



Bruna Campos Guedes



MÉTODOS DE CORRELAÇÃO APLICADOS A IMAGENS DE RESSONÂNCIA MAGNÉTICA FUNCIONAL

Uberlândia, Brasil

2022

Bruna Campos Guedes

MÉTODOS DE CORRELAÇÃO APLICADOS A IMAGENS DE RESSONÂNCIA MAGNÉTICA FUNCIONAL

Trabalho de conclusão de curso apresentado à Faculdade de Engenharia Elétrica da Universidade Federal de Uberlândia, como parte dos requisitos exigidos para a obtenção título de Bacharel em Engenharia Biomédica.

Universidade Federal de Uberlândia – UFU

Faculdade de Engenharia Elétrica

Bacharelado em Engenharia Biomédica

Orientador: Prof. Dr. Pedro Cunha Carneiro

Uberlândia, Brasil

2022

Agradecimentos

Primeiramente aos meus pais, Júlio e Marcela, que sempre me proporcionaram o necessário e me apoiaram para chegar até aqui, e são minha maior inspiração.

Ao meu irmão Vitor, meu maior suporte, exemplo e aliado da vida, e também por todos os conselhos que me fizeram questionar a realidade.

À Mariana, minha pessoa, por me apoiar e estar sempre ao meu lado em todos os momentos.

À minha grande família de avós, primos, tios e todos os outros que também me apoiaram, incentivaram e foram grandes exemplos de amparo, amor e força.

Aos meus amigos de todos os momentos da minha vida, os melhores, que tornaram toda a jornada mais leve, agradável e feliz. Em especial ao Gabriel Felipe, que esteve sempre ao meu lado durante grande parte da nossa jornada de aprendizado.

A todos os professores que compartilharam comigo seus conhecimentos e foram tão inspiradores, em especial ao professor Dr. Pedro Cunha Carneiro, ao professor Dr. Fernando Pasquini Santos e ao professor Dr. Naoki Masuda pelo incentivo e orientação para a realização dessa pesquisa, e à professora Dra. Ana Cláudia Patrocínio por me inspirar a procurar conhecimento sobre imagens médicas e apaixonar pelo assunto.

À educação pública e órgãos de incentivo à pesquisa que tornaram possível chegar até aqui.

E por fim à unico IDTech, em especial ao time Machine Learning, pelo apoio e incentivo ao conhecimento.

*"Um navio no porto é seguro,
mas não é para isso que os navios existem.
Navegue para o mar e faça coisas novas."
(Grace Hopper)*

Resumo

Com o avanço da tecnologia, as imagens de Ressonância Magnética Funcional (fMRIs) começaram a ser amplamente usadas, principalmente por permitir a visualização da função cerebral a partir da diferença de contraste dos tecidos vizinhos, devido à mudança do nível de oxigênio da hemoglobina. Utilizou-se fMRIs de grupos de indivíduos com transtornos neurológicos pela possibilidade de encontrar padrões e diferenças nas conectividades das regiões do cérebro, o que pode facilitar no diagnóstico ou ajudar a encontrar tratamento para os indivíduos com tais transtornos. Assim, o objetivo deste trabalho é comparar qual método de correlação pode ser mais eficaz na análise de fMRIs, especificamente calculando a Correlação Pearson, Correlação Parcial e Informação Mútua. Tais correlações foram comparadas utilizando a correlação Spearman. Foram utilizadas imagens da base de dados pública criada pelo centro médico Langone da Universidade de Nova Iorque, distintos em dois grupos: participantes com Transtorno do Espectro Autista e um grupo controle. Como principais resultados, foi possível notar que a diferença entre os métodos de correlação foi muito pequena, não podendo fazer qualquer inferência direta com os distúrbios mentais e psicológicos dos indivíduos. Além disso, neste trabalho, a utilização da correlação Parcial não foi eficiente, visto que o aumento das regiões de interesse do cérebro analisadas fez com que os valores diminuíssem muito, além do fato do tipo dos dados analisados neste trabalho apresentarem uma distribuição normal. Com resultados similares entre Correlação Pearson e Informação Mútua (correlação moderada), a escolha do método dependerá do objetivo da pesquisa, bem como dos dados em questão. A partir deste trabalho foi possível selecionar e dividir dados de fMRI de pacientes dos dois grupos analisados, bem como implementar algoritmos para análise desse tipo de imagem baseado em teoria de grafos, com base na visualização de gráficos.

Palavras-chave: fMRI. correlação. regiões de interesse.

Abstract

With the advance of technology, Functional Magnetic Resonance (fMRI) images began to be widely used, mainly for allowing the visualization of brain functions from the contrast difference of neighboring tissues, due to the change in the oxygen level of hemoglobin. fMRIs of groups of individuals with neurological disorders were used because of the possibility of finding patterns and differences in the connectivity of brain regions, which may facilitate diagnosis or help find treatment for individuals with such disorders. Thus, the objective of this work is to compare which correlation method can be more effective in the analysis of fMRIs, specifically calculating Pearson Correlation, Partial Correlation and Mutual Information. Such correlations were compared using the Spearman correlation. Images from the public database created by the Langone Medical Center at New York University were used, divided into two groups: participants with Autism Spectrum Disorder and a control group. As main results, it was possible to notice that the difference between the correlation methods was very small, not being able to make any direct inference with the mental and psychological disorders of the individuals. Furthermore, in this work, the use of Partial correlation was not efficient, since the increase in the regions of interest in the brain analyzed caused the values to decrease a lot, in addition to the fact that the type of data analyzed in this work presented a normal distribution. With similar results between Pearson Correlation and Mutual Information (moderate correlation), the choice of method will depend on the research objective, as well as the data in question. From this work, it was possible to select and divide fMRI data from patients of the two groups analyzed, as well as implement algorithms for analysis of this type of image based on graph theory, using graphics visualization.

Keywords: fMRI. correlation. regions of interest.

Lista de ilustrações

Figura 1 – Corte de um aparelho de ressonância magnética.	13
Figura 2 – Exemplo de fMRI. Resultados dos estudos ALFF e ReHo. Amplitude de Flutuações de Baixa Frequência (ALFF) e homogeneidade regional (ReHo) são métodos comumente usados em avaliações de segregação funcional. A, resultados ALFF. B, resultados ReHo. Os resultados de ALFF e ReHo refletem atividades neurais regionais. ALFF está focado em medir a força da atividade, enquanto ReHo é mais específico para coerência e centralidade da atividade regional. T indica intensidade de pico.	14
Figura 3 – Exemplo de um grafo simples.	16
Figura 4 – Exemplo de uma rede complexa, tráfego aéreo global	16
Figura 5 – Diagrama de blocos sequencial do trabalho.	18
Figura 6 – Recorte de uma matriz gerada no pré-processamento das imagens.	20
Figura 7 – Correlações de um indivíduo do grupo ASD com 264 ROIs.	23
Figura 8 – Correlações de um indivíduo do grupo TD com 264 ROIs.	23
Figura 9 – Correlações de um indivíduo do grupo ASD com 30 ROIs.	24
Figura 10 – Correlações de um indivíduo do grupo TD com 30 ROIs.	24
Figura 11 – Média e desvio padrão da Correlação Spearman aplicada à Correlação Pearson.	25
Figura 12 – Média e desvio padrão da Correlação Spearman aplicada à Correlação Parcial.	25
Figura 13 – Média e desvio padrão da Correlação Spearman aplicada à Informação Mútua.	25

Lista de tabelas

Tabela 1 – Informações sobre os dados coletados	19
Tabela 2 – Média e Desvio Padrão das correlações Spearman.	26

