

Detecção da Doença de Alzheimer através de Parâmetros Não-Lineares de Sinais de Fala

Martim G. Silva, Pedro Ribeiro, Bruno C. Bispo e Pedro M. Rodrigues

Resumo—Este trabalho tem como objetivo a detecção da doença de Alzheimer (DA) através de parâmetros não-lineares de sinais de fala. Os parâmetros são extraídos de sub-bandas dos sinais, as quais são obtidas por meio da transformada Wavelet, e algumas das suas estatísticas descritivas são utilizadas como entrada para vários classificadores. Acurácias de 100, 77,8 e 85,2% são obtidas na detecção da DA entre mulheres, homens e todos, respectivamente, utilizando classificadores de regressão logística.

Palavras-Chave—Doença de Alzheimer, sinal de fala, análise não-linear, wavelet, classificadores.

Abstract—This work aims to detect Alzheimer’s disease (AD) through non-linear features of speech signals. The features are extracted from signal subbands, which are obtained through the wavelet transform, and some of its descriptive statistics are used as input for several classifiers. Accuracies of 100, 77.8 and 85.2% are obtained in detecting AD among women, men and all, respectively, using logistic regression classifiers.

Keywords—Alzheimer’s disease, speech signal, non-linear analysis, wavelet, classifiers.

I. INTRODUÇÃO

Os casos de demência estão aumentando ao redor do mundo e é previsto que tripliquem até 2050 [1]. Também é estimado que entre 50 e 75% dos casos atuais são devidos à Doença de Alzheimer (DA) [2], a qual atualmente é a doença neurodegenerativa mais comum e não tem cura [3]. Seu diagnóstico numa fase inicial ou assintomática pode permitir a medicamentos inovadores retardarem os seus efeitos nefastos.

Como meios de diagnóstico da DA, clínicos e pesquisadores utilizam atualmente testes cognitivos, especialmente o MMSE (do inglês *Mini-Mental State Examination*) [4], testes sanguíneos [5], imagem por ressonância magnética (MRI) [6], tomografia computadorizada (CT) [7], tomografia por emissão de pósitrons (PET) [8], exame de eletroencefalograma [9] e exames neurológicos [10]. Estes métodos são utilizados para avaliar a presença de biomarcadores específicos, como as proteínas β -amiloide e τ , e possíveis anomalias em determinadas regiões do cérebro. Porém, as taxas de acurácia de diagnóstico nesta fase inicial ainda se encontram longe do satisfatório.

Por ser rico em informações como conteúdo semântico, emoção, entoação e características individuais, o sinal de fala tem sido recentemente explorado para auxiliar o diagnóstico da DA. O fato do sinal de fala ser relativamente fácil de ser gravado e digitalizado, sendo um procedimento menos invasivo, de menor consumo de tempo e mais conveniente para um primeira triagem da doença do que os exames de

diagnósticos mais habituais em clínica, tem contribuído para o aumento de pesquisa nessa área [11].

No que diz respeito a classificação entre pessoas saudáveis e pacientes diagnosticados com DA a partir de sinais de fala, acurácias entre 73 e 88% foram obtidas em [12]–[18]. Os parâmetros e os classificadores utilizadas nesses trabalhos, assim como as acurácias obtidas, estão resumidas na Tabela I.

TABELA I

ESTADO DA ARTE - SUJEITOS SAÚDAVEIS VS. PACIENTES COM DA

Ref.	Métricas	Classificadores	Acurácia (%)
[12]	Espectogramas	Regressão Logística CV	75 - 81
[13]	Tempo médio de silêncio, tempo médio de conteúdo vocal, rácio de silêncio, média de momentos de silêncio, número de fonemas e de palavras.	Análise Discriminante Linear	80 - 88
[14]	Traços de fala espontânea (SE) - vocais, linguísticos e acústicos	Ensemble	87
[15]	Tempo de reação vocal, quantidade de silêncio e irregularidade do sinal	Support-vector-machines	60 - 84,5
[16]	Complexidade Linguística e fluência	Regressão Logística	76 - 83
[17]	Coefficientes Mel Cepstrais (MFCC)	Regressor de aumento de gradiente e Support-vector-machines	81,5
[18]	Caraterísticas Acústicas - Emobase e IS10	Ensemble Support-vector-machines	76-80

Este trabalho tem como objetivo avaliar a aplicação de parâmetros não-lineares de sinais de fala para detectar DA entre mulheres, homens e todos. Com o intuito de explorar uma possível maior capacidade discriminante de bandas frequenciais, os parâmetros são obtidos, por meio da transformada wavelet, de sub-bandas dos sinais um algoritmo de suporte à detecção da DA por meio de uma análise não linear do sinal de voz.

