

# Autenticação Biométrica de Locutores Drasticamente Disfônicos Aprimorada pela *Imagined Speech*

- - - Projeto de Pesquisa de Doutorado - - -

- - Conexo ao Auxílio à Pesquisa FAPESP 2021/12407-4 - -

Candidato: André Furlan, *MSc*

Orientador: Prof. Dr. Eng. Rodrigo Capobianco Guido

Programa de Pós-Graduação em Ciência da Computação da UNESP

**Resumo:** A partir do momento em que uma alteração severa na fala dificulta a sua compreensão, e conseqüentemente a vida social do indivíduo, faz-se presente uma significativa disfonia de origem orgânica, funcional ou orgânico-funcional. Desse modo, a sua identificação acústica, não somente pelos seres humanos mas principalmente pelas máquinas, torna-se difícil ou até mesmo impossível. Assim, como este é um tema ainda modestamente explorado, a intenção do autor deste projeto de pesquisa é a de analisar as implicações das disfonias laríngeas severas (DLSs) na autenticação biométrica de locutores (ABLs), criando algoritmos robustos para esta finalidade frente à existência daqueles problemas. Para isso, explorar-se-ão, em conjunto com os sinais de voz degradados, sinais de eletroencefalograma (EEG) visando decodificar a *imagined speech* associada. Após uma ampla revisão bibliográfica envolvendo os conceitos pertinentes, será iniciado o procedimento investigativo. Na etapa de extração de características, a intenção é a de comparar o potencial das estratégias de *feature learning* baseadas em *autoencoders* com a análise provida por abordagens clássicas, tal como a Transformada *Wavelet-Packet* de Tempo Discreto (DTWPT), à luz da engenharia paraconsistente de características (EPCs). Em seguida, objetivando autenticar corretamente os locutores registrados na base de dados pública a ser utilizada nos experimentos, assim como naquela que será criada em parceria com profissionais da área da saúde, a acurácia e o desempenho de estratégias recentes, tais como as *Residual Neural Networks* (RNNs) e as *Deep Spiking Neural Networks* (DSNNs), serão avaliados e comparados em duas modalidades: *text-dependent* e *text-independent*. Finalmente, pretende-se documentar e publicar os resultados obtidos em renomados veículos literários.

# 1 Introdução e Enunciado do Problema

Em conformidade com os frutos colhidos em investigações pregressas na área de processamento de sinais de voz, tais como a dissertação [1], observa-se a existência de significativa demanda pelos mecanismos de autenticação biométrica de locutores (ABLS). Destacadas obras científicas, tais como [2] e [3], sedimentam o referido entendimento, o qual tem também sido perceptível diante do volume de artigos científicos publicados em periódicos de impacto, tais como aqueles veiculados no *IEEE Signal Processing Magazine* [4], no *Elsevier Neurocomputing* [5] e no *Elsevier Computer Speech and Language* [6].

Também no âmbito da análise de sinais de voz, percebe-se a demanda por uma linha de pesquisa diversa, que tem atraído a atenção da comunidade científica internacional, conforme se constata em diversas publicações recentes, tais como [7] e [8]: as estratégias acústico-computacionais para pré-diagnóstico e classificação não invasiva de alterações laríngeas e demais irregularidades que prejudiquem a fonação. Registre-se que expressivo número de trabalhos envolvendo este tema tem sido, inclusive, baseado em metodologias mais recentes para reconhecimento de padrões, tais como as *deep neural networks* [9], a exemplo daqueles documentados nas referências [10] e [11], além da lógica paraconsistente, tal como aquele descrito em [12].

Notavelmente, a associação de ambas as linhas de pesquisa mencionadas constitui um problema desafiador e que tem recebido atenção, apesar das dificuldades para solucioná-lo: autenticar locutores acometidos por distúrbios da voz, sejam eles de origem orgânica, funcional ou orgânico-funcional [13] [14], os quais modificam, de alguma maneira, a impressão acústica da fala.

