioned autistic brain problem.
To find a minimum spanning tree, various algorithms can be employed. These algorithms systematically select edges to include in the minimum spanning tree based on their weights, gradually growing the tree until all vertices are connected. In this work, the minimum spanning trees are extracted using the Kruskal algorithm [114] which is a Greedy Algorithm. The Greedy Choice is to pick the smallest

weight edge in order to create the minimum spanning tree. To achieve that, the algorithm sorts the edges based on their weight. Then, it began building the tree by adding increasing cost arcs at each step while respecting the acyclic proprety of the tree and keeping the total weight of all the edges to the minimum. This process is represented on the right side of figure 2.1.

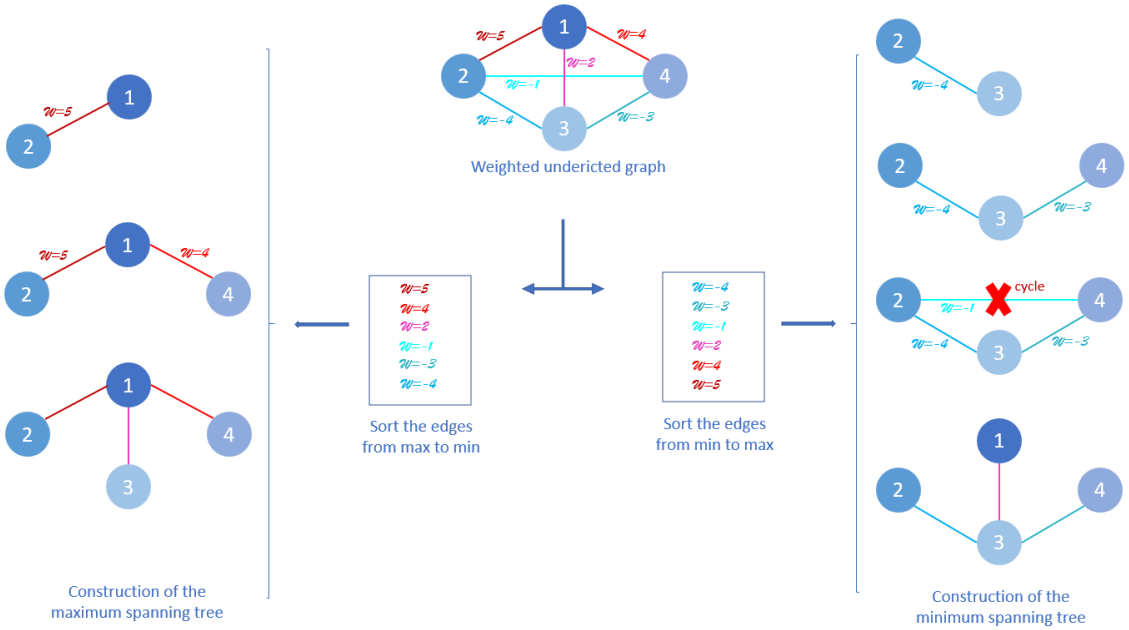

FIGURE 2.1: Construction of the minimum and maximum spanning trees.

2.4.1.2 Maximum spanning tree

A maximum spanning tree (MaxST) is a concept in graph theory that refers to a tree that spans or connects all the vertices of a connected, weighted graph while maximizing the sum of the weights of its edges.

In other words, in a weighted graph, where each edge is assigned a weight or cost value, the maximum spanning tree seeks to find a spanning tree with the highest possible sum of edge weights.

This process is achieved in the same way as for the minimum spanning tree and is represented on the left side of figure 2.1. To find a maximum spanning tree, various algorithms can be employed, such as Kruskal's algorithm or Prim's algorithm, which are commonly used for the minimum spanning tree calculations. However, the approach for finding a maximum spanning tree is slightly different. Instead of selecting the edges with the smallest weight as in the minimum spanning tree, the maximum spanning tree algorithm chooses the edges with the largest weight, prioritizing higher-weighted edges for inclusion in the tree.

The resulting maximum spanning tree has the following properties:

It is a tree: which is a connected acyclic subgraph that spans all the vertices of the original graph.

It maximizes the sum of edge weights: among all possible spanning trees, the maximum spanning tree has the highest total weight.

Maximum spanning trees find applications in various fields, such as network design, transportation planning, and communication networks. They can help identify critical edges or connections in a

network based on their weights, determining the most influential or significant components of the graph. In the case of autism detection, the MaxST permits to extract the brain connections that have the highest correlation values. Hence, it can be used to detect overconnectivity.

2.4.2 Clustering

In addition to examining the over and under-connectivity, investigating the connectivity patterns in the autistic brain involves considering both the local and long-range connectivity deficits.
To assess the validity of these theories, it is necessary to cluster the regions of interest (ROIs) within the brain into groups and analyze the connectivity within and between these groups, thus evaluating the intra and inter-connectivity.
The hierarchical clustering (HC) algorithm [115] is a widely used approach for this purpose. It constructs a cluster hierarchy, often visualized as a dendrogram, which is a tree-like diagram. Initially, each object or ROI is assigned to its own separate cluster. Then, in a step-by-step process, the dissimilarity distances between every pair of clusters are computed, and the two most similar clusters are merged into a new cluster. This process continues iteratively until a final cluster is formed that contains all the elements.
To calculate the dissimilarity distances, an appropriate metric is used, which measures the distance between pairs of objects or ROIs. In the context of brain connectivity, the Euclidean distance is often advantageous when working with ROIs defined in a 3-dimensional space. Using this distance metric ensures that the resulting clusters are spatially homogeneous, meaning that the ROIs within each cluster exhibit similar spatial characteristics.
Employing hierarchical clustering with the Euclidean distance as the dissimilarity measure, allows for the investigation of both local and long-range connectivity deficits in autism and provides insights into the organization of neural networks in the autistic brain. The Euclidean distance between two points $a = (a_1, a_2, \ldots, a_n)$ and $b = (b_1, b_2, ..., b_n)$ of an Euclidean n-space is given by:

$$d(a, b) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \tag{2.3}$$

When performing hierarchical clustering to compute the dissimilarity of sets, the Ward's method is often utilized. As mentioned in [116], this method is known to provide higher accuracy compared to other clustering methods and aims to minimize the variance within each cluster.
The Ward's method selects clusters to merge at each step based on optimizing an objective function. The objective is to minimize the total within-cluster variance or the sum of squares of the differences between each element and the centroid of its cluster. By minimizing the variance, the Ward's method aims to create clusters that are as compact and homogeneous as possible, ensuring that the elements

within each cluster are similar to one another. Mathematically, its formula is represented as:

$$\begin{aligned} d(C_i \cup C_j, C_k) = \quad & \frac{n_i+n_k}{n_i+n_j+n_k} d(C_i, C_k) \\ & + \frac{n_j+n_k}{n_i+n_j+n_k} d(C_j, C_k) \\ & - \frac{n_k}{n_i+n_j+n_k} d(C_i, C_j) \end{aligned} \tag{2.4}$$

Where, C_i, C_j and C_k are disjoint clusters with sizes n_i, n_j and n_k respectively. A summary of the hierarchical clustering process is represented in figure 2.2.

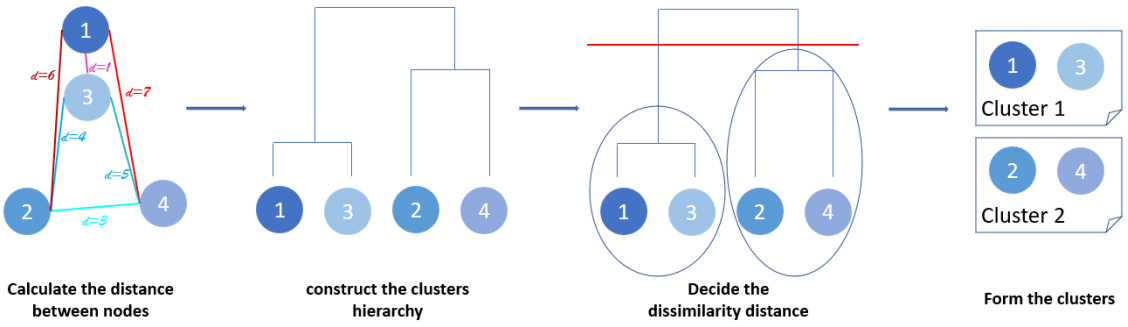

FIGURE 2.2: Summary of the hierarchical clustering process.

2.5 Classification

Classification plays a crucial role in categorizing subjects based on their features, and it can be used to evaluate the validity of theories or the effectiveness of specific methods. In our case, we employ classification for both purposes: to examine the theories regarding autistic connectivity deficits and to assess the performance of the aforementioned methods in enhancing autism detection.
In this work, there are two official groups: group 0 representing neurotypical subjects and group 1 representing individuals with ASD. The classifier used in the first part of this work is SVC, which is an implementation of Support Vector Machine (SVM) available in the scikit-learn library [117]. The objective is to approximate the decision boundary that distinguishes ASD patients from control subjects. This is done in three different spaces:

- The space of feature vectors obtained from the connectivity matrices.
- The space of features generated by applying autism-tailored methods on the connectivity matrices.
- The space of feature vectors extracted using deep learning techniques.

The classification process follows a 10-fold cross-validation procedure, where the data is divided into 10 subsets, and the classification model is trained and evaluated multiple times, with each subset serving as the validation set once.

To assess the performance of the classification, three metrics are used, namely, the accuracy, the sensitivity and the specificity [118].
The accuracy measures the overall correctness of the classification, indicating the proportion of correctly classified samples. It is defined by:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.5}$$

The sensitivity, also known as the true positive rate, quantifies the ability of the classifier to correctly identify individuals with ASD and is defined by:

$$Sensitivity = \frac{TP}{TP+FN} \tag{2.6}$$

The specificity metric, also called the true negative rate, evaluates the ability of the classifier to correctly identify neurotypical individuals using the equation:

$$Specificity = \frac{TN}{TN+FP} \tag{2.7}$$

Where:
True positive (TP) = the number of cases correctly identified as ASD.
False positive (FP) = the number of cases incorrectly identified as ASD.
True negative (TN) = the number of cases correctly identified as TDC.
False negative (FN) = the number of cases incorrectly identified as TDC.
Employing these metrics, help evaluate the effectiveness of the classification approach in distinguishing between individuals with ASD and neurotypical individuals, providing insights into the performance and applicability of the methods used for autism detection.
In the last part, we use deep learning for classification also with a 10-fold cross validation and assess it with the accuracy value.

2.6 Conclusion

In this chapter, we embarked on a comprehensive exploration of the materials and methods employed in our investigation of the Autism Spectrum Disorder through brain connectivity.
We contextualized our investigation by providing a historical background of ASD, analyzing its defining characteristics, and examining its prevalence across populations, its diagnostic methods and treatment strategies.
With the goal of finding a new mean of efficient diagnosis, we explored different brain imaging techniques. Then taking into consideration the availability of data, its fit to be harmlessly used with young children and the comparative advantage it gives in the state of art, we decided to use fMRIs extracted from ABIDE I preprocessed.
Later on, we presented the components extracted from these images, namely the connectivity matrices and graphs. Considering, the analysis of these latter can present a unique lens to explore network-level changes in the brain. Corroborating that visualizing brain connectivity matrices as graphs can

illuminate the potential disruptions in brain network properties related to ASD, furthering our understanding of the altered brain connectivity in the disorder.
Moreover, to boost the efficiency of this autism investigation, we took a new path of adapting theories and tools related to ASD, such as spanning trees and clustering methods and explained these methods in the Autism theories adapted tools section.
In conclusion, this chapter serves as a critical cornerstone of our research journey, paving the way for a deeper understanding of the intricate link between ASD and brain connectivity. Through the integration of diverse data sources and advanced methodologies, we set out in the remaining chapters to shed light on the unique neural signatures characterizing individuals with ASD.

Chapter 3

Elimination approach for autism identification

In the state of art, autism is the subject of multiple theories. Hence, to make our work autism tailored, using some of these theories can be pivotal. The objective behind investigating existing theories about autism have not only an aim to verify the veracity of these theories but also, to leverage them in order to improve the accuracy of diagnosis.
Our concern here is twofold:
On the one hand, to conduct the investigation on brain images of subjects with different backgrounds, age and other characteristics to abide by the heterogeneity resulting of real case scenarios.
On the other hand, to find and apply methods that allow to extract information related to the studied theories to efficiently model an autism adaptive system and improve the results of the study.
The most known theories about autism are the under-connectivity reflected by a lack of connection between some of the regions of the brain, and the over-connectivity theory reported in [109, 119, 120] where an excess of signal exchange results in over-communication. The over-connectivity is usually reported between regions that are physically close, and the under-connectivity between long-range regions.
In this chapter, we investigate both theories based on a novel mathematical elimination approach that corroborates the autism tailored methods.
Technically two important methods can be said to be adapted to these theories, namely, Hierarchical clustering [115] and the Spanning trees [111]. HC permits to extract what we would call local-ranged connections, i.e. connections of physically close brain regions, and long range connections, while the spanning trees permit to extract brain connections with specific criteria, with: the maximum spanning tree extracting connections with specificity of over-connectivity and the minimum spanning tree connections with specificity of under-connectivity.

3.1 The elimination approach

The concept of eliminating connections in neural networks is often related to network pruning, which aims to reduce the number of parameters and computational complexity while maintaining or improving the network performance. However, it is important to note that simply eliminating connections

based on mathematical theory without careful consideration can have reverse consequences [121].
Network pruning techniques typically involve evaluating the importance or relevance of connections using different criteria. They can help identify and remove connections that contribute less to the overall network performance [122].
Removing connections arbitrarily or without proper analysis can lead to loss of important information and degrade the model's performance. Pruning should be performed with careful consideration, validation, and possibly retraining to ensure that the pruned network still achieves the desired level of performance.
Hence, to avoid the huge loss of information that can result from eliminating all connections by only keeping the spanning trees, we decided to use a reverse approach based on the Mathematical law of double negation [123] where one can negate certain information in order to prove its veracity.
The mathematical law of double negation states that a double negation of a proposition is equivalent to the original proposition [124]. In other words, if you negate a proposition twice, you get back the original proposition.
Symbolically, if P represents a proposition, the law of double negation can be expressed as:

$$\neg(\neg P) = P \tag{3.1}$$

This law holds true in classical logic, where a proposition can be either true (T) or false (F). When you negate a proposition, you switch its truth value. Negating it again brings it back to its original truth value.
The law of double negation is often used in logic and mathematics as a tool to simplify expressions or proofs.

3.1.1 Connections elimination

As aforementioned, the theories of under-connectivity and over-connectivity in autism represent the pillar of our approach. Assuming a strong relation between over-connectivity and high connectivity values, as well as between under-connectivity and weak connectivity values, we utilize spanning trees to explore these connectivity patterns.
A spanning tree is a subgraph of a graph that includes all the vertices of the original graph, connecting them together without forming any cycles. To extract the spanning tree, we employ the Kruskal algorithm, which is a greedy algorithm that orders the connections between vertices based on their weights. It constructs the spanning tree by selecting edges from the ordered list of connections according to the desired criteria.
Since spanning trees can only be extracted from graphs and not connectivity matrices, we construct our connectivity graphs using the regions of interest (ROIs) from the chosen atlases as vertices. The connectivity values stored in the computed connectivity matrices are then used as the weights of the edges connecting these vertices. This results in the creation of undirected graphs that model the interactions between brain regions.
However, instead of using the spanning trees as descriptors or representation of the network, the uniqueness of our approach stems from the fact that we remove the spanning trees from the complex networks. The reason behind this removal is to prevent the substantial loss of information that would

occur if we were to focus solely on the spanning tree structure.
By removing the spanning trees, we aim to explore the remaining connectivity patterns in the brain, beyond the highly organized structure represented by the spanning trees. This allows us to delve deeper into the complexity of the network and investigate additional aspects of connectivity beyond the primary tree-like structure. Two main advantages are deduced from the elimination strategy:
The first resides in the alleviation of the network, which is of a great help when dealing with complex structures such as the brain connections.
The second is that the elimination helps bringing out the impact of the suppressed phenomenon (the over and under connections). The effect of the elimination will help to prove the existence of the studied theories.
The elimination process proceeds as follows:
Firstly, the spanning trees are extracted from the constructed graphs for each test. Next, the extracted STs are subtracted from the initial graphs, resulting in graphs that retain all connections between regions of interest except for the connections with the specification based on the ST criteria.
Subsequently, the matrices of these modified graphs are utilized for classification using a Support Vector Machine (SVM) with a 10-fold cross-validation. The classification quality is assessed using three metrics, namely accuracy, sensitivity, and specificity [125].
The cross-validation is a technique used to assess the performance and generalization ability of a machine learning model. It involves dividing the available dataset into multiple subsets or "folds" to train and evaluate the model iteratively. It involves multiple steps:

- Data Split: The dataset is divided into k equal-sized folds. Typically, 10, but it can vary depending on the specific application.
- Iterative Training and Evaluation: The model is trained and evaluated k times. In each iteration, one fold is used as the test set, while the remaining k-1 folds are combined to form the training set. The model is trained on the training set and then evaluated on the test set.
- Performance Metrics: Performance metrics are computed for each iteration based on the model's predictions on the test set.
- Average Performance: The performance metrics obtained from each iteration are averaged to provide an overall performance estimate of the model.

The advantage of using cross-validation stems from the fact that it helps in assessing how well a model generalizes to unseen data by simulating the process of training and testing on different data subsets. It reduces the impact of data partitioning and provides a more reliable estimate of the model's performance. Overall, cross-validation is a valuable technique for assessing the performance and robustness of machine learning models by providing more reliable estimates of their performance on unseen data.

3.2 Investigation of the overconnectivity deficit in autism

3.2.1 Data

As mentioned before, our objective is to study the brain connectivity. Hence, we compute connectivity matrices from the rs-fMRI datasets of the C-PAC pipeline. Then, construct the brain connectivity networks on which we apply our method. Since the launch of the ABIDE preprocessed, C-PAC has been the most used in the state of art. Hence, we use the time series provided for this pipeline using the AAL atlas ROIs. Then, we compute the connectivity matrices 2.3.5 that are afterward, used for classification.

First, the correlation matrices that have been the most used for brain connectivity investigation [126], [127]. They are computed by comparing the regions activities and giving a value within -1 and 1. With 1 highly correlated.

Later, the correlation matrices are compared with the tangent connectivity matrices. The tangent space matrix approach is based on the concept of Riemannian geometry [128] and is used to analyze connectivity patterns in a space called the tangent space. It involves transforming the correlation matrix into a tangent space matrix, which is a more suitable representation for capturing the geometry of the connectivity patterns. This matrix has been used previously in the context of brain analysis [129], [130] and was tested in [104] where it was proven to provide the best results for autism detection when compared to other connectivity computation methods. The figure 3.1 draws the steps needed to extract the connectivity matrices from the fMRI data.

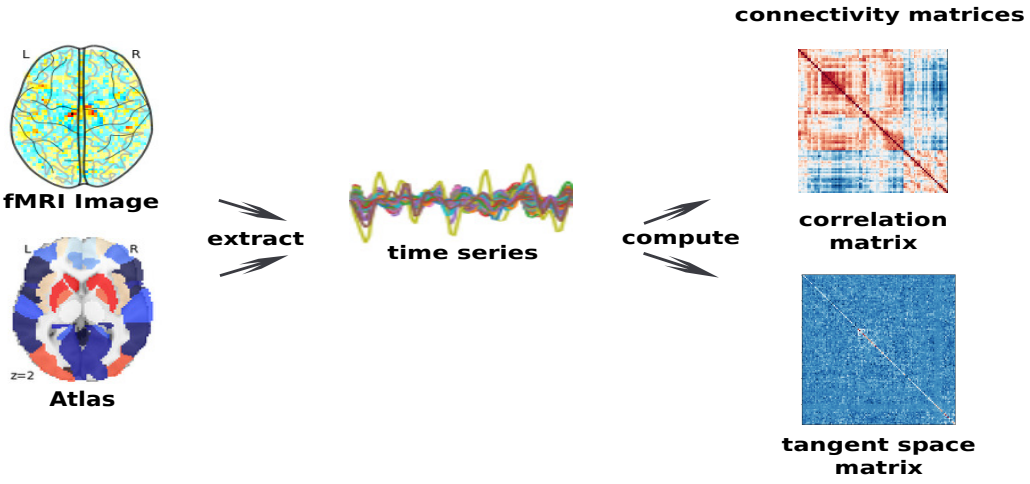

FIGURE 3.1: Connectivity matrices extraction from fMRI data.

3.2.2 Elimination of the strong connections

Overconnectivity of the autistic brain refers to the concept that individuals with autism spectrum disorder may exhibit enhanced neural connections or activity in certain regions of the brain compared to neurotypical individuals. This idea stems from various neuroimaging studies that have suggested atypical connectivity patterns in the brains of individuals with ASD.

Typically, the brain exhibits a balance between excitation (increased activity) and inhibition (decreased activity) to regulate information processing. However, in the autistic brain, there is evidence to suggest that this balance is disrupted, leading to overconnectivity in specific neural circuits.

In terms of connectivity values, overconnectivity can be understood in terms of the relative strength or magnitude of connectivity observed in specific brain regions or networks. Overconnectivity implies that the connectivity values in certain regions or networks may be higher in individuals with ASD compared to neurotypical individuals. These higher connectivity values may indicate stronger or more synchronized neural communication within these circuits.
Hence, in order to explore the potential overconnectivity deficit in the brain, it is necessary to scrutinize connections characterized by notably strong connectivity values. These robust connections can be isolated using the maximum spanning tree. Consequently, we employ the elimination approach, utilizing the MaxSt method, on the connectivity matrices derived from brain images of both autistic and non-autistic patients. These modified matrices are subsequently employed for the classification process. The steps that begins with a connectivity matrix and finish with the elimination resultant matrix, using the MaxST elimination, are showcased in figure 3.2. It is to note that we are the first

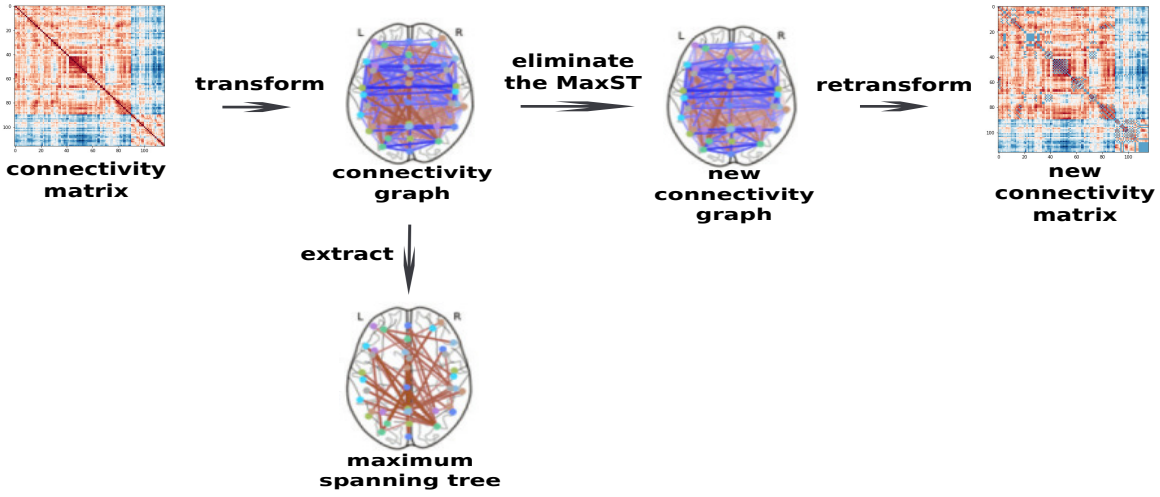

FIGURE 3.2: elimination of the maximum spanning tree from the connectivity matrices

to employ the maximum spanning tree in the context of Autism identification.

3.2.3 Elimination of the local-ranged strong connections

In Autism, there is evidence to suggest that overconnectivity may occur within certain neighboring brain regions or circuits. These areas often involve regions responsible for sensory processing and integration, as well as regions associated with attention, memory, and social cognition. To verify this theory, there is a need to create clusters of brain regions and examine the connectivity inside the clusters. To achieve this, we use the Hierarchical clustering to group the neighbouring ROIs into clusters on which we will focus our investigation. As mentioned in chapter 2.4, clusters are constructed using the Euclidean distance since it provides enforced spatial contiguity and also because we are dealing with ROIs that are defined in a 3-dimensional space. For this step, we use Scipy (a Python-based ecosystem of open-source software for mathematics) [131]. It first allocates each ROI to a separate cluster, then merges them using the equations described in chapter 2.4 till the creation of the final cluster that contains all the ROIs.
(Figure.3.3 represents the dendrogram created in this step.)

Then, a dissimilarity value is fixed to decide the clusters used for creating the local-ranged connectivity matrices that we use for classification.

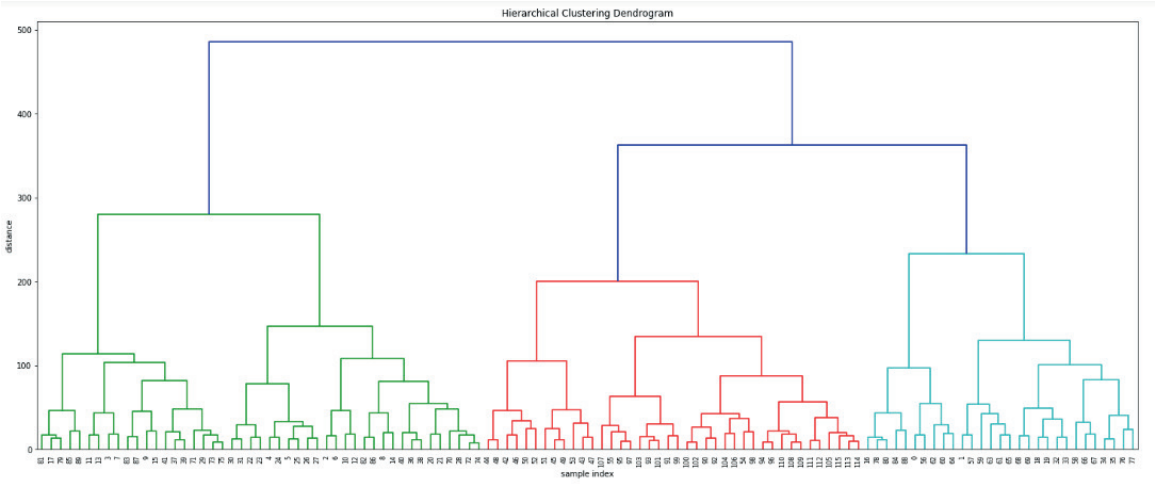

FIGURE 3.3: Dendrogram of the AAL Atlas ROIs.