O artigo está organizado da seguinte maneira: a Seção II descreve a base de sinais de fala; a Seção III explica a decomposição multibanda realizada; a Seção IV apresenta os parâmetros não-lineares utilizados; a Seção V descreve a metodologia empregada; a Seção VI apresenta e discute os resultados obtidos; a Seção VII conclui o trabalho

II. BASE DE DADOS DE FALA

A base de sinais de fala utilizada foi a *DementiaBank Spanish Ivanova Corpus* [19], [20]. Ela contém 269 áudios, sendo um áudio por indivíduo, amostrados a uma frequência de amostragem de 44100 Hz e com durações entre 25 e 45 s [2]. Os conteúdos são frases em língua espanhola de pessoas

TABELA II

RESUMO DOS GRUPOS DE ESTUDO PROVENIENTES DA BASE DE DADOS.

Grupo	Gênero	Número	Idade	MMSE
MCS	Mulher	137	76 ± 7,4	28,2 ± 2,0
MDA	Mulher	44	80 ± 7,9	18,9 ± 4,7
HCS	Homem	58	75 ± 9,0	28,4 ± 1,6
HDA	Homem	30	79 ± 7,7	21,5 ± 5,4
TCS	Todos	195	75 ± 7,9	28,3 ± 1,9
TDA	Todos	74	79 ± 7,8	20,1 ± 5,1

sem diagnóstico de DA (grupo de controle) e de pessoas previamente diagnosticadas com DA.

Neste trabalho, a base de dados foi dividida por gênero, resultando em 88 áudios de homens e 181 de mulheres. Cada um desses grupos foi subdividido entre controle (CS) e pacientes com DA, com uma distribuição de 137 áudios de mulheres de controle (MCS), 58 de homens de controle (HCS), 44 de mulheres com DA (MDA) e 30 de homens com DA (HDA). Maiores informações sobre os grupos de estudos provenientes da base de dados são mostradas na Tabela II.

Artefatos presentes nos sinais de fala foram removidos ao aplicar um filtro Butterworth de segunda ordem, passa-baixas e com frequência de corte de 1500 Hz. Em seguida, os trechos de silêncio foram removidos ao descartar as amostras com magnitude inferior a 0,001, considerando os sinais de fala com magnitude máxima de 1. Por fim, os sinais foram segmentados em trechos com duração de 20 ms e sobreposição de 50% usando a janela de Hamming.

III. DECOMPOSIÇÃO MULTIBANDA

A transformada wavelet discreta (DWT, do inglês *Discrete Wavelet Transform*) de um sinal de tempo discreto e energia finita é a sua decomposição em um conjunto de funções base obtidas a partir de um número finito de sequências protótipas e suas versões deslocadas no tempo [21].

Essa expansão estruturada é implementada por meio de um banco de filtros em bandas de oitavas e criticamente decimado [22], [23]. Considerando apenas as frequências positivas, a m -ésima sub-banda está limitada a

$$W_m = \begin{cases} [0, \pi/2^S], & m = 0, \\ [\pi/2^{S-m+1}, \pi/2^{S-m}], & m = 1, 2, \dots, S, \end{cases} \quad (1)$$

onde S é o número de sub-bandas ou níveis de decomposição e π é a frequência angular normalizada [23].