O tópico em apreço foi objeto de investigação do orientador deste projeto de doutorado, com financiamento da FAPESP e do CNPq, tendo sido registrados, em função delas, alguns avanços [15]. Visando aprimorar tais investigações, uma possibilidade mais recente que tem sido considerada é a análise da *imagined speech* [16] associada às locuções degradadas.

Buscando pelo tema em questão nos registros do *Web of Science* e outras bases de dados conexas, tais como o *ieeexplore.org*, *springernature.com* e o *sciencedirect.com*, pode-se perceber uma nítida escassez de artigos científicos. Por um lado, a expressiva maioria constitui um conjunto de propostas genéricas envolvendo metodologias para viabilizar *brain-computer interfaces* (BCIs), tal como descrito nos artigos [17] e [18], ou que almejam o reconhecimento da fala imaginada em vez da autenticação de locutores, tal como é possível observar nos exemplos constantes nas referências de [19] até [24]. Por outro lado, os mais relevantes e particularmente dedicados à biometria são aqueles que constam nas referências [16], [25], [26], [27], [28] e [29]. Entretanto, nenhum destes trabalhos fornece qualquer solução absoluta para o problema em questão, nem mesmo associando a análise da *imagined speech* com a da voz degradada, principalmente quando a disfonia laríngea existente, seja ela de origem orgânica, funcional ou orgânico-funcional, é severa. Portanto, existindo ainda amplo espaço para futuras investigações, delimita-se o problema-alvo deste projeto de pesquisa: conceber algoritmos biométricos para autenticar, em princípio por meio da fala, indivíduos que sejam capazes de produzir somente locuções severamente degradadas, complementando as informações oriundas da voz com aquelas concomitantemente extraídas dos sinais cerebrais durante a fonação, isto é, advindas da *imagined speech*.

Claramente, conforme detalhado adiante, esta proposta insere-se na fronteira entre as sub-áreas de processamento de sinais biológicos, eletrônica,

sistemas inteligentes e neurociência. Os detalhes estão descritos ao longo deste documento, o qual está organizado da seguinte forma: os objetivos geral e específicos constam na Seção 2, a metodologia e o cronograma estão descritos na Seção 3, os resultados esperados foram incluídos na Seção 4 e, suplementarmente, a Seção 5 foi reservada para tratar da relevância e do impacto deste projeto.

## 2 Objetivos e Metas

O objetivo geral deste trabalho é o de projetar e implementar algoritmos biométricos capazes de autenticar, em princípio por meio da fala, indivíduos capazes de produzir somente locuções potencialmente degradadas, adicionando informações extraídas dos sinais cerebrais durante a fonação, isto é, a *imagined speech*, ao conjunto daquelas que são acústicas e oriundas da voz prejudicada. Em um nível mais específico, os propósitos são os seguintes:

1. realizar um amplo levantamento bibliográfico sobre o estado-da-arte na área de ABLs com disfonias laríngeas severas (DLSs), complementando-as com o estudo das estratégias do tipo BCI para decodificação de *imagined speech*;
2. estudar as bases de dados públicas de *imagined speech*, tal como a referida no artigo científico [30], visando possibilitar experimentos iniciais;
3. criar, por meio do fortalecimento de parcerias com profissionais da área da saúde, uma base de dados contendo sinais de voz oriundos de indivíduos com DLS e os respectivos sinais de *imagined speech*, a qual deverá ser disponibilizada publicamente. Pretende-se incluir locuções pré-definidas e outras diversificadas, de centenas de indivíduos;
4. extrair, com vistas à ABLs, características representativas dos sinais de voz e de *imagined speech* de cada locutor, experimentando e compa-

rando as modalidades *handcrafted extraction*, considerando diferentes conceitos físicos para a análise tempo-frequência das locuções, e *feature learning*, com base em *autoencoders*;