The creation of the local-ranged connectivity matrices is achieved by nullifying all connections of a cluster ROIs with the other cluster ROIs, keeping only the intra-connectivity within each cluster. Figure 3.4 summarizes the steps to the creation of the local ranged connectivity matrices.

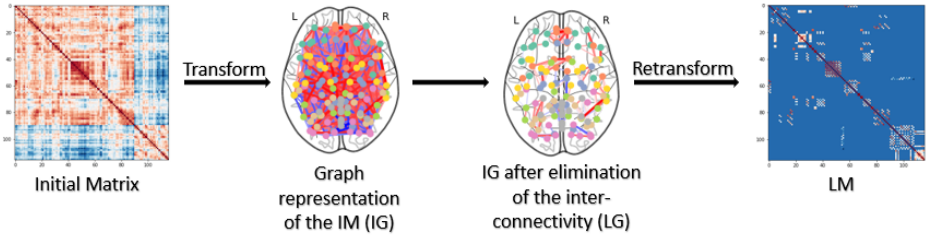

FIGURE 3.4: Summary of the creation of the local ranged connectivity matrices.

Constructing the local-ranged connectivity matrices constitute the first step to verifying the local-ranged overconnectity.

The second step consists in extracting the MaxSts of the connectivity matrices, which are a composition of all the MaxSTs of the clusters created with HC.

Then, using the elimination approach we eliminate these STs from the local-ranged connectivity matrices.

Finally, we feed the resultant matrices into the SVC classifier described in chapter 2.4.2.

Since the local-ranged overconnectivity theory suggests a deficit between neighbouring regions, examining the local-ranged matrices classification results might be insightful.

A brief description of the abbreviations used in the figures and tables of this experiment are given as follows:

- IM (Initial Matrices)

- HC-LM (Hierarchical Clustering- Local connectivity Matrices): This method extracts the ROIs clusters local connections. Then it constructs Local connectivity matrices for the classification.
- HCLM-MaxST (Hierarchical Clustering- Local connectivity Matrices Minus MaxST): In this method the MaxST is removed from the HC-LM matrices before classification. This MaxSt is a combination of the different MaxSts of every cluster intra-connectivity.

3.2.4 Results and Discussion

3.2.4.1 Experimental Results

As aforementioned, the most used preprocessing pipeline in the state of art is C-PAC. This pipeline comes with four strategies. That are:

- TrueTrue: is a strategy that includes band-pass filtering and global signal regression.
- TrueFalse: includes only the band-pass filtering.
- FalseFalse: is a strategy where neither the band-pass filtering nor the global signal regression is included.
- FalseTrue: adds the global signal regression and no band-pass filtering.

To bring out the difference these strategies can have on the analysis, we first conduct tests using the different strategies using SVC with a 10-fold cross validation and different p-values.
The choice of the p-value is important since it is a statistical measure used in hypothesis testing to determine the strength of evidence against the null hypothesis. It helps in making a decision about the null hypothesis. If the p-value is small (typically less than a predetermined significance level, often denoted as α), it suggests that the observed data is unlikely to have occurred by chance under the null hypothesis. In such cases, the null hypothesis is rejected in favor of the alternative hypothesis.
Hence, a small p-value reduces the possibility of over-fitting of the data. Thus, for classification, we use a p-value of 0.1 and 0.01 separately. Then, the classification performance is measured by the accuracy, the sensitivity and the specificity.
Tables 3.1, 3.2, 3.3 and 3.4 show performances of the classification conducted on connectivity matrices of the AAL ROIs.

From the values contained in the four tables, the first point that jump to attention is that the

TABLE 3.1: Classification results of the AAL atlas correlation matrices with a p-value=0.1

	Accuracy	**Sensitivity**	**Specificity**
TrueTrue	63	58	68
TrueFalse	63	59	67
FalseFalse	65	59	70
FalseTrue	64	60	67

tangent matrices improve the classification results when compared to the correlation. We also remark

TABLE 3.2: Classification results of the AAL atlas correlation matrices with a p-value=0.01

	Accuracy	**Sensitivity**	**Specificity**
TrueTrue	65	60	70
TrueFalse	64	59	68
FalseFalse	65	57	71
FalseTrue	66	60	71

TABLE 3.3: Classification results of the AAL atlas Tangent matrices with a p-value=0.1

	Accuracy	**Sensitivity**	**Specificity**
TrueTrue	65	59	71
TrueFalse	67	61	72
FalseFalse	66	60	70
FalseTrue	66	59	71

TABLE 3.4: Classification results of the AAL atlas Tangent matrices with a p-value=0.01

	Accuracy	**Sensitivity**	**Specificity**
TrueTrue	68	61	74
TrueFalse	69	62	75
FalseFalse	67	59	74
FalseTrue	66	57	73

that the four strategies have no significant difference in result values. Moreover, no pattern can be extracted from the different tables since no strategy keeps the lead in the tests. Concerning the p-value, it is visible that 0.01 is better. Hence, we continue the tests with the FalseFalse strategy alone where no additional preprocessing is added.

Once the strategy and p-value decided, we proceed to evaluate the proposed methods using different sub datasets extracted from this initial ABIDE dataset considering different criteria.

ABIDE is a heterogeneous database with different criteria. Thus, we decided to broaden the investigation by not only testing the elimination approach on the initial dataset but also on subdatasets extracted from the initial data as shown in table 3.5. The choice of these subdatasets considered some of the autism characteristics and brain function as the reported male dominance and handedness as well as age. As mentioned in chapter 2.1 autism has a four times increased prevalence in males over girls. Moreover, autism affects children and its manifestation is reported to be more severe in the first years of childhood. Another interesting criteria that can be considered when dealing with brain function is the handedness. Handedness is often related to hemispheric dominance, with the majority of right-handed individuals having left-hemisphere dominance and vice versa for left-handed individuals [132]. Furthermore, handedness is strongly connected to motor control, as it influences the preference for using either the right or left hand for various tasks. Motor control involves complex interactions between the brain's motor cortex, basal ganglia, and other regions. The connections and coordination

between these areas may differ depending on handedness [133].

TABLE 3.5: Sub-datasets of ABIDE used in evaluation

Sub-datasets	**cardinal**	**Selection criteria**
D1	871	All subjects
D2	596	Right handed males
D3	56	subjects with age inferior to 9
D4	12	Right handed males with age inferior to 9

For the tests, we used the aforementioned correlation and tangent connectivity matrices with the ROIs derived from the parcellation of the AAL atlas. Table 3.6 contains the results of the classification of the AAL correlation graphs and Table 3.7 the results of the classification of the AAL tangent space connectivity graphs.

TABLE 3.6: Results of the classification of the AAL correlation graphs of the subjects before and after the elimination of the Maximum spanning tree.

		Accuracy	**Sensitivity**	**Specificity**
D1	Initial matrices classification	66.51	61.25	70.98
	Classification after elimination of MaxST	66.47	60.35	71.64
D2	Initial matrices classification	62.55	54.44	69.16
	Classification after elimination of MaxST	64.12	56.27	70.63
D3	Initial matrices classification	69.58	65.14	73.46
	Classification after elimination of MaxST	70.21	66.03	73.87
D4	Initial matrices classification	65.35	58.16	70.73
	Classification after elimination of MaxST	64.58	59.27	68.74

The first visible point in table 3.6 and table 3.7 is the difference in values between the correlation and tangent matrices classification. There is a visible lead of the tangent matrices over the correlation in almost all datasets, especially the initial dataset containing all subjects.
Looking at the correlation matrices classification results, we remark that the specificity that represent the system ability to classify control subjects is always better than the sensitivity, which is the ability to correctly identify autistic subjects. The sensitivity achieved for all subjects classification is of 61.25%. After elimination of the MaxSt the value diminishes to 60.35%. This decrease in sensitivity can be interpreted to be a result of eliminating connections that might contain some autism biomarkers.
Nevertheless, when examining the results of the subdatasets, the sensitivity does not decrease; in fact,

TABLE 3.7: Results of the classification of the AAL tangent connectivity graphs of the subjects before and after the elimination of the Maximum spanning tree.

		Accuracy	**Sensitivity**	**Specificity**
D1	Initial matrices classification	68.13	61.90	73.42
	Classification after elimination of MaxST	66.95	60.13	72.79
D2	Initial matrices classification	63.88	53.43	72.14
	Classification after elimination of MaxST	63.66	51.52	73.11
D3	Initial matrices classification	67.30	61.57	72.62
	Classification after elimination of MaxST	66.98	61.28	72.43
D4	Initial matrices classification	67.04	55.59	76.52
	Classification after elimination of MaxST	65.08	54.65	73.55

it exhibits an increase. This disparity in outcomes opens the door to multiple potential conclusions. The first, is that the use of the spanning trees might not be compatible with the correlation matrices. Or in the substructed datasets the strongest connections might not be related to the overconnectivity deficit in autism. But it can also be due to the fact that the overconnectivity is not general in the autistic brain but specific to some regions.

The results of the tangent connectivity matrices classification show a slight difference in the sensitivity value of all subjects when compared to the correlation. Correlation matrices 3.6 lead to a sensitivity of 61.25%. This value increased to 61.90% when considering the tangent matrices 3.7. Added to the overall increase in accuracy of the tangent classification, we can say that the tangent computation method is more compatible with autism detection.

Looking more closely at the results of the subdatasets results, we remark a contradiction in the form of a decrease in sensitivity between the correlation and tangent matrices.

However, after elimination of the MaxSt from the tangent matrices, the sensitivity always decrases.

Going back to the stipulated theories above, we can not reach any conclusive decision about their veracity.

Using the same subdatasets described in table 3.5, we continue our investigation of the overconnectivity, going to the local-ranged overconnectivity examination.

In this part, we extract the local ranged connectivity matrices, then, we eliminate the MaxSts of every cluster of connections.

Table. 3.8 and table. 3.9 contain the results of classification of the initial matrices, the local ranged connectivity matrices and the local ranged connectivity matrices after elimination of the MaxSts, according to the three metrics using the correlation and tangent connectivity matrices, respectively.

Both tables show an undeniable humongous decrease in accuracy in the intra-connectivity matrices

TABLE 3.8: Results of the classification of initial matrices, the HC-LM and HCLM-MaxST of the correlation graphs.

		Accuracy	**Sensitivity**	**Specificity**
D1	Initial matrices classification	66.51	61.25	70.98
	HC-LM classification	60.88	49.19	70.94
	HCLM-MaxST classification	61.43	48.72	72.37
D2	Initial matrices classification	62.55	54.44	69.16
	HC-LM classification	58.34	43.52	70.34
	HCLM-MaxST classification	59.62	43.29	73.32
D3	Initial matrices classification	69.58	65.14	73.46
	HC-LM classification	64.09	54.94	72.27
	HCLM-MaxST classification	64.07	63.37	69.60
D4	Initial matrices classification	65.35	58.16	70.73
	HC-LM classification	61.63	48.88	72.73
	HCLM-MaxST classification	61.04	47.82	71.89

classification, referred to in the tables as HC-LM classification.

To this result we can give two interpretations:

First, the decrease can be due to the big loss of information resultant from the inter-connectivity nullification, (by examining the figure 3.4, we can see that the intra-connectivity matrices contain very few connections when compared to the initial matrices.)

But from another angle, taking into consideration the elimination approach theory, we can say that this decrease is a proof that the inter-connections of the brain have more conclusive autism biomarkers, hence the big decrease in accuracy and sensitivity.

The results also show that in some cases, the accuracy gets slightly better when MaxSTs are suppressed from the initial correlation matrices, compared to classification before the MaxSt elimination. Here also, the atlases shows different patterns. In the correlation matrices classification results, D1 and D2 show an improvement of the accuracy of classification after elimination of the MaxST. While in the other subdatasets the classification accuracy decreases. For the tangent the datasets showing improvement in accuracy after elimination of the MaxST are D2 and D4.

As for the sensitivity, there is also heterogeneity in the results, for both the correlation and tangent local ranged connectivity matrices classification.

TABLE 3.9: Results of the classification of initial matrices, the HC-ML and HCLM-MaxST of the tangent connectivity graphs.

		Accuracy	**Sensitivity**	**Specificity**
D1	Initial matrices classification	68.13	61.90	73.42
	HC-LM classification	60.96	51.33	69.24
	HCLM-MaxST classification	60.74	49.41	70.44
D2	Initial matrices classification	63.88	53.43	72.14
	HC-LM classification	59.67	47.25	69.62
	HCLM-MaxST classification	60.15	48.93	69.23
D3	Initial matrices classification	67.30	61.57	72.62
	HC-LM classification	60.27	53.49	66.31
	HCLM-MaxST classification	56.81	48.02	64.69
D4	Initial matrices classification	67.04	55.59	76.52
	HC-LM classification	58.02	47.02	67.21
	HCLM-MaxST classification	59.55	49.38	68.04

3.2.4.2 Discussion

In this chapter, we introduced the elimination strategy and used it with multiple settings to verify the overconnectivity deficit theories of the autistic brain. First we verified the overconnectivity using the correlation matrix as one of the most used matrices to investigate brain connectivity and the tangent space connectivity matrix which has been proven to be efficient in brain connectivity analysis. Then verified the local ranged overconnectivity of the autistic brain using a combination of methods in association with the elimination approach. To test the efficiency of the elimination method we extracted different sub-datasets with different criteria from the initial dataset that contain all subjects.
Before investigating the elimination strategy we began by comparing the different preprocessing strategies proposed with every preprocessing pipeline. The results we got from all these tests led us to different conclusions. The first, by comparing the impact of the p-value on the classification results, we were able to see an evident improvement of the classification results using a p-value of 0.01. Reposing on the explanation of the impact of the p-value introduced above, a smaller value is best fit to decrease the risk of overfitting of the data during classification. From the results achieved, this small p-value not only help in the overfitting issue, but it also improves the classification results. Second, from comparing the results of all tables we found that the preprocessing strategies impact the results

with a difference that can reach up to 3% which is quite important.
However, no strategy kept the lead in all the tests using the different connectivity matrices. In table 3.2, the FalseTrue strategy achieved the best accuracy value, but in table 3.4, it achieved the least accuracy value in comparison to the other strategies. Same goes for the TrueFalse strategy that achieved the highest value in table 3.4 and the smallest in table 3.2. Thus, we decided to continue the tests with the FalseFalse strategy that does not add anymore preprocessing other than the basic preprocessing contained in the pipeline.
When applying the elimination approach on the correlation and tangent matrices, we remarked a heterogeneity in the results of the subdatasets of the correlation matrices classification. Where only the complete dataset show a decrease in sensitivity after the elimination of the MaxST. For the other datasets the sensitivity increases.
Considering the tangent matrices classification, the elimination of the MaxST always leads to a decrease in sensitivity which imply leaving out some of the autistic subjects previously correctly classified. From these results we can conclude that the eliminated strong connections have some autism biomarkers. However, this conclusion is not general.
Going deeper in the elimination approach and eliminating inter-connections of the clusters of the regions of the brain, the results showed a huge decrease in accuracy. This decrease, can imply the loss of important autism biomarkers. But it can also be related to the big loss of information, since the eliminated connections are a lot more when compared to the remaining ones. However, the fact that the elimination of the maximum spanning tree affected the sensitivity results, implies a loss of autism related information. This suggests the presence of over-connectivity in the autistic brain. However, the fact that this loss is not consistently observed across all datasets can be interpreted as the fact that over-connectivity may not affect all regions of interest (ROIs) in the brain uniformly.

3.3 Investigation of the underconnectivity theories

3.3.1 Data

In the same light as the experiments conducted in the previous sections of this chapter, the elimination approach will be used to investigate the underconnectivity of the autistic brain. The goal is to suppress connections with specifications of the examined theories to highlight their importance.
From a technical perspective, underconnectivity in autism refers to a hypothesis that there are reduced or weakened connections between brain regions in individuals with autism spectrum disorder. To verify these conclusions, there is a need to assess and quantify brain connectivity. This process is achieved by measuring the correlation or synchronization of neural activity between the different brain regions.
Hence, in this part, we will be employing the tangent space method as a means for computing connectivity matrices. This decision is grounded in the extensive testing conducted previously. Notably, the utilization of the tangent matrices not only enhanced classification results within our study but also demonstrated similar improvements as reported in [104], thereby solidifying its effectiveness in our research context.
The experiments are conducted on the ABIDE database. To enrich the investigation even more, we

widened the experimental data by creating a group of sub datasets using multiple criteria as shown in table 3.10.

TABLE 3.10: Sub-datasets of ABIDE used in the evaluation process

Sub-datasets	**cardinal**	**Selection criteria**
D1	871	All subjects
D2	726	Males subjects
D3	719	Right handed subjects
D4	556	Subjects with age between 9 and 18
D5	56	Subjects with age inferior to 9

3.3.2 Elimination of the weak connections

As aforementioned the under-connectivity theory stipulates that some regions in the autistic brain are under-connected, or in other words, cannot communicate efficiently. Therefore, to verify this theory, we use the elimination approach to remove the weakest connections between the regions of the brain. The elimination process involves many steps beginning with transforming the connectivity matrices into graphs. Then extracting the minimum spanning trees from the initial connectivity graphs using the algorithm described in section 2.4. Followed by eliminating the constructed tree by suppressing all its edges from the original graph from which it was extracted. Finally, retransforming the graphs weights into new correlation matrices for the classification. Figure.3.5 presents all the important steps of this elimination process.

It is to note that the minimum spanning tree (MST) has been already used in the context of brain

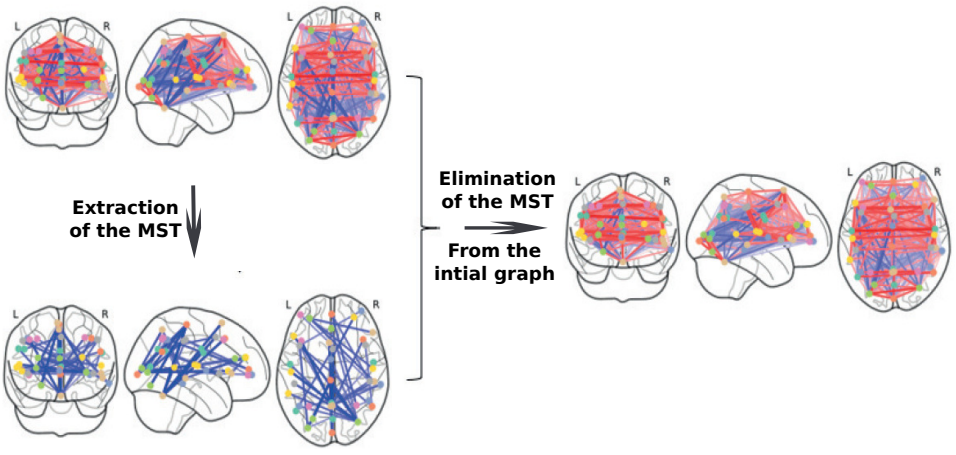

FIGURE 3.5: Elimination of the minimum spanning tree. The blue connections' weights are inferior to 0 and the red's superior to 0. The strong degrees of color represent the extremes

imaging considering other diseases as in [134] where its potential as a powerful method has been demonstrated in analyzing the structural brain network in schizophrenia. Moreover, its application greatly improved the diagnostic accuracy for Alzheimers disease [135].

3.3.3 Elimination of the Long-range connections

The investigation of long-range underconnectivity in autism can involve an elimination approach based on eliminating connections with a specific underconnectivity specification from connectivity brain maps. This approach aims to identify and analyze the disrupted or weakened connections between distant brain regions in individuals with autism spectrum disorder (ASD).
The elimination process suppresses connections extracted using the tools also described in section 2.4. It is also to note, that the approach is ASD specific since the used tools are technically tailored to autism theories. It first separates the brain connectivity map into long-range and local connectivity. Then, weak connections are eliminated from the long-range connectivity matrices to test the long-range-under-connectivity.
Figure 3.6 summarizes all the steps of the proposed approach.

The process typically involves transforming brain connectivity maps that represent the initial con-

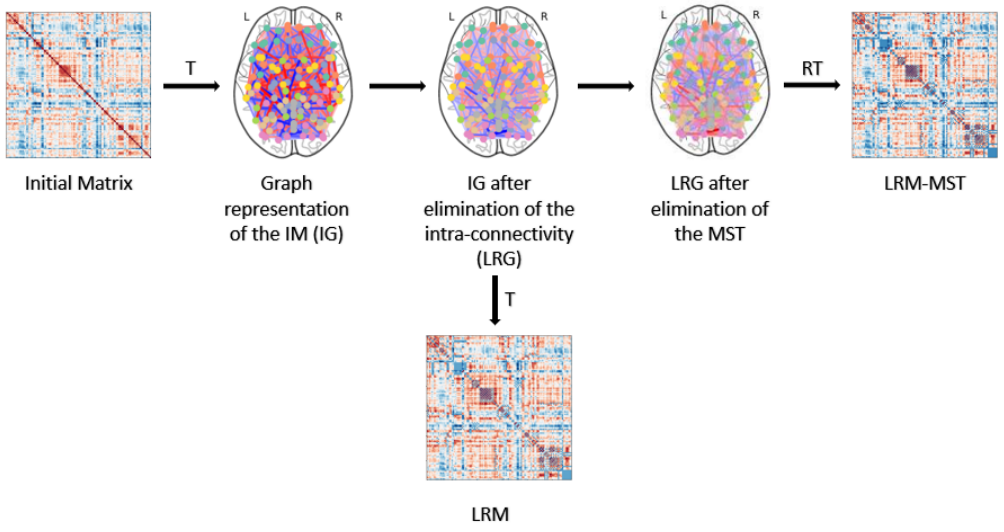

FIGURE 3.6: Summary of the proposed approach.

nectivity patterns of both the autistic group and control group into graphs. These gaphs serves as a base to apply the different tools that permit to investigate the long-range overconnectivity.
First, separating the brain connectivity map into long-range and local connectivity is a form of application of the elimination approach. In other words, to create the long range connectivity matrices, the eliminated connections are those constituting the intraconnectivity of the HC clusters.
Next, applying the elimination approach on the inter-connectivity graphs (the initial graphs after nullifying the inta-connections), where we systematically eliminate connections from the connectivity brain maps based on the inter-connectivity graphs MSTs, results in the final graphs representing the long-range connectivity minus the strongest connections.
By eliminating the connections that meet the underconnectivity criterion, we were able to generate modified connectivity brain maps that highlight the disrupted long-range connections associated with autism. These modified maps allow for a focused analysis of the remaining connections and their potential implications for the observed cognitive and behavioral characteristics in individuals with ASD.

This approach can provide insights into the specific patterns of underconnectivity in the brains of individuals with autism, helping to identify regions or networks that may be critical for understanding the disorder. It may also contribute to a better understanding of the neurobiological basis of autism and guide the development of targeted interventions and treatments.
In the same light as for the creation of the intraconnectivity matrices, the technical steps of this approach, begin firstly by creating the clusters of the AAL atlas ROIs using Scipy. Then, a dissimilarity value is fixed to decide the clusters, before extracting the connections outside these clusters. In other words by eliminating the inta-connections we construct the long-range connectivity matrices that we use for classification.
Depending on the value of the dissimilarity measure, the number of ROIs in the clusters changes. This can be seen in Figure 3.3. Therefore, multiple dissimilarity measures have been tested to bring out their impact on the results. Their values and results are reported below in the experimental results section.

3.3.4 Results and Discussion

3.3.4.1 Experimental results

In this chapter, we investigate the connectivity patterns of individuals with autism in a highly diverse setting, using large test datasets. Our approach involves applying various methods to the entire ABIDE I dataset to examine the overall connectivity of autistic brains. Additionally, we analyze different sub-datasets to assess the influence of specific criteria such as age, gender, and handedness on the results as provided in Table 3.10.
Furthermore, to highlight the effectiveness of this approach, we compare its performance with that of the most influential state-of-the-art approaches, as described in [136] and [104].
Hence, we end up with multiple classifications that we summarize as follows:

- Nielsen and al.: Where a leave-one-subject-out cross-validation approach is applied for classification using the correlation matrices of the ABIDE subjects.
- Abraham and al.: utilizes a leave-one-site-out cross-validation on participant-specific functional connectivity matrices (connectomes).
- IM (Initial Matrices)[Proposed]: Similar to Abraham et al., this classification matrices are computed using the regions of interest (ROIs) from the AAL atlas.
- M-MST (Minus Minimum Spanning Tree)[Proposed]: In this classification, the minimum spanning trees (MST) are removed from the initial matrices obtained in the IM method.
- HC-LRM (Hierarchical Clustering- Long Range connectivity Matrices)[Proposed]: Employs hierarchical clustering on the AAL ROIs to form clusters. It then extracts long-range connections from the initial matrices and constructs long-range connectivity matrices for classification.
- HC-LM (Hierarchical Clustering- Local connectivity Matrices)[Proposed]: Is comparable to HC-LRM, however, here it extracts clusters local connections. Then it constructs Local connectivity matrices.

- HCM-MST (Hierarchical Clustering Minus MST)[Proposed]: Uses the same strategy as M-MST, where MST is removed from the HC-LRM matrices before classification.