A DWT utiliza uma função escala de análise $\tilde{\phi}_1(n)$ e uma função wavelet de análise $\tilde{\psi}_1(n)$ que são respostas ao impulso dos filtros de análise passa-baixa e passa-alta de meia banda, respectivamente. Definindo as seguintes fórmulas de recursão

$$\tilde{\phi}_{i+1}(n) = \tilde{\phi}_i(n/2) * \tilde{\phi}_1(n) \quad (2)$$

$$\tilde{\psi}_{i+1}(n) = \tilde{\phi}_i(n) * \tilde{\psi}_1(n/2^i), \quad (3)$$

o filtro de análise equivalente da m -ésima sub-banda é dado por

$$h_m(n) = \begin{cases} \tilde{\phi}_S(n), & m = 0 \\ \tilde{\psi}_{S+1-m}(n), & m = 1, 2, \dots, S. \end{cases} \quad (4)$$

E o sinal da m -ésima sub-banda do sinal $x(n)$ é dado por

$$x_m(n) = \begin{cases} \sum_{k=-\infty}^{\infty} x(k)h_m(2^S n - k), & m = 0, \\ \sum_{k=-\infty}^{\infty} x(k)h_m(2^{S-m+1} n - k), & m = 1, 2, \dots, S. \end{cases} \quad (5)$$

Neste trabalho, a decomposição em sub-banda foi aplicada até o quinto nível, $S = 5$, aos segmentos dos sinais de fala utilizando a wavelet-mãe Biortogonal 3.3. Em trabalhos anteriores [24], esta wavelet emergiu com uma excelente opção para análise de sinais da voz. Os sinais das sub-bandas foram reamostrados para a frequência de amostragem original através do método de interpolação wavelet [25].

IV. ANÁLISE NÃO-LINEAR

Os parâmetros não-lineares utilizados neste trabalho serão brevemente em seguida descritos, assumindo que o sinal $x(n)$ possui tamanho N e média \bar{x} . Maiores detalhes sobre os parâmetros podem ser encontrados na literatura [26], [27].

A. Medidas do Atrator Reconstruído

Em cada instante de tempo, o estado de um sistema dinâmico definido por m variáveis pode ser representado por um ponto no espaço m -dimensional. Nesse espaço, denominado espaço de estado ou de fase, a sequência de estados ao longo do tempo define curvas denominadas trajetórias. Estas trajetórias, quando observadas por longos períodos, podem convergir para uma estrutura geométrica específica e independente das condições iniciais do sistema, denominada atrator [26].

Entre as técnicas para representar o espaço de estado de um sistema dinâmico a partir de uma única série temporal (sinal), processo denominado reconstrução, utilizou-se a técnica de incorporação de atraso. Nesse caso, uma reconstrução do espaço de estado é dada por

$$\mathbf{x}_i = [x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)], \quad (6)$$

onde τ é o atraso de incorporação [26]. Os valores de τ e m foram obtidos de $x(n)$ por meio dos métodos descritos em [26], [28].

A sequência de vetores $\mathbf{x}_i, i = 1, 2, \dots, M$ onde $M = N - (m - 1)\tau$, formam o atrator reconstruído.

1) *Dimensão da Correlação*: Medida que caracteriza a distribuição dos pontos do atrator, refletindo a complexidade do sistema dinâmico, estimada como

$$D_2 = \lim_{r \rightarrow 0} \frac{\log(C(r, M))}{\log(r)}, \quad (7)$$

onde

$$C(r, M) = \frac{2}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|) \quad (8)$$

é a probabilidade do par de pontos $\{\mathbf{x}_i, \mathbf{x}_j\}$ no atrator estar separado por uma distância menor ou igual a r e Θ é a função Heaviside [26], [28].

2) *Expoentes de Lyapunov*: Medidas que fornecem informação sobre a evolução das trajetórias ao longo do tempo [28], refletindo a estabilidade do sistema dinâmico [26]. Coeficientes positivos e negativos indicam divergência e convergência exponencial de trajetórias próximas no atrator, respectivamente [29]. Um sistema caótico tem pelo menos um expoente positivo [28], [29]. O maior expoente de Lyapunov (LLE, do inglês *Largest Lyapunov Exponent*) pode ser estimado ao, para cada estado \mathbf{x}_i , encontrar o estado \mathbf{x}_j que satisfaça $\min_j \|\mathbf{x}_i - \mathbf{x}_j\|$ tal que $|i - j| > T_m$, onde T_m é o período médio de $x(n)$. As estimativas são dadas por

$$\lambda(i) = \frac{1}{M+2} \sum_{k=1}^M \frac{1}{kT_s} \ln \frac{\|\mathbf{x}_{i+k} - \mathbf{x}_{j+k}\|}{\|\mathbf{x}_i - \mathbf{x}_j\|}, \quad (9)$$

onde T_s é o período de amostragem. E o LLE é definido pela inclinação da melhor aproximação linear de $\lambda(i)$ [30].