Lista de abreviaturas e siglas

MR	Ressonância Magnética, do inglês <i>Magnetic Resonance</i>
MRI	Imagem de Ressonância Magnética, do inglês <i>Magnetic Resonance Imaging</i>
Rf	Rádio frequência, do inglês <i>Radio frequency</i>
BOLD	Dependente do nível de oxigênio no sangue, do inglês <i>Blood oxygen level dependent</i>
fMRI	Imagem de Ressonância Magnética Funcional, do inglês <i>Functional Magnetic Resonance Imaging</i>
ROIs	Regiões de interesse, do inglês <i>Regions of interest</i>
ASD	Transtorno do Espectro Autista, do inglês <i>Autism Spectrum Disorder</i>
TD	Dado de teste, do inglês <i>Test data</i>
KS	Kolmogorov-Smirnov

Lista de símbolos

Γ	Letra grega Gama (função gama)
ρ	Letra grega rho
Σ	Somatório

Sumário

1	INTRODUÇÃO	11
1.1	Objetivo geral	12
1.1.1	Objetivos específicos	12
2	FUNDAMENTAÇÃO TEÓRICA	13
2.1	Ressonância Magnética Funcional	13
2.2	Python para análise de dados	15
2.3	Redes cerebrais funcionais baseadas na Teoria de Grafos	15
3	METODOLOGIA	18
3.1	Dados utilizados	19
3.2	Métodos de correlação	20
3.2.1	Correlação Pearson	20
3.2.2	Correlação Parcial	21
3.2.3	Informação Mútua	21
3.2.4	Correlação Spearman	22
4	RESULTADOS E DISCUSSÕES	23
5	CONCLUSÃO	28
	REFERÊNCIAS	29

1 Introdução

O avanço da tecnologia fez com que as imagens de Ressonância Magnética passassem a ser o método de investigação de diagnóstico para diversos problemas clínicos mais utilizado dos últimos 30 anos (KATTI; ARA; SHIREEN, 2011). Isso dado à sua alta capacidade de diferenciar tecidos e explorar aspectos tanto anatômicos quanto funcionais (MAZZOLA, 2009).

Em 1990, ao observar o diagnóstico por imagem de Ressonância Magnética de ratos, cientistas perceberam um aumento do contraste de vasos sanguíneos do cérebro na imagem em relação ao tecido vizinho de acordo com o aumento da concentração de oxigênio. Essa diferença de contraste dos vasos sanguíneos pode permitir a visualização da função cerebral (MAZZOLA, 2009).

A discussão sobre a relação entre a função das conectividades e as estruturas cerebrais vem crescendo devido à sua possibilidade de investigar patologias, já que podemos comparar essas conexões em cérebros de pessoas com determinados transtornos neurológicos e mentais aos de pessoas saudáveis (POWER et al., 2011). Estes, podem estar diretamente ligados à disfunção defeituosa de neurotransmissores (SPRONK et al., 2021). Estudos anteriores estimam que imagens de Ressonância Magnética Funcional (fMRI) em estado de repouso estabelecem uma alta conectividade com as dadas redes funcionais cerebrais, ou seja, as conexões (COLE et al., 2011).

Estes fatores podem ser utilizados para comparar e diagnosticar a natureza e severidade das mudanças neurais subjacentes aos transtornos mentais, até mesmo ajudar no desenvolvimento de tratamentos (SPRONK et al., 2021). Essas conexões do cérebro a serem comparadas são regiões que podem estar envolvidas na regulação de determinadas tarefas, e, para encontrar semelhanças e diferenças são utilizadas as fMRI em repouso como imagens padrão (FAIR et al., 2009).

A ciência de redes é amplamente usada para a análise de padrões de conexões, podendo elas serem comportamentais, estruturais, funcionais, entre outras. São utilizadas propriedades estatísticas para caracterizá-las e sugerir maneiras apropriadas para medir tais propriedades (NEWMAN, 2003). Na neurociência as regiões cerebrais e suas conexões são consideradas uma rede complexa (RUBINOV; SPORNS, 2010) e, para quantificar mudanças nas conexões funcionais, a medida de distância de grafos (correlação) é o método mais utilizado (SPRONK et al., 2021).

1.1 Objetivo geral

O objetivo deste trabalho é comparar qual método de correlação pode ser mais eficaz na análise de fMRIs. Para isso, serão calculadas correlações entre as regiões de interesse do cérebro utilizando três métodos diferentes: correlação Pearson, correlação Parcial e informação mútua. Ademais, para comparar estes métodos, será utilizada a correlação Spearman.

1.1.1 Objetivos específicos

Dentre os objetivos específicos destacam-se:

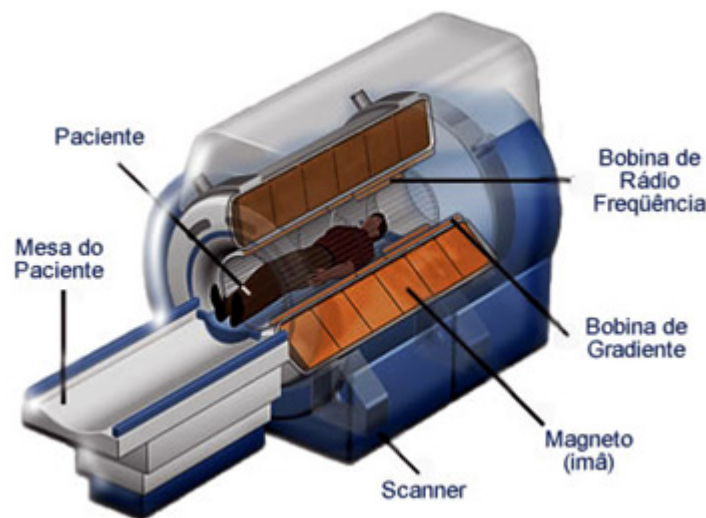
- Selecionar dados utilizados buscando um grupo com algum transtorno neurológico e um grupo controle;
- Elaborar algoritmos para análise das fMRI baseado na teoria de grafos;
- Analisar os padrões para comparar os métodos de correlação para os diferentes grupos a partir da visualização de gráficos e comparação de médias e desvios padrão;
- Verificar se algum dos métodos de correlação testados pode ser mais eficaz na análise de fMRIs.

2 Fundamentação Teórica

2.1 Ressonância Magnética Funcional

A história da imagem de Ressonância Magnética (MRI - *magnetic resonance imaging*) começou em 1857 com a equação de Larmor, que refere-se à taxa de precessão do momento magnético do próton em torno do campo magnético externo (KATTI; ARA; SHIREEN, 2011). É um método não-invasivo que permite mapear estruturas internas e certos aspectos funcionais do corpo. A Ressonância Magnética (MR - *magnetic resonance*) é um aparelho com um grande ímã, como o da Figura 1, que induz na criação de um campo magnético relativamente forte, fazendo com que o núcleo dos átomos de Hidrogênio se alinhem ao campo magnético e à radiação de radiofrequência induzida (RF - *radio frequency*) liberando energia do corpo. Esta é detectada e utilizada por um computador para construir a imagem (KATTI; ARA; SHIREEN, 2011).

Figura 1 – Corte de um aparelho de ressonância magnética.