The initial method of the proposed approach, called M-MST, involves removing the connections extracted from the minimum spanning tree (MST) from the initial connectivity matrices. After performing this elimination, we observe a slight increase in the accuracy values for sub-datasets D1, D2, and D5. However, in sub-datasets D3 and D4, the accuracy decreases.
Interestingly, a noteworthy finding emerges from these results. The sensitivity decreases across all datasets, regardless of the change in accuracy. This indicates that the connections removed from the MST contain valuable information regarding the connectivity deficits in the autistic brain.
These findings are summarized in Table 3.11, which presents the classification results of the initial matrices compared to the results after eliminating the MST values. The table provides insights into the impact of removing the MST connections on the classification performance, highlighting the trade-off between accuracy and sensitivity in different sub-datasets. In the second method, where we classify

TABLE 3.11: Classification results of the subjects before and after eliminating the minimum spanning tree. IM here refers to the initial matrices and M-MST the matrices after elimination of the MST

		Accuracy	**Sensitivity**	**Specificity**
D1	IM	68.13	61.90	73.42
	M-MST	68.78	60.69	75.60
D2	IM	66.33	62.82	69.70
	M-MST	67.48	62.45	72.30
D3	IM	66.21	53.86	75.80
	M-MST	65.93	50.71	77.89
D4	IM	67.31	61.58	72.62
	M-MST	66.76	60.03	72.83
D5	IM	53.63	75.42	37.00
	M-MST	67.62	65.42	65.50

the HC-LRM (Hierarchical Clustering - Long Range connectivity Matrices), we divide the brain into multiple clusters. The hierarchical clustering technique allows us to create distinct groups of clusters based on a chosen dissimilarity measure. Through empirical testing of different dissimilarity values, we determine the most suitable value to extract different clusters of regions of interest (ROIs).
Using these clusters, we construct both long-range and local connectivity matrices for the first dataset, D1, and perform separate classifications. The classification performance metrics for this method are presented in Table 3.12. Notably, we observe an increased accuracy in the classification of inter-connectivity matrices when utilizing dissimilarity values of 80, 90, and 100, with an average of 6-8 ROIs per cluster.
Following tests on the other datasets, we find that a dissimilarity value of 80 remains the most appropriate choice for tuning the clustering algorithm. This value consistently yields optimal classification results, highlighting the importance of selecting an appropriate dissimilarity threshold for effective cluster formation and subsequent analysis.
One interesting observation from Table 3.12 is that the inter-connectivity matrices (HC-LRM) exhibit

TABLE 3.12: Accuracy results of inter- and intra-connectivity on D1 using different values of the dissimilarity distance. HC-LRM are the long-range connectivity matrices and HC-LM the local connectivity matrices

d	**50**	**80**	**90**	**100**	**150**	**250**	**300**
HC-LRM	68.90	69.39	69.34	69.61	67.51	66.94	64.86
HC-LM	61.09	60.96	61.43	62.01	63.09	63.04	65.70

higher accuracy results compared to the intra-connectivity matrices (HC-LM). In fact, the highest accuracy achieved for intra-connectivity classification in the entire dataset (D1) is 65.09%, which was obtained using a dissimilarity distance of 300 (resulting in 3 clusters of ROIs). However, even with this value, the accuracy remains lower than the classification accuracy without clustering, which is 68.13%. Moreover, the dissimilarity distance of 80 achieved the second best accuracy in the HC-LRM classification but the lowest in HC-LM classification but from the elimination approach view angle, this means that the eliminated connections contain more autistic biomarkers. Hence, moving forward, the results are obtained from clusters constructed using a dissimilarity value of 80.
In Table 3.13 the results are based on the classification of inter-connectivity matrices. These results demonstrate an improvement in all three classification metrics (accuracy, sensitivity, and specificity) when compared to the initial matrices across all datasets. This improvement in classification performance even with different dataset criteria underscores the significance of eliminating local connections within the clusters. This finding provides further support for the presence of communication deficits, particularly long-range deficits, in the autistic brain. The results presented in Table 3.11 and Table

TABLE 3.13: Results of the inter-connectivity matrices classification compared to the initial matrices classification. IM are the initial matrices and HC-LRM the long-range matrices

		Accuracy	**Sensitivity**	**Specificity**
D1	IM	68.13	61.90	73.42
	HC-LRM	69.40	63.07	74.78
D2	IM	66.33	62.82	69.70
	HC-LRM	67.24	63.65	70.73
D3	IM	66.21	53.86	75.80
	HC-LRM	68.81	56.63	78.34
D4	IM	67.31	61.58	72.62
	HC-LRM	69.88	63.45	75.71
D5	IM	53.63	75.42	37.00
	HC-LRM	56.41	78.33	39.00

3.13 allow us to draw several conclusions. Firstly, both the elimination of the minimum spanning tree (MST) and the removal of the local connections within the hierarchical clustering (HC) clusters have an impact on the classification process. When these two approaches are combined, i.e., eliminating the local connections within the clusters and then removing the MST from the resulting inter-cluster connections, a similar decrease in sensitivity is observed compared to the results in Table 3.11 after removing the MST from the initial matrices.

TABLE 3.14: Classification results of the inter-connectivity matrices before and after eliminating the spanning tree. Here HC-LRM are the long range connectivity matrices and HCM-MST the long range matrices after eliminating the MST

		Accuracy	**Sensitivity**	**Specificity**
D1	HC-LRM	69.40	63.07	74.78
	HCM-MST	69.76	61.06	77.11
D2	HC-LRM	67.24	63.65	70.73
	HCM-MST	67.80	62.71	72.64
D3	HC-LRM	68.81	56.63	78.34
	HCM-MST	68.42	54.94	78.98
D4	HC-LRM	69.88	63.45	75.71
	HCM-MST	69.88	60.46	78.45
D5	HC-LRM	56.41	78.33	39.00
	HCM-MST	61.28	64.17	59.50

This finding indicates that the removal of the minimum spanning tree directly affects the sensitivity metric, which represents the ability to predict autism. It suggests that the weakest connections identified by the MST contain valuable information about autism, and their elimination results in a reduction in sensitivity. Consequently, the sensitivity consistently decreases when the MST is removed from the connectivity matrices.

Table 3.14 provides a summary of these results, highlighting the impact of removing the minimum spanning tree and the local connections on the sensitivity metric. The table underscores the importance of considering the weaker connections and their role in capturing important information related to autism in connectivity matrices.

By comparing our approach's methods with the state-of-the-art approaches applied to ABIDE I, as presented in Table 3.15, we observe an increase in the accuracy values. Our approach achieves an accuracy of 69.76%, which is 3% higher than the best result reported in [104] where the highest accuracy achieved was 66.80%. This improvement in accuracy highlights the effectiveness of our approach in autism classification.

Furthermore, when examining the sensitivity metric in the long-range connectivity classification, our approach outperforms the highest result achieved in [136]. This demonstrates the impact and relevance of our proposed method in capturing the connectivity deficits in the autistic brain, particularly in terms of long-range connections.

Importantly, the significance of eliminating the minimum spanning tree (MST) is underscored when comparing the sensitivity values across all methods. The sensitivity consistently decreases after removing the MST, emphasizing the valuable information contained within the weaker connections captured by the MST.

These findings demonstrate the enhanced performance of our approach compared to the state-of-the-art methods applied to ABIDE I, specifically in terms of increased accuracy and sensitivity, and the informative role of the MST removal.

TABLE 3.15: Classification performance of the proposed approach compared to state-of-the-art methods applied to ABIDE I

	Accuracy	Sensitivity	Specificity
HCM-MST	69.76	61.06	77.11
HC-LRM	69.40	63.07	74.78
M-MST	68.78	60.69	75.60
Abraham and al.	66.8	61.0	72.3
Nielsen and al.	60.0	62.0	58.0

3.3.4.2 Discussion

In this section, we conducted a comprehensive study on the connectivity of the autistic brain using a highly heterogeneous dataset from ABIDE I. Given that autism primarily affects children during early childhood, we sought to identify the most common and non-invasive imaging acquisition methods capable of extracting brain connectivity without causing harm. We aimed to utilize an imaging method that is readily available and cost-effective compared to other techniques.

Using ABIDE data comes with the challenge of uncontrolled heterogeneity due to the multi-site fMRI data acquisition, which can influence the results. It is important to note that uncontrolled variations can arise across sites, such as differences in scanner type, pulse sequence, and sample composition. However, considering our objective of simplifying the process of autism detection and the fact that heterogeneity is inherent in real-world settings, we chose to use the datasets with minimal changes. Only images with quality issues were eliminated to ensure the reliability of the analysis. This approach allowed us to address the challenges posed by the uncontrolled heterogeneity and potentially provide more robust findings.

Our primary objective was twofold: to investigate the theories surrounding autism spectrum disorder (ASD) and to improve the detection process of autism. To achieve this, our approach involved selectively removing brain connections that were associated with specific theories. We employed a combination of specific tools to extract and analyze these connections. The use of the minimum spanning tree (MST) was particularly significant, as it enabled us to simplify the complex network into a more straightforward representation based on connection weights. Moreover, the MST aligned with the under-connectivity theory of the autistic brain, as it emphasized the extraction of the weakest connections.

By employing these methods, we aimed to contribute to a better understanding of the underlying connectivity patterns in autism and enhance the accuracy of autism detection. By selectively removing specific connections, we were able to focus on the key connectivity deficits associated with ASD and potentially uncover novel insights into the disorder.

Additionally, we incorporated hierarchical clustering to divide the brain regions into clusters, enabling us to examine both intra- and inter-connectivity within the brain, similar to the approach in [109]. The use of clustering allowed us to explore deficits in ASD brain connectivity and better understand the organization of connections within and between clusters.

Consequently, we tested multiple ASD-specific methods and combined them to enhance the accuracy of ASD detection. The results demonstrated an accuracy of 69.76% on the whole dataset D1, as shown

in Table 3.14, surpassing previously published findings in ABIDE [136], [104].
The elimination of the minimum spanning tree resulted in a decrease in sensitivity across all datasets. This reduction suggests that the weak connections identified by the minimum spanning tree contain valuable biomarkers related to autism and provides support for the under-connectivity theories of autism. This outcome is highlighted in Table 3.11.
Conversely, the elimination of intra-connectivity, specifically the removal of local connections within the clusters constructed by hierarchical clustering, led to improvements in both accuracy and sensitivity. Sensitivity increased from 61.9% to 63.07%, indicating a deficit in long-range connectivity in individuals with ASD. These results can be observed in Table 3.13, which showcases the elimination of intra-connectivity across all datasets.
The choice of dissimilarity distance used to construct the clusters also emerged as an important factor, as different values yielded varying degrees of prediction accuracy and sensitivity. Optimal selection of the dissimilarity distance greatly influenced the performance of the classification, as demonstrated in Table 3.12.
When the minimum spanning tree was once again eliminated from the long-range connectivity matrices, we observed the same findings as reported in Table 3.11, which confirmed the decrease in sensitivity after removing the weak connections extracted from the inter-connectivity matrices. This provides further evidence for the under-connectivity deficit in autism, particularly in the long-range connections between brain regions. Therefore, we can confidently conclude the existence of a long-range under-connectivity deficit in the autistic brain.
Moreover, by eliminating inter-connectivity and the minimum spanning tree, we were able to highlight the long-range under-connectivity deficit and improve the results compared to existing approaches, as depicted in Table 3.15.

3.4 Conclusion

In this chapter we investigated the autistic brain connectivity using an approach of elimination.
The analysis of classification approaches involving the elimination of MaxST and inter-connections among brain regions reveals intriguing insights into Autism Spectrum Disorder. When considering tangent classification, removing MaxST consistently results in reduced sensitivity, indicating the exclusion of correctly classified autistic subjects. This suggests the presence of autism biomarkers associated with the strong connections that are eliminated. However, this conclusion might not universally apply.
Delving further into the elimination strategy by targeting inter-connections within brain region clusters, a significant drop in accuracy is observed. This decline could signify the loss of vital autism biomarkers, but it could also stem from the considerable loss of information due to the large number of removed connections compared to the retained ones.
The noteworthy observation is that the elimination of the maximum spanning tree influences sensitivity results, pointing to the forfeiture of autism-related information. This implies a potential over-connectivity pattern in the brains of individuals with autism. Nonetheless, the inconsistency in this loss across different datasets suggests that the impact of over-connectivity might not uniformly affect all regions of interest in the brain. Thus, while patterns of connectivity alteration are evident

in ASD, the precise nature of these changes appears to be context-dependent and potentially varied among different brain regions.

The second investigation, concerning the long range and underconnectivity, offers compelling support for the existence of impaired long-range connections between different regions of the autistic brain. Remarkably, this deficit remains consistent across various factors such as age, gender, and handedness, reinforcing the reliability of our findings. The uniformity of this discovery across different subsets of data underscores the robustness of our analytical approach.

With an achieved accuracy rate of 70%, our classification method demonstrates its potential for accurate prediction of autism. Additionally, our innovative strategy, involving the elimination of inter-connectivity and the minimum spanning tree, has allowed us to pinpoint and enhance our understanding of the long-range under-connectivity deficit in autism.

Chapter 4

Machine and deep learning combination for ASD identification

In Deep learning, we go beyond what is visible to extract new hidden information. It can be achieved through different networks as Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN). All share the same testing structure with a learning phase where the system matches information from multiple data to extract new markers, and a test phase where the performance of the network is evaluated. Convolutional Neural Networks are a particular type of Artificial Neural Networks that comprises many types of layers. Their use in image analysis, can uncover hidden information and enhance the success of categorizing in classification. This made them very popular especially in the medical field [137]. For ASD detection, various approaches have been explored. One notable study by [138] applied deep learning algorithms to identify patients from large datasets based solely on brain activation patterns. They achieved an accuracy of 70% and discovered an anticorrelation of brain function between anterior and posterior areas in individuals with ASD.

Another study by [109] utilized a multi-atlas functional connectivity approach for ASD detection. They calculated multiple functional connectivity measures based on different brain atlases from the ABIDE fMRI database. To extract discriminative features, they proposed a multi-atlas deep feature representation method using stacked denoising autoencoders. The final identification was performed using a multilayer perceptron and ensemble learning, resulting in an accuracy of 74.52%, sensitivity of 80.69%, specificity of 66.71%.

However, it is important to note that many existing approaches in the literature do not explicitly incorporate the known theories about autistic brains into the detection process. They may not consider the specific characteristics and connectivity deficits associated with ASD. In contrast, basing our knowledge on the findings of the previous investigation, we take into account the theories of under and over-connectivity in autism, providing a theoretically adapted framework for ASD detection.

Hence, in this chapter, in the same light as in the previous chapter 3, we focus on the autistic connectivity deficit theories. After proving their existence through the elimination approach and the negation theory, we study them in a new context coupled with deep learning with a three-fold objective:

- 1- To address the issue of early diagnosis based on resting state fMRI images of subjects with different backgrounds, ages and other characteristics.
- 2- To consolidate the finding of the elimination approach concerning both the over-connectivity, and under-connectivity.
- 3- To improve the ASD identification by using recent deep learning techniques and by incorporating the autism theories into the process of detection.

For this, we propose an ASD adaptive transfer learning approach that identifies the disorder and tests different atlases, different connectivity matrices and different preprocessing strategies separately. Transfer learning permits to reduce the processing time, because the weights the network uses are transferred from other systems that have been trained on huge data and given the best possible results with minimum errors. Here, the model used for transfer learning is a CNN based on Tensorflow [139] proposed by Keras [140]. We simulate the 3D RGB format which is the base entry of the CNN model through 3D connectivity matrices that we compose by dividing the original connectivity matrices extracted from fMRI time series via two approaches:

- The first divides the connectivity matrices information by 3 and puts every third of the information on a layer.
- The second is based on the aforementioned maximum and minimum spanning trees. In this approach, the first layer incorporates the original connectivity matrix minus the highest and lowest weights extracted using the maximum and minimum spanning trees, respectively. The second layer contains the high weights, while the third layer contains the lowest weights (under-connectivity information).

4.1 Machine and deep learning techniques

4.1.1 Machine learning

Machine learning is a subset of artificial intelligence (AI) that focuses on the development of algorithms and statistical models that enable computer systems to learn from data and improve their performance on a specific task without being explicitly programmed. Their goal is to allow computers to identify patterns, make predictions, or take actions based on the data they have been exposed to.
Machine learning can be categorized into three main categories:

- Supervised Learning, where the algorithm is trained on a labeled dataset, where the input data and the corresponding correct output are provided. The goal is for the algorithm to learn a mapping function from the input to the output so that it can make accurate predictions on new, unseen data. Common supervised learning tasks include classification (assigning inputs to predefined classes) and regression (predicting continuous values).
- Unsupervised learning that involves training the algorithm on an unlabeled dataset, where the input data is not labeled with the correct output. The algorithm's objective is to discover patterns, structures, or representations within the data without specific guidance. Clustering,

where the algorithm groups similar data points together, and dimensionality reduction, where the algorithm reduces the data's complexity while retaining essential information, are examples of unsupervised learning tasks.

- Reinforcement learning which is concerned with training agents to interact with an environment and learn from the feedback it receives in the form of rewards or penalties. The agent's goal is to maximize cumulative rewards over time, leading to better decision-making and actions in the given environment. Reinforcement learning has been widely used in various applications, including game playing, robotics, and autonomous vehicles.

The machine learning process typically involves the following steps:

- Data Collection: Gathering relevant and representative data that will be used for training and evaluating the model. The quality and quantity of the data are crucial factors in the success of a machine learning model.
- Data Preprocessing: Cleaning, transforming, and preparing the data to be in a suitable format for the learning algorithm. This step may involve dealing with missing values, normalizing features, or encoding categorical variables.
- Model Selection: Choosing an appropriate algorithm or model architecture that best suits the problem at hand and the nature of the data. Different algorithms have different strengths and are suitable for different types of tasks.
- Training: In this step, the model is fed with the training data to learn the underlying patterns and relationships. The learning process typically involves adjusting the model's parameters iteratively to minimize errors or maximize performance.
- Evaluation: The trained model is evaluated on a separate dataset (the test set) to assess its performance and generalization ability. Various metrics are used to measure how well the model performs on unseen data.
- Deployment: After successful training and evaluation, the model is deployed into production to make predictions or take actions in real-world scenarios.

In the case of our data, the algorithm best fit is the Support Vector Machines that falls under the category of supervised machine learning.

4.1.2 Deep learning

Deep learning is a specialized subset of machine learning that focuses on training artificial neural networks to mimic the human brain's structure and function. It is a powerful and advanced form of machine learning that has gained significant popularity and success in various domains, particularly in areas that involve complex patterns and large amounts of data, such as computer vision, natural language processing, and speech recognition.
At the core of deep learning are artificial neural networks, which are inspired by the interconnected neurons in the human brain. These networks are composed of multiple layers of interconnected nodes

(artificial neurons) that work together to process data and extract patterns. Each layer progressively refines the representations of the input data, leading to hierarchical feature learning, allowing the network to capture intricate relationships and abstract features.
Key concepts and components of deep learning include:

- Neural Networks: Deep learning models are built using neural networks, which consist of input, hidden, and output layers. The hidden layers, also known as "deep" layers, give deep learning its name. The connections between neurons in different layers are represented by weights that are learned during the training process.
- Backpropagation: Deep learning models use a technique called backpropagation to adjust the network's weights during training. It involves computing the gradients of the model's error with respect to each weight and using these gradients to update the weights, thereby minimizing the error and improving the model's performance.
- Activation Functions: Activation functions introduce non-linearity into the neural network, allowing it to approximate complex functions. Common activation functions include sigmoid, tanh, and rectified linear units (ReLU).
- Convolutional Neural Networks (CNNs): CNNs are a type of deep neural network specifically designed for processing grid-like data, such as images and videos. They use convolutional layers to automatically learn local patterns and spatial hierarchies in the data.
- Recurrent Neural Networks (RNNs): RNNs are designed to process sequential data, such as time series or natural language. They have recurrent connections that allow information to persist across time steps, making them suitable for tasks like speech recognition and language translation.
- Long Short-Term Memory (LSTM): LSTM is a type of RNN that addresses the vanishing gradient problem, allowing it to capture long-term dependencies in sequential data.
- Generative Adversarial Networks (GANs): GANs are a special type of deep learning model consisting of two neural networks, a generator, and a discriminator, which compete against each other in a game-theoretic setting. GANs have been successful in generating realistic data, such as images, and have various creative applications.

Deep learning has achieved remarkable breakthroughs in various fields, including computer vision (e.g., image and object recognition), natural language processing (e.g., language translation and sentiment analysis), autonomous vehicles, drug discovery, and recommendation systems. Its ability to automatically learn and extract complex patterns from data has made it one of the most exciting and promising areas of artificial intelligence research. However, deep learning models often require a significant amount of data and computational resources for training, making them more computationally demanding than traditional machine learning algorithms.

4.1.3 Combining machine and deep learning

Combining machine learning and deep learning techniques is a common and effective approach to solving complex problems in the field of artificial intelligence. Both machine learning and deep learning are subsets of AI, but they have different strengths and weaknesses. By integrating the two, we can leverage the best of both worlds and achieve more robust and accurate models for various tasks.
This combination can take multiple forms as:

- Feature Engineering: Machine learning often relies on handcrafted feature engineering to extract relevant information from the raw data. Deep learning, on the other hand, can automatically learn relevant features from the data. By combining the two, you can use deep learning to automatically extract useful features, which can then be fed into a traditional machine learning algorithm for further processing and decision-making.

- Transfer Learning: Transfer learning is a powerful technique in deep learning where pre-trained neural network models on large datasets are used as a starting point for solving a new task. By fine-tuning these pre-trained models with smaller amounts of task-specific data, you can achieve better results than training from scratch. Transfer learning can be combined with various traditional machine learning algorithms to improve their performance.

- Ensemble Methods: Ensemble methods, such as stacking, bagging, and boosting, combine multiple machine learning or deep learning models to make more accurate predictions. For instance, you can stack together different deep learning models and traditional machine learning algorithms to create a more powerful ensemble model that leverages their individual strengths.

- Data Augmentation: Data augmentation is commonly used in deep learning to artificially expand the training dataset by applying transformations to the existing data, such as flipping, rotating, or cropping images. These augmented datasets can then be used to train a machine learning model, leading to improved generalization and performance.

- Hybrid Models: You can design hybrid models that combine both deep learning and traditional machine learning layers. For example, you could use a deep neural network as a feature extractor and then use those extracted features as inputs to a machine learning algorithm, such as support vector machines (SVMs) or random forests, for the final prediction.

- Improve Training with Machine Learning: Machine learning techniques can be used to optimize the hyperparameters of deep learning models, helping to speed up convergence and improve generalization. Additionally, machine learning can assist in handling imbalanced datasets or generating synthetic data to supplement small training sets.

- Active Learning: Active learning is a machine learning approach that uses human oracles to label a small subset of the most informative data points, which are then used to train the model. You can apply active learning to deep learning models to make the data labeling process more efficient, as labeling large datasets for deep learning can be time-consuming and expensive.

By combining machine learning and deep learning techniques, powerful AI systems that address a wide range of real-world problems can be created, making the most of both methodologies' unique

strengths.
In brain anomalies detection combining machine learning and deep learning techniques is a promising approach that can leverage the strengths of both methodologies to achieve more accurate and robust results.
While deep learning models can achieve high accuracy, they are often considered black boxes, making it challenging to understand their decision-making process. Integrating interpretable machine learning techniques, can help provide insights into the features contributing to the anomaly detection. Hence, Deep Learning can be used for Feature Extraction and machine learning for classifying these features. Deep learning models, especially Convolutional Neural Networks (CNNs), have shown remarkable success in image feature extraction. However, the drawback of using deep learning manifest in the size of the data. A database with limited number of samples can not be used for deep learning. The solustion is using Pre-trained CNNs, such as VGG, ResNet, or Inception, through transfer learning. These models have been trained on large image datasets, and their convolutional layers have learned to detect hierarchical features that are useful for various image recognition tasks.

4.2 Data preparation

Continuing our investigation of the autism deficits, we use the preprocessed version of the ABIDE I dataset after elimination of the defective data. As mentioned before, the most used pipeline in autism detection is the C-PAC which is also the pipeline used in this work, considering the FalseFalse strategy.
For parcellation, in addition to the AAL atlas, we use two other atlases, namely, the DosenBach160 with 161 ROIs and the CC200 atlas that divide the brain into 200 regions.
For the connectivity matrices, in addition to the correlation matrices and the tangent space matrices, we also test the covariance matrices that we compute using the tools proposed by Nilearn [141].
Finally, we construct the connectivity graphs based on the information stored in the connectivity matrices. The graphs constructed take for nodes the regions of the used atlas and for edges the connectivity weights of the connectivity matrices.

4.2.1 Extraction of the strongest and weakest brain connections

As mentioned before, the spanning tree is considered as an unbiased graph theory tool that simplifies high-order network structure while preserving its core framework. A graph can have many spanning trees, however, when a condition is met, we can extract a specific tree as the widely known minimum spanning tree (MST). The MST permits to extract the brain connections that have the weakest connections. As aforementioned, these latter can be related to the underconnectivity deficit, and through chapter 3, where an elimination approach was used, the connections extracted using this tree proved their importance.
The other spanning tree used also in this chapter is the maximum spanning tree (MaxST). The MaxST permits to extract the brain connections that have the highest correlation values. Again, these latter can be related to the overconnectivity. Both spanning trees are subgraphs that are trees which include all of the vertices of the graph they were applied to, and connect them together without any cycles.

To extract our spanning trees, we use the Kruskal algorithm which orders the vertices connections according to their weight. Then, it constructs the spanning tree by extracting information from the ordered vertices connections list. The important steps of this extraction are illustrated in figure 4.1.

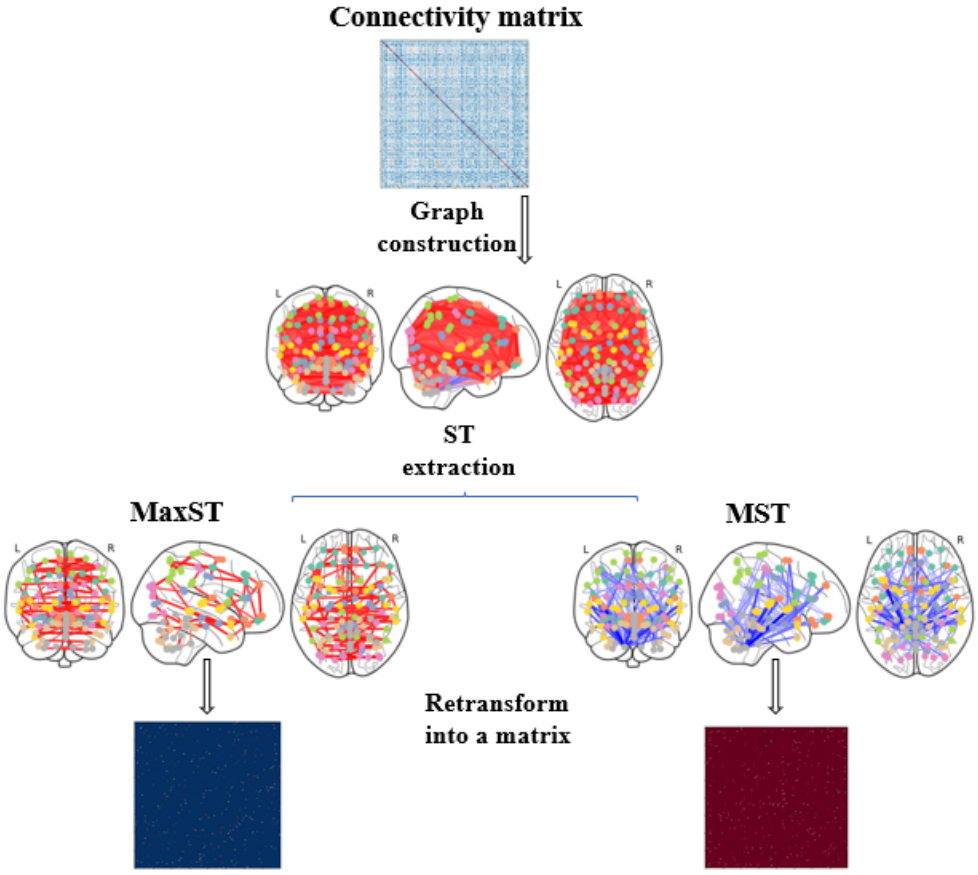

FIGURE 4.1: Steps to extract the Spanning Trees matrices from a connectivity matrix.