B. Medidas de Memória Longa

1) *Expoente Hurst*: Medida da dependência estatística de longo alcance de uma série temporal, utilizada para avaliar a sua suavidade, propriedades de correlação e grau de auto-similaridade [26], [28]. Pode ser definida como

$$H = \frac{\log(R/S)}{\log(N)}, \quad (10)$$

onde R/S é uma medida estatística denominada intervalo reescalado [28]. Na prática, H é estimado pela inclinação da melhor aproximação linear de $\log[R(n)/S(n)]$ em função de $\log(n)$, onde detalhes sobre o cálculo de $R(n)/S(n)$ podem ser encontrados em [31].

2) *Análise de Flutuação Destendenciada*: Método de análise que fornece um parâmetro para quantificar as correlações de longo alcance (auto-similaridade) de séries temporais aparentemente não-estacionárias [27], [32]. A partir de $x(n)$, calcula-se a série de desvio cumulativo $y(k) = \sum_{i=1}^k [x(i) - \bar{x}]$. Em seguida, para cada trecho de comprimento m de $y(k)$, calcula-se uma aproximação linear denotada por $y_m(k)$. A flutuação média do sinal é definida como

$$F(m) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_m(k)]^2}. \quad (11)$$

A inclinação da melhor aproximação linear de $\log[F(m)]$ em função de $\log(m)$ é o expoente de escala α que representa as propriedades de correlação do sinal $x(n)$ [27].

C. Dimensão Fractal

A dimensão fractal (FD, do inglês *Fractal Dimension*) de um objeto geométrico é uma medida de sua complexidade - extensão espacial, complicação ou propensão ao preenchimento do espaço - e está intimamente ligada à sua forma [33].

Numa forma de onda, a FD é uma ferramenta poderosa para a detecção de eventos transitórios [28]. Ela pode ser calculada diretamente a partir dos sinais, não havendo necessidade de reconstruir o atrator. Entre os algoritmos para determinar a FD de uma forma de onda, os brevemente descritos abaixo são conhecidos por serem adequados para análise de EEG [28].

1) *Algoritmo de Higuchi*: Para $m = 1, \dots, k$ e $k = 1, \dots, k_{\max}$, onde k_{\max} é obtido experimentalmente apesar de $k_{\max} = 8$ ter sido inicialmente proposto, calcula-se a seguinte medida de distância

$$L_m(k) = \frac{N-1}{[a]k} \sum_{i=1}^{[a]} |x(m+ik) - x(m+(i-1)k)|, \quad (12)$$

onde $a = (N-m)/k$ e $[a]$ representa o maior número inteiro igual ou menor que a . A distância média é calculada como $L(k) = \sum_{m=1}^k L_m(k)/k$ para $k = 1, \dots, k_{\max}$. A estimativa da FD, denotada por FD_H , é então dada pela inclinação da melhor aproximação linear de $\ln[L(k)]$ em função de $\ln(1/k)$ [26], [28].

2) *Algoritmo de Katz*: Segundo [33], a FD de uma forma de onda $x(n)$ pode ser definida como

$$FD_K = \frac{\log(L/a)}{\log(d/a)}, \quad (13)$$

onde L é a soma das distâncias entre os pontos sucessivos de $x(n)$, a é a distância média entre os pontos sucessivos e d é a maior distância entre $x(1)$ e os restantes pontos de $x(n)$.