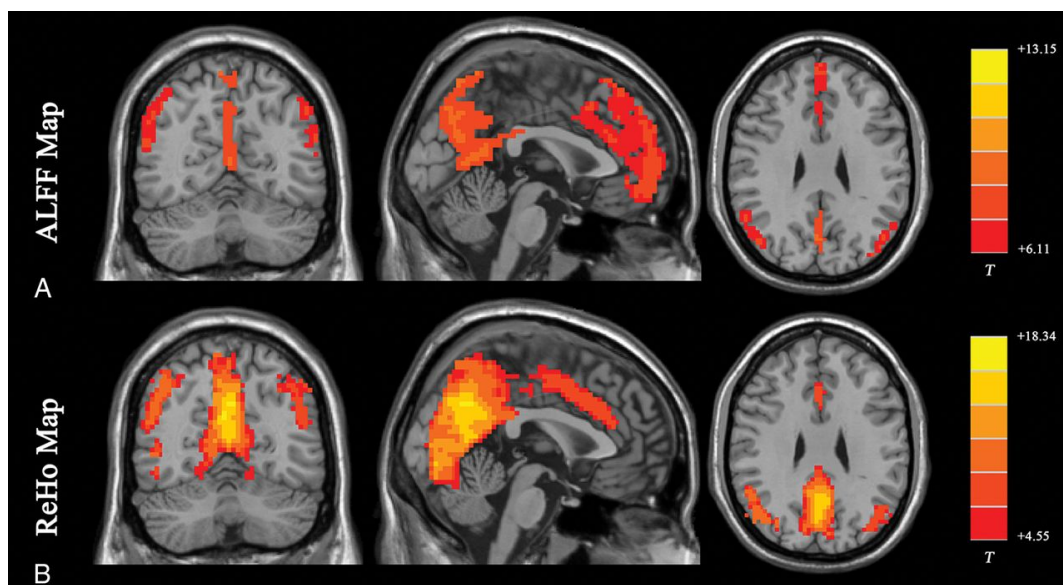
Fonte: (SOUZA, 2021)

Foi observado em 1990 que os vasos sanguíneos do cérebro de um rato, quando respirava ar normalmente, eram originalmente pretos e apresentavam maior sinal na MRI. Quando a respiração atingia 100% da concentração de oxigênio, as imagens apresentavam maior sinal e perdiam o contraste com tecidos vizinhos. Isso se dá devido à mudança nas propriedades magnéticas do sangue, em que a hemoglobina sem oxigênio apresenta efeito de susceptibilidade magnética em relação ao tecido vizinho, enquanto o sangue com maior concentração de hemoglobina com oxigênio deixa de apresentar diferenças em relação

ao tecido vizinho. Com isso existe a hipótese de que este mecanismo de contraste pode permitir a visualização da função cerebral, pois o aumento do nível de hemoglobina sem oxigênio com o aumento do consumo de oxigênio faria surgir pontos escuros na imagem, podendo significar uma ativação em dada região do cérebro (MAZZOLA, 2009).

Assim, o chamado contraste dependente do nível de oxigênio no sangue (BOLD - *blood oxygen level dependent*) é usado para descrever atividades cerebrais e respostas cognitivas das chamadas imagens de Ressonância Magnética Funcional (fMRI - *functional magnetic resonance imaging*), como da Figura 2, sendo essa uma técnica da MRI que analisa as imagens durante um determinado tempo (NORRIS, 2006). Esse tipo de imagem associada a análises estatísticas entre regiões de interesse (ROIs - *regions of interest*) podem mostrar a dependência entre séries temporais extraídas dessas ROIs distintas, sem levar em consideração a causalidade ou direção da conexão. Comparar esta dependência entre grupos de indivíduos pode mostrar padrões espaciais de conectividade (SALMON; LEONI, 2019).

Figura 2 – Exemplo de fMRI. Resultados dos estudos ALFF e ReHo. Amplitude de Flutuações de Baixa Frequência (ALFF) e homogeneidade regional (ReHo) são métodos comumente usados em avaliações de segregação funcional. A, resultados ALFF. B, resultados ReHo. Os resultados de ALFF e ReHo refletem atividades neurais regionais. ALFF está focado em medir a força da atividade, enquanto ReHo é mais específico para coerência e centralidade da atividade regional. T indica intensidade de pico.



Fonte: (LV et al., 2018)

2.2 Python para análise de dados

Sendo uma linguagem de programação *open source*, Python é considerada atualmente a segunda mais utilizada no mundo (O'GRADY, 2021). Ela se tornou uma linguagem famosa por sua facilidade em criar *scripts*, programas curtos e funcionais, com bibliotecas, pacotes e ferramentas que resolvem problemas de álgebra linear, integração, entre outros. Além disso, desde o início dos anos 2000 possui uma comunidade científica bem ativa, e sua capacidade de permitir criações científicas e tecnológicas utilizando a mesma linguagem (MCKINNEY, 2012).

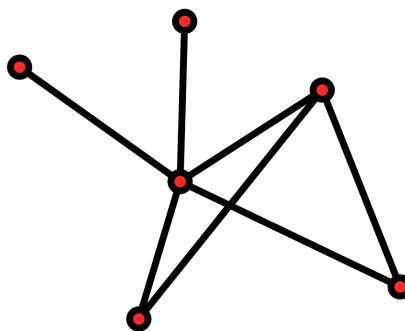
A criação de novas bibliotecas de suporte para a linguagem permitiu que fosse possível analisar e manipular dados de tipos variados. Entre elas, as principais e mais utilizadas pela comunidade científica são (MCKINNEY, 2012):

- a) Numpy: este pacote permite construir, manipular e visualizar *arrays*, realizar operações matemáticas entre *arrays*, entre outros;
- b) pandas: esta ferramenta utiliza da alta performance do Numpy com funções que permitem trabalhar com dados estruturados;
- c) matplotlib: é uma biblioteca para visualizar e plotar gráficos e outros tipos de visualização de dados;
- d) Scipy: é uma coleção de pacotes de diversos tipos de problemas padrões da computação científica, como a realização de integrais, derivadas, ferramentas de processamento de sinais, entre outros.

2.3 Redes cerebrais funcionais baseadas na Teoria de Grafos

A Teoria de Grafos é um ramo da matemática que lida com a descrição formal e análise de grafos. Um grafo (Figura 3) é um conjunto de nós (vértices) ligados por conexões (arestas), podendo ser ou não diretos. Quando trazemos para o mundo real, um grafo pode ser a representação de um sistema ou rede complexa (BULLMORE; SPORNS, 2009).

Figura 3 – Exemplo de um grafo simples.



Fonte: (SÁNCHEZ, 2011)

A internet, as relações entre indivíduos ou empresas, as redes neurais, rotas de entregas, são alguns exemplos de redes complexas (Figura 4). Seu estudo é um dos pilares fundamentais da matemática discreta, concretizada durante o século XX. A primeira grande prova da teoria de redes foi a solução dada por Euler para o problema da ponte de Königsberg (NEWMAN, 2003).

Figura 4 – Exemplo de uma rede complexa, tráfego aéreo global



Fonte: (GRANDJEAN, 2016)

As redes funcionais são um tipo de rede complexa, aplicadas principalmente em neurociência. Existe a discussão de que as oscilações neurais podem mostrar a dinâmica no cérebro. É possível que essas oscilações reflitam as flutuações da excitabilidade rítmica de conjuntos de neurônios, e que podem prover o fluxo de informações entre esses conjuntos. Com a alteração da força, padrão ou frequência dessas oscilações o cérebro poderia coordenar dinamicamente o fluxo de informações entre diferentes áreas do cérebro envolvidas