5. autenticar os locutores matriculados na base de dados, representados pelos seus respectivos vetores de características (VsCs) que contemplam informações das vozes degradadas e das *imagined speeches*, utilizando estruturas profundas de aprendizagem, tais como as *Residual Neural Networks* (RNNs) e as *Deep Spiking Neural Networks* (DSNNs), comparando-as em termos de custo computacional para treinamento e acurácia. À medida que os dados disponíveis permitirem, considerar-se-ão duas modalidades: autenticação *text-dependent* e *text-independent*;
6. implementar os algoritmos criados usando a linguagem de programação C/C++ para funcionamento *off-line* e, na medida do possível, em dispositivos que permitam a execução em tempo-real.
7. disseminar os resultados na literatura científica, por meio da apresentação de trabalhos em congressos e publicações em periódicos de alto impacto.

Assim, estarão cumpridos os objetivos do presente projeto, o qual conta com consolidada experiência do candidato e do seu orientador.

### 3 Metodologia e Cronograma

Com base no cronograma incluso na Tabela 1, ao final desta seção, proceder-se-á de acordo com a metodologia descrita adiante, a qual está organizada por meio das etapas de  $E_1$  até  $E_5$ .

A etapa inicial,  $E_1$ , será dedicada ao levantamento bibliográfico sobre o estado-da-arte nas sub-áreas de ABLs com DLSs e de aquisição e análise

de *imagined speech* por meio de eletroencefalografia (EEG) [31]. Particularmente, além do estudo detalhado dos trabalhos publicados nas referências de [2] até [25], complementados por outros que possam surgir, diálogos serão estabelecidos com profissionais das áreas de fonoaudiologia e neurologia vinculados à Faculdade de Medicina de São José do Rio Preto - SP (FAMERP), para intercâmbio de conceitos. Adicionalmente, estudar-se-á a base de dados pública descrita em [30], objetivando compreender os padrões corriqueiramente encontrados nos sinais em apreço.

$E_2$ , que se desenvolverá parcialmente em paralelo com  $E_1$ , será dedicada à criação de uma base de dados própria contendo sinais de voz verbalizados por locutores com DLSs assim como sinais de EEG que possivelmente contenham representações das respectivas *imagined speeches*. Para tal, será definido um conjunto foneticamente rico de frases que serão lidas em voz alta pelos locutores voluntários, todos portadores de anomalias que degradem a voz, os quais, concomitantemente ou em momento imediatamente anterior ou posterior, imaginarão as mesmas locuções. Por um lado, os sinais acústicos serão capturados com um mesmo modelo de gravador eletrônico, utilizando uma taxa de 22050 amostras por segundo e quantização de 16 bits, consideradas condizentes com os objetivos [32]. Por outro lado, os sinais de EEG serão capturados com eletrodos específicos, devidamente posicionados sobre o couro cabeludo de cada voluntário, de modo não invasivo e seguindo o padrão internacional 10-20 aprimorado [33]. Possivelmente, serão utilizados 14 canais de entrada [16], além de um circuito específico de aquisição e digitalização, a ser projetado e implementado com base em condicionadores de sinais, filtros para as bandas *alfa*, *beta*, *delta* e *teta* de EEG [31], amplificadores de instrumentação [34] e conversores analógico-digitais [35]. Além das atividades de pesquisa da fase  $E_2$ , o referido período também será reservado para o cumprimento dos créditos necessários em disciplinas para o Doutorado.

Assim que a base de dados a ser criada esteja com um número de sinais considerado suficiente para a análise, dar-se-á início aos trabalhos envolvendo a etapa  $E_3$ , centrada no pré-processamento e na extração das características (EC) representativas para o objetivo proposto. Após as normalizações clássicas, tais como padronização das amplitudes, remoção das médias e atenuação das irradiações labiais [36], duas serão as modalidades para EC: *handcrafted extraction* e *feature learning*. Por um lado, aquela abordagem, por meio da qual o projetista do sistema escolhe quais e quantos devem ser os descritores, será guiada pela engenharia paraconsistente de características (EPCs) [37][38]. Por outro lado, nesta técnica, as características serão aprendidas com base em *autoencoders*, experimentando qual a topologia neural e qual a dimensão dos VsCs são as mais adequadas. Em ambos os casos, serão contempladas tanto as características oriundas dos sinais de voz quanto aquelas originadas das *imagined speeches*.