D. Energia

A energia das sub-bandas está entre os parâmetros mais utilizados na detecção de disfonias vocais e é definida como

$$EN = \sum_{n=1}^N |x(n)|^2. \quad (14)$$

E. Entropia

A entropia é uma medida que descreve a quantidade de desordem, incerteza, em um sistema [28]. No caso de uma variável aleatória, a entropia quantifica a informação média associada às suas observações. As entropias de Shannon e Logarítmica podem ser estimadas como

$$ET_S = - \sum_{n=1}^N |x(n)|^2 \log[|x(n)|^2] \quad (15)$$

e

$$ET_L = \sum_{n=1}^N \log[|x(n)|^2], \quad (16)$$

respectivamente [34], [35].

A entropia aproximada (ApET) é uma modificação na entropia de Kolmogorov-Sinai, que descreve a taxa na qual informações sobre o estado do sistema dinâmico são perdidas ao longo do tempo, para evitar divergências numéricas quando o sinal é contaminado por ruído, sendo definida como

$$ApET(m, r, N) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \log[C_i^m(r)] - \frac{1}{N-m} \sum_{i=1}^{N-m} \log[C_i^{m+1}(r)], \quad (17)$$

onde

$$C_i^m(r) = \frac{1}{N-m+1} \sum_{j=1}^{N-m+1} \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|) \quad (18)$$

é a probabilidade do ponto \mathbf{x}_i no atrator estar separado dos outros pontos por uma distância menor ou igual a r [28].

V. METODOLOGIA

Este trabalho foi desenvolvido num computador com 8 GB de memória RAM, processador Intel Core i5-9300HF e placa gráfica NVIDIA GeForce GTX 1650, e utilizou os softwares Matlab R2022b 64-bits e Python 3.10.10 64-bits.

Para cada um dos 269 indivíduos, os 10 parâmetros não-lineares (D_2 , LLE , α , FD_H , FD_K , EN , ET_S , ET_L e $ApET$) foram calculados para os sinais das 6 sub-bandas de todos os segmentos. Para cada série temporal de características, as seguintes estatísticas descritivas foram calculadas: média, desvio padrão, assimetria, curtose, variância e raiz quadrada média. Isso resultou em 360 parâmetros (6 sub-bandas \times 10 parâmetros \times 6 estatísticas) por indivíduo. Essa etapa foi executada no software Matlab.

Em seguida, no ambiente Python, os dados foram aplicados a métodos de aprendizado de máquina para diferenciar dentro dos pares de grupos MCS-MDA, HCS-HDS e TCS-TDA.

A. Processos de Classificação

Dentro de cada par de grupo de estudo, nomeadamente MCS-MDA, HCS-HDA e TCS-TDA, os parâmetros foram normalizados por F-score. E, aleatoriamente, 80% do total de dados disponíveis foram utilizados para treinar e os restantes 20% para testar 19 classificadores. A lista dos classificadores *scikit-learn* [36] e suas configurações é apresentada abaixo:

- *AdaBoostClassifier* (AdaBoost) - Parâmetros padrão
- *BaggingClassifier* (BaggC) - Parâmetros padrão
- *DecisionTreeClassifier* (DeTreeC) - Profundidade máxima = 5
- *ExtraTreesClassifier* (ExTreeC) - Número de estimadores = 300
- *GaussianNB* (GauNB) - Parâmetros padrão
- *GradientBoostingClassifier* (GradBoost) - Parâmetros padrão
- *GaussianProcessClassifier* (GauPro) - $1.0 \times RBF(1.0)$
- *KNearestNeighbors* (KNN) - Parâmetros padrão
- *LogisticRegression* (LogReg) - Solver Limited-memory Broyden–Fletcher–Goldfarb–Shanno
- *LogisticRegressionCV* (LogRegCV) - CV = 3
- *LinearDiscriminantAnalysis* (LinDis) - Parâmetros padrão
- *LinearSVC* (LinSVC) - Parâmetros padrão
- *MLPClassifier* (MLP) - $\alpha = 1$ e iteração máxima = 1000
- *OneVsRestClassifier* (OvsR) - Estado aleatório = 0
- *QuadraticDiscriminantAnalysis* (QuadDis) - Parâmetros padrão
- *RandomForestClassifier* (RF) - Profundidade máxima = 300, número de estimadores = 5, número de características máximas = 1
- *SGDClassifier* (SGD) - Iteração máxima = 100 e tolerância = 0,001
- *SGDClassifier* (SGDC) - Parâmetros padrão
- *Support vector machines* (SVM) - γ automático

B. Métricas de Classificação

As métricas de classificação utilizadas para determinar a eficácia do modelos de aprendizagem foram acurácia (AC), precisão (PR), *recall* (Rec) e F1-score, as quais são definidas como

$$AC = \frac{VP + VN}{VP + VN + FP + FN} \times 100, \quad (19)$$

TABELA III
MELHORES RESULTADOS DE CLASSIFICAÇÃO.