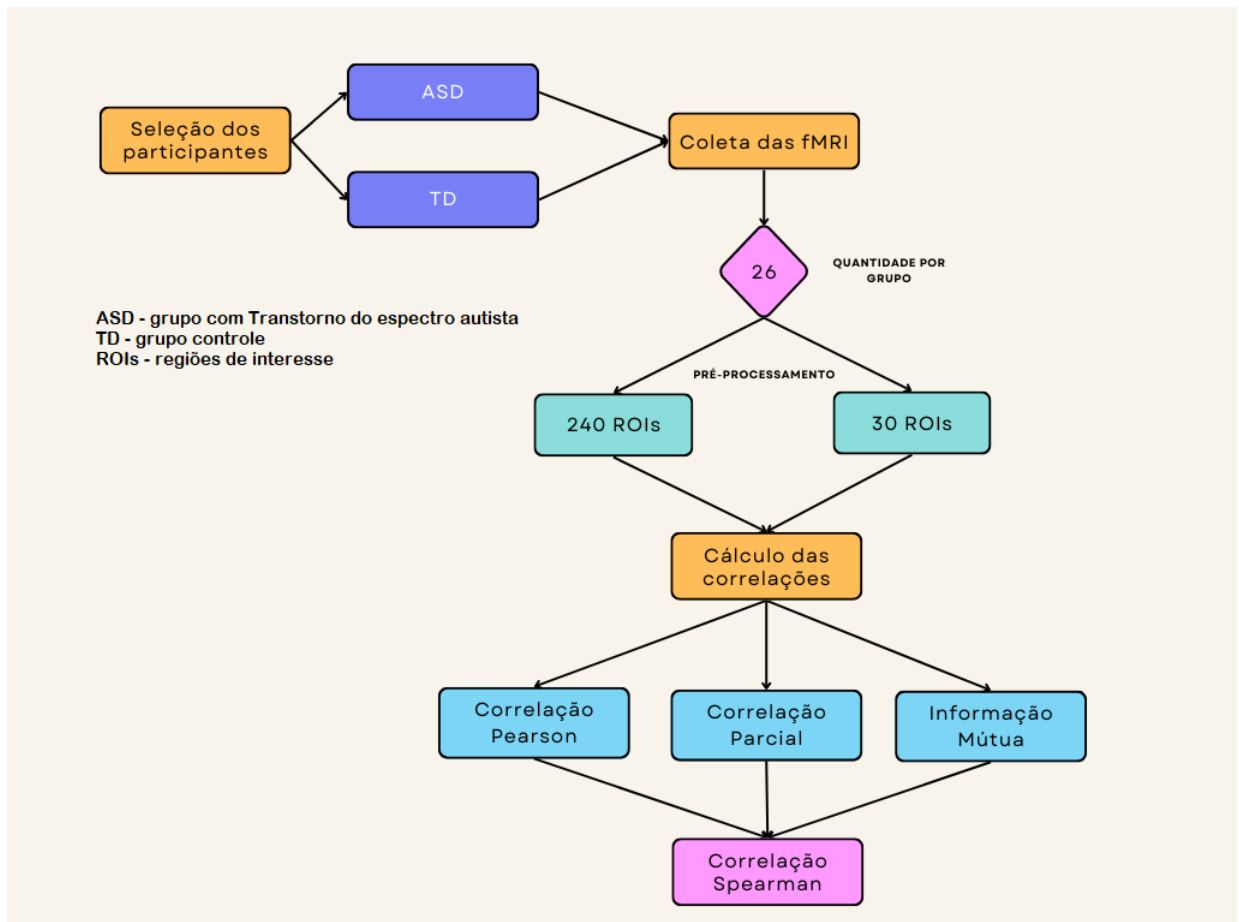
na sincronia oscilatória (BASTOS; SCHOFFELEN, 2016). Teorias de como essas podem estar ligadas a padrões de comportamento ou patologias são amplamente discutidas na comunidade científica (BASSETT; SPORNS, 2017).

A análise temporal é base para a construção dos vértices nas redes funcionais. Normalmente são utilizadas técnicas de covariância ou correlação para estimar a direção e intensidade da influência entre os nós. Em neurociência, diferentes nós (regiões do cérebro) podem ter distintas funcionalidades dependendo da resposta sensorial a algum ou nenhum estímulo (repouso), pois quando comparamos a rede funcional com a rede estrutural do cérebro podemos supor em quais estruturas ou ROIs estão ocorrendo as atividades (BASSETT; SPORNS, 2017).

3 Metodologia

A Figura 5 apresenta um diagrama de blocos resumindo de forma sequencial as etapas efetuadas na metodologia do trabalho. Inicialmente foi realizada a seleção dos dados a serem utilizados, os dividindo em dois grupos (ASD e TD). Em seguida realizou-se a coleta das imagens dos 26 participantes de cada grupo. O pré-processamento das fMRIs aconteceu de duas maneiras diferentes, a primeira com 30 ROIs e a segunda com 264 ROIs. Depois, foram calculadas as correlações (Pearson, parcial e informação mútua) e, por fim, para compará-las utilizamos a correlação Spearman.

Figura 5 – Diagrama de blocos sequencial do trabalho.



Fonte: Autora.

3.1 Dados utilizados

Neste trabalho foram utilizadas imagens da base de dados pública criada pelo centro médico Langone da Universidade de Nova Iorque ([NYU Langone Medical Center, 2012](#)). Os 29 participantes incluídos no estudo possuíam Transtorno do Espectro Autista (ASD - *Autism Spectrum Disorder*) e idades entre 7 e 40 anos. Além desses, também foi selecionado o grupo controle (TD - *test data*) contendo 26 participantes com o mesmo intervalo de idade do grupo anterior (ASD).

A aquisição das imagens de ambos os grupos (ASD e TD) foram realizadas em um equipamento da fabricante Siemens modelo Magnetom Allegra 3.0 Tesla. Durante a coleta, foi pedido aos participantes que permanecessem quietos e com os olhos fechados, mas sem dormir. Essa técnica é conhecida como estado de repouso e é amplamente usada em neurociência ([SPRONK et al., 2021](#)).

As imagens foram pré-processadas ([NYU Langone Medical Center, 2012](#)) de duas maneiras diferentes. A primeira, contendo 30 ROIs ([FAIR et al., 2009](#)) e a outra, 264 ROIs ([POWER et al., 2011](#)). Foi necessário eliminar as imagens de três participantes do grupo ASD, pois estavam corrompidas, restando 26 para ambos os ROIs ([Tabela 1](#)). Essas regiões são consideradas os nós da rede complexa que é o cérebro, separadas por partes e funções já conhecidas pela ciência ([FAIR et al., 2009](#); [POWER et al., 2011](#)).

Tabela 1 – Informações sobre os dados coletados

ROIs	Grupo	Número de participantes
30	ASD	26
30	TD	26
264	ASD	26
264	TD	26

Fonte: Autora.

O pré-processamento das imagens gerou matrizes, como a da [Figura 6](#), com os valores obtidos no tempo para cada ROI, utilizadas neste trabalho para calcular as correlações entre regiões.

Figura 6 – Recorte de uma matriz gerada no pré-processamento das imagens.

```

0.3132107630506005    0.1635387549858874    0.03387070012114275
0.1635387549858874    0.4482099365602559    0.023084775486334894
0.03387070012114275    0.023084775486334894    0.216621624247716
-0.020376405445387828  0.04776835128035362    0.029742621207995205
-0.007508639481647378  -0.01111867657117835    -0.009838468931379697
0.09627873521375813    0.07515563063251389    -0.030209478029110257
0.009319045225993505    0.19042548433677517    0.026091948621609913
0.12143998304549879    0.07753306574888286    -0.018816109041898454
0.010715365177798256    0.10480741180165375    0.05798109679853145
-0.05171333063729516    -0.10317114253407612    -0.03874231565437939
-0.08525959873236606    0.02029429846881513    -0.0012806655852188142
0.037430826342105386    0.025370781574481718    0.03936592244358122
-0.10173432825305283    -0.028576969757813597    0.029647387890201556
-0.05162930482486455    -0.010326009989350758    0.05079726443293495
-0.036652981182463525  -0.002252051036233964    0.035575364207142336 ...
...