Considerando que os VsCs oriundos de um mesmo indivíduo constituem uma classe, independentemente da locução e da respectiva *imagined speech*, então, por um lado, a sistemática *handcrafted extraction* dar-se-á pela busca de VsCs similares para uma mesma classe e distintos para classes diferentes, sendo estas caracterizadas por um *universal background model* [2]. Experimentando combinações de características tais como energia [39], taxa de cruzamentos por zero [40], entropia [41], operador não-linear de Teager [42], conteúdo de sub-bandas *wavelet-packet* (DTWPT) [43, 44], *jitter*, *shimmer*, medidas diversas de perturbação, entre outras [36], almeja-se obter VsCs que impliquem a ausência de indefinições e ambiguidades [37]. Por outro lado, o procedimento de *feature learning* implicará a investigação de arquiteturas de redes neurais artificiais profundas [9], projetadas para autocodificação, nas quais uma camada oculta central possui dimensão fixa e relativamente redu-

zida em comparação com a dimensão dos sinais sob análise. Obtendo-se uma estrutura que alcance convergência, os códigos gerados mediante a aplicação de um par de sinais de voz e *imagined speech* de entrada serão usados como VsCs.

Concomitantemente ao desenvolvimento da etapa  $E_3$ , os procedimentos referentes à  $E_4$  serão realizados visando integrar as etapas de EC e de autenticação. Pretende-se empregar as estruturas inteligentes de aprendizado profundo dos tipos RNNs e DSNs, as quais têm se mostrado propícias para o reconhecimento de padrões [9]. Os resultados serão mensurados e comparados em termos de acurácia, levando-se também em conta questões relativas ao tempo de treinamento. Adicionalmente, a relação entre a convergência das estruturas neurais, o tempo para aprendizado e o número de voluntários matriculados no sistema serão considerados.

Complementarmente, será realizada uma comparação em termos de acurácia para as modalidades *text-dependent*, em que uma mesma frase pronunciada e imaginada é considerada para treinamento e testes, e *text-independent*, a qual implica o uso de locuções e *imagined speeches* distintas no treinamento e nos testes. De uma maneira geral, cada resultado obtido será avaliado sistematicamente por meio de testes consecutivos e refinamentos na técnica proposta, à medida que se perceber tal necessidade. Na Figura 1, é possível observar a estrutura geral das etapas  $E_3$  e  $E_4$ .

Finalmente,  $E_5$ , que constitui a apresentação dos experimentos em eventos científicos nacionais e internacionais, assim como a publicação dos resultados em periódicos de impacto e também a confecção da monografia de qualificação e posteriormente da tese de doutorado, se concretizará ao longo de todo o trabalho, paralelamente com as demais etapas. Diante do exposto,



conclui-se que o projeto possui objetivos claros e bem definidos, possibilitando contribuições originais.

Tabela 1: *Cronograma proposto para o projeto, com duração de quatro anos.*

$\frac{Mês}{Ano}$	$E_1$	$E_2$	$E_3$	$E_4$	$E_5$	$\frac{Mês}{Ano}$	$E_1$	$E_2$	$E_3$	$E_4$	$E_5$	$\frac{Mês}{Ano}$	$E_1$	$E_2$	$E_3$	$E_4$	$E_5$
09/22	X	X				01/24			X	X	X	05/25				X	X
10/22	X	X				02/24			X	X	X	06/25				X	X
11/22	X	X				03/24			X	X	X	07/25				X	X
12/22	X	X				04/24			X	X	X	08/25				X	X
01/23		X	X		X	05/24			X	X	X	09/25				X	X
02/23		X	X		X	06/24			X	X	X	10/25				X	X
03/23		X	X		X	07/24				X	X	11/25				X	X
04/23		X	X		X	08/24				X	X	12/25				X	X
05/23		X	X		X	09/24				X	X	01/26				X	X
06/23		X	X		X	10/24				X	X	02/26				X	X
07/23			X		X	11/24				X	X	03/26				X	X
08/23			X		X	12/24				X	X	04/26				X	X
09/23			X		X	01/25				X	X	05/26					X
10/23			X		X	02/25				X	X	06/26					X
11/23			X		X	03/25				X	X	07/26					X
12/23			X		X	04/25				X	X	08/26					X