4.2.2 RGB images alike matrices construction

In this chapter, we propose a new approach, that consists of constructing RGB images alike matrices to use as base entry for deep learning features extraction. These matrices are constructed by dividing the connectivity matrix into three layers:

- Strong weighted connections layer.
- Weak connections layer.
- Remaining connections of the connectivity matrix.

Our method introduces a creative approach to convert matrices into visually informative RGB-like images. By leveraging the inherent structure of matrices, we layer the matrix information onto three color channels, namely Red, Green, and Blue. Each channel represents specific features or characteristics of the data within the matrix. For instance, the Red channel signifies the strong

weighted connectivity patterns, the Green channel encodes the weak connectivity information, and the Blue channel captures the remaining matrix properties as shown in Figure 4.2.
By mapping the matrix elements to appropriate color intensities, we create RGB-alike images that visually represent the underlying matrix data. This technique proves particularly useful in visualizing complex patterns into an image, enabling the use of CNN.
Furthermore, the method is versatile and applicable with numerous variations. Hence, in addition to

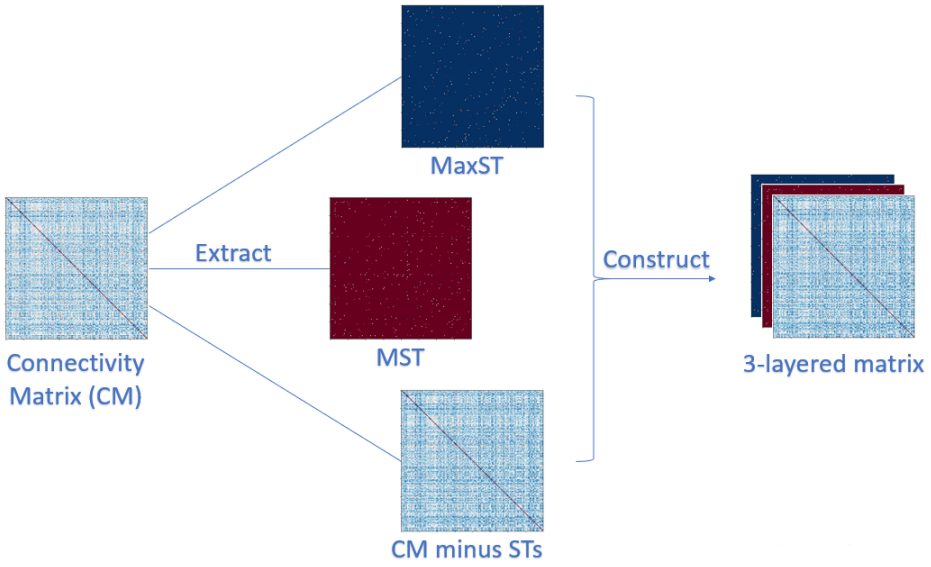

FIGURE 4.2: 3D Connectivity matrices construction with the Spanning trees strategy.

the RGB-image alike based on the MaxSt and MST, we also construct RGB-image alike images by simply dividing the information of the whole connectivity matrix by three and putting every third of the information on a layer as shown in Figure 4.3.

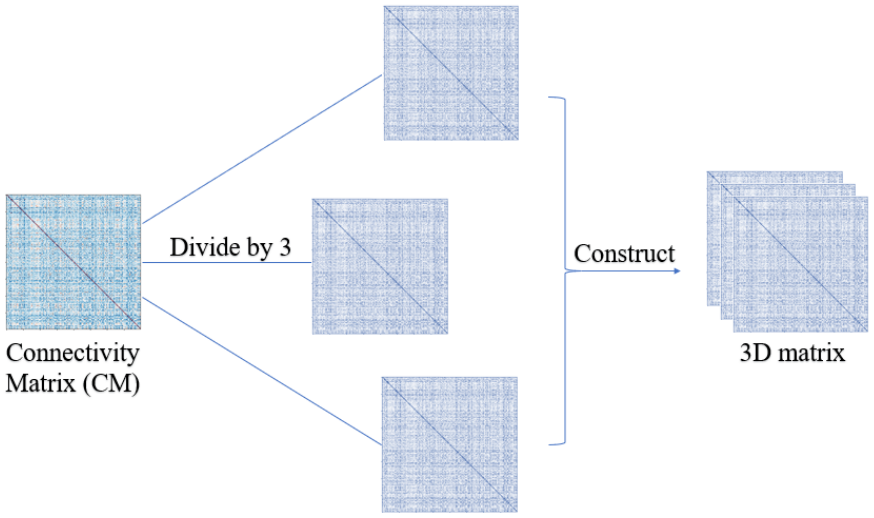

FIGURE 4.3: 3D Connectivity matrices construction with the division strategy.

4.3 Transfer learning for feature extraction

4.3.1 Transfer learning

In deep learning, to get a high accuracy with little error, the deep learning models have to train on huge databases. Which is not always possible. For instance, even though ABIDE is considered a big database, but for deep learning it is still far from the norm. Hence, the solutions to using deep learning on ABIDE is transfer learning. Transfer learning is a good alternative to creating a network from scratch, because it permits to use characteristics from previously trained networks that have given good results on big datasets. It also has a big advantage of time saving, especially in the case of big sized data. For 2-dimensional data classification, object of this work, Convolutional neural networks are the most adapted. CNNs take images as entries and are composed of many types of layers, from which we can cite: the convolution layers, max pooling layers and the fully connected layer.
Keras [140], an open-source neural-network library written in Python, provides many trained CNNs with different combinations of the aforementioned layers. Keras can run on top of different machine learning libraries, such as TensorFlow and Theano.
In this chapter, we use ResNet-50 (Residual Network-50) which is a convolutional neural network architecture that was introduced in 2015 by Kaiming He et al. It is part of the ResNet family, which includes various network depths (e.g., ResNet-18, ResNet-34, ResNet-101) with increasing numbers of layers.
The key innovation in the ResNet architecture is the introduction of residual connections, also known as skip connections or shortcut connections. These connections allow the network to learn residual functions, which help address the degradation problem encountered in very deep neural networks. The degradation problem refers to the issue of increasing network depth leading to reduced accuracy due to the vanishing gradient problem and difficulty in optimization.
By using residual connections, ResNet enables the network to effectively learn the difference between the desired mapping and the identity mapping. This is achieved by adding the original input to the output of one or more stacked convolutional layers. The skip connections allow the gradient to flow directly from later layers to earlier layers during backpropagation, which helps alleviate the vanishing gradient problem and facilitates the training of deeper networks.
ResNet-50 consists of 50 layers, including convolutional layers, pooling layers, fully connected layers, and shortcut connections. The architecture follows a "bottleneck" design, which reduces the computational cost by using 1x1 convolutions to reduce the dimensionality before applying 3x3 convolutions. It also incorporates batch normalization for improved training efficiency and utilizes average pooling instead of fully connected layers at the end of the network.
ResNet-50 has achieved remarkable performance in various image classification competitions and benchmarks, demonstrating its effectiveness in handling complex visual tasks. It has been widely adopted as a pre-trained model for transfer learning and has also influenced the design of subsequent CNN architectures.

4.3.2 Features extraction

Feature extraction via deep learning is a fundamental process where a deep neural network is used to automatically learn relevant and informative patterns from raw data. Feature extraction has revolutionized various fields, such as computer vision, natural language processing, and speech recognition. It allows the automatic learning of discriminative and abstract features, eliminating the need for manual feature engineering and significantly improving the performance of AI systems on a wide range of tasks.

The process of feature extraction in deep learning begins by the choice of the Neural Network Architecture: The deep learning model should be appropriately chosen based on the nature of the data and the task. For images, CNNs are the best fit, as they can automatically learn hierarchical features from images and extract meaningful features.

For the Input Data: Deep learning models require raw data as input. In our case, image processing, the input data is a collection of images. Since the trained deep learning models take 3-dimensional images as entries, we organized our connectivity matrices into three layers as figure 4.2 and figure 4.3 show.

Then, we feed these three-layered matrices into the Keras ResNet50 model trained with imageNet [142], an image database with millions of images.

However, using these models is only to extract features. Hence, the execution stops before the fully connected layer, to extract the features to be used later. As the data flows through the layers of the deep neural network, each layer learns to extract increasingly abstract and informative features. In the early layers, the model detects simple patterns, like edges and corners in images, while in the deeper layers, it captures complex and high-level representations, such as object shapes and semantic features.

To extend the test and verify the efficiency of the tested approach we used 3 different types of connectivity matrices as a base to create the 3-layered matrices used to extract features. The Correlation matrices, the Tangent-space matrices and the Covariance matrices proposed by Nilearn. For these latter, to stay in the image context we divided the matrices with values exceeding 1 by the maximum value to get matrices with values between -1 and 1.

Then after feeding the 3D matrices into the deep learning models we extracted different arrays of features. These features sizes go from 4 x 4 to 7 x 7 depending on the fed matrices size.

An example of extracted features is presented in Figure 4.4

4.4 Classification and validation

For the classification of autistic subjects, we employ the widely recognized Support Vector Machine algorithm, known as SVC and provided by the scikit-learn library. The suitability of SVC for binary classification tasks, along with its established effectiveness in diverse domains, including medical diagnostics, motivated our choice to use it for classification in this chapter.

Our primary research objective is to distinguish between individuals with Autism Spectrum Disorder and neurotypical control subjects based on their brain imaging data. To accomplish this, we adopt a two-step approach. Initially, we extract features from the raw brain imaging data through the application of transfer learning. This method enables our deep learning model to leverage pre-trained

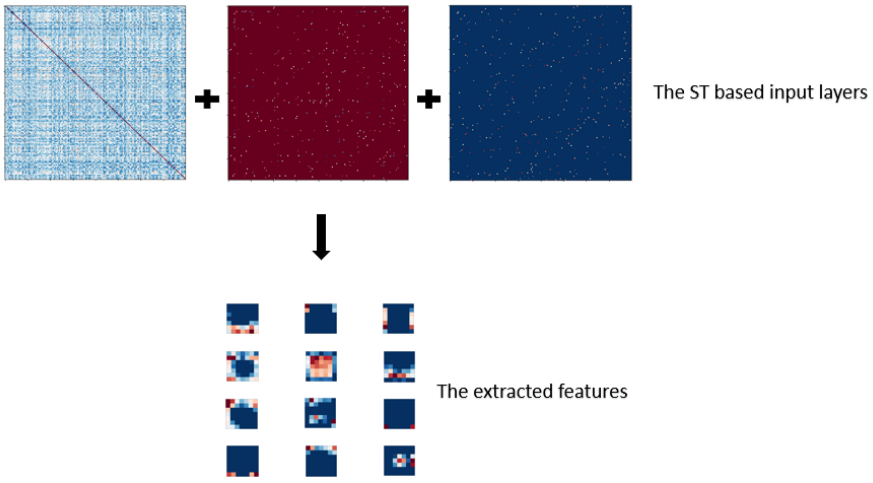

FIGURE 4.4: A sample of extracted features using the spanning trees based RGB images-alike matrices.

neural networks, fine-tuned on extensive image datasets, for the extraction of high-level features from the 3D matrices. Subsequently, these extracted features are input into the SVC for the classification task.

Then, based on the features vectors derived, SVC determines the optimal decision boundary that separates the two classes (ASD and neurotypical). This decision boundary maximizes the margin between the two groups, allowing for robust classification even in the presence of noisy or overlapping data points.

In this classification, the approximation of the decision boundary between the ASD patients and the control subjects is calculated in the space of the feature vectors extracted using transfer learning using each 3D matrices' features separately as shown in figure 4.5. Then the data is administered a label of 0 for neurotypical subjects and 1 for ASD subjects.

To evaluate the effectiveness of the proposed method, a validation process is conducted using three

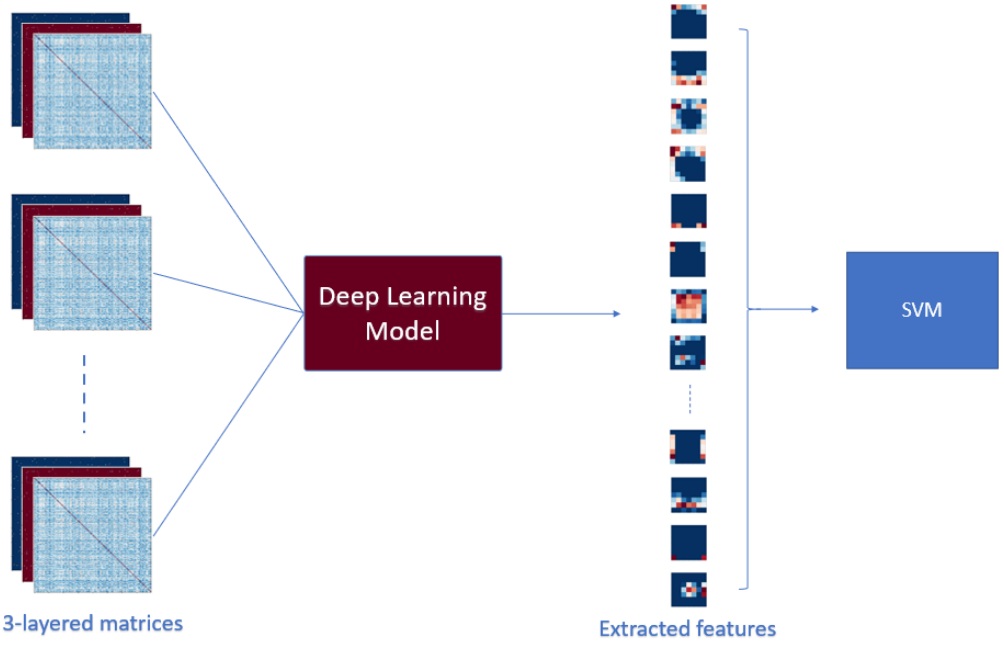

FIGURE 4.5: summary schema of the extraction and classification of the 3D matrices deep features.

performance metrics that are the accuracy, sensitivity and specificity. Additionally, cross-validation techniques are employed to assess the model's generalization ability and mitigate potential overfitting issues.

For the p-value choice, since the size of the features matrices are small we decided to test both p-value. The reduction effect of the use of the p-value is represented in Figure 4.6

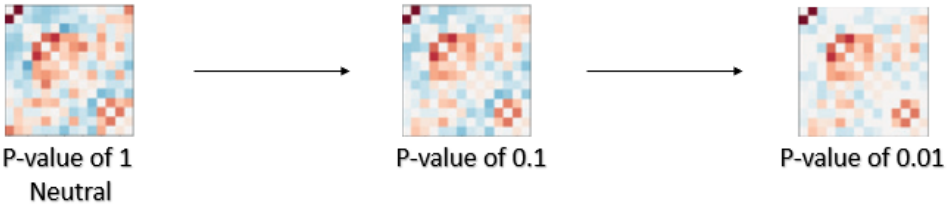

FIGURE 4.6: Effect of the p-value on a matrix representation.

4.5 Results

After processing the connectivity matrices to obtain their corresponding RGB-alike images, we fed these matrices into the transfer learning model, enabling it to extract essential features. These extracted features were then flattened into one-dimensional vectors to prepare them for classification using the Support Vector Classifier (SVC).

To ensure robust evaluation, we adopted a 10-fold cross-validation strategy for the classification process. This approach involved dividing the dataset into ten subsets (folds), where one fold was used for validation, while the model learned from the remaining nine folds. This process was repeated ten times, with each fold taking a turn as the validation set, thus providing a comprehensive assessment of the model's performance.

Furthermore, we used two different p-values, specifically 0.1 and 0.01 to determine the statistical significance of the results.

In the first Table 4.1, we present the detailed outcomes of the classification process for the two types of covariance 3D matrices, which were generated using the division and the spanning trees strategies. We specifically used the p-value of 0.1 for this analysis.

In Table 4.2, we present the results of the classification process, focusing on the two distinct strategies

TABLE 4.1: Accuracy of the classification of the covariance features extracted with ResNet50 with a p-value of 0.1

	Covar with STs			Covar with division		
	ACC	Sen	Spe	**ACC**	Sen	Spe
aal	**73.59%**	62.39%	86.60%	**35.02%**	65.17%	0.00%
db	**65.10%**	65.60%	64.52%	**57.18%**	64.10%	49.13%
cc200	**36.39%**	47.65%	23.33%	**59.24%**	46.15%	74.44%

employed to construct correlation 3D matrices. These strategies involve the creation of correlation matrices using both the division strategy and the spanning trees strategy. As for the p-value it is maintained to 0.1.

In Table 4.3, we present the final set of results for the classification process with a p-value of 0.1. This

TABLE 4.2: Accuracy of the classification of the correlation features extracted with ResNet50 with a p-value of 0.1

	Cor with STs			Cor with division		
	ACC	Sen	Spe	**ACC**	Sen	Spe
aal	**54.54%**	66.24%	40.94%	**56.49%**	86.11%	22.08%
db	**47.53%**	12.82%	87.84%	**47.19%**	87.82%	0.00%
cc200	**29.51%**	54.91%	0.00%	**61.42%**	83.55%	35.73%

table results concern the tangent 3D matrices. In the same light as the previous tests, the tangent matrices were created using both the division strategy and the spanning trees strategy.

In Table 4.1, one immediately notices the remarkable high accuracy achieved by the classification

TABLE 4.3: Accuracy of the classification of the tangent features extracted with ResNet50 with a p-value of 0.1

	Tan with STs			Tan with division		
	ACC	Sen	Spe	**ACC**	Sen	Spe
aal	**31.57%**	26.92%	36.97%	**32.84%**	25.85%	40.94%
db	**43.97%**	24.57%	66.50%	**43.17%**	48.50%	36.97%
cc200	**44.89%**	63.46%	23.33%	**18.83%**	14.74%	23.57%

of features extracted from the AAL atlas covariance 3D matrices created using the spanning trees strategy, reaching 73.59%. The DosenBach features, generated with the same strategy, follow closely with an accuracy of 65.10%. In contrast, the classification performance for CC200 features was comparatively mediocre, with an accuracy of 36.39%. The remaining classification results in this table did not exhibit noteworthy accuracy values.

Moving to the correlation features, as presented in Table 4.2, we observe average accuracy values across both strategies and all atlases, without any striking variations in performance.

Similarly, the tangent features' classification results, shown in Table 4.3, indicate below-average performance compared to the other methods explored.

Moreover, Tables 4.2 and 4.3 reveal that when employing a p-value of 0.1, the AAL and Dosenbach atlases exhibit minimal differences in accuracy between the two strategies.

Changing the p-value to 0.01, to further investigate the impact of the p-value on the classification results, we conducted additional tests. The first results pertaining to the classification of covariance features with the p-value of 0.01 are detailed in Table 4.4, shedding light on the changes in accuracy

with the adjustment in the p-value.

TABLE 4.4: Accuracy of the classification of the covariance features extracted from ResNet50 with a p-value of 0.01

	Covar with STs			Covar with division		
	ACC	Sen	Spe	**ACC**	Sen	Spe
aal	**25.83%**	48.08%	0.00%	**13.78%**	25.64%	0.00%
db	**66.02%**	65.60%	66.50%	**41.10%**	10.68%	76.43%
cc200	**33.41%**	22.86%	45.66%	**57.98%**	76.07%	36.97%

TABLE 4.5: Accuracy of the classification of the correlation features extracted from ResNet50 with a p-value of 0.01

	Cor with STs			Cor with division		
	ACC	Sen	Spe	**ACC**	Sen	Spe
aal	**19.75%**	22.44%	16.63%	**26.64%**	31.84%	20.60%
db	**38.81%**	63.68%	9.93%	**80.48%**	63.68%	100.00%
cc200	**12.74%**	3.42%	23.57%	**60.73%**	62.61%	58.56%

TABLE 4.6: Accuracy of the classification of the tangent features extracted from ResNet50 with a p-value of 0.01

	Tan with STs			Tan with division		
	ACC	Sen	Spe	**ACC**	Sen	Spe
aal	**8.27%**	15.17%	0.25%	**19.29%**	35.47%	0.50%
db	**47.65%**	34.40%	63.03%	**25.72%**	36.32%	13.40%
cc200	**77.61%**	87.18%	66.50%	**25.37%**	36.54%	12.41%

When the p-value was changed to 0.01 for the covariance classification (results of Table 4.4), the DosenBach and CC200 atlases showed nearly identical accuracy values, but the sensitivity and specificity measures switched values. In contrast, the AAL atlas's classification performance deteriorated significantly, resulting in poor accuracy compared to the other atlases.

In Table 4.5, which displays the classification results for correlation features with a p-value of 0.01, the division strategy took the lead for the first time, achieving the highest accuracy value. Notably, the DosenBach features achieved an outstanding accuracy of 80.48%, demonstrating their discriminative power. The CC200 atlas followed in this test, obtaining an average accuracy of 60.73%.

Finally, Table 4.6 showcased the results of classification using tangent matrices with the spanning trees

strategy at a p-value of 0.01. Surprisingly, the CC200 atlas performed exceptionally well, achieving its highest result of 77.61% accuracy. On the other hand, the AAL and DosenBach features once again demonstrated suboptimal performance with both strategies.
These findings underscore the sensitivity of the classification results to changes in the p-value and the construction strategies of the 3D matrices. The switch in ranking for atlases and strategies emphasizes the importance of carefully selecting appropriate parameters and methodologies for accurate and reliable brain anomaly detection.
The results presented in Tables 4.4, 4.5, and 4.6 collectively contribute to our understanding of the strengths and limitations of various approaches in neuroimaging-based classification. These findings prompt further exploration and refinement of our method, aiming to optimize the classification accuracy and to enhance our ability to distinguish between Autism Spectrum Disorder and neurotypical individuals effectively.

4.6 Discussion

Previously, we verified the over and under-connectivity of the autistic brain and found proof of their existence using an approach of elimination, where connections with specifications related to over and under-connectivity were eliminated to bring out their importance.
Hence, to explore these theories more in depth, we used a new approach that combines deep learning with machine learning to classify autistic subjects with the goal of enhancing the ability of detecting autism through fMRI images information.
In this approach we used two strategies to construct RGB images alike 3D matrices to give as entry to deep learning models and extract features that will permit the detection of autism.
After identifying the significant connections associated with autism disorder in our previous analysis, we took a strategic approach to enhance our connectivity matrices. Dividing the matrices into layers, we incorporated the specific connections that exhibited over-connectivity and under-connectivity patterns, as extracted from the original connectivity matrices computed using diverse atlas Regions of Interest (ROIs).
To achieve this, we introduced two additional layers to our connectivity matrices, each dedicated to preserving the connections with distinct specifications. One layer accommodated the strongest connections, while the other encompassed the weakest connections. This careful layering allowed us to concentrate on the crucial connections that bear potential diagnostic significance for autism.
Subsequently, we harnessed the power of deep neural networks in combination with the Support Vector Classifier (SVC) to execute our classification task. Leveraging transfer learning, we extracted features from the connectivity matrices using deep neural networks. These extracted features, representing the discriminative patterns found in the data, were then subjected to separate classification with SVC.
The incorporation of deep learning in feature extraction bestowed the model with the capacity to capture intricate and high-level representations, enabling it to discern subtle differences between autism and neurotypical cases. Simultaneously, SVC, a proven and effective classification algorithm, facilitated the determination of the optimal decision boundary to distinguish between the two classes.
For the sake of comparison, we constructed a basic 3D connectivity matrices where to preserve the original information within the connectivity weights, we divided the connectivity weights by three.

This strategic approach ensured that the fundamental information encoded in the connectivity weights remained intact, while effectively populating the layers of the 3D matrices. By doing so, we aimed to strike a balance between retaining the essential features of the connectivity data and ensuring the successful formation of the 3D matrices, setting the stage for subsequent analyses and investigations.
The results achieved were greatly different from those achieved in the previous chapter.
First, the AAL Tangent combination that have been taking the lead in all the previous test plummeted greatly achieving the worst results in this chapter (see Table 4.6).
Moreover, every atlas performed its best only once in all tests with different combinations.
In these tests, the most remarkable performance was observed, achieving an impressive 80.48% accuracy, 63.68% sensitivity, and a perfect 100% specificity. These exceptional results were attained through the utilization of a p-value of 0.01 and the classification of features extracted from the 3D correlation matrices constructed with the DosenBach atlas, combined with the division strategy. This winning combination demonstrated the robustness and discriminative power of the chosen approach, highlighting its potential significance in accurately identifying autism cases from the neurotypical population.
With this same p-value, the CC200 atlas also showed its best reaching 77.61%. However, the 3D matrices were constructed from the tangent matrices with the spanning trees strategy.
The AAL atlas reached only 73.59% with a p-value of 0.1 and the covariance 3D matrices constructed using the spanning trees strategy.
From the wealth of results presented in this chapter, several distinct conclusions emerge. Foremost, the comparisons drawn from all the tests underscore the multifaceted nature of working with heterogeneous data. It becomes evident that numerous factors can significantly influence the accuracy of detection. Notably, the findings revealed that the accuracy of detection can differ substantially when using various connectivity matrices, even when keeping the same atlas constant. This observation underscores the importance of meticulous consideration and thorough exploration of data features when striving for accurate and reliable detection in neuroimaging-based studies. The varying outcomes serve as a poignant reminder of the intricacies involved in analyzing complex brain connectivity data and the need for a tailored approach to meet the specific demands of each analysis scenario.
Also, testing the two 3D matrices construction strategies showed the huge impact they can have on the results. Which is quite logical, since one is based on the autism theories while the other is based on a simple division that add no autism specification to the 3D matrices. Moreover, the combination of the atlases with the 3D construction strategies also matters. Since the AAL and CC200 best results were achieved with the spanning trees strategy while only the DosenBach was better with the division strategy. This last reported result was also the best which is quite interesting.
In the other hand, not all results were above average, which means that the features extracted were not enough to identify autism with high accuracy. This result might be related to the depth and structure of the deep learning model. But it can also be related to the choice of the connectivity matrices computation method as well as the choice of the atlas and the number of its ROIs.

4.7 Conclusion

In the mathematical field, the negation is used to prove the veracity of a theory by demonstrating that if its negation is false, automatically the theory is true. In the previous chapter, we used this concept and were able to detect some changes that proved the existence of communication deficits in the autistic brain.

Going further, we integrated these findings in a new approach using deep learning. In this approach we introduced a new vision to using deep learning with connectivity matrices, by pursuing a path of image mimicking.

This implied constructing a new data based on the connectivity matrices and using a layering strategy that permitted to construct two sets of RGB images alike matrices for every data source.

Notably, the approach of dividing the information into three equal layers proved to be particularly well-suited, leading to a significant increase in the accuracy of detection. The integration of autism theories in this test introduced modifications to the structure of the matrices, resulting in enhanced detection accuracy compared to the existing state of the art.