```

Fonte: (NYU Langone Medical Center, 2012)

3.2 Métodos de correlação

Ao definir as imagens e os nós, cada um tem seu próprio curso de tempo associado, que são usados para estimar a conexão dos nós. Quanto mais semelhantes são os períodos entre qualquer par de nós, mais provável que haja uma conexão funcional entre eles. Neste trabalho, serão avaliados diferentes métodos para estimar a conexão entre os nós. Serão utilizados métodos de correlação entre duas séries temporais, ou seja, não será avaliado a direção do fluxo de informações, ou se existem conexões indiretas (SMITH et al., 2011).

Utilizando a linguagem Python, foram criados algoritmos para analisar três tipos de métodos (correlação Pearson, correlação parcial e informação mútua) a fim de obter as matrizes de correlação entre os nós. Para comparar as dependências dos nós entre os participantes, utilizamos a correlação Spearman e, em seguida, calculamos o valor médio e o desvio padrão desses valores para cada tipo de correlação, afim de comparar tanto os métodos quanto a correlação entre os grupos (ASD e TD, ASD e ASD, TD e TD).

3.2.1 Correlação Pearson

Na correlação Pearson entre duas séries temporais (x_1, x_2) , seu coeficiente de correlação (ρ) é dado entre os valores -1 e 1. Quanto mais próximo de 1 temos um aumento simultâneo do valor das variáveis, quanto mais próximo de -1 o valor de uma variável aumenta enquanto o da outra diminui, porém em ambos os casos podemos dizer que os nós estão correlacionados. Caso esteja próximo de 0 podemos dizer que não há relação entre os nós (BENESTY et al., 2009). O coeficiente é dado pela Equação 3.1.

$$\rho = \frac{Cov(x_1, x_2)}{\sqrt{Cov(x_1, x_1) \times Cov(x_2, x_2)}} \quad (3.1)$$

Nela, ρ é o coeficiente de correlação e Cov indica a covariância entre as séries temporais x_1 e x_2 .

É importante lembrar que as variáveis devem ser aproximadamente distribuídas (curva normal), com uma associação linear. Para confirmar isso, foi realizado o teste de Kolmogorov-Smirnov (KS) (DANIEL; CROSS, 2018) e o p -value, que indica se os dados possuem distribuição normal, ou não. O valor obtido para as variáveis deste trabalho foram muito próximas de zero, o que indica que o conjunto de dados está em distribuição normal.

3.2.2 Correlação Parcial

Assim como a correlação Pearson, o coeficiente de correlação parcial (r) varia entre -1 e 1, se próximo de 1, as séries temporais aumentam concomitantemente, se próximo de -1, o valor de uma série temporal aumenta e o valor da outra diminui, se próximo de 0 as séries temporais não estão relacionadas. Dado duas séries temporais x_1 e x_2 , o coeficiente é dado pela Equação 3.2 (WATANABE et al., 2013).

$$r_{x_1x_2} = \frac{-\Gamma_{x_1x_2}^{-1}}{\sqrt{\Gamma_{x_1x_1}^{-1} \Gamma_{x_2x_2}^{-1}}} \quad (3.2)$$

Na qual r indica a correlação parcial entre duas séries temporais dadas, x_1 e x_2 .

As variáveis na correlação Parcial devem ser aleatórias, e a associação entre elas pode gerar resultados enganadores caso haja outra variável, de confusão, numericamente relacionada com ambas as variáveis de interesse.

3.2.3 Informação Mútua

A informação mútua (I) também é usada para calcular a correlação entre dois nós (variáveis aleatórias), comparando os histogramas individuais e conjuntos (SMITH et al., 2011). Quando zero, os nós são independentes, e quanto maior o valor de I , maior a relação entre eles (GIERLICH et al., 2008). I é dada pela Equação 3.3.

$$I(X; Y) = H(Y) - H(Y|X) \quad (3.3)$$

Sendo $H(Y)$ uma entropia marginal, $H(Y|X)$ uma entropia condicional e, X e Y , os nós a serem comparados.

Diferente dos coeficientes de correlação, a Informação Mútua não é limitada a variáveis aleatórias com valores reais. Sendo assim, essa métrica é considerada mais genérica, definindo a "quantidade de informação", isto é, a informação compartilhada entre as variáveis aleatórias. Por exemplo, se caso duas variáveis A e B forem independentes,

conhecer a variável A não irá fornecer nenhuma informação sobre B, e vice-versa, o que ocasiona em um valor de Informação Mútua igual a zero.

3.2.4 Correlação Spearman

Assim como as correlações Pearson e parcial, a correlação Spearman (ρ) varia entre -1 e 1. Quando próximo de 0, consideramos que não há dependência entre as variáveis. Se positivo, os dois valores estão aumentando, caso negativo, o valor de um tende a aumentar enquanto o outro tende a diminuir. Ela é utilizada para calcular a dependência estatística entre a classificação de duas variáveis (podendo ou não terem relações lineares), chamados de postos (ZAR, 2005). Considerando d a diferença entre dois postos de cada série temporal e n o número de séries temporais, temos a [Equação 3.4](#).

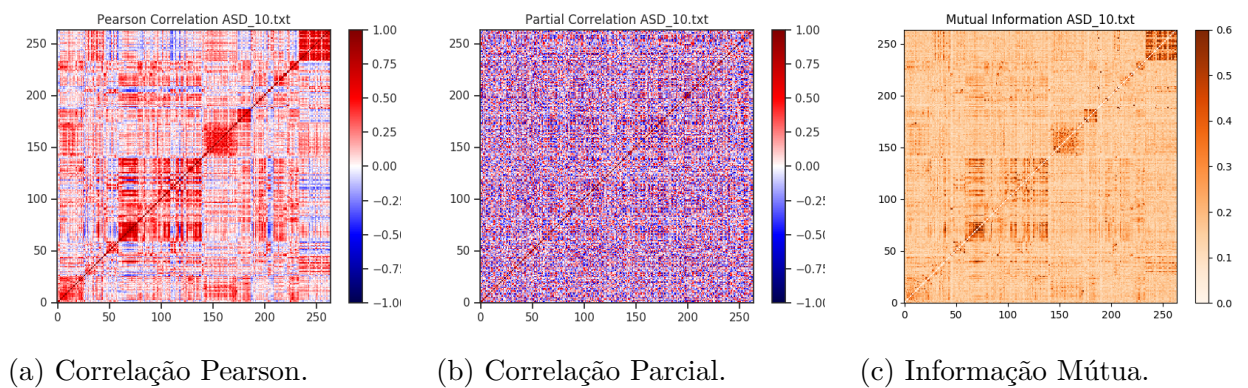
$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (3.4)$$

Onde ρ indica a correlação Spearman, d a diferença entre dois postos e n o número de séries temporais.

4 Resultados e Discussões

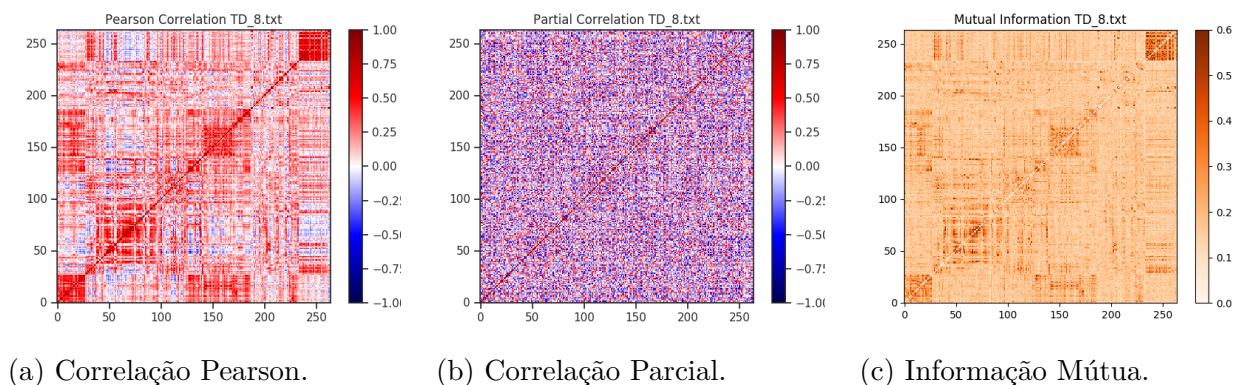
Ao utilizar os algoritmos desenvolvidos em Python para criar as matrizes de correlação das imagens de ambos os 30 e 264 ROIs, foram obtidos 52 gráficos para cada método. A [Figura 7](#) mostra as correlações Pearson, parcial e informação mútua, respectivamente, de um indivíduo do grupo ASD com 264 ROIs, e a [Figura 8](#) as de um indivíduo do grupo TD com 264 ROIs. Comparando essas imagens, podemos observar certa similaridade entre os grupos ASD e TD, observando as regiões de 0 a 25, ou 50 a 100, 150 a 175 e 250 a 264, além disso essas são regiões apresentam um coeficiente alto de correlação nos três métodos.