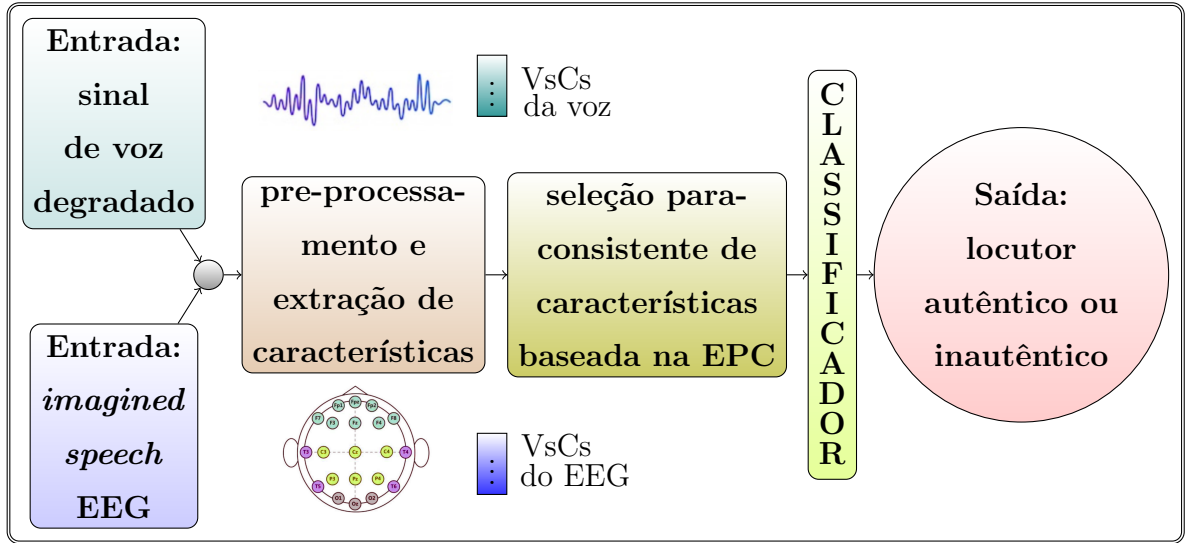


Figura 1: Visão geral da estrutura das etapas  $E_3$  e  $E_4$ . A extração de características dar-se-á nos modos *handcrafted extraction* e *feature learning*, para fins de comparação. RNNs e DSNNs serão testadas como opções para o classificador.

## 4 Resultados Esperados

Os principais resultados esperados frente ao andamento e à plena execução deste projeto são os seguintes:

1. alcançar implementações *off-line* e, possivelmente também de tempo-real, que permitam utilizar a estratégia proposta em situações cotidianas de autenticação biométrica, associando voz e *imagined speech*;
2. tornar pública a base de dados de sinais a ser criada, combinando vozes e *imagined speeches*, especialmente construída para fins de ABLs severamente disfônicos. Isso permitirá que outros pesquisadores utilizem os mesmos dados e reportem os resultados dos seus experimentos, possibilitando comparações;
3. investigar e obter conclusões relativas ao potencial das representações *handcrafted extraction* e *feature learning* para formar os VsCs que melhor representem os indivíduos, visando autenticá-los com acurácia;
4. analisar e obter conclusões relativas à adequabilidade para reconhecimento de padrões das estruturas RNNs e DSNNs associadas aos processos de *handcrafted feture extraction* e *feature learning*;
5. divulgar os resultados obtidos em palestras e apresentá-los em congressos científicos nacionais e internacionais, tais como a *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), o *Interspeech* e o *IEEE International Workshop on Machine Learning for Signal Processing* (MLSP);
6. publicar os resultados das pesquisas em periódicos de alto fator de impacto, tais como o *IEEE Signal Processing Magazine*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *Neurocomputing*, *Computers in Biology and Medicine*, entre outros;

7. consolidar parcerias nacionais com profissionais da área da saúde, principalmente vinculados á FAMERP, que proporcionarão suporte para realizar os testes e implementar melhorias no sistema proposto, contribuindo significativamente com aprimoramentos;

## 5 Relevância e Impacto do Projeto

Por um lado, uma ampla busca na literatura especializada demonstra claramente que os sistemas para ABLs são utilizados com frequência e têm atraído a atenção das comunidades acadêmica e industrial ao redor do mundo. Por outro lado, autenticar locutores severamente disfônicos é tarefa difícil e que demanda considerável nível de pesquisa, o que constitui um desafio que ainda não foi vencido ou, pelo menos, não de maneira adequada.

Nesse sentido, percebe-se que inexistente registro na literatura sobre o emprego de *imagined speech* em associação com sinais de voz para autenticar biometricamente locutores severamente disfônicos. Complementarmente, estima-se que o sistema proposto, pela associação de dois tipos de sinais no mecanismo autenticador, isto é, sinais de voz e neuronais, seja menos susceptível aos ataques de *voice spoofing* [45], tais como os *replay attacks* [46]. Registre-se, assim, a principal contribuição e tendência inovadora deste projeto, enriquecendo a área de pesquisa na qual ele está inserido e impactando positivamente os trabalhos correspondentes.

Adicionalmente, estudos, desenvolvimentos e fortalecimento de parcerias na área deste projeto, tais como as mencionadas ao longo do texto, certamente proporcionarão inegável expansão no sentido de prover maior aplicabilidade das tecnologias em apreço. Assim, ressalta-se a contribuição social proporcionada ao possibilitar maior inserção de indivíduos disfônicos nas ati-

vidades cotidianas. Portanto, é nítida a relevância deste projeto, que possui um promissor potencial de impacto nos meios científico e tecnológico, com alcance geográfico em níveis nacional e internacional assegurado pelo seu caráter inovador e pela contribuição que proverá.

Finalmente, registre-se que este projeto conta com o suporte proporcionado pelo auxílio à pesquisa FAPESP 2021/12407-4, centrado na mesma área, possuindo o orientador como beneficiário direto. Além disso, todos os recursos, materiais, acesso à bases de dados, entre outros detalhes necessários, estão disponíveis. Assim, garante-se a exequibilidade.

## Referências

- [1] Furlan, A. **Caracterização de *voice spoofing* para fins de verificação de locutores com base na Transformada *Wavelet* e na Análise Paraconsistente de Características**. 2021. Dissertação (Mestrado em Ciência da Computação) - Universidade Estadual Paulista - campus de São José do Rio Preto-SP. Orientador: Prof Dr Rodrigo Capobianco Guido.
- [2] Beigi, H. **Fundamentals of Speaker Recognition**. Springer, 2011.
- [3] Neustein, A.; Patil, H.A. **Forensic Speaker Recognition: Law Enforcement and Counter-Terrorism**. Springer, 2012.
- [4] Hansen, J. H. L.; Hasan, T. Speaker Recognition by Machines and Humans. *IEEE Signal Processing Magazine*, v. 32, n. 6, pp. 74-99, (2015).
- [5] Wang, X. M. et. al. RACP: A Network with Attention Corrected Prototype for Few-shot Speaker Recognition Using Indefinite Distance Metric. *Neurocomputing*, v. 490, pp. 283-294, (2022).

- [6] Lee, K. A. et. al. Two Decades into Speaker Recognition Evaluation - are we there yet? *Computer Speech & Language*, v. 61, article 101058, (2020).
- [7] Chaiani, M. Voice disorder classification using speech enhancement and deep learning models. *Biocybernetics and Biomedical Engineering*, v. 42, n. 2, pp. 463-480, (2022).
- [8] Fujimura, S. Classification of Voice Disorders Using a One-Dimensional Convolutional Neural Network. *Journal of Voice*, v. 36, n. 1, pp. 15-20, (2022).
- [9] Goodfellow, I.; Bengio, Y.; Courville, A. **Deep Learning**. The MIT Press, 2016.
- [10] Mittal, V.; Sharma, R. K. Deep Learning Approach for Voice Pathology Detection and Classification. *International Journal of Healthcare Information Systems and Informatics*, v. 16, n. 4, article oa28, (2021).
- [11] Miliarezi, I.; Poutos, K.; Pikrakis, A. Combining Acoustic Features and Medical Data in Deep Learning Networks for Voice Pathology Classification. *In: 28th European Signal Processing Conference (EUSIPCO)*, pp. 1190-1194, (2021).
- [12] Fonseca, E., et. al. Linear Prediction and Discrete Wavelet Transform to Identify Pathologies in Voice Signals. *In: Proc. IEEE Signal Processing Symposium (IEEE SPSympo 2017)*, Jachranka Village, Poland, v.1, pp. 1-4, (2017).
- [13] Le Huche, F.; Allali, A. **A Voz - patologia vocal de origem funcional**. v.2, ArtMed, 2005.
- [14] Le Huche, F.; Allali, A. **A Voz - patologia vocal de origem orgânica**. v.3, ArtMed, 2005.

- [15] Gupta, S. et. al. Residual Neural Network precisely quantifies dysarthria severity-level based on short-duration speech segments. *Neural Networks*, n. 139, pp. 105–117, (2021).
- [16] Moctezuma, L. A. et. al. Subjects Identification Using EEG-recorded Imagined Speech. *Expert Systems with Applications*, v. 118, pp. 201-208, (2019).
- [17] Rusnac, A. L.; Grigore, O. EEG Preprocessing Methods for BCI Imagined Speech Signals. In: *9th International Conference on E-health and Bioengineering (EHB)*, Romania, (2021).
- [18] Brigham, K. ; Kumar, B. V. K. V. Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy. In: *Proceeding of the 4th IEEE International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, Chengdu, China, (2010).
- [19] Pawar, D.; Dhage, S. Wavelet-based imagined speech classification using electroencephalography. *International Journal of Biomedical Engineering and Technology*, v. 38, n. 3, pp. 215-224, (2022).
- [20] Cooney, C.; Folli, R.; Coyle, D. A Bimodal Deep Learning Architecture for EEG-fNIRS Decoding of Overt and Imagined Speech. *IEEE Transactions on Biomedical Engineering*, v. 69, n. 6, pp. 1983-1994, (2022).
- [21] Lopez-Bernal, D. et. al. A State-of-the-Art Review of EEG-Based Imagined Speech Decoding. *Frontiers in Human Neuroscience*, v. 16, article 867281, (2022).
- [22] Bakhshali, M. A.; Khademi, M.; Ebrahimi-Moghadam, A. Investigating the neural correlates of imagined speech: an EEG-based connectivity analysis. *Digital Signal Processing*, v. 123, article 103435, (2022).

- [23] Lee, D. Y.; Lee, M.; Lee, S. W. Decoding Imagined Speech Based on Deep Metric Learning for Intuitive BCI Communication. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, v. 29, pp. 1363-1374, (2021).
- [24] Tamm, M. O.; Muhammad, Y.; Muhammad, N. Classification of Vowels from Imagined Speech with Convolutional Neural Networks. *Computers*, v. 9, n. 2, article 46, (2020