By leveraging this division strategy, we were able to capitalize on the inherent information contained within the connectivity data, resulting in improved performance. This successful approach holds great promise for advancing the field of autism detection and underscores the importance of incorporating domain-specific knowledge to optimize classification outcomes in neuroimaging studies.

Chapter 5

Improving Autism Detection through an enhanced Deep Learning-based approach

The previous chapter showed that in a heterogeneous setting using the ABIDE dataset, machine learning achieved an accuracy range of 60% to 70% for autism detection. The application of deep learning further improved detection accuracy, reaching 80%.
Hence, after investigating autism using machine learning then by combining machine learning and deep learning, and continuing on the same track of autism tailored detection, in this chapter, we propose a two-phase deep system with an enhancement approach for identifying autistic patients based on deep learning.
Deep learning usually requires a large database for the learning phase. However, our data is limited. To overcome this limitation, similar to the approach in the previous chapter, we adopt transfer learning. Transfer learning leverages pre-trained models on well-known databases, such as ImageNet, to improve performance and address the challenge of training deep learning models.
Our approach also addresses the inherent complexity of the brain by leveraging functional magnetic resonance imaging (fMRI) data to extract connectivity maps that capture the intricate interactions between different regions of the brain. To overcome the challenges posed by the sheer number of neural units and enhance interpretability, we utilize atlases to reduce the complexity of the connectivity maps. What sets our approach apart is the introduction of a novel enhancement method based on a specific layering of the data within the convolutional neural network. This layering is tailored to ASD and draws inspiration from previous autism theories. By incorporating layers that represent the properties of over-connectivity and under-connectivity, we aim to highlight connectivity patterns that have been identified as relevant autism biomarkers. This approach aims to improve the accuracy of the deep learning system in detecting autism.

5.1 Data Construction

Since this is a continuation of the previous autism investigation, the data is preprocessed with C-PAC. The data construction involves two important steps. First, the computation of the connectivity

matrices that include the covariance, the correlation and the tangent methods with the time series of the AAL, DosenBach and CC200 atlases.
For the second step, similar to the approach outlined in Chapter 4, we mimic RGB images by constructing three-dimensional connectivity maps that contain different interactions of the brain regions. This step includes a novel enhancement method highlighting the connectivity information crucial to autism detection.

5.1.1 Creation of the 3D Connectivity Matrices

To generate the 3D (RGB images alike) connectivity matrices, following the same methodology from the previous Chapter 4, we begin by computing the connectivity matrices for subjects in the ABIDE I preprocessed database using three distinct atlases: AAL, DosenBach, and CC200 atlases. These atlases offer diverse sets of clusters of Regions of Interest (ROIs) and allow for the construction of connectivity matrices with dimensions of 116×116, 161×161, and 200×200 connections, respectively. The connectivity matrices are then derived using three connectivity computation methods:

- The covariance method [143], which calculates coefficients reflecting direct and indirect connections between each pair of regions.
- The correlation method, which assesses the likelihood of communication between the ROIs.
- The tangent space embedding [144], which goes beyond the individual level and couples information from all interactions into a single group connectome using a geometrical framework. This approach allows the measurement of interactions in a common space known as the tangent space [145].

By combining the three time series and the connectivity computing methods, a total of nine different connectivity matrices (CM) are obtained for each subject.
Subsequently, each CM is divided separately by 3 to extract a third of its connectivity weights. The 3D connectivity matrix of this CM is then constructed by transposing the extracted thirds, as illustrated on the right side of Figure 5.1.

5.1.2 Enhanced 3D Matrices

The enhanced matrices exhibit specific highlighted connections for each layer, incorporating information relevant to autism theories. This process builds upon the earlier division technique while emphasizing connections aligned with autism-related hypotheses. To achieve this, we first transform each connectivity matrix into a graph $g = (V, E)$, where V is a finite set of vertices representing the ROIs from the atlases, and $E \subseteq V \times V$ is a finite set of weighted edges representing connections between pairs of ROIs.
Next, we extract the layers of the enhanced matrices based on high-weight and low-weight connections using the maximum spanning tree and the minimum spanning tree, respectively, as depicted on the left side of Figure 5.1.
After extracting both trees, we construct the enhanced 3D matrices for each subject by overlaying the

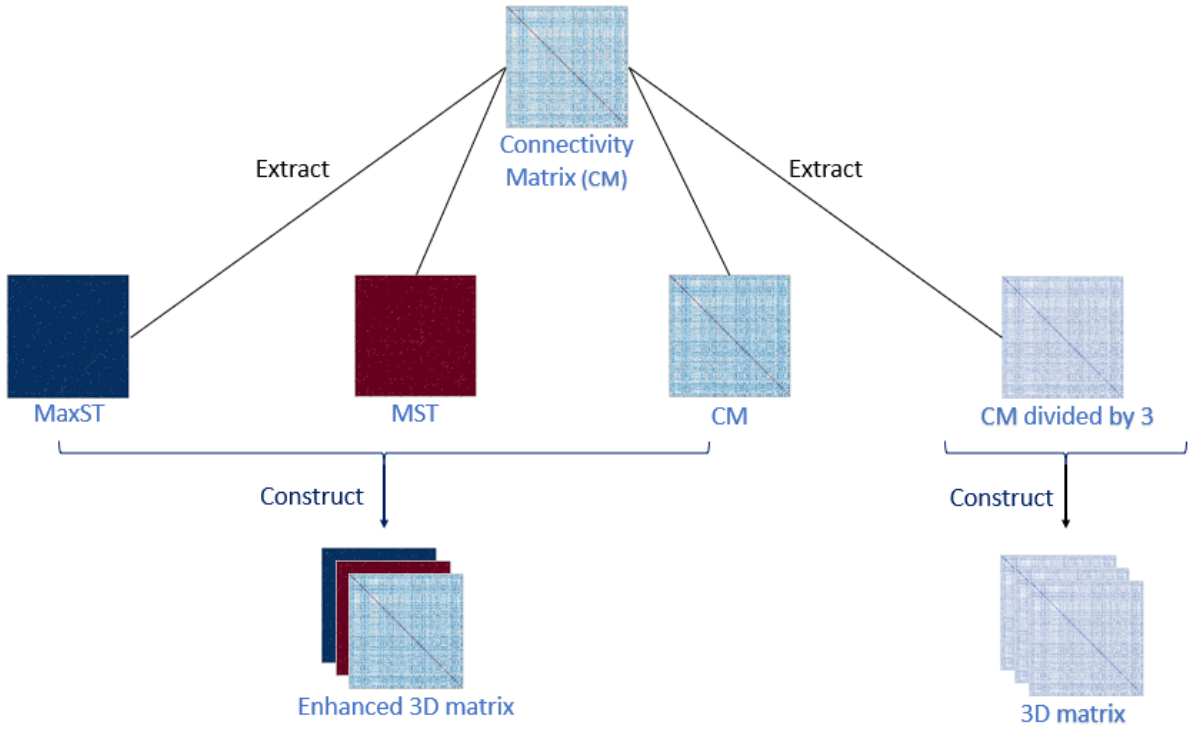

FIGURE 5.1: The 3D matrices construction process. MST and MaxST represent the minimum spanning tree and the maximum spanning tree, respectively.

original connectivity matrix on the first layer, the maximum spanning tree on the second layer, and the minimum spanning tree on the third layer (see Figure 5.1).
Thus, the enhanced matrices embody the concepts of over-connectivity and under-connectivity, which have been previously evidenced in the brain regions in the earlier chapters.

5.2 Classification with CNN

In this stage of our study, we leverage the power of convolutional neural networks and transfer learning to classify the constructed 3D matrices and overcome the limitation of limited data size. The main steps of our classification process are illustrated in Figure 5.2, providing a visual representation of our approach. To begin, we follow a typical transfer learning workflow, where we employ a pre-trained base model and load its weights. By doing so, we benefit from the knowledge acquired through the training of the base model on a large-scale dataset. To preserve this valuable information, we freeze all layers of the base model, ensuring that the learned weights remain unchanged during the training process.
Building upon the output of the base model, we create a new model that incorporates additional layers tailored to our specific task. This allows us to extract meaningful features from the connectivity matrices and capture the distinctive patterns associated with autism.
As the model trains and converges on the extracted features, we unfreeze all layers of the base model and initiate an end-to-end re-training phase with a very low learning rate. This step aims to further refine the model's performance and potentially improve the classification results.
To enhance the reliability of our findings, we adopt a 10-fold cross-validation approach. Where we

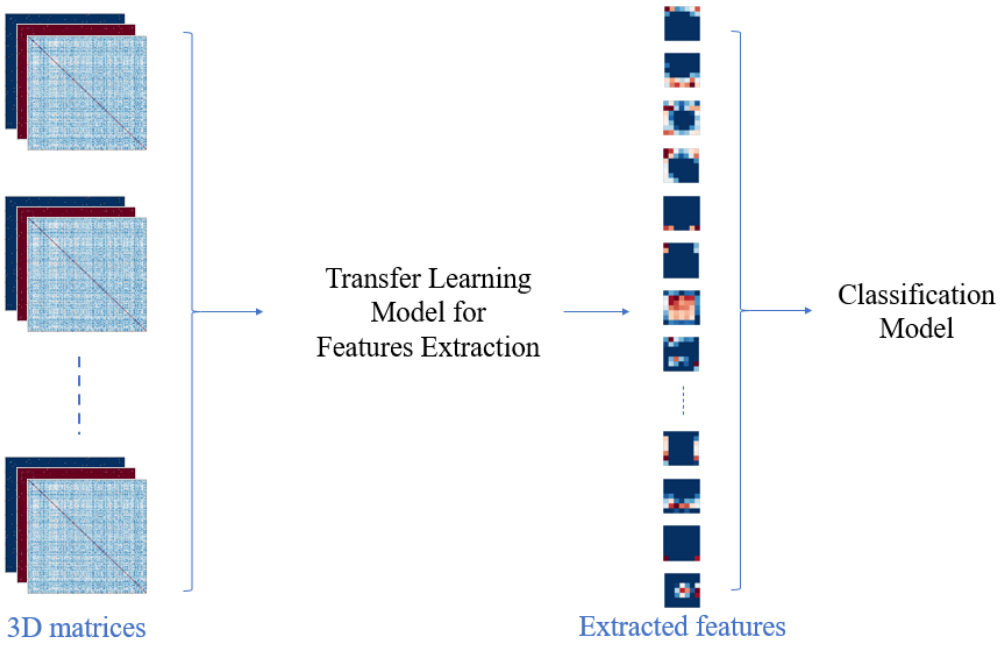

FIGURE 5.2: The main steps of deep learning classification.

perform ten separate tests, each time reserving one fold as a test set while utilizing the remaining folds for training and validation. This comprehensive evaluation strategy provides a robust assessment of our model's performance.

Then to evaluate the effectiveness of our classification approach, we employ the accuracy metric, which measures the model's ability to correctly differentiate between individuals with ASD and control subjects [146]. We compute the mean accuracy across all ten tests of the cross-validation, ensuring a reliable estimation of our model's performance.

By combining transfer learning, CNNs, and cross-validation, we aim to develop a robust and accurate classification system that effectively distinguishes between individuals with autism and neurotypical individuals. The utilization of these techniques allows us to leverage existing knowledge while fine-tuning the model to capture the unique connectivity patterns associated with autism.

The models used are VGG16, VGG19, ResNet50, Inception, InceptionResNet, ResNet152V2, and Xception with the weights of imageNet.

5.2.1 VGG16

VGG16 is a convolutional neural network architecture that belongs to the VGG (Visual Geometry Group) family of models. It was introduced by Karen Simonyan and Andrew Zisserman in their research paper titled "Very Deep Convolutional Networks for Large-Scale Image Recognition" in 2014 [147].

The VGG16 architecture is known for its simplicity and effectiveness in image classification tasks. The "16" in VGG16 refers to the number of layers in the network, including convolutional layers, pooling layers, and fully connected layers.

The key feature of VGG16 is its use of a series of stacked convolutional layers with small 3x3 filters, along with max pooling layers, throughout the network. The architecture maintains a consistent filter size and pooling size across all layers, which contributes to its simplicity and uniformity.

VGG16 consists of multiple convolutional blocks, each containing two or more convolutional layers with the same filter size, followed by a max pooling layer. The number of filters in each block increases as we go deeper into the network, allowing for the extraction of increasingly complex features.
The final part of VGG16 comprises fully connected layers that are responsible for the classification task. These layers take the output of the convolutional layers, flatten it, and pass it through fully connected layers with progressively reducing dimensions. The final layer uses softmax activation to produce class probabilities.
With its straightforward and uniform architecture, it serves as a strong baseline for image classification tasks and has been widely adopted and used as a foundation for many subsequent CNN architectures.
VGG16 is characterized by its depth, as it has 16 weight layers, including 13 convolutional layers and 3 fully connected layers. It has a homogeneous architecture with a consistent configuration of convolutional and pooling layers throughout the network that can summarized as follows:

- The input of 224x224 RGB images.
- It begins with a series of convolutional layers, each followed by a rectified linear unit (ReLU) activation function. The convolutional layers use small 3x3 filters with a stride of 1 and padding to preserve spatial dimensions. These layers are stacked multiple times to capture increasingly complex features.
- After each set of convolutional layers, VGG16 applies max pooling layers with a 2x2 window and a stride of 2. These pooling layers progressively reduce the spatial dimensions of the feature maps while retaining important features.
- After the convolutional and pooling layers, VGG16 consists of three fully connected layers. Each fully connected layer has a large number of neurons and is followed by a ReLU activation function. These layers perform high-level feature extraction and mapping.
- The final layer of VGG16 is a softmax activation function, which produces the predicted class probabilities for image classification. The number of neurons in the output layer corresponds to the number of classes in the classification task.

5.2.2 VGG19

VGG19 is also part of the VGG family of models. Its architecture is also simple with "19" in VGG19 refering to the number of layers in the network, including convolutional layers, pooling layers, and fully connected layers.
In the same way as VGG16, VGG19 uses the stack of convolutional layers with small 3x3 filters, along with max pooling layers, throughout the network. The architecture maintains a consistent filter size and pooling size across all layers, which contributes to its simplicity and makes the model easier to understand and replicate.
Its architecture is similar to VGG16 with multiple convolutional blocks that contain multiple convolutional layers, followed by a max pooling layer. The final part of VGG19 consists of fully connected layers, which are responsible for the classification task. These layers take the output of the convolutional layers, flatten it, and pass it through fully connected layers with progressively reducing

dimensions. The final layer uses softmax activation to produce class probabilities.
Although VGG19 is deeper and more computationally expensive than some other architectures, it provides a strong baseline for image classification tasks. Its straightforward and uniform structure makes it easier to understand and implement, which has contributed to its popularity and widespread adoption in the computer vision community.
Its architecture can be summarized in the following:

- The input to the VGG19 model is a 224x224 RGB image.
- VGG19 begins with a series of convolutional layers, each followed by a rectified linear unit activation function. The convolutional layers use small 3x3 filters with a stride of 1 and padding to preserve spatial dimensions. These layers are stacked multiple times to capture increasingly complex features.
- After each set of convolutional layers, VGG19 applies max pooling layers with a 2x2 window and a stride of 2. These pooling layers progressively reduce the spatial dimensions of the feature maps while retaining important features.
- After the convolutional and pooling layers, VGG19 consists of three fully connected layers. Each fully connected layer has a large number of neurons and is followed by a ReLU activation function. These layers perform high-level feature extraction and mapping.
- The final layer of VGG19 is a softmax activation function, which produces the predicted class probabilities for image classification. The number of neurons in the output layer corresponds to the number of classes in the classification task.

5.2.3 Inception

Inception refers to a deep convolutional neural network architecture that was introduced in the research paper titled "Going Deeper with Convolutions" [148]. The Inception architecture was designed to address the challenge of constructing deeper networks while managing computational complexity and improving accuracy.
The key idea behind the Inception architecture is the use of "Inception modules" that consist of multiple parallel convolutional layers with different filter sizes. This allows the network to capture features at different spatial scales and learn diverse representations. The parallel convolutional layers within an Inception module are followed by 1x1 convolutions that act as dimensionality reduction layers to reduce computational cost.
Inception networks have been widely used in computer vision tasks, particularly in image classification, object detection, and image segmentation. They have achieved remarkable performance on benchmark datasets, such as ImageNet, and have been influential in the development of subsequent architectures.
It's worth mentioning that Inception is a foundational concept in CNN architectures, and its principles have inspired the design of many subsequent models. While the original Inception architecture has made significant contributions, researchers continue to explore and develop new architectures that build upon its ideas to further improve the performance of deep neural networks.
The InceptionV3 architecture is a deep convolutional neural network that was introduced by Christian

Szegedy et al. as part of the Inception family of models. It is designed for image classification and feature extraction tasks. Its architecture is defined by six important points:

- The input to the InceptionV3 model is a 299x299 RGB image.
- The initial layers of InceptionV3 consist of multiple convolutional layers with different filter sizes (1x1, 3x3, 5x5) and strides. These layers extract low-level features from the input image.
- The core building blocks of InceptionV3 are the Inception modules. These modules are designed to capture features at different spatial scales. Each Inception module contains parallel branches of convolutions with different filter sizes. These branches are combined using concatenation along the channel dimension. InceptionV3 includes multiple stacked Inception modules, each with different numbers and sizes of convolutional filters.
- InceptionV3 incorporates auxiliary classifiers at intermediate layers of the network. These auxiliary classifiers have their own softmax layer and are used during training to combat the vanishing gradient problem. They encourage the network to learn more useful and discriminative features.
- The final part of InceptionV3 consists of fully connected layers that take the output of the convolutional layers and perform classification. The fully connected layers gradually reduce the spatial dimensions and ultimately produce the predicted class probabilities using softmax activation.
- InceptionV3 employs various training techniques, including batch normalization, dropout regularization, and gradient normalization. These techniques help in stabilizing and improving the training process, leading to better generalization and performance.

5.2.4 RESNET50

ResNet50 is a convolutional neural network architecture that belongs to the ResNet (Residual Network) family. It was introduced by Kaiming He et al. in their research paper titled "Deep Residual Learning for Image Recognition" in 2015 [149].
Again the "50" in ResNet50 refers to the number of layers in the network, including convolutional layers, pooling layers, fully connected layers, and skip connections. ResNet50 is a variant of the original ResNet architecture that was specifically designed for image classification tasks.
The key innovation in ResNet50 is the use of residual connections or skip connections, which address the problem of vanishing gradients in deep networks. These skip connections allow the gradient to flow directly from one layer to another, enabling the training of much deeper networks without degradation in performance. This helps in mitigating the challenges of training very deep neural networks and allows for better feature propagation.
ResNet50 consists of several residual blocks, each containing multiple convolutional layers and skip connections. It also incorporates global average pooling, which replaces traditional fully connected layers, and a softmax layer for classification.
Due to its architectural simplicity and effectiveness, ResNet50 has become widely adopted in computer vision applications. It serves as a powerful tool for tasks such as image recognition, object detection,

and image segmentation. Researchers and practitioners often use ResNet50 as a benchmark or starting point for developing more advanced network architectures.
Its architecture key points are:

- The input to the ResNet50 model is a 224x224 RGB image.
- The initial layer of ResNet50 is a 7x7 convolutional layer with stride 2, followed by a max pooling layer. This reduces the spatial dimensions of the input image.
- The core building blocks of ResNet50 are the residual blocks. A residual block consists of multiple convolutional layers with shortcut connections. The convolutional layers learn the residual mapping, which represents the difference between the desired output and the current input. The shortcut connections allow the gradients to flow directly through the block, mitigating the vanishing gradient problem and making it easier to train very deep networks. ResNet50 has 16 residual blocks grouped into four stages, each with a different number of filters. The first stage has a single convolutional layer, and subsequent stages have three convolutional layers per block. The number of filters increases as we go deeper into the network, allowing for the extraction of more complex features.
- After the residual blocks, ResNet50 applies global average pooling to reduce the spatial dimensions to 1x1x2048. This operation summarizes the learned features for each channel.
- The final part of ResNet50 consists of a fully connected layer with 1000 neurons, followed by a softmax activation function. This layer produces the predicted class probabilities.

5.2.5 InceptionResNet

InceptionResNet refers to a hybrid convolutional neural network architecture that combines the concepts of the Inception architecture and the ResNet architecture. It was introduced in the research paper titled "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning" [150].
The InceptionResNet architecture aims to leverage the advantages of both Inception and ResNet networks, namely the ability to capture multi-scale features and the ability to train very deep networks with skip connections.
The main building block of the InceptionResNet architecture is the Inception-ResNet module, which incorporates the residual connections from the ResNet architecture into the Inception module of the Inception architecture.
The Inception-ResNet module follows a similar structure as the Inception module, with parallel branches of different convolutions, but also includes residual connections. These residual connections allow the gradient to flow directly through shortcut connections, similar to the original ResNet architecture. This enables the network to mitigate the vanishing gradient problem and make it easier to train deeper networks.
By combining the benefits of both architectures, InceptionResNet models can effectively capture multi-scale features with improved training and optimization. They have demonstrated excellent performance on various computer vision tasks, including image classification, object detection, and image

segmentation.
InceptionResNet models are typically named with a combination of the terms "Inception" and "ResNet" followed by a number indicating the number of layers or variations. Examples include InceptionResNetV1 and InceptionResNetV2, which represent different versions of the architecture with varying modifications and improvements.
The InceptionResNetV2 was introduced by Christian Szegedy et al. in their research paper titled "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning" in 2016. InceptionResNetV2 aims to leverage the benefits of both architectures for improved performance in image classification tasks.
InceptionResNetV2 architecture is defined by:

- The initial layers of InceptionResNetV2 form the stem. This includes a combination of convolutional layers, pooling layers, and batch normalization to extract low-level features from the input image.
- Similar to the Inception architecture, InceptionResNetV2 incorporates Inception modules to capture features at different scales. These modules consist of parallel branches of convolutions with varying filter sizes. However, unlike the original Inception modules, InceptionResNetV2 also includes residual connections to address the vanishing gradient problem and improve gradient flow.
- InceptionResNetV2 includes residual blocks, inspired by the ResNet architecture, to allow for the training of very deep networks. These residual blocks contain multiple convolutional layers and incorporate residual connections that directly pass the input through the block. This enables the network to learn residual mappings and makes it easier to train deeper networks.
- InceptionResNetV2 incorporates reduction blocks that include a combination of convolutional layers, pooling layers, and dimensionality reduction techniques. These blocks reduce the spatial dimensions of the feature maps while increasing the number of filters, enabling the extraction of more complex features.
- After the convolutional layers and residual blocks, InceptionResNetV2 applies global average pooling to reduce the spatial dimensions to 1x1x2048, summarizing the learned features for each channel.
- The final part of InceptionResNetV2 consists of a fully connected layer with a softmax activation function to produce the predicted class probabilities.

5.2.6 RESNET152V2

ResNet152V2, also known as Residual Network-152 Version 2, as it name suggests, is a convolutional neural network architecture that belongs to the ResNet family.
The ResNet152V2 architecture is an extension of the original ResNet152 model, introducing some additional enhancements to improve its performance. It utilizes residual connections.
ResNet152V2 consists of 152 layers, making it a very deep model. It has a similar basic structure as other ResNet models, which is based on the concept of residual blocks. Each residual block contains

a stack of convolutional layers, batch normalization, and ReLU activation functions, with residual connections added to the output of each block.
The key improvements in ResNet152V2 include the use of bottleneck residual blocks and a combination of other techniques such as batch normalization, weight decay, and adaptive learning rate. These enhancements help to reduce the number of parameters and computational complexity while maintaining or improving accuracy.
The ResNet152V2 architecture builds upon the ResNet architecture and introduces some additional improvements, resulting in an architecture with seven important points:

- The input Layer of ResNet152V2 takes an input image of size 224x224x3 (RGB color channels).
- The initial layer applies a convolution operation with a 7x7 kernel and 64 filters, followed by batch normalization and ReLU activation. A max pooling layer with a 3x3 kernel and a stride of 2 is then applied to downsample the feature maps.
- ResNet152V2 consists of a stack of residual blocks. Each block is composed of several convolutional layers, batch normalization, and ReLU activation functions. The key difference compared to earlier ResNet versions is the use of "bottleneck" blocks, which reduce the number of parameters and computational complexity. A bottleneck block consists of three convolutional layers with 1x1, 3x3, and 1x1 filter sizes, respectively.
- ResNet152V2 uses skip connections or shortcut connections to address the vanishing gradient problem. These connections allow the gradient to flow directly across layers, improving gradient propagation and easing the training of deep networks.
- Instead of using fully connected layers, ResNet152V2 employs global average pooling after the residual blocks. This operation calculates the average value of each feature map, resulting in a fixed-length feature vector.
- A dense layer with softmax activation is added on top of the global average pooling layer to perform the final classification. The number of neurons in this layer depends on the specific task and dataset being used.
- ResNet152V2 is typically pretrained on large-scale image classification datasets like ImageNet. After pretraining, the model can be fine-tuned on a smaller task-specific dataset to adapt to the specific classification task.

5.2.7 Xception

Xception is a convolutional neural network architecture that was introduced in the research paper titled "Xception: Deep Learning with Depthwise Separable Convolutions" [151]. The name "Xception" stands for "Extreme Inception," indicating that it is an extension or an extreme version of the Inception architecture.
The Xception architecture is based on the idea of depthwise separable convolutions, which aim to improve the efficiency and effectiveness of convolutional layers in deep neural networks. Traditional

convolutions apply a series of filters to input channels, resulting in a high computational cost. In contrast, depthwise separable convolutions split the convolution operation into two separate stages: depthwise convolutions and pointwise convolutions.
In the depthwise convolution stage, each input channel is convolved with its own set of filters independently. This step captures spatial information within each channel. In the pointwise convolution stage, 1x1 convolutions are applied to combine the output channels of the depthwise convolution. This step allows for cross-channel interactions and the extraction of higher-level features.
By separating the spatial and cross-channel operations, Xception reduces the number of parameters and computations required compared to traditional convolutions. This enables deeper networks and more efficient learning.
The Xception architecture follows a similar overall structure as the Inception architecture, with multiple parallel branches of different convolutions, pooling layers, and skip connections. However, the key difference lies in the use of depthwise separable convolutions instead of traditional convolutions.
In summary, the Xception architecture is characterized by:

- An initial part, called the entry flow, which consists of a series of convolutional layers with depthwise separable convolutions. Unlike traditional convolutions that operate on both the spatial and channel dimensions together, depthwise separable convolutions decouple the two operations. They first apply spatial convolutions (1x1 convolutions) to each channel separately and then combine the outputs with depthwise convolutions (3x3 convolutions). This reduces computational complexity while preserving representational power.

- The middle flow of Xception repeats a set of residual blocks that further deepen the network. These residual blocks utilize depthwise separable convolutions and residual connections, similar to the ResNet architecture, to capture and propagate information through the network effectively.