Grupos	Acurácia (%)	Precisão (%)	Recall (%)	F1-score	Classificador
MCS-MDA	100	100	100	1,00	LogRegCV
HCS-HDA	77,8	86,7	71,4	0,723	LogReg
TCS-TDA	85,2	79,3	75,6	0,772	LogReg

$$PR = \frac{VP}{VP + FP} \times 100, \quad (20)$$

$$Rec = \frac{VP}{VP + FN} \times 100, \quad (21)$$

$$F1-score = \frac{2 \times PR \times Rec}{PR + Rec}, \quad (22)$$

onde, VP , VN , FP e FN representam o número de casos de verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo, respetivamente.

VI. RESULTADOS E DISCUSSÃO

O melhor resultado de classificação obtido para cada par de grupo de estudo analisado é mostrado na Tabela III, onde os classificadores que geraram esses resultados são indicados.

Nota-se que, para o par MCS-MDA, valores máximos de classificação foram obtidos em todas as métricas, indicando que o método proposto é capaz de identificar entre sinais de fala de mulheres saudáveis e de mulheres com DA. No par HCS-HDA, valores acima de 70% foram obtidos em todas as métricas, com destaque para 86,7% de precisão. A maior eficácia do método proposto em detectar DA a partir da fala de mulheres pode estar relacionada com a discrepância do número de mulheres e homens na base de dados, onde 67% dos sinais foram proferidos por mulheres.

Em relação ao par TCS-TDA, onde todos os indivíduos de controle e pacientes com DA são incluídos independentemente do gênero, resultados de classificação superiores a 75% foram obtidos, com destaque para 85,2% de acurácia. Observa-se que, nos três pares de grupos de estudo, o melhor resultado de classificação foi obtido utilizando o classificador de regressão logística, com validação cruzada no caso MCS-MDA.

Em comparação com os métodos do estado-da-arte para detectar DA no par de grupos TCS-TDA, a acurácia obtida no presente trabalho supera as alcançadas em [12], [14]–[18], sendo superada apenas pelo 88% alcançado em [13], como pode ser visto na Tabela I. No entanto, é importante ressaltar que bases de dados diferentes foram utilizadas nos trabalhos mencionados e, por conseguinte, comparações diretas de desempenho devem ser feitas com cautela. De qualquer forma, os resultados obtidos no presente trabalho indicam que a combinação de parâmetros não-lineares multibanda de sinais fala pode ser adequada para auxiliar na detecção da DA.

VII. CONCLUSÕES

Este trabalho investigou a detecção da doença de Alzheimer (DA) através de parâmetros não-lineares de sinais de fala. Dez parâmetros foram extraídos de seis sub-bandas dos sinais, as quais foram obtidas por meio da transformada Wavelet, e seis

estatísticas descritivas dos parâmetros foram utilizadas como entrada para dezenove classificadores. Acurácias de 100, 77,8 e 85,2% foram obtidas na detecção de DA entre mulheres, homens e todos, respectivamente, utilizando classificadores de regressão logística.

A maior eficácia do método proposto em detectar DA a partir da fala de mulheres pode estar relacionada com a discrepância do número de mulheres e homens na base de dados, onde 67% dos sinais foram proferidos por mulheres. Apesar de indicarem que a combinação de parâmetros não-lineares multibanda de sinais fala é adequada para auxiliar na detecção da DA, futuramente os resultados devem ser atualizados com uma população maior e mais equilibrada.