Figura 7 – Correlações de um indivíduo do grupo ASD com 264 ROIs.



Fonte: Autora

Figura 8 – Correlações de um indivíduo do grupo TD com 264 ROIs.

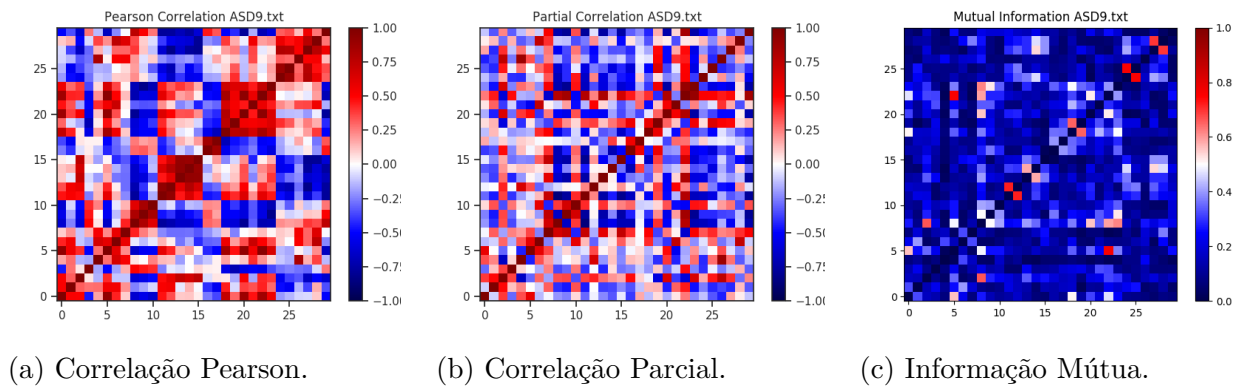


Fonte: Autora

O mesmo pode ser observado na [Figura 9](#), que mostra as correlações Pearson, parcial e informação mútua, respectivamente, de um indivíduo do grupo ASD com 30 ROIs, e a [Figura 10](#) as de um indivíduo do grupo TD com 30 ROIs. As regiões de 10 a 15

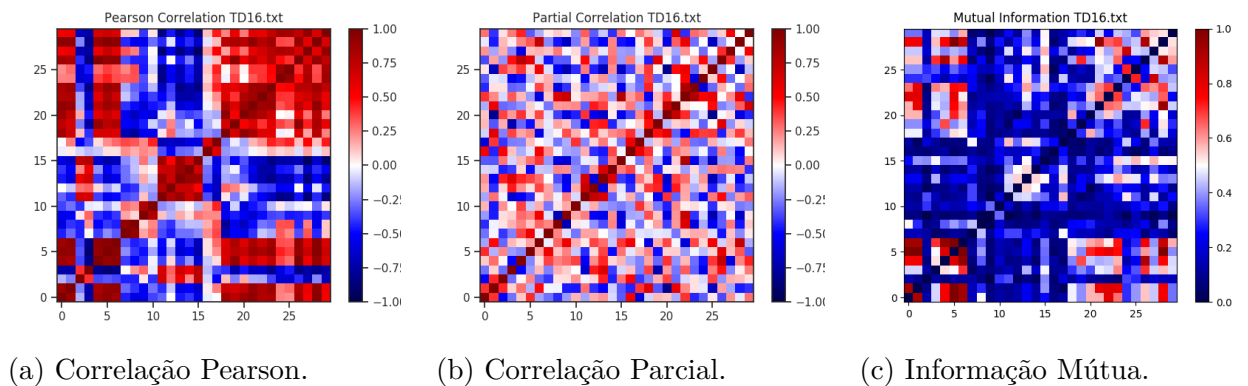
e 25 a 30, por exemplo, apresentam maior coeficiente de correlação nos três métodos, para ambas as quantidades de ROIs.

Figura 9 – Correlações de um indivíduo do grupo ASD com 30 ROIs.



Fonte: Autora

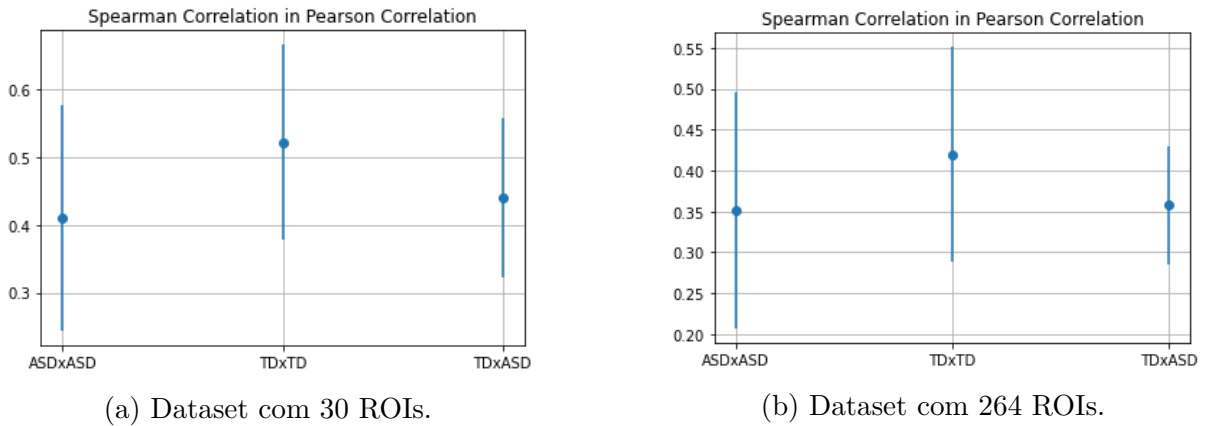
Figura 10 – Correlações de um indivíduo do grupo TD com 30 ROIs.



Fonte: Autora

Em seguida, foram calculadas as correlações Spearman a fim de analisar a dependência das regiões entre os participantes, e, para observar melhor o resultado, foram calculadas as médias e desvios padrão, comparando os grupos tanto de 264 ROIs quanto de 30 ROIs. Os resultados obtidos para todos os métodos podem ser vistos nos gráficos das Figuras 11, 12 e 13.

Figura 11 – Média e desvio padrão da Correlação Spearman aplicada à Correlação Pearson.