- The exit flow of Xception contains a combination of residual blocks and average pooling layers. It gradually reduces the spatial dimensions of the feature maps and increases the number of filters. The exit flow concludes with a global average pooling layer to summarize the learned features.

- The final part of Xception consists of a fully connected layer with a softmax activation function. This layer maps the learned features to the predicted class probabilities.

5.3 Results

The combination of the three atlases and the three connectivity methods yields a total of nine distinct connectivity matrices for each subject. This number doubles after the application of 3D creation, resulting in nine basic 3D matrices and nine enhanced 3D matrices. Consequently, for every individual in our study, we have a total of 18 different connectivity matrices to analyze.
These matrices are then fed separately into the seven deep learning models, leading to the extraction of 128 groups of features. The sizes of these extracted features vary depending on the depth of the model and the specific atlas used, ranging from 2×2 to 7×7.

By employing multiple atlases, connectivity methods, and deep learning models, we aim to capture the diverse aspects of brain connectivity in individuals with autism. This comprehensive approach allows us to explore various patterns and representations within the connectivity matrices, providing a rich set of features for our classification task.
Table 5.1 groups all the different features sizes in correspondence with the deep learning model and matrices type.
Then, for every atlas, by using a 10-fold cross-validation we ended up with 42 groups of 10 accuracy values. The mean of these accuracy values are displayed in the tables below.
First, Table 5.2 contains the classification results of the AAL atlas.

TABLE 5.1: Sizes of the features extracted by the different models from the three atlases' connectivity matrices

	RN152V2	Inception	RN50	IRN	X	VGG19	VGG16
Dosenbach	6 x 6	3 x 3	6 x 6	3 x 3	5 x 5	5 x 5	5 x 5
CC200	7 x 7	4 x 4	7 x 7	4 x 4	7 x 7	6 x 6	6 x 6
AAL	4 x 4	2 x 2	4 x 4	2 x 2	4 x 4	3 x 3	3 x 3

TABLE 5.2: Accuracy of the classification of the AAL 3D matrices with and without enhancement.

	Correlation		Covariance		Tangent	
	3D Matrices	Enhanced 3D Matrices	3D Matrices	Enhanced 3D Matrices	3D Matrices	Enhanced 3D Matrices
ResNet152	72.12%	**90.01%**	76.14%	87.97%	87.77%	**91.85%**
Inception	**92.68%**	**93.59%**	82.81%	85.80%	84.99%	**94.39%**
ResNet50	46.26%	46.26%	85.44%	**93.69%**	46.26%	46.27%
InResNet	**92.22%**	**93.70%**	47.75%	46.85%	85.58%	**92.68%**
Xeption	89.93%	**95.41%**	64.42%	78.09%	**90.04%**	**93.58%**
VGG16	82.35%	**96.10%**	73.84%	**91.51%**	61.91%	**92.89%**
VGG19	80.39%	**95.98%**	70.75%	**90.48%**	68.57%	**94.04%**

The matrices of the AAL atlas have a shape of (116, 116, 3). When fed into the models, these matrices result in feature sizes of 2×2 for Inception and InceptionResNet models, 3×3 for the VGG models, and 4×4 for the remaining models.
An important observation from the results in Table 5.2 is the high number of accuracy values exceeding 90%, which are highlighted in bold font. Many of these values surpass the state-of-the-art achievements in fMRI-based autism detection. Furthermore, the utilization of the enhancement approach improves the classification results. Specifically, the classification of enhanced 3D matrices achieves accuracy values of 96.10% and 95.98% using the correlation matrices with VGG16 and VGG19, respectively.
For the other models, the enhancement approach consistently leads to high accuracy values with all connectivity calculation methods (correlation, covariance, and tangent), except for ResNet50, which demonstrates high accuracy only with the covariance method. However, in general, VGG16 and VGG19 remain the top choices for modeling the enhanced matrices of the AAL atlas, as all their accuracy values exceed 90%.

These results indicate the effectiveness of the proposed approach in achieving high accuracy in fMRI-based autism detection, particularly with the AAL atlas enhanced matrices. Through these results the enhancement approach, combined with deep learning models, has proven to be successful in capturing relevant features and improving classification performance.

Moving on to the DosenBach atlas and its matrices with a shape of 161 × 161 × 3, the resulting features also have varying sizes. Specifically, the features are 3 × 3 for Inception and InceptionResNet models, 5 × 5 for the VGG models and Xception, and 6 × 6 for the ResNet models.

When comparing the results of the DosenBach atlas to those of the AAL atlas in Table 5.3, the overall performance is lower. This suggests that the deep learning models struggle to effectively extract meaningful features from the matrices of this atlas. Interestingly, the VGG models, which performed well with the AAL atlas features, show poor performance with all strategies and types of connectivity calculations when applied to the DosenBach atlas. Only Inception and ResNet152 achieved an accuracy of over 90%, but only with the correlation and tangent methods, respectively.

In this case, the poor results obtained with the DosenBach atlas may be attributed to the increased number of ROIs and the delineation of the ROIs, which could have been the reason of a loss of critical information necessary for accurate classification.

TABLE 5.3: Accuracy of the classification of the DosenBach 3D matrices with and without enhancement.

	Correlation		Covariance		Tangent	
	3D matrices	Enhanced 3D matrices	3D matrices	Enhanced 3D matrices	3D matrices	Enhanced 3D matrices
ResNet152	53.28%	88.42%	58.67%	66.95%	65.12%	**91.63%**
Inception	70.07%	**92.44%**	60.16%	64.08%	49.94%	48.22%
ResNet50	46.27%	46.27%	60.41%	77.98%	46.27%	46.27%
InResNet	57.41%	72.23%	49.25%	47.42%	60.41%	81.20%
Xeption	66.83%	79.94%	60.97%	67.18%	63.85%	76.82%
VGG16	46.27%	46.28%	50.07%	59.24%	46.27%	46.27%
VGG19	46.27%	47.07%	57.98%	58.80%	46.26%	46.26%

The classification results of the CC200 atlas in Table 5.4 do not align with the theory that an increased number of ROIs leads to poor accuracy results. Many accuracy values exceeded 90%. Although the CC200 atlas did not achieve the same level of performance as the AAL atlas, it still ranked in the second position. The matrices of the CC200 atlas have a size of 200 × 200 × 3. Notably, the feature sizes vary between 4 × 4 for Inception and InceptionResNet models, 6 × 6 for VGG16 and VGG19 models, and 7 × 7 for all other models.

the Xception model achieved the highest accuracy of 95.52%, followed by VGG16 with 95.30%. Once again, the enhancement approach improved the results, although the overall performance was lower than with the AAL atlas. Additionally, the larger size of the CC200 matrices results in the longest execution time among all the atlases. As the feature sizes are the largest compared to the other atlases, they contribute to an extended analysis time. Therefore, considering both accuracy and time, the AAL atlas using the enhanced 3D correlation matrices remains the best combination.

TABLE 5.4: Accuracy of the classification of the CC200 3D matrices with and without enhancement.

	Correlation		Covariance		Tangent	
	3D Matrices	Enhanced 3D Matrices	3D Matrices	Enhanced 3D Matrices	3D Matrices	Enhanced 3D Matrices
ResNet152	86.72%	89.93%	61.56%	89.23%	78.90%	**91.52%**
Inception	88.89%	88.55%	72.34%	87.39%	87.40%	**92.78%**
ResNet50	46.27%	46.28%	84.99%	**94.38%**	46.28%	46.27%
InResNet	88.56%	**94.95%**	49.96%	52.93%	84.78%	**93.70%**
Xeption	88.33%	**95.52%**	73.27%	75.21%	87.86%	**93.01%**
VGG16	53.86%	**90.15%**	**93.68%**	**95.30%**	46.26%	51.67%
VGG19	69.60%	53.97%	73.85%	**90.13%**	46.04%	46.27%

Upon comparing the results from the three tables, several observations can be made. Firstly, ResNet50 consistently performed poorly in the classification of correlation and tangent matrices, but showed improvement when using the covariance method. On the other hand, ResNet152 consistently achieved high accuracy in classifying tangent matrices across all three atlases, with only a slight decrease in performance from the AAL atlas to the CC200 atlas. Inception also exhibited some consistency in performance across the atlases, particularly with the correlation method. However, a notable decrease in accuracy was observed when transitioning from the AAL atlas to the CC200 atlas.

These findings highlight the importance of the deep learning model structure in achieving accurate classification. However, it is evident that a single model structure cannot effectively handle all atlases and connectivity calculation methods.

Additionally, Figure 5.3 illustrates the classification results before and after fine-tuning. It is clear that fine-tuning improves the classification accuracy. Furthermore, the enhancement technique positively impacts the results even without fine-tuning, which aligns with the findings in Tables 5.2–5.4.

5.4 Discussion

The primary objective in this chapter was to assess the efficiency of the enhancement approach in detecting autism through a classification task. The approach involved two key steps: data preparation and application of deep learning for classification.

Staying in the same context as the other tests, we utilized the ABIDE I dataset as a base to construct our data. We explored various options, resulting in nine combinations derived from three atlases and three connectivity computation methods. Subsequently, we constructed 3D matrices using two strategies: one that preserved the original information divided into three parts and fed into a CNN model, and another that incorporated enhancement to address the connectivity deficits associated with autism, aiming to assist deep learning models in distinguishing between autistic and non-autistic subjects.

For the deep learning phase, the utilization of transfer learning in this study enabled the application of deep learning models even with a relatively small dataset of 871 subjects. Hence allowing us to leverage the advantages of deep learning.

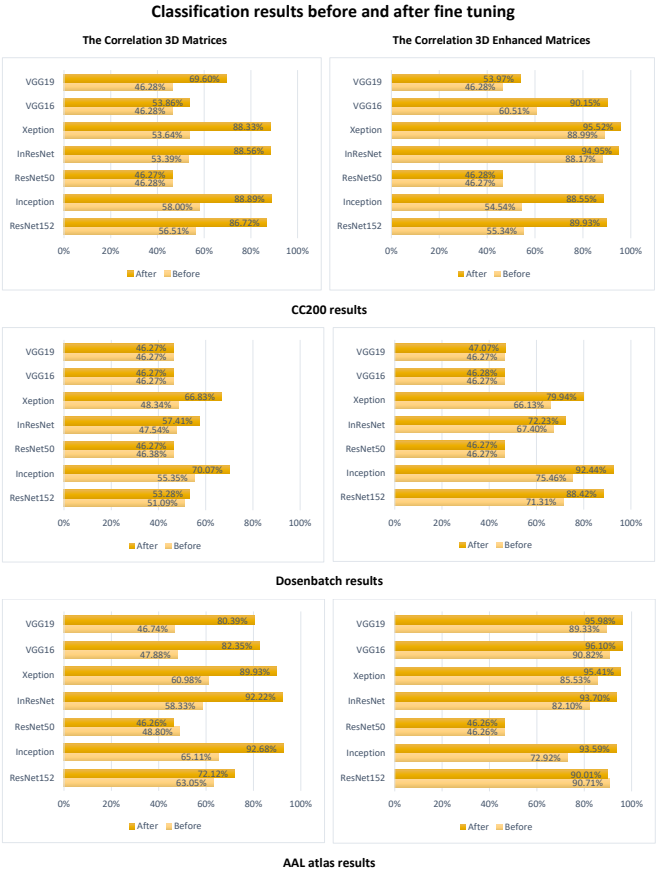

FIGURE 5.3: Classification before and after tuning.

Then, by considering the different combinations of atlases, connectivity methods, 3D creation approaches, and the seven CNN models, a total of 126 classification strategies were evaluated, resulting in 126 different accuracy values.

To enhance the reliability of the results and mitigate the risk of overfitting, we employed cross-validation and introduced fine-tuning by incorporating a dropout layer in the transfer learning models. This fine-tuning step further improved the accuracy of autism detection, as evidenced in Figure 5.3. Among the various combinations tested, the enhancement strategy proved to be particularly effective in increasing the accuracy of detection for most models, resulting in notable improvements in autism detection compared to previous studies. Without deep learning, the highest accuracy achieved in the literature was only 70% with the entire ABIDE I dataset [104, 144]. In contrast, our work achieved an accuracy of 90% by employing the RGB-mimicking approach to construct 3D matrices. Furthermore, the incorporation of the enhancement strategy further boosted the accuracy, with the highest achieved accuracy reaching 96%.

The combination of VGG models with the connectivity matrices of the AAL atlas, including correlation, covariance, and tangent, consistently yielded high accuracy results. This can be attributed to the structure of the VGG models, which do not have a large number of layers, and the relatively small number of ROIs in the AAL atlas. Consequently, the execution time for these combinations was notably faster compared to others.

Interestingly, the number of ROIs in an atlas did not show a direct correlation with classification accuracy. Both the CC200 atlas, with the highest number of ROIs, and the AAL atlas, with the smallest number of ROIs, yielded high accuracy values with most deep learning models. On the other hand, the Dosenbach atlas, which falls in between in terms of the number of ROIs, did not achieve exceptional results. This suggests that the choice and delineation of ROIs can significantly impact the classification outcomes, as even slight variations in the borders of these regions can lead to different findings or the exclusion of crucial information.
ResNet152V2 stood out as a model capable of achieving accuracy values above 91% when used with tangent enhanced matrices across all atlases. It consistently performed well in various enhancement scenarios and showcased versatility in its accuracy results.
In terms of the connectivity calculation methods, the tangent method yielded favorable outcomes in over half of the models across all atlases, except for the Dosenbach atlas, which showed good results only with three out of the seven models. The correlation method resulted in accuracy values above 80% in more than 70% of the cases, particularly excelling with the AAL atlas, where accuracy often surpassed 90%.
Furthermore, the results displayed in Figure 5.3 indicate that combining correlation with the AAL atlas already achieved high accuracy values even before fine-tuning. However, fine-tuning further improved performance in most cases, highlighting its effectiveness in enhancing the classification results.
These results highlight the significant impact of deep learning and the enhancement strategy in improving the accuracy of autism detection, even with a limited dataset size. The findings demonstrate the potential of leveraging advanced machine learning techniques to enhance our understanding of autism and improve early detection methodologies.
Moreover, the results obtained in this study highlight the importance of testing multiple combinations of models and methods to achieve accurate autism classification. No single model or method proved to be universally superior, emphasizing the need for careful exploration and evaluation.
With an unprecedented accuracy of 96.10%, this work surpasses previous achievements using the ABIDE dataset. The heterogeneity of the dataset, which is often seen as a challenge, actually became an advantage in identifying common characteristics among individuals with autism. Furthermore, by focusing on functional images rather than structural images, potential confounding factors such as differences in brain growth rates were minimized. Structural abnormalities in brain volume observed in children with autism tend to normalize as they reach adulthood. Therefore, investigating functional abnormalities provides more decisive and age-independent insights.
Farther more, the utilization of a large and diverse database like ABIDE I enhances the generalizability of the results, avoiding limitations typically associated with small datasets.
Overall, this study contributes to the ongoing efforts in understanding and detecting autism by demonstrating the significance of leveraging existing knowledge and conducting further investigations to enable early diagnosis of autism and mitigate its impact while awaiting a definitive cure.

5.5 Conclusion

This chapter has delved into the exploration of autism detection through an innovative enhancement approach. By seamlessly integrating deep learning methodologies and meticulous data curation from

the ABIDE I dataset, our study has yielded remarkable accuracy rates, peaking at an impressive 96%. The incorporation of enhancement strategies has proven pivotal in significantly elevating accuracy levels, eclipsing previous benchmarks. This achievement underscores the remarkable potential of advanced machine learning techniques, even in the face of constrained datasets, as well as the autism theories integration. Our findings shows the power of collaborative research and innovative methods in advancing autism detection, offering a promising path towards early diagnosis and intervention.
The diversity within the dataset, once viewed as a challenge, has served as an asset in identifying common characteristics among individuals with autism. This underscores the importance of leveraging large and varied datasets to enhance the reliability and generalizability of our findings. Our extensive assessment of different models, methods, and strategies underscores the nuanced nature of this task. As our understanding of autism continues to evolve, the collaborative synergy of advanced machine learning and comprehensive datasets will undoubtedly remain a cornerstone in the quest for earlier diagnosis, tailored interventions, and ultimately, improved quality of life for individuals on the autism spectrum.
In conclusion, this chapter has been dedicated to evaluating the effectiveness of an enhancement approach for detecting autism through a classification task. With an unprecedented accuracy of 96.10%, this work surpasses prior achievements using the ABIDE dataset, showcasing the significance of building on existing knowledge and conducting further investigations. This paves the way for early autism diagnosis and mitigation, contributing to the broader goal while awaiting definitive cures.

Chapter 6

Conclusion

Autism is a disorder that impacts a person's whole life and can lead to many problems as the difficulty to deal with the outside world, which may finally lead to isolation. Therefore, studying and understanding this disorder is very important. Especially, since specialists declared that early discovery of the disorder can be a good alternative in the present circumstances where no cure is available. Because an early diagnosis permits to put the child through an adapted program that can help lessening the disorder's severity. So, the most important task in parallel to finding a cure is to find a method of early diagnosis. This task is the objective of our study.

To achieve early diagnosis, there is a need for a mean that can be applied as early as possible to detect this disorder, such as the brain images. Taking into consideration the development pace that vary for every child and the fact that imaging techniques used for diagnosis should be non-invasive and safe for use on children, we decided to focus on the functional brain by studying the connectivity activity of the brain's regions to extract biomarkers that can permit the detection of autism. But, in the same time we wanted to take advantage of the existing autism theories, especially those tested on small data that prevented the generalization of their findings, to tailor the detection process more specifically to autism.

Neuroimaging studies have revealed both overconnectivity and underconnectivity in various brain regions and networks associated with social cognition, sensory processing, and executive functions.

Overconnectivity refers to the idea that certain brain regions or circuits in individuals with ASD exhibit heightened neural connections or activity compared to neurotypical individuals. It suggests an increased level of connectivity within specific circuits.

Going deeper into the specifics of this theory, individuals with ASD may exhibit stronger or more numerous connections between neighboring brain regions compared to typically developing individuals. This local overconnectivity theory suggests that there is an increased functional connectivity within specific local brain circuits in individuals with autism.

The idea behind local overconnectivity is that heightened connectivity within specific brain circuits may lead to an increased focus on local details and difficulties in integrating information across larger brain networks. This hyperconnectivity hypothesis suggests that individuals with autism may have a tendency to process and perceive information at a more local level, rather than integrating it into a broader context.

Underconnectivity, on the other hand, refers to reduced connectivity or weaker neural connections between brain regions in individuals with ASD compared to neurotypical individuals.

This theory also includes a variant, which is the Long-range underconnectivity, refering to a hypothesis that suggests individuals with ASD may exhibit reduced functional connectivity between distant brain regions. It is one of several proposed theories aiming to explain the neural basis of autism.
Long-range underconnectivity specifically refers to weaker functional connectivity between brain regions that are anatomically distant from each other. In typical brain functioning, distant brain regions often communicate and coordinate their activities through long-range connections, allowing for efficient information processing and integration. However, in individuals with autism, these long-range connections may be weaker or less efficient.
The hypothesis of long-range underconnectivity suggests that reduced connectivity between distant brain regions could lead to difficulties in integrating and coordinating information across different brain regions, resulting in the atypical cognitive and behavioral characteristics observed in individuals with autism. These difficulties may manifest as challenges in social interaction, communication, and the processing of complex information.
It is worth noting that both long-range underconnectivity and local overconnectivity are not mutually exclusive and may coexist in individuals with ASD. These connectivity abnormalities are part of a complex network-level dysregulation that can vary among individuals. The exact relationship between local overconnectivity and long-range underconnectivity, as well as their impact on the cognitive and behavioral features of autism, are still topics of ongoing research and investigation.
However, understanding the patterns of both local overconnectivity and long-range underconnectivity is essential for gaining insights into the neural mechanisms underlying autism and potentially developing targeted interventions or therapies in the future.
In fMRI, neural connections are identified by measuring the temporal correlation between fluctuations in the Blood-Oxygen-Level-Dependent (BOLD) signal of distinct anatomical regions [152]. The outcomes are 4-dimensional images that facilitate the visualization of brain structure and the study of functional activities in different brain regions. Two types of fMRI settings are commonly used: resting-state and task-based fMRI.
While task-based fMRI might be more suitable for achieving high classification accuracy, it faces challenges in acquiring large datasets. Typically, the sizes of task-based datasets consist of fewer than 50 subjects [153]. Moreover, task-based fMRI might not be well-suited for certain individuals with ASD, particularly those considered low-functioning or young children [24].
On the other hand, resting-state fMRI does not require a constrained experimental setup or the active and focused participation of the subjects. This makes it more accessible for a broader range of individuals, including those with ASD. Resting-state fMRI has also demonstrated its ability to capture interactions between brain regions that could serve as diagnostic biomarkers for neuropathology [25]. Various methods have been employed to study connectivity networks in resting-state fMRI images, including multi-echo independent component analysis (multi-echo ICA) [154], Group-ICA [155], and neighborhood one-class SVM [156]. These methods aim to derive descriptors for classifying subjects based on their connectivity patterns.
However, a significant challenge in previous studies lies in the limited size of the data, which hinders the generalization of descriptors, as highlighted by [157] and [158]. Researchers in [136] and [104] addressed this using the extensive ABIDE database.
In [136], the authors used fMRI images of 964 subjects to compute 7266×7266 connectivity matrices

by calculating the pair-wise correlation between each ROI. Then applied a leave-one-out approach, resulting in a classification accuracy of 60%.

Subsequently, in [104], the authors adopted a specific intra and inter site classification approach using cross-validation, achieving an accuracy of 67%. However, the pipeline did not specifically focus on autism-related connections, as the connectomes were utilized without any intermediate processing before classification.

Contrastingly, [159] highlighted the importance of considering autistic traits to achieve improved results in autism identification. In support of this notion, [160] demonstrated significant differences in connectivity patterns between healthy children and those with autism when local and long-distance deficits were taken into account. This underscores the significance of focusing on autism-specific connections and features to enhance the accuracy of classification methodologies.

Therefore, in this work, we focused on integrating autism tailored methods that have technical similarities with the above mentioned theories and conducted multiple tests based on multiple approaches with the objective of improving autism identification.

In doing so, we focused on the functional brain to construct brain communication networks which we first analyzed and compared to detect the autism-related over and under-connectivity deficits. In other words, we examined the brain connectivity activity to validate existing autism theories, while also extending them to encompass large and diverse datasets. To achieve this, we employed spanning trees, an unbiased graph theory tool that permits to extract specific information based on predefined conditions. In our case, we applied this tool to identify overconnectivity conditions associated with strong connections and underconnectivity conditions associated with weak connection values.

Following that, in our pursuit of validating and exploring the variations of these theories, we introduced novel methods and experimented with different combinations of tools. One such method was hierarchical clustering, which allowed us to test the hypotheses of local over-connectivity and long-range under-connectivity in ASD. This involved nullifying the connections both within and between clusters in the computed connectivity graphs. Additionally, we incorporated spanning trees into the investigation process using the elimination approach to further analyze these theories and their implications. These additional methods and approaches provided us with valuable insights and expanded our understanding of the neural connectivity patterns in individuals with autism spectrum disorder.

However, prior to delving into the investigation of the elimination strategy, we recognized the significance of the preprocessing strategy and its potential impact on the analysis. Therefore, we initiated our research by comparing various preprocessing strategies proposed within the preprocessing pipeline. Notably, we discovered that these preprocessing approaches significantly influenced the results, with differences of up to 3 %. This finding aligns with previous studies [31], emphasizing the crucial role of preprocessing in influencing test outcomes. The observed disparities in results among preprocessing strategies underscore the potential of this choice to enhance the identification of autism through imaging-based approaches.

After establishing the presence of overconnectivity and underconnectivity deficits in autism through the elimination approach, we subsequently modified our methodology by incorporating deep learning techniques.

Staying in the realm of machine learning, our initial approach employing deep learning involved extracting features from the data, which were subsequently used for classification purposes. As we

progressed, we further developed our methodology by constructing a comprehensive system based on deep learning. This system incorporated a novel enhancement approach that specifically targeted the identification of overconnectivity and underconnectivity patterns associated with autism. Through this approach, we achieved remarkable results, attaining the highest accuracy of 96%. This significant improvement in accuracy demonstrates the effectiveness of our enhanced deep learning system in accurately identifying autism based on the identified connectivity patterns.

In conclusion, based on our extensive experimentation and exploration of various tests and methods combinations in this study, we have come to the realization that there is no universally perfect method or tool for detecting autism. Our findings consistently demonstrated that the results varied significantly each time we altered our approach, even when working with the same underlying data and methods. Our research has demonstrated that methods initially deemed incompatible with the optimal combination can still yield remarkable results when thoroughly investigated and tested in different contexts and with different strategies. This underscores the importance of open-mindedness and thorough exploration in the field of autism detection. By embracing a diverse range of approaches and adapting them to specific contexts, we increase the likelihood of uncovering valuable insights and achieving accurate results in the identification of autism.

Hence, based on the insights gained from our research, pursuing multiple strategies and continuously exploring the theories of autism will be crucial for future advancements in early autism diagnosis. By incorporating a diverse range of tools and methodologies into the detection system, we can enhance our ability to identify and diagnose autism at an early stage. This approach will enable us to capture a broader range of potential indicators and manifestations of autism, leading to more accurate and timely diagnoses. Ultimately, the goal of early autism identification can be better achieved by continually expanding our knowledge, refining our detection systems, and leveraging multiple strategies to comprehensively address the complexity of autism spectrum disorder.

Expending a targeted investigation of over and under-connectivity, long range and short ranged connectivity, can provide valuable insights into the neural underpinnings of autism, elucidating the specific regions or circuits involved and their potential role in the development and manifestation of autism-related symptoms. It can contribute to a more nuanced understanding of the complexity of brain connectivity in autism and guide future research and interventions aimed at addressing the connectivity deficits of the autistic brain in a targeted and effective manner.

Bibliography

[1] Martha R. Herbert. 52 - autism. In Sid Gilman, editor, *Neurobiology of Disease*, pages 581–591. Academic Press, Burlington, 2007. ISBN 978-0-12-088592-3. doi: https://doi.org/10.1016/B978-012088592-3/50054-2. URL https://www.sciencedirect.com/science/article/pii/B9780120885923500542.

[2] Joan Stiles and Terry L. Jernigan. The basics of brain development. *Neuropsychology Review*, 20(4):327–348, November 2010. doi: 10.1007/s11065-010-9148-4. URL https://doi.org/10.1007/s11065-010-9148-4.