AGRADECIMENTOS

Esse trabalho foi apoiado pelos Fundos Nacionais da Fundação para a Ciência e a Tecnologia (FCT), Portugal, através do projeto UIDB/50016/2020.

REFERÊNCIAS

- [1] E. Nichols and T. Vos, "The estimation of the global prevalence of dementia from 1990-2019 and forecasted prevalence through 2050: An analysis for the global burden of disease (GBD) study 2019," *Alzheimer's & Dementia*, vol. 17, no. S10, dec 2021.
- [2] O. Ivanova, "Dementiabank spanish ivanova corpus," 2022.
- [3] J. M. Long and D. M. Holtzman, "Alzheimer disease: an update on pathobiology and treatment strategies," *Cell*, vol. 179, no. 2, pp. 312–339, 2019.
- [4] A. Brugnolo, F. Nobili, M. Barbieri, B. Dessi, A. Ferro, N. Girtler, E. Palummeri, D. Partinico, U. Raiteri, G. Regesta *et al.*, "The factorial structure of the mini mental state examination (mmse) in alzheimer's disease," *Archives of Gerontology and Geriatrics*, vol. 49, no. 1, pp. 180–185, 2009.
- [5] T. G. Beach, S. E. Monsell, L. E. Phillips, and W. Kukull, "Accuracy of the clinical diagnosis of alzheimer disease at national institute on aging alzheimer disease centers, 2005–2010," *Journal of neuropathology and experimental neurology*, vol. 71, no. 4, pp. 266–273, 2012.
- [6] J. Silva, B. C. Bispo, and P. M. Rodrigues, "Structural MRI texture analysis for detecting alzheimer's disease," *Journal of Medical and Biological Engineering*, apr 2023.
- [7] A. Smith, K. Jobst, M. Szatmari, A. Jaskowski, E. King, A. Smith, A. Molyneux, M. Esiri, B. McDonald, and N. Wald, "Detection in life of confirmed alzheimer's disease using a simple measurement of medial temporal lobe atrophy by computed tomography," *The Lancet*, vol. 340, no. 8829, pp. 1179–1183, 1992.
- [8] R. Duara, C. Grady, J. Haxby, M. Sundaram, N. Cutler, L. Heston, A. Moore, N. Schlageter, S. Larson, and S. I. Rapoport, "Positron emission tomography in alzheimer's disease," *Neurology*, vol. 36, no. 7, pp. 879–879, 1986.
- [9] P. M. Rodrigues, B. C. Bispo, C. Garrett, D. Alves, J. P. Teixeira, and D. Freitas, "Lacsogram: A new EEG tool to diagnose alzheimer's disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3384–3395, sep 2021.
- [10] B. Dubois, G. Picard, and M. Sarazin, "Early detection of alzheimer's disease: new diagnostic criteria," *Dialogues in clinical neuroscience*, 2022.
- [11] R. BROWN, "A first language," *The Early Stages*, 1973. [Online]. Available: <https://cir.nii.ac.jp/crid/1570291225247558400>
- [12] L. Liu, S. Zhao, H. Chen, and A. Wang, "A new machine learning method for identifying alzheimer's disease," *Simulation Modelling Practice and Theory*, vol. 99, p. 102023, feb 2020.
- [13] T. S. Jochen Weiner, Christian Herff, "Speech-based detection of alzheimer's disease in conversational german," *INTER_SPEECH 2016 / Cognitive Systems Lab, University of Bremen, Germany*, 2016.
- [14] A. S. Gulapalli and V. K. Mittal, "Detection of alzheimer's disease through speech features and machine learning classifiers," in *Intelligent Sustainable Systems*. Springer Nature Singapore, 2022, pp. 627–639.
- [15] A. König, A. Satt, A. Sorin, R. Hoory, O. Toledo-Ronen, A. Derreumaux, V. Manera, F. Verhey, P. Aalten, P. H. Robert *et al.*, "Automatic speech analysis for the assessment of patients with predementia and alzheimer's disease," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 1, no. 1, pp. 112–124, 2015.
- [16] Y. Qiao, X. Yin, D. Wiechmann, and E. Kerz, "Alzheimer's disease detection from spontaneous speech through combining linguistic complexity and (dis) fluency features with pretrained language models," *arXiv preprint arXiv:2106.08689*, 2021.