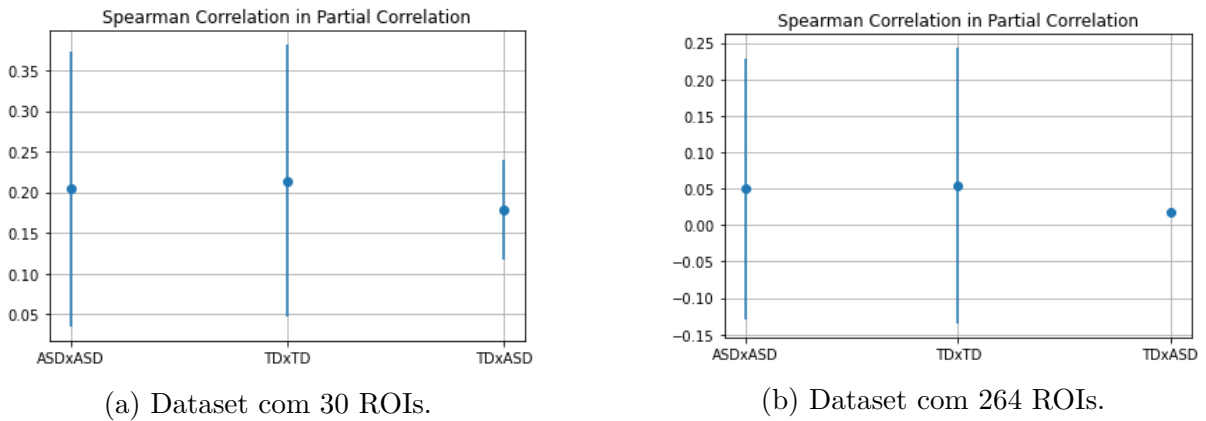


(a) Dataset com 30 ROIs.

(b) Dataset com 264 ROIs.

Fonte: Autora

Figura 12 – Média e desvio padrão da Correlação Spearman aplicada à Correlação Parcial.

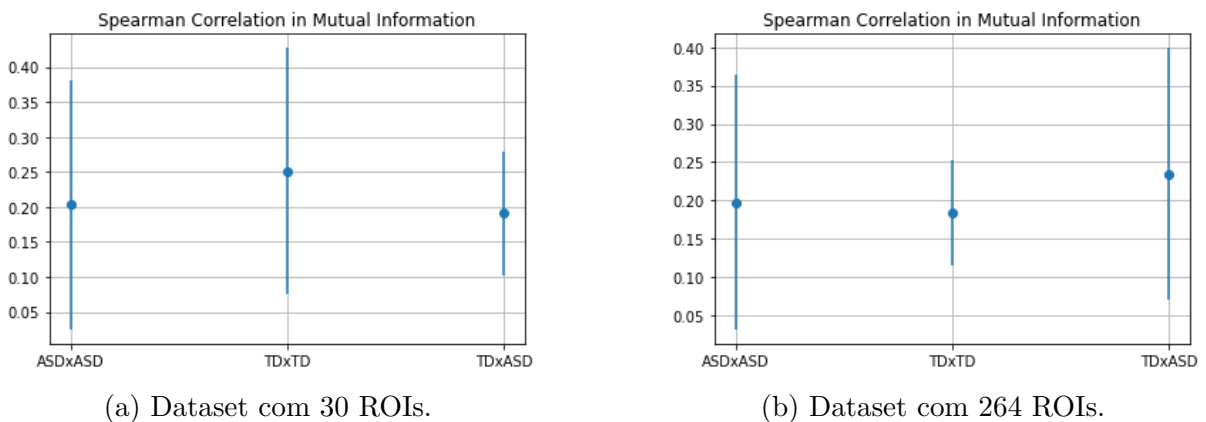


(a) Dataset com 30 ROIs.

(b) Dataset com 264 ROIs.

Fonte: Autora

Figura 13 – Média e desvio padrão da Correlação Spearman aplicada à Informação Mútua.



(a) Dataset com 30 ROIs.

(b) Dataset com 264 ROIs.

Fonte: Autora

Pode-se observar a partir dos gráficos que os valores de correlação são muito próximos ao comparar os grupos (ASDxASD, TDxTD e TDxASD), o que pode indicar

que não há tanta diferença nas correlações dos diferentes grupos, porém estudos anteriores mostram que mesmo essas pequenas diferenças podem apresentar informações importantes em relação ao diagnóstico de distúrbios mentais e psicológicos (SPRONK et al., 2021).

Idade, estilo de vida, rotina, são fatores que podem influenciar nas diferenças de correlação entre os grupos (SPRONK et al., 2021), portanto não é possível dizer com certeza que as diferenças de correlação estão diretamente ligadas aos distúrbios dos indivíduos. Observa-se na Tabela 2 que a correlação Parcial não apresentou grandes quantidades de correlação nos grupos com 264 ROIs, e valores baixos de correlação dos grupos com 30 ROIs, sendo assim pode-se descartar essa como melhor maneira de analisar as conexões de fMRIs. Isso pode ter acontecido pelo fator de que os dados apresentam distribuição normal.

Tabela 2 – Média e Desvio Padrão das correlações Spearman.

ROIs	Grupos	C. Pearson	C. Parcial	Informação Mútua
30	ASD x ASD	0,410 ± 0,166	0,204 ± 0,169	0,203 ± 0,178
30	TD x TD	0,521 ± 0,144	0,214 ± 0,167	0,251 ± 0,176
30	TD x ASD	0,439 ± 0,117	0,178 ± 0,061	0,191 ± 0,088
264	ASD x ASD	0,351 ± 0,144	0,050 ± 0,179	0,198 ± 0,166
264	TD x TD	0,420 ± 0,131	0,054 ± 0,189	0,184 ± 0,068
264	TD x ASD	0,358 ± 0,072	0,017 ± 0,006	0,235 ± 0,164

Fonte: Autora.

Ao comparar tanto as imagens quanto os valores de correlação Spearman das correlações Pearson e Informação Mútua nota-se muitas semelhanças. Apesar de apresentar maiores valores, não é possível afirmar que a correlação Pearson é melhor que a Informação Mútua, uma vez que considerando a média e desvio padrão da correlação Spearman os valores podem estar na mesma faixa. Para isso, seria necessário analisar a acurácia das correlações.

Cohen (COHEN, 1992) sugere os seguintes tamanhos de efeito:

- $r = |0, 10| \implies$ correlação fraca.
- $r = |0, 30| \implies$ correlação moderada.
- $r = |0, 50| \implies$ correlação forte.

Já o artigo “How to Interpret a Correlation Coefficient r – dummies” (RUMSEY, 2016) sugere que:

- $r = |0,30| \implies$ correlação fraca.
- $r = |0,50| \implies$ correlação moderada.
- $r = |0,70| \implies$ correlação forte.

Mukaka ([MUKAKA, 2012](#)) sugere que:

- 0,9 para mais ou para menos indica uma correlação muito forte.
- 0,7 a 0,9 positivo ou negativo indica uma correlação forte.
- 0,5 a 0,7 positivo ou negativo indica uma correlação moderada.
- 0,3 a 0,5 positivo ou negativo indica uma correlação fraca.
- 0 a 0,3 positivo ou negativo indica uma correlação desprezível.

Seguindo a sugestão do primeiro, as correlações Pearson seriam moderadas, e as demais (correlação Parcial e Informação Mútua) seriam fracas. Se considerada a segunda ou a terceira, todas as correlações seriam fracas. Observa-se também, que nenhuma correlação teve resultado negativo, o que indica que todas as variáveis estão diretamente correlacionadas.

5 Conclusão

A partir deste trabalho, foi possível selecionar e dividir dados de fMRI de pacientes com transtorno neurológico e de um grupo controle. Além disso, implementar algoritmos para análise dessas imagens baseado em teorias de grafos, analisando padrões a partir da visualização de gráficos e de comparação dos valores de média e desvio padrão obtidos.