[3] Eric Courchesne, Peter R. Mouton, Michael E. Calhoun, Katerina Semendeferi, Clelia Ahrens-Barbeau, Melodie J. Hallet, Cynthia Carter Barnes, and Karen Pierce. Neuron number and size in prefrontal cortex of children with autism. *JAMA*, 306(18):2001, November 2011. doi: 10.1001/jama.2011.1638. URL https://doi.org/10.1001/jama.2011.1638.

[4] Heather Cody Hazlett, , Hongbin Gu, Brent C. Munsell, Sun Hyung Kim, Martin Styner, Jason J. Wolff, Jed T. Elison, Meghan R. Swanson, Hongtu Zhu, Kelly N. Botteron, D. Louis Collins, John N. Constantino, Stephen R. Dager, Annette M. Estes, Alan C. Evans, Vladimir S. Fonov, Guido Gerig, Penelope Kostopoulos, Robert C. McKinstry, Juhi Pandey, Sarah Paterson, John R. Pruett, Robert T. Schultz, Dennis W. Shaw, Lonnie Zwaigenbaum, and Joseph Piven. Early brain development in infants at high risk for autism spectrum disorder. *Nature*, 542 (7641):348–351, February 2017. doi: 10.1038/nature21369. URL https://doi.org/10.1038/nature21369.

[5] Elliot M. Tucker-Drob and K. Paige Harden. Gene-by-preschool interaction on the development of early externalizing problems. *Journal of Child Psychology and Psychiatry*, 54(1):77–85, July 2012. doi: 10.1111/j.1469-7610.2012.02578.x. URL https://doi.org/10.1111/j.1469-7610.2012.02578.x.

[6] Kimm J.E. van Hulzen, Claus J. Scholz, Barbara Franke, Stephan Ripke, Marieke Klein, Andrew McQuillin, Edmund J. Sonuga-Barke, John R. Kelsoe, Mikael Landén, Ole A. Andreassen, Klaus-Peter Lesch, Heike Weber, Stephen V. Faraone, Alejandro Arias-Vasquez, and Andreas Reif. Genetic overlap between attention-deficit/hyperactivity disorder and bipolar disorder: Evidence from genome-wide association study meta-analysis. *Biological Psychiatry*, 82(9):634–641, November 2017. doi: 10.1016/j.biopsych.2016.08.040. URL https://doi.org/10.1016/j.biopsych.2016.08.040.

[7] Claire Amiet, Isabelle Gourfinkel-An, Anissa Bouzamondo, Sylvie Tordjman, Michel Baulac, Philippe Lechat, Laurent Mottron, and David Cohen. Epilepsy in autism is associated with intellectual disability and gender: Evidence from a meta-analysis. *Biological Psychiatry*, 64(7): 577–582, October 2008. doi: 10.1016/j.biopsych.2008.04.030. URL https://doi.org/10.1016/j.biopsych.2008.04.030.

[8] Emily Simonoff, Andrew Pickles, Tony Charman, Susie Chandler, Tom Loucas, and Gillian Baird. Psychiatric disorders in children with autism spectrum disorders: Prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(8):921–929, August 2008. doi: 10.1097/chi.0b013e318179964f. URL https://doi.org/10.1097/chi.0b013e318179964f.

[9] Barbara Maughan. Editorial: Counting the cost: estimating the burden of child mental health. *Journal of Child Psychology and Psychiatry*, 54(12):1261–1262, November 2013. doi: 10.1111/jcpp.12179. URL https://doi.org/10.1111/jcpp.12179.

[10] Ovsanna T. Leyfer, Susan E. Folstein, Susan Bacalman, Naomi O. Davis, Elena Dinh, Jubel Morgan, Helen Tager-Flusberg, and Janet E. Lainhart. Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, 36(7):849–861, July 2006. doi: 10.1007/s10803-006-0123-0. URL https://doi.org/10.1007/s10803-006-0123-0.

[11] Eva Billstedt, I. Carina Gillberg, and Christopher Gillberg. Autism in adults: symptom patterns and early childhood predictors. use of the DISCO in a community sample followed from childhood. *Journal of Child Psychology and Psychiatry*, 48(11):1102–1110, November 2007. doi: 10.1111/j.1469-7610.2007.01774.x. URL https://doi.org/10.1111/j.1469-7610.2007.01774.x.

[12] Julie Lounds Taylor and Marsha Mailick Seltzer. Changes in the autism behavioral phenotype during the transition to adulthood. *Journal of Autism and Developmental Disorders*, 40(12): 1431–1446, April 2010. doi: 10.1007/s10803-010-1005-z. URL https://doi.org/10.1007/s10803-010-1005-z.

[13] Amy M Daniels and David S Mandell. Explaining differences in age at autism spectrum disorder diagnosis: A critical review. *Autism*, 18(5):583–597, June 2013. doi: 10.1177/1362361313480277. URL https://doi.org/10.1177/1362361313480277.

[14] Fred R Volkmar and Brian Reichow. Autism in DSM-5: progress and challenges. *Molecular Autism*, 4(1):13, 2013. doi: 10.1186/2040-2392-4-13. URL https://doi.org/10.1186/2040-2392-4-13.

[15] Lynn Waterhouse. Chapter 7 - increasing prevalence and the problem of diagnosis. In Lynn Waterhouse, editor, *Rethinking Autism*, pages 345–398. Academic Press, San Diego, 2013. ISBN 978-0-12-415961-7. doi: https://doi.org/10.1016/B978-0-12-415961-7.00007-1. URL https://www.sciencedirect.com/science/article/pii/B9780124159617000071.

[16] Josephine Barbaro and Cheryl Dissanayake. Diagnostic stability of autism spectrum disorder in toddlers prospectively identified in a community-based setting: Behavioural characteristics and predictors of change over time. *Autism*, 21(7):830–840, July 2016. doi: 10.1177/1362361316654084. URL https://doi.org/10.1177/1362361316654084.

[17] Chris PlauchÉ Johnson and Scott M. Myers. Autism spectrum disorders. In *Developmental-Behavioral Pediatrics*, pages 519–577. Elsevier, 2008. doi: 10.1016/b978-0-323-04025-9.50018-0. URL https://doi.org/10.1016/b978-0-323-04025-9.50018-0.

[18] Joan A. Camprodon and Theodore A. Stern. Selecting neuroimaging techniques. *The Primary Care Companion For CNS Disorders*, August 2013. doi: 10.4088/pcc.12f01490. URL https://doi.org/10.4088/pcc.12f01490.

[19] Alessandro Grecucci, Roma Siugzdaite, and Remo Job. Editorial: Advanced neuroimaging methods for studying autism disorder. *Frontiers in Neuroscience*, 11, September 2017. doi: 10.3389/fnins.2017.00533. URL https://doi.org/10.3389/fnins.2017.00533.

[20] Yuhui Du, Zening Fu, and Vince D. Calhoun. Classification and prediction of brain disorders using functional connectivity: Promising but challenging. *Frontiers in Neuroscience*, 12, August 2018. doi: 10.3389/fnins.2018.00525. URL https://doi.org/10.3389/fnins.2018.00525.

[21] Abul Hussain. Growth and development of children. *SSRN Electronic Journal*, 2021. doi: 10.2139/ssrn.3771757. URL https://doi.org/10.2139/ssrn.3771757.

[22] Till Nierhaus, Daniel Margulies, Xiangyu Long, and Arno Villringer. fMRI for the assessment of functional connectivity. In *Neuroimaging - Methods*. InTech, February 2012. doi: 10.5772/23864. URL https://doi.org/10.5772/23864.

[23] U. Frith. *Autism: A Very Short Introduction*. EBSCO ebook academic collection. OUP Oxford, 2008. ISBN 9780199207565. URL https://books.google.co.ma/books?id=j40SDAAAQBAJ.

[24] Way K. W. Lau, Mei-Kei Leung, and Benson W. M. Lau. Resting-state abnormalities in autism spectrum disorders: A meta-analysis. *Scientific Reports*, 9(1), March 2019. doi: 10.1038/s41598-019-40427-7. URL https://doi.org/10.1038/s41598-019-40427-7.

[25] Michael Greicius. Resting-state functional connectivity in neuropsychiatric disorders. *Current Opinion in Neurology*, 21(4):424–430, August 2008. doi: 10.1097/wco.0b013e328306f2c5. URL https://doi.org/10.1097/wco.0b013e328306f2c5.

[26] A. Havdahl, M. Niarchou, A. Starnawska, M. Uddin, C. van der Merwe, and V. Warrier. Genetic contributions to autism spectrum disorder. *Psychological Medicine*, 51(13):2260–2273, February 2021. doi: 10.1017/s0033291721000192. URL https://doi.org/10.1017/s0033291721000192.

[27] Sherlly Xie, Håkan Karlsson, Christina Dalman, Linnea Widman, Dheeraj Rai, Renee M. Gardner, Cecilia Magnusson, Sven Sandin, Loni P. Tabb, Craig J. Newschaffer, and Brian K. Lee. The familial risk of autism spectrum disorder with and without intellectual disability. *Autism Research*, 13(12):2242–2250, October 2020. doi: 10.1002/aur.2417. URL https://doi.org/10.1002/aur.2417.

[28] Kealan Pugsley, Stephen W. Scherer, Mark A. Bellgrove, and Ziarih Hawi. Environmental exposures associated with elevated risk for autism spectrum disorder may augment the burden of deleterious de novo mutations among probands. *Molecular Psychiatry*, 27(1):710–730, May 2021. doi: 10.1038/s41380-021-01142-w. URL https://doi.org/10.1038/s41380-021-01142-w.

[29] Mojgan Karahmadi, Padideh Karimi, Elahe Kamali, and SeyyedMohammad Mousavi. Environmental factors influencing the risk of autism. *Journal of Research in Medical Sciences*, 22(1):27, 2017. doi: 10.4103/1735-1995.200272. URL https://doi.org/10.4103/1735-1995.200272.

[30] John D. Lewis, Alan C. Evans, John R. Pruett, Kelly N. Botteron, Robert C. McKinstry, Lonnie Zwaigenbaum, Annette M. Estes, D. Louis Collins, Penelope Kostopoulos, Guido Gerig, Stephen R. Dager, Sarah Paterson, Robert T. Schultz, Martin A. Styner, Heather C. Hazlett, Joseph Piven, J. Piven, H.C. Hazlett, C. Chappell, S.R. Dager, A.M. Estes, D. Shaw, K.N. Botteron, R.C. McKinstry, J. Constantino, J.R. Pruett, R.T. Schultz, S. Paterson, L. Zwaigenbaum, J.T. Elison, A.C. Evans, D.L. Collins, G.B. Pike, V. Fonov, P. Kostopoulos, S. Das, G. Gerig, M.A. Styner, and H. Gu. The emergence of network inefficiencies in infants with autism spectrum disorder. *Biological Psychiatry*, 82(3):176–185, 2017. ISSN 0006-3223. doi: https://doi.org/10.1016/j.biopsych.2017.03.006. URL https://www.sciencedirect.com/science/article/pii/S0006322317313616. Alterations in Cortical Development in Autism Spectrum Disorder.

[31] S.C. Strother. Evaluating fMRI preprocessing pipelines. *IEEE Engineering in Medicine and Biology Magazine*, 25(2):27–41, March 2006. doi: 10.1109/memb.2006.1607667. URL https://doi.org/10.1109/memb.2006.1607667.

[32] Jocelyn V. Hull, Lisa B. Dokovna, Zachary J. Jacokes, Carinna M. Torgerson, Andrei Irimia, and John Darrell Van Horn. Resting-state functional connectivity in autism spectrum disorders: A review. *Frontiers in Psychiatry*, 7, jan 2017. doi: 10.3389/fpsyt.2016.00205. URL https://doi.org/10.3389%2Ffpsyt.2016.00205.

[33] E. BLEULER and A. VIALLARD. *Dementia Praecox ou Groupe des schizophrénies*. Essais. Epel Éditions, 2012. ISBN 9782354270414. URL https://books.google.co.ma/books?id=jCy1DgAAQBAJ.

[34] James Harris. Leo kanner and autism: a 75-year perspective. *International Review of Psychiatry*, 30(1):3–17, January 2018. doi: 10.1080/09540261.2018.1455646. URL https://doi.org/10.1080/09540261.2018.1455646.

[35] Nicole E. Rosen, Catherine Lord, and Fred R. Volkmar. The diagnosis of autism: From kanner to DSM-III to DSM-5 and beyond. *Journal of Autism and Developmental Disorders*, 51(12):4253–4270, February 2021. doi: 10.1007/s10803-021-04904-1. URL https://doi.org/10.1007/s10803-021-04904-1.

[36] Roberto Canitano and Mauro Pallagrosi. Autism spectrum disorders and schizophrenia spectrum disorders: Excitation/inhibition imbalance and developmental trajectories. *Frontiers in Psychiatry*, 8, May 2017. doi: 10.3389/fpsyt.2017.00069. URL https://doi.org/10.3389/fpsyt.2017.00069.

[37] I. Kolvin. Studies in the childhood psychoses i. diagnostic criteria and classification. *British Journal of Psychiatry*, 118(545):381–384, April 1971. doi: 10.1192/bjp.118.545.381. URL https://doi.org/10.1192/bjp.118.545.381.

[38] Stefano Barlati, Giacomo Deste, Cassandra Ariu, and Antonio Vita. Autism spectrum disorder and schizophrenia: Do they overlap? *International Journal of Emergency Mental Health and Human Resilience*, 18:760–763, 04 2016. doi: 10.4172/1522-4821.1000318.

[39] Stephen D. Luke. National dissemination center for children with disabilities. In *Encyclopedia of Clinical Neuropsychology*, pages 1707–1707. Springer New York, 2011. doi: 10.1007/978-0-387-79948-3_1573. URL https://doi.org/10.1007/978-0-387-79948-3_1573.

[40] June Sprock. *DSM-III and DSM-III-R*. 01 2015. doi: 10.1002/9781118625392.wbecp240.

[41] Glyn Lewis. DSM-IV. diagnostic and statistical manual of mental disorders, 4th edn. by the american psychiatric association. (pp. 886 34.95.) APA: Washington, DC.1994. *Psychological Medicine*, 26(3):651–652, May 1996. doi: 10.1017/s0033291700035765. URL https://doi.org/10.1017/s0033291700035765.

[42] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, May 2013. doi: 10.1176/appi.books.9780890425596. URL https://doi.org/10.1176/appi.books.9780890425596.

[43] Edward R. Ritvo and Betty Jo Freeman. Introduction: The national society for autistic children's definition of the syndrome of autism. *Journal of the American Academy of Child Psychiatry*, 17(4):565–575, September 1978. doi: 10.1016/s0002-7138(09)61011-6. URL https://doi.org/10.1016/s0002-7138(09)61011-6.

[44] I. Kolvin. Studies in the childhood psychoses i. diagnostic criteria and classification. *British Journal of Psychiatry*, 118(545):381–384, April 1971. doi: 10.1192/bjp.118.545.381. URL https://doi.org/10.1192/bjp.118.545.381.

[45] Michael Rutter. Maternal deprivation, 1972-1978: New findings, new concepts, new approaches. *Child Development*, 50(2):283, June 1979. doi: 10.2307/1129404. URL https://doi.org/10.2307/1129404.

[46] Lawrence Bartak and Michael Rutter. SPECIAL EDUCATIONAL TREATMENT OF AUTISTIC CHILDREN: A COMPARATIVE STUDY?1. DESIGN OF STUDY AND CHARACTERISTICS OF UNITS. *Journal of Child Psychology and Psychiatry*, 14(3):161–179, September 1973. doi: 10.1111/j.1469-7610.1973.tb01185.x. URL https://doi.org/10.1111/j.1469-7610.1973.tb01185.x.

[47] Susan Folstein and Michael Rutter. INFANTILE AUTISM: A GENETIC STUDY OF 21 TWIN PAIRS. *Journal of Child Psychology and Psychiatry*, 18(4):297–321, September 1977. doi: 10.1111/j.1469-7610.1977.tb00443.x. URL https://doi.org/10.1111/j.1469-7610.1977.tb00443.x.

[48] Naustin Deslauriers. The empty fortress: Infantile autism and the birth of the self. *Archives of General Psychiatry*, 17(4):510, October 1967. doi: 10.1001/archpsyc.1967.01730280126018. URL https://doi.org/10.1001/archpsyc.1967.01730280126018.

[49] P. Bebbington. Welcome to ICD-10. *Social Psychiatry and Psychiatric Epidemiology*, 27(6): 255–257, nov 1992. doi: 10.1007/bf00788894. URL https://doi.org/10.1007%2Fbf00788894.

[50] Lorna Wing and Judith Gould. Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, 9(1):11–29, March 1979. doi: 10.1007/bf01531288. URL https://doi.org/10.1007/bf01531288.

[51] Stephan Ehlers and Christopher Gillberg. The epidemiology of asperger syndrome. *Journal of Child Psychology and Psychiatry*, 34(8):1327–1350, November 1993. doi: 10.1111/j.1469-7610.1993.tb02094.x. URL https://doi.org/10.1111/j.1469-7610.1993.tb02094.x.

[52] Sula Wolff. The history of autism. *European Child and Adolescent Psychiatry*, 13(4), August 2004. doi: 10.1007/s00787-004-0363-5. URL https://doi.org/10.1007/s00787-004-0363-5.

[53] Julie Green, Faline Howes, Elizabeth Waters, Elise Maher, and Frank Oberklaid. Promoting the social and emotional health of primary school-aged children: Reviewing the evidence base for school-based interventions. *The International Journal of Mental Health Promotion*, 7:30–36, 08 2005. doi: 10.1080/14623730.2005.9721872.

[54] Gillian Baird, Emily Simonoff, Andrew Pickles, Susie Chandler, Tom Loucas, David Meldrum, and Tony Charman. Prevalence of disorders of the autism spectrum in a population cohort of children in south thames: the special needs and autism project (SNAP). *The Lancet*, 368(9531): 210–215, July 2006. doi: 10.1016/s0140-6736(06)69041-7. URL https://doi.org/10.1016/s0140-6736(06)69041-7.

[55] Stephen Blumberg, Matthew Bramlett, Michael Kogan, Laura Schieve, Jessica Jones, and Michael Lu. Changes in prevalence of parent-reported autism spectrum disorder in school-aged u.s. children: 2007 to 2011-2012. *National health statistics reports*, 65:1–11, 1 p following 11, 03 2013.

[56] Young Shin Kim, Bennett L. Leventhal, Yun-Joo Koh, Eric Fombonne, Eugene Laska, Eun-Chung Lim, Keun-Ah Cheon, Soo-Jeong Kim, Young-Key Kim, HyunKyung Lee, Dong-Ho Song, and Roy Richard Grinker. Prevalence of autism spectrum disorders in a total population sample. *American Journal of Psychiatry*, 168(9):904–912, September 2011. doi: 10.1176/appi.ajp.2011.10101532. URL https://doi.org/10.1176/appi.ajp.2011.10101532.

[57] Mayada Elsabbagh, Gauri Divan, Yun-Joo Koh, Yongbaek Kim, Shuaib Kauchali, Carlos Marcin, Cecilia Montiel-nava, Vikram Patel, Cristiane Paula, Chongying Wang, Mohammad Taghi Yasamy, and Eric Fombonne. Global prevalence of autism and other pervasive developmental disorders. *Autism Research*, 04 2012. doi: 10.1002/aur.239.

[58] Matthew J. Maenner, Zachary Warren, Ashley Robinson Williams, Esther Amoakohene, Amanda V. Bakian, Deborah A. Bilder, Maureen S. Durkin, Robert T. Fitzgerald, Sarah M. Furnier, Michelle M. Hughes, Christine M. Ladd-Acosta, Dedria McArthur, Elise T. Pas, Angelica Salinas, Alison Vehorn, Susan Williams, Amy Esler, Andrea Grzybowski, Jennifer Hall-Lande, Ruby H.N. Nguyen, Karen Pierce, Walter Zahorodny, Allison Hudson, Libby Hallas, Kristen Clancy Mancilla, Mary Patrick, Josephine Shenouda, Kate Sidwell, Monica DiRienzo, Johanna Gutierrez, Margaret H. Spivey, Maya Lopez, Sydney Pettygrove, Yvette D. Schwenk, Anita Washington, and Kelly A. Shaw. Prevalence and characteristics of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, united states, 2020. *MMWR. Surveillance Summaries*, 72(2):1–14, March 2023. ISSN 1545-8636. doi: 10.15585/mmwr.ss7202a1. URL http://dx.doi.org/10.15585/mmwr.ss7202a1.

[59] Ilse Noens and G. Nygren. Comparisons between the disco and the adi-r and the ados. 05 2012.

[60] Geerte Slappendel, William Mandy, Jan van der Ende, Frank C. Verhulst, Ad van der Sijde, Jorieke Duvekot, David Skuse, and Kirstin Greaves-Lord. Utility of the 3di short version for the diagnostic assessment of autism spectrum disorder and compatibility with DSM-5. *Journal of Autism and Developmental Disorders*, 46(5):1834–1846, jan 2016. doi: 10.1007/s10803-016-2713-9. URL https://doi.org/10.1007%2Fs10803-016-2713-9.

[61] Tamara Dawkins, Allison T. Meyer, and Mary E. Van Bourgondien. The relationship between the childhood autism rating scale: Second edition and clinical diagnosis utilizing the DSM-IV-TR and the DSM-5. *Journal of Autism and Developmental Disorders*, 46(10):3361–3368, jul 2016. doi: 10.1007/s10803-016-2860-z. URL https://doi.org/10.1007%2Fs10803-016-2860-z.

[62] Daniel H. Geschwind. Genetics of autism spectrum disorders. *Trends in Cognitive Sciences*, 15 (9):409–416, sep 2011. doi: 10.1016/j.tics.2011.07.003. URL https://doi.org/10.1016%2Fj.tics.2011.07.003.

[63] Rachel K. Schuck, Ryan E. Flores, and Lawrence K. Fung. Brief report: Sex/gender differences in symptomology and camouflaging in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 49(6):2597–2604, apr 2019. doi: 10.1007/s10803-019-03998-y. URL https://doi.org/10.1007%2Fs10803-019-03998-y.

[64] Niamh Dooley, Amber Ruigrok, Rosemary Holt, Carrie Allison, Alexandros Tsompanidis, Jack Waldman, Bonnie Auyeung, Michael V. Lombardo, and Simon Baron-Cohen. Is there an association between prenatal testosterone and autistic traits in adolescents? *Psychoneuroendocrinology*, 136:105623, feb 2022. doi: 10.1016/j.psyneuen.2021.105623. URL https://doi.org/10.1016%2Fj.psyneuen.2021.105623.

[65] Kholoud N. Bastaki, Sura Alwan, and Farah R. Zahir. Maternal prenatal exposures in pregnancy and autism spectrum disorder: An insight into the epigenetics of drugs and diet as key environmental influences. In *Advances in Neurobiology*, pages 143–162. Springer International Publishing, 2020. doi: 10.1007/978-3-030-30402-7_5. URL https://doi.org/10.1007%2F978-3-030-30402-7_5.

[66] Manoj Kumar, Marwa Saadaoui, and Souhaila Al Khodor. Infections and pregnancy: Effects on maternal and child health. *Frontiers in Cellular and Infection Microbiology*, 12, jun 2022. doi: 10.3389/fcimb.2022.873253. URL https://doi.org/10.3389%2Ffcimb.2022.873253.

[67] Md Mostafijur Rahman, Yu-Hsiang Shu, Ting Chow, Frederick W. Lurmann, Xin Yu, Mayra P. Martinez, Sarah A. Carter, Sandrah P. Eckel, Jiu-Chiuan Chen, Zhanghua Chen, Pat Levitt, Joel Schwartz, Rob McConnell, and Anny H. Xiang. Prenatal exposure to air pollution and autism spectrum disorder: Sensitive windows of exposure and sex differences. *Environmental Health Perspectives*, 130(1), jan 2022. doi: 10.1289/ehp9509. URL https://doi.org/10.1289%2Fehp9509.

[68] PreetiA Devnani and AnaitaU Hegde. Autism and sleep disorders. *Journal of Pediatric Neurosciences*, 10(4):304, 2015. doi: 10.4103/1817-1745.174438. URL https://doi.org/10.4103/1817-1745.174438.

[69] Camille Hours, Christophe Recasens, and Jean-Marc Baleyte. ASD and ADHD comorbidity: What are we talking about? *Frontiers in Psychiatry*, 13, February 2022. doi: 10.3389/fpsyt.2022.837424. URL https://doi.org/10.3389/fpsyt.2022.837424.

[70] Frank Besag. Epilepsy in patients with autism: links, risks and treatment challenges. *Neuropsychiatric Disease and Treatment*, Volume 14:1–10, December 2017. doi: 10.2147/ndt.s120509. URL https://doi.org/10.2147/ndt.s120509.

[71] Brian A Zaboski and Eric A Storch. Comorbid autism spectrum disorder and anxiety disorders: a brief review. *Future Neurology*, 13(1):31–37, feb 2018. doi: 10.2217/fnl-2017-0030. URL https://doi.org/10.2217%2Ffnl-2017-0030.

[72] Katherine M. Magnuson and John N. Constantino. Characterization of depression in children with autism spectrum disorders. *Journal of Developmental and Behavioral Pediatrics*, 32(4):332–340, may 2011. doi: 10.1097/dbp.0b013e318213f56c. URL https://doi.org/10.1097%2Fdbp.0b013e318213f56c.

[73] Bert Müller, Anna Khimchenko, Griffin Rodgers, Melissa Osterwalder, Christine Tanner, and Georg Schulz. X-ray imaging of human brain tissue down to the molecule level. In Davide Bleiner, editor, *International Conference on X-Ray Lasers 2020*. SPIE, jul 2021. doi: 10.1117/12.2592603. URL https://doi.org/10.1117%2F12.2592603.

[74] *Computed Tomography*. Springer Berlin Heidelberg, 2008. doi: 10.1007/978-3-540-39408-2. URL https://doi.org/10.1007%2F978-3-540-39408-2.

[75] Vijay P.B. Grover, Joshua M. Tognarelli, Mary M.E. Crossey, I. Jane Cox, Simon D. Taylor-Robinson, and Mark J.W. McPhail. Magnetic resonance imaging: Principles and techniques: Lessons for clinicians. *Journal of Clinical and Experimental Hepatology*, 5(3):246–255, sep 2015. doi: 10.1016/j.jceh.2015.08.001. URL https://doi.org/10.1016%2Fj.jceh.2015.08.001.

[76] Ruth A. Carper and Eric Courchesne. Localized enlargement of the frontal cortex in early autism. *Biological Psychiatry*, 57(2):126–133, jan 2005. doi: 10.1016/j.biopsych.2004.11.005. URL https://doi.org/10.1016%2Fj.biopsych.2004.11.005.