- [17] R. Pappagari, J. Cho, L. Moro-Velazquez, and N. Dehak, "Using state of the art speaker recognition and natural language processing technologies to detect alzheimer's disease and assess its severity," in *INTER_SPEECH*, 2020, pp. 2177–2181.
- [18] N. Wang, Y. Cao, S. Hao, Z. Shao, and K. Subbalakshmi, "Modular multi-modal attention network for alzheimer's disease detection using patient audio and language data," in *INTER_SPEECH*, 2021, pp. 3835–3839.
- [19] O. Ivanova, J. J. G. Meilán, F. Martínez-Sánchez, I. Martínez-Nicolás, T. E. Llorente, and N. C. González, "Discriminating speech traits of Alzheimer's disease assessed through a corpus of reading task for Spanish language," *Computer Speech & Language*, vol. 73, p. 101341, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0885230821001340>
- [20] A. M. Lanzi, A. K. Saylor, D. Fromm, H. Liu, B. MacWhinney, and M. L. Cohen, "DementiaBank: Theoretical rationale, protocol, and illustrative analyses," *American Journal of Speech-Language Pathology*, vol. 32, no. 2, pp. 426–438, mar 2023.
- [21] M. Vetterli and J. Kovacevic, *Wavelets and subband coding*. Prentice-hall, 1995, no. BOOK.
- [22] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, New Jersey: Prentice Hall, 1995.
- [23] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, Massachusetts: Artech House, 1992.
- [24] M. Alves, G. Silva, B. C. Bispo, M. E. Dajer, and P. M. Rodrigues, "Voice disorders detection through multiband cepstral features of sustained vowel," *Journal of Voice*, vol. 37, no. 3, pp. 322–331, may 2023.
- [25] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14–38, October 1991.
- [26] G. R. Bermúdez and P. J. G. Laencina, "Analysis of EEG signals using nonlinear dynamics and chaos - a review," *Applied Mathematics & Information Sciences*, vol. 9, pp. 1–13, 01 2015.
- [27] J.-M. Lee, D.-J. Kim, I.-Y. Kim, K.-S. Park, and S. I. Kim, "Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data," *Computers in biology and medicine*, vol. 32, no. 1, pp. 37–47, 2002.
- [28] O. Faust and M. G. Bairy, "Nonlinear analysis of physiological signals: a review," *Journal of Mechanics in Medicine and Biology*, vol. 12, no. 04, pp. 1–21, Sep. 2012.
- [29] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, pp. 1388–2457, Oct. 2005.
- [30] M. T. Rosenstein, J. J. Collins, and C. J. D. Lucaa, "A practical method for calculating largest Lyapunov exponents from small data sets," *Physica D: Nonlinear Phenomena*, vol. 65, no. 1, pp. 117–134, May 1993.
- [31] B. Qian and K. Rasheed, "Hurst exponent and financial market predictability," in *Proceedings of IASTED International Conference on Financial Engineering and Applications*, Berkeley, USA, September 2007, pp. 203–209.
- [32] P. Gifani, H. R. Rabiee, M. Hashemi, P. Taslimi, and M. Ghanbari, "Optimal fractal-scaling analysis of human EEG dynamic for depth of anesthesia quantification," *Journal of the Franklin Institute*, vol. 344, no. 3-4, pp. 212–229, 2007.
- [33] M. J. Katz, "Fractals and the analysis of waveforms," *Computers in Biology and Medicine*, vol. 18, no. 3, pp. 145–156, January 1988.
- [34] J. Rodríguez-Sotelo, A. Osorio-Forero, A. Jiménez-Rodríguez, D. Cuesta-Frau, E. Cirugeda-Roldán, and D. Peluffo, "Automatic sleep stages classification using EEG entropy features and unsupervised pattern analysis techniques," *Entropy*, vol. 16, no. 12, pp. 6573–6589, Dec. 2014.
- [35] A. B. Das and M. I. H. Bhuiyan, "Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain," *Biomedical Signal Processing and Control*, vol. 29, pp. 11–21, August 2016.
- [36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.