Buscou-se utilizar fMRIs de grupos de indivíduos com transtornos neurológicos, em razão da possibilidade de encontrar padrões e diferenças nas conectividades das regiões do cérebro, o que pode facilitar no diagnóstico ou ajudar a encontrar tratamento para os indivíduos com tais transtornos. Para isso, foi utilizada a linguagem Python como ferramenta de análise dos dados, utilizando de ferramentas da computação e teoria de Grafos para analisar a correlação entre essas regiões.

Os métodos escolhidos para estimar as correlações entre os nós (ROIs) são amplamente usados na literatura para análises de fMRIs (SMITH et al., 2011). Cada um deles apresenta sua particularidade e podem, ou não, serem bons. Com os resultados obtidos a correlação Parcial foi descartada como um dos métodos a serem considerados bons, porque ele apresentou uma média de correlação Spearman muito próxima de zero, mostrando que as regiões não estão correlacionadas. Essa exclusão foi devido à semelhança dos valores de correlação Spearman e dos gráficos de correlação dos outros dois métodos (Correlação Pearson e Informação Mútua).

Com os valores e gráficos obtidos nas correlações Pearson e Informação Mútua não pode-se afirmar que um método seja melhor que o outro, pois eles apresentam faixas de correlação semelhantes, podendo assim serem ambos usados para esse tipo de análise. A escolha do método depende do objetivo da pesquisa e do tipo de dado em questão.

Como trabalhos futuros propõe-se analisar a acurácia das correlações Pearson e Informação Mútua, utilizar dados de indivíduos com outros tipos de transtornos neurológicos a fim de encontrar diferenças e semelhanças nas conexões, comparar os métodos escolhidos com outros tipos de métodos de correlação e analisar fatores como gênero, idade e rotina na hora de comparar os valores de correlações.

Referências

- BASSETT, D. S.; SPORNS, O. Network neuroscience. *Nature neuroscience*, Nature Publishing Group, v. 20, n. 3, p. 353–364, 2017. Citado na página 17.
- BASTOS, A. M.; SCHOFFELEN, J.-M. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in systems neuroscience*, Frontiers, v. 9, p. 175, 2016. Citado na página 17.
- BENESTY, J. et al. Pearson correlation coefficient. In: *Noise reduction in speech processing*. [S.l.]: Springer, 2009. p. 1–4. Citado na página 20.
- BULLMORE, E.; SPORNS, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews neuroscience*, Nature Publishing Group, v. 10, n. 3, p. 186–198, 2009. Citado na página 15.
- COHEN, J. Statistical power analysis. *Current directions in psychological science*, Sage Publications Sage CA: Los Angeles, CA, v. 1, n. 3, p. 98–101, 1992. Citado na página 26.
- COLE, M. W. et al. Variable global dysconnectivity and individual differences in schizophrenia. *Biological psychiatry*, Elsevier, v. 70, n. 1, p. 43–50, 2011. Citado na página 11.
- DANIEL, W. W.; CROSS, C. L. *Biostatistics: a foundation for analysis in the health sciences*. [S.l.]: Wiley, 2018. Citado na página 21.
- FAIR, D. A. et al. Functional brain networks develop from a “local to distributed” organization. *PLoS computational biology*, Public Library of Science San Francisco, USA, v. 5, n. 5, p. e1000381, 2009. Citado 2 vezes nas páginas 11 e 19.
- GIERLICH, B. et al. Mutual information analysis. In: SPRINGER. *International Workshop on Cryptographic Hardware and Embedded Systems*. [S.l.], 2008. p. 426–442. Citado na página 21.
- GRANDJEAN, M. *Connected World: Untangling the Air Traffic Network*. 2016. Disponível em: <<http://www.martingrandjean.ch/connected-world-air-traffic-network/>>. Acesso em: 18 de outubro de 2021. Citado na página 16.
- KATTI, G.; ARA, S. A.; SHIREEN, A. Magnetic resonance imaging (mri)—a review. *International journal of dental clinics*, Citeseer, v. 3, n. 1, p. 65–70, 2011. Citado 2 vezes nas páginas 11 e 13.
- LV, H. et al. Resting-state functional mri: everything that nonexperts have always wanted to know. *American Journal of Neuroradiology*, Am Soc Neuroradiology, v. 39, n. 8, p. 1390–1399, 2018. Citado na página 14.
- MAZZOLA, A. A. Ressonância magnética: princípios de formação da imagem e aplicações em imagem funcional. *Revista Brasileira de Física Médica*, v. 3, n. 1, p. 117–129, 2009. Citado 2 vezes nas páginas 11 e 14.

- MCKINNEY, W. *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. [S.l.]: "O'Reilly Media, Inc.", 2012. Citado na página 15.
- MUKAKA, M. M. A guide to appropriate use of correlation coefficient in medical research. *Malawi medical journal*, v. 24, n. 3, p. 69–71, 2012. Citado na página 27.
- NEWMAN, M. E. The structure and function of complex networks. *SIAM review*, SIAM, v. 45, n. 2, p. 167–256, 2003. Citado 2 vezes nas páginas 11 e 16.
- NORRIS, D. G. Principles of magnetic resonance assessment of brain function. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Wiley Online Library, v. 23, n. 6, p. 794–807, 2006. Citado na página 14.
- NYU Langone Medical Center. *Autism Brain Imaging Data Exchange*. 2012. Disponível em: <http://fcon_1000.projects.nitrc.org/indi/abide/manuscripts.html>. Acesso em: 18 de outubro de 2021. Citado 2 vezes nas páginas 19 e 20.
- O'GRADY, S. *The RedMonk Programming Language Rankings: June 2021*. 2021. Disponível em: <<https://redmonk.com/sogrady/2021/08/05/language-rankings-6-21/>>. Acesso em: 18 de outubro de 2021. Citado na página 15.
- POWER, J. D. et al. Functional network organization of the human brain. *Neuron*, Elsevier, v. 72, n. 4, p. 665–678, 2011. Citado 2 vezes nas páginas 11 e 19.
- RUBINOV, M.; SPORNS, O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, Elsevier, v. 52, n. 3, p. 1059–1069, 2010. Citado na página 11.
- RUMSEY, D. J. How to interpret a correlation coefficient r. *Statistics For Dummies*, 2016. Citado na página 26.
- SALMON, C. E. G.; LEONI, R. F. Conectividade funcional cerebral utilizando técnicas de imagens por ressonância magnética. *Revista Brasileira De Física Médica*, v. 13, n. 1, p. 66–75, 2019. Citado na página 14.
- SÁNCHEZ, P. *Grafo Simples*. 2011. Disponível em: <https://commons.wikimedia.org/wiki/File:Grafo_simple.svg>. Acesso em: 18 de outubro de 2021. Citado na página 16.
- SMITH, S. M. et al. Network modelling methods for fmri. *Neuroimage*, Elsevier, v. 54, n. 2, p. 875–891, 2011. Citado 3 vezes nas páginas 20, 21 e 28.
- SOUZA, I. T. S. de. *Ressonância magnética*. 2021. Disponível em: <https://www.professorinterativo.com.br/aval_on_line/02_AI2_text_quest/Trab_Inter47/texto.htm>. Acesso em: 18 de outubro de 2021. Citado na página 13.
- SPRONK, M. et al. A whole-brain and cross-diagnostic perspective on functional brain network dysfunction. *Cerebral Cortex*, Oxford University Press, v. 31, n. 1, p. 547–561, 2021. Citado 3 vezes nas páginas 11, 19 e 26.
- WATANABE, T. et al. A pairwise maximum entropy model accurately describes resting-state human brain networks. *Nature communications*, Nature Publishing Group, v. 4, n. 1, p. 1–10, 2013. Citado na página 21.

ZAR, J. H. Spearman rank correlation. *Encyclopedia of biostatistics*, Wiley Online Library, v. 7, 2005. Citado na página [22](#).