[77] Antonio Y. Hardan, Sri Muddasani, Madhuri Vemulapalli, Matcheri S. Keshavan, and Nancy J. Minshew. An MRI study of increased cortical thickness in autism. *American Journal of Psychiatry*, 163(7):1290–1292, jul 2006. doi: 10.1176/ajp.2006.163.7.1290. URL https://doi.org/10.1176%2Fajp.2006.163.7.1290.

[78] Thomas A. Holly, Brian G. Abbott, Mouaz Al-Mallah, Dennis A. Calnon, Mylan C. Cohen, Frank P. DiFilippo, Edward P. Ficaro, Michael R. Freeman, Robert C. Hendel, Diwakar Jain, Scott M. Leonard, Kenneth J. Nichols, Donna M. Polk, and Prem Soman. Single photon-emission computed tomography. *Journal of Nuclear Cardiology*, 17(5):941–973, jun 2010. doi: 10.1007/s12350-010-9246-y. URL https://doi.org/10.1007%2Fs12350-010-9246-y.

[79] Katherine Lameka, Michael D. Farwell, and Masanori Ichise. Positron emission tomography. In *Handbook of Clinical Neurology*, pages 209–227. Elsevier, 2016. doi: 10.1016/b978-0-444-53485-9.00011-8. URL https://doi.org/10.1016%2Fb978-0-444-53485-9.00011-8.

[80] Sibylle I. Ziegler. Positron emission tomography: Principles, technology, and recent developments. *Nuclear Physics A*, 752:679–687, April 2005. doi: 10.1016/j.nuclphysa.2005.02.067. URL https://doi.org/10.1016/j.nuclphysa.2005.02.067.

[81] Joachim Gross. Magnetoencephalography in cognitive neuroscience: A primer. *Neuron*, 104(2):189–204, oct 2019. doi: 10.1016/j.neuron.2019.07.001. URL https://doi.org/10.1016%2Fj.neuron.2019.07.001.

[82] Bradley R. Buchbinder. Functional magnetic resonance imaging. In *Handbook of Clinical Neurology*, pages 61–92. Elsevier, 2016. doi: 10.1016/b978-0-444-53485-9.00004-0. URL https://doi.org/10.1016%2Fb978-0-444-53485-9.00004-0.

[83] Mehdi Poustchi-Amin, Scott A. Mirowitz, Jeffrey J. Brown, Robert C. McKinstry, and Tao Li. Principles and applications of echo-planar imaging: A review for the general radiologist. *RadioGraphics*, 21(3):767–779, may 2001. doi: 10.1148/radiographics.21.3.g01ma23767. URL https://doi.org/10.1148%2Fradiographics.21.3.g01ma23767.

[84] Megha Rajendran, Eric Dane, Jason Conley, and Mathew Tantama. Imaging adenosine triphosphate (ATP). *The Biological Bulletin*, 231(1):73–84, aug 2016. doi: 10.1086/689592. URL https://doi.org/10.1086%2F689592.

[85] K Vincent, J Moore, S Kennedy, and I Tracey. Blood oxygenation level dependent functional magnetic resonance imaging: current and potential uses in obstetrics and gynaecology. *BJOG: An International Journal of Obstetrics $ Gynaecology*, 116(2):240–246, dec 2008. doi: 10.1111/j.1471-0528.2008.01993.x. URL https://doi.org/10.1111%2Fj.1471-0528.2008.01993.x.

[86] A Di Martino, C-G Yan, Q Li, E Denio, F X Castellanos, K Alaerts, J S Anderson, M Assaf, S Y Bookheimer, M Dapretto, B Deen, S Delmonte, I Dinstein, B Ertl-Wagner, D A Fair, L Gallagher, D P Kennedy, C L Keown, C Keysers, J E Lainhart, C Lord, B Luna, V Menon, N J Minshew, C S Monk, S Mueller, R-A Müller, M B Nebel, J T Nigg, K O'Hearn, K A Pelphrey, S J Peltier, J D Rudie, S Sunaert, M Thioux, J M Tyszka, L Q Uddin, J S Verhoeven,

N Wenderoth, J L Wiggins, S H Mostofsky, and M P Milham. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 19(6):659–667, June 2013. doi: 10.1038/mp.2013.78. URL https://doi.org/10.1038/mp.2013.78.

[87] Ebubeogu Amarachukwu Felix and Sai Peck Lee. Systematic literature review of preprocessing techniques for imbalanced data. *IET Software*, 13(6):479–496, December 2019. doi: 10.1049/iet-sen.2018.5193. URL https://doi.org/10.1049/iet-sen.2018.5193.

[88] Ross M. Lawrence, Eric W. Bridgeford, Patrick E. Myers, Ganesh C. Arvapalli, Sandhya C. Ramachandran, Derek A. Pisner, Paige F. Frank, Allison D. Lemmer, Aki Nikolaidis, and Joshua T. Vogelstein. Standardizing human brain parcellations. *Scientific Data*, 8(1), March 2021. doi: 10.1038/s41597-021-00849-3. URL https://doi.org/10.1038/s41597-021-00849-3.

[89] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15(1):273–289, January 2002. doi: 10.1006/nimg.2001.0978. URL https://doi.org/10.1006/nimg.2001.0978.

[90] Nico UF Dosenbach, Binyam Nardos, Alexander L Cohen, Damien A Fair, Jonathan D Power, Jessica A Church, Steven M Nelson, Gagan S Wig, Alecia C Vogel, Christina N Lessov-Schlaggar, et al. Prediction of individual brain maturity using fmri. *Science*, 329(5997):1358–1361, 2010.

[91] R. Cameron Craddock, G.Andrew James, Paul E. Holtzheimer, Xiaoping P. Hu, and Helen S. Mayberg. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, 33(8):1914–1928, July 2011. doi: 10.1002/hbm.21333. URL https://doi.org/10.1002/hbm.21333.

[92] K. Amunts, M.J. Hawrylycz, D.C. Van Essen, J.D. Van Horn, N. Harel, J.-B. Poline, F. De Martino, J.G. Bjaalie, G. Dehaene-Lambertz, S. Dehaene, P. Valdes-Sosa, B. Thirion, K. Zilles, S.L. Hill, M.B. Abrams, P.A. Tass, W. Vanduffel, A.C. Evans, and S.B. Eickhoff. Interoperable atlases of the human brain. *NeuroImage*, 99:525–532, 2014. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2014.06.010. URL https://www.sciencedirect.com/science/article/pii/S105381191400487X.

[93] Nikos Makris, Jill M. Goldstein, David Kennedy, Steven M. Hodge, Verne S. Caviness, Stephen V. Faraone, Ming T. Tsuang, and Larry J. Seidman. Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophrenia Research*, 83(2-3):155–171, April 2006. doi: 10.1016/j.schres.2005.11.020. URL https://doi.org/10.1016/j.schres.2005.11.020.

[94] Wieslaw L Nowinski and Dmitry Belov. The cerefy neuroradiology atlas: a talairach–tournoux atlas-based tool for analysis of neuroimages available over the internet. *NeuroImage*, 20(1):50–57, 2003. ISSN 1053-8119. doi: https://doi.org/10.1016/S1053-8119(03)00252-0. URL https://www.sciencedirect.com/science/article/pii/S1053811903002520.

[95] Robert H. Shumway and David S. Stoffer. *Time Series Analysis and Its Applications*. Springer International Publishing, 2017. doi: 10.1007/978-3-319-52452-8. URL https://doi.org/10.1007/978-3-319-52452-8.

[96] Ben D. Fulcher, Max A. Little, and Nick S. Jones. Highly comparative time-series analysis: the empirical structure of time series and their methods. *Journal of The Royal Society Interface*, 10 (83):20130048, June 2013. doi: 10.1098/rsif.2013.0048. URL https://doi.org/10.1098/rsif.2013.0048.

[97] Takeru Matsuda and Fumiyasu Komaki. Time Series Decomposition into Oscillation Components and Phase Estimation. *Neural Computation*, 29(2):332–367, 02 2017. ISSN 0899-7667. doi: 10.1162/NECO_a_00916. URL https://doi.org/10.1162/NECO_a_00916.

[98] Alan V Oppenheim. Applications of digital signal processing. *Englewood Cliffs*, 1978.

[99] Time series analysis. In *The Data Science Handbook*, pages 243–260. John Wiley and Sons, Inc., January 2017. doi: 10.1002/9781119092919.ch17. URL https://doi.org/10.1002/9781119092919.ch17.

[100] Oliver Schmitt, Peter Eipert, Frauke Ruß, Julia Beier, Kanar Kadir, and Anja Horn. The brainstem connectome database. *Scientific Data*, 9(1), April 2022. doi: 10.1038/s41597-022-01219-3. URL https://doi.org/10.1038/s41597-022-01219-3.

[101] Andrew M. Nemeth. Biological systems: Recognition of covariance. *Pharmacology*, 29(4):181–184, 1984. doi: 10.1159/000138010. URL https://doi.org/10.1159/000138010.

[102] Joan Daemen, René Govaerts, and Joos Vandewalle. Correlation matrices. In Bart Preneel, editor, *Fast Software Encryption*, pages 275–285, Berlin, Heidelberg, 1995. Springer Berlin Heidelberg. ISBN 978-3-540-47809-6.

[103] Gaël Varoquaux, Flore Baronnet, Andreas Kleinschmidt, Pierre Fillard, and Bertrand Thirion. Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling. In Tianzi Jiang, Nassir Navab, Josien P. W. Pluim, and Max A. Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, pages 200–208, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. ISBN 978-3-642-15705-9.

[104] Alexandre Abraham, Michael Milham, Adriana Di Martino, Cameron Craddock, Dimitris Samaras, Bertrand Thirion, and Gael Varoquaux. Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. *NeuroImage*, 147, 11 2016. doi: 10.1016/j.neuroimage.2016.10.045.

[105] Gaël Varoquaux, Flore Baronnet, Andreas Kleinschmidt, Pierre Fillard, and Bertrand Thirion. Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, pages 200–208. Springer Berlin Heidelberg, 2010. doi: 10.1007/978-3-642-15705-9_25. URL https://doi.org/10.1007%2F978-3-642-15705-9_25.

[106] Eric Courchesne, Karen Pierce, Cynthia M. Schumann, Elizabeth Redcay, Joseph A. Buckwalter, Daniel P. Kennedy, and John Morgan. Mapping early brain development in autism. *Neuron*, 56(2):399–413, October 2007. doi: 10.1016/j.neuron.2007.10.016. URL https://doi.org/10.1016/j.neuron.2007.10.016.

[107] Francesca Happé. *Weak Central Coherence*, pages 3344–3346. Springer New York, New York, NY, 2013. ISBN 978-1-4419-1698-3. doi: 10.1007/978-1-4419-1698-3_1744. URL https://doi.org/10.1007/978-1-4419-1698-3_1744.

[108] Takamitsu Watanabe and Geraint Rees. Anatomical imbalance between cortical networks in autism. *Scientific Reports*, 6(1), August 2016. doi: 10.1038/srep31114. URL https://doi.org/10.1038/srep31114.

[109] Yasser Ghanbari, Luke Bloy, Birkan Tunc, Varsha Shankar, Timothy P.L. Roberts, J. Christopher Edgar, Robert T. Schultz, and Ragini Verma. On characterizing population commonalities and subject variations in brain networks. *Medical Image Analysis*, 38:215–229, May 2017. doi: 10.1016/j.media.2015.10.009. URL https://doi.org/10.1016/j.media.2015.10.009.

[110] Tony W. Wilson, Donald C. Rojas, Martin L. Reite, Peter D. Teale, and Sally J. Rogers. Children and adolescents with autism exhibit reduced MEG steady-state gamma responses. *Biological Psychiatry*, 62(3):192–197, August 2007. doi: 10.1016/j.biopsych.2006.07.002. URL https://doi.org/10.1016/j.biopsych.2006.07.002.

[111] John Daintith and Edmund Wright. *A Dictionary of Computing*. Oxford University Press, January 2008. doi: 10.1093/acref/9780199234004.001.0001. URL https://doi.org/10.1093/acref/9780199234004.001.0001.

[112] Xiaowei Li, Zhuang Jing, Bin Hu, Jing Zhu, Ning Zhong, Mi Li, Zhijie Ding, Jing Yang, Lan Zhang, Lei Feng, and Dennis Majoe. A resting-state brain functional network study in MDD based on minimum spanning tree analysis and the hierarchical clustering. *Complexity*, 2017: 1–11, 2017. doi: 10.1155/2017/9514369. URL https://doi.org/10.1155/2017/9514369.

[113] Hao Guo, Lei Liu, Junjie Chen, Yong Xu, and Xiang Jie. Alzheimer classification using a minimum spanning tree of high-order functional network on fMRI dataset. *Frontiers in Neuroscience*, 11, December 2017. doi: 10.3389/fnins.2017.00639. URL https://doi.org/10.3389/fnins.2017.00639.

[114] Haiming Li, Qiyang Xia, and Yong Wang. Research and improvement of kruskal algorithm. *Journal of Computer and Communications*, 05:63–69, 01 2017. doi: 10.4236/jcc.2017.512007.

[115] Zhaoyuan Fang. The methods and tools for clustering analysis. In *Systems Medicine*, pages 9–13. Elsevier, 2021. doi: 10.1016/b978-0-12-801238-3.11463-1. URL https://doi.org/10.1016/b978-0-12-801238-3.11463-1.

[116] Domokos Eszergár-Kiss and Bálint Caesar. Definition of user groups applying ward's method. *Transportation Research Procedia*, 22:25–34, 2017. doi: 10.1016/j.trpro.2017.03.004. URL https://doi.org/10.1016/j.trpro.2017.03.004.

[117] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Andreas Müller, Joel Nothman, Gilles Louppe, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in python. 2012. doi: 10.48550/ARXIV.1201.0490. URL https://arxiv.org/abs/1201.0490.

[118] Robert Trevethan. Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health*, 5, 2017. ISSN 2296-2565. doi: 10.3389/fpubh.2017.00307. URL https://www.frontiersin.org/articles/10.3389/fpubh.2017.00307.

[119] Kaustubh Supekar, Lucina Q. Uddin, Amirah Khouzam, Jennifer Phillips, William D. Gaillard, Lauren E. Kenworthy, Benjamin E. Yerys, Chandan J. Vaidya, and Vinod Menon. Brain hyperconnectivity in children with autism and its links to social deficits. *Cell Reports*, 5(3): 738–747, November 2013. doi: 10.1016/j.celrep.2013.10.001. URL https://doi.org/10.1016/j.celrep.2013.10.001.

[120] Christopher Lee Keown, Patricia Shih, Aarti Nair, Nick Peterson, Mark Edward Mulvey, and Ralph-Axel Müller. Local functional overconnectivity in posterior brain regions is associated with symptom severity in autism spectrum disorders. *Cell Reports*, 5(3):567–572, November 2013. doi: 10.1016/j.celrep.2013.10.003. URL https://doi.org/10.1016/j.celrep.2013.10.003.

[121] Network science. *Contemporary Sociology*, 46(1):120–120, 2017. doi: 10.1177/0094306116681814. URL https://doi.org/10.1177/0094306116681814.

[122] Maryam Parsa, John P. Mitchell, Catherine D. Schuman, Robert M. Patton, Thomas E. Potok, and Kaushik Roy. Bayesian multi-objective hyperparameter optimization for accurate, fast, and efficient neural network accelerator design. *Frontiers in Neuroscience*, 14, 2020. ISSN 1662-453X. doi: 10.3389/fnins.2020.00667. URL https://www.frontiersin.org/articles/10.3389/fnins.2020.00667.

[123] M. Hazewinkel. *Encyclopaedia of Mathematics: Volume 10*. Encyclopaedia of Mathematics. Springer Netherlands, 2013. ISBN 9789400959835. URL https://books.google.co.ma/books?id=_YPtCAAAQBAJ.

[124] Stephen Cole Kleene. *Mathematical Logic*. Mineola, N.Y.: Dover Publications, 1967.

[125] Alireza Baratloo, Mostafa Hosseini, Ahmed Negida, and Gehad El Ashal. Evidence based emergency medicine; part 1: Simple definition and calculation of accuracy, sensitivity and specificity. *Emergency*, 3:48–49, 05 2015.

[126] Bharat Biswal, F. Zerrin Yetkin, Victor M. Haughton, and James S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine*, 34(4):537–541, October 1995. doi: 10.1002/mrm.1910340409. URL https://doi.org/10.1002/mrm.1910340409.

[127] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, February 2009. doi: 10.1038/nrn2575. URL https://doi.org/10.1038/nrn2575.

[128] Kisung You and Hae-Jeong Park. Re-visiting riemannian geometry of symmetric positive definite matrices for the analysis of functional connectivity. *NeuroImage*, 225:117464, 2021. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2020.117464. URL https://www.sciencedirect.com/science/article/pii/S1053811920309496.

[129] Stephen M. Smith, Karla L. Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F. Beckmann, Thomas E. Nichols, Joseph D. Ramsey, and Mark W. Woolrich. Network modelling methods for FMRI. *NeuroImage*, 54(2):875–891, January 2011. doi: 10.1016/j.neuroimage.2010.08.063. URL https://doi.org/10.1016/j.neuroimage.2010.08.063.

[130] Javier M. Buldú and Mason A. Porter. Frequency-based brain networks: From a multiplex framework to a full multilayer description. *Network Neuroscience*, 2(4):418–441, October 2018. doi: 10.1162/netn_a_00033. URL https://doi.org/10.1162/netn_a_00033.

[131] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, Aditya Vijaykumar, Alessandro Pietro Bardelli, Alex Rothberg, Andreas Hilboll, Andreas Kloeckner, Anthony Scopatz, Antony Lee, Ariel Rokem, C. Nathan Woods, Chad Fulton, Charles Masson, Christian Häggström, Clark Fitzgerald, David A. Nicholson, David R. Hagen, Dmitrii V. Pasechnik, Emanuele Olivetti, Eric Martin, Eric Wieser, Fabrice Silva, Felix Lenders, Florian Wilhelm, G. Young, Gavin A. Price, Gert-Ludwig Ingold, Gregory E. Allen, Gregory R. Lee, Hervé Audren, Irvin Probst, Jörg P. Dietrich, Jacob Silterra, James T Webber, Janko Slavič, Joel Nothman, Johannes Buchner, Johannes Kulick, Johannes L. Schönberger, José Vinícius de Miranda Cardoso, Joscha Reimer, Joseph Harrington, Juan Luis Cano Rodríguez, Juan Nunez-Iglesias, Justin Kuczynski, Kevin Tritz, Martin Thoma, Matthew Newville, Matthias Kümmerer, Maximilian Bolingbroke, Michael Tartre, Mikhail Pak, Nathaniel J. Smith, Nikolai Nowaczyk, Nikolay Shebanov, Oleksandr Pavlyk, Per A. Brodtkorb, Perry Lee, Robert T. McGibbon, Roman Feldbauer, Sam Lewis, Sam Tygier, Scott Sievert, Sebastiano Vigna, Stefan Peterson, Surhud More, Tadeusz Pudlik, Takuya Oshima, Thomas J. Pingel, Thomas P. Robitaille, Thomas Spura, Thouis R. Jones, Tim Cera, Tim Leslie, Tiziano Zito, Tom Krauss, Utkarsh Upadhyay, Yaroslav O. Halchenko, and Yoshiki Vázquez-Baeza and. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods*, 17(3):261–272, February 2020. doi: 10.1038/s41592-019-0686-2. URL https://doi.org/10.1038/s41592-019-0686-2.

[132] Marietta Papadatou-Pastou, Eleni Ntolka, Judith Schmitz, Maryanne Martin, Marcus Munafò, Sebastian Ocklenburg, and Silvia Paracchini. Human handedness: A meta-analysis. *Psychological Bulletin*, 146, 04 2020. doi: 10.1037/bul0000229.

[133] Zhiqiang Sha, Antonietta Pepe, Dick Schijven, Amaia Carrión-Castillo, James M. Roe, René Westerhausen, Marc Joliot, Simon E. Fisher, Fabrice Crivello, and Clyde Francks. Handedness and its genetic influences are associated with structural asymmetries of the cerebral cortex in 31,864 individuals. *Proceedings of the National Academy of Sciences*, 118(47):e2113095118, 2021. doi: 10.1073/pnas.2113095118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2113095118.

[134] Ali Anjomshoa, Mahsa Dolatshahi, Fatemeh Amirkhani, Farzaneh Rahmani, M.M. Mirbagheri, and Mohammad Aarabi. Structural brain network analysis in schizophrenia using minimum spanning tree. volume 2016, pages 4075–4078, 08 2016. doi: 10.1109/EMBC.2016.7591622.

[135] Hao Guo, Lei Liu, Junjie Chen, Yong Xu, and Xiang Jie. Alzheimer classification using a minimum spanning tree of high-order functional network on fmri dataset. *Frontiers in Neuroscience*, 11:639, 12 2017. doi: 10.3389/fnins.2017.00639.

[136] Jared A. Nielsen, Brandon A. Zielinski, P. Thomas Fletcher, Andrew L. Alexander, Nicholas Lange, Erin D. Bigler, Janet E. Lainhart, and Jeffrey S. Anderson. Multisite functional connectivity MRI classification of autism: ABIDE results. *Frontiers in Human Neuroscience*, 7, 2013. doi: 10.3389/fnhum.2013.00599. URL https://doi.org/10.3389/fnhum.2013.00599.

[137] Yen-Wei Chen, Satoshi Tanaka, Robert J. Howlett, and Lakhmi C. Jain, editors. *Innovation in Medicine and Healthcare 2016*. Springer International Publishing, 2016. doi: 10.1007/978-3-319-39687-3. URL https://doi.org/10.1007/978-3-319-39687-3.

[138] Anibal Sólon Heinsfeld, Alexandre Rosa Franco, R. Cameron Craddock, Augusto Buchweitz, and Felipe Meneguzzi. Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, 17:16–23, 2018. doi: 10.1016/j.nicl.2017.08.017. URL https://doi.org/10.1016/j.nicl.2017.08.017.

[139] TensorFlow Developers. Tensorflow, 2023. URL https://zenodo.org/record/4724125.

[140] François Chollet et al. Keras. https://github.com/fchollet/keras, 2015.

[141] Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gaël Varoquaux. Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8, 2014. doi: 10.3389/fninf.2014.00014. URL https://doi.org/10.3389/fninf.2014.00014.

[142] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. doi: 10.1109/CVPR.2009.5206848.

[143] Scott Chamberlin. *Covariation*, pages 83–95. 09 2021. ISBN 9781003238201. doi: 10.4324/9781003238201-7.

[144] Xin Yang, Ning Zhang, and Paul Schrader. A study of brain networks for autism spectrum disorder classification using resting-state functional connectivity. *Machine Learning with Applications*, 8:100290, 2022. ISSN 2666-8270. doi: https://doi.org/10.1016/j.mlwa.2022.100290. URL https://www.sciencedirect.com/science/article/pii/S2666827022000226.

[145] Gaël Varoquaux, Flore Baronnet, Andreas Kleinschmidt, Pierre Fillard, and Bertrand Thirion. Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling. In Dinggang Shen, Alejandro Frangi, and Gabor Szekely, editors, *Medical Image Computing and Computer Added Intervention*, Lecture Notes in Computer Science, Beijing, China, September 2010. Tianzi Jiang, Springer. doi: 10.1007/978-3-642-15705-9_25. URL https://hal.inria.fr/inria-00512417.

[146] Alireza Baratloo, Mostafa Hosseini, Ahmed Negida, and Gehad El Ashal. Evidence-based emergency medicine; part 1: Simple definition and calculation of accuracy, sensitivity and specificity. *Emergency (Tehran, Iran)*, 3:48–9, 10 2015.

[147] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv 1409.1556*, 09 2014.

[148] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. pages 1–9, 06 2015. doi: 10.1109/CVPR.2015.7298594.

[149] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 7, 12 2015.

[150] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. *AAAI Conference on Artificial Intelligence*, 31, 02 2016. doi: 10.1609/aaai.v31i1.11231.

[151] Francois Chollet. Xception: Deep learning with depthwise separable convolutions. pages 1800–1807, 07 2017. doi: 10.1109/CVPR.2017.195.

[152] Till Nierhaus, Daniel Margulies, Xiangyu Long, and Arno Villringer. fMRI for the assessment of functional connectivity. In *Neuroimaging - Methods*. InTech, February 2012. doi: 10.5772/23864. URL https://doi.org/10.5772/23864.

[153] Meijie Liu, Baojuan Li, and Dewen Hu. Autism spectrum disorder studies using fMRI data and machine learning: A review. *Frontiers in Neuroscience*, 15, September 2021. doi: 10.3389/fnins.2021.697870. URL https://doi.org/10.3389/fnins.2021.697870.

[154] Prantik Kundu, Valerie Voon, Priti Balchandani, Michael V. Lombardo, Benedikt A. Poser, and Peter A. Bandettini. Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. *NeuroImage*, 154:59–80, July 2017. doi: 10.1016/j.neuroimage.2017.03.033. URL https://doi.org/10.1016/j.neuroimage.2017.03.033.

[155] Xiao-Yan Tang, Wei-Ming Zeng, Ni-Zhuan Wang, Yu-Hu Shi, and Le Zhao. A novel layered data reduction mechanism for clustering fMRI data. *Biomedical Signal Processing and Control*, 33:48–65, March 2017. doi: 10.1016/j.bspc.2016.11.014. URL https://doi.org/10.1016/j.bspc.2016.11.014.

[156] Jian Yang, Ning Zhong, Peipeng Liang, Jue Wang, Yiyu Yao, and Shengfu Lu. Brain activation detection by neighborhood one-class SVM. *Cognitive Systems Research*, 11(1):16–24, March

2010. doi: 10.1016/j.cogsys.2008.08.001. URL https://doi.org/10.1016/j.cogsys.2008.08.001.

[157] Katherine S. Button, John P. A. Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan Flint, Emma S. J. Robinson, and Marcus R. Munafò. Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5):365–376, April 2013. doi: 10.1038/nrn3475. URL https://doi.org/10.1038/nrn3475.

[158] Jeanette A. Mumford. A power calculation guide for fMRI studies. *Social Cognitive and Affective Neuroscience*, 7(6):738–742, May 2012. doi: 10.1093/scan/nss059. URL https://doi.org/10.1093/scan/nss059.

[159] Sven Bölte, Sonya Girdler, and Peter B. Marschik. The contribution of environmental exposure to the etiology of autism spectrum disorder. *Cellular and Molecular Life Sciences*, 76 (7):1275–1297, December 2018. doi: 10.1007/s00018-018-2988-4. URL https://doi.org/10.1007/s00018-018-2988-4.

[160] Chang Liu, Jie Xue, Xu Cheng, Weiwei Zhan, Xin Xiong, and Bin Wang. Tracking the brain state transition process of dynamic function connectivity based on resting state fMRI. *Computational Intelligence and Neuroscience*, 2019:1–12, October 2019. doi: 10.1155/2019/9027803. URL https://doi.org/10.1155/2019/9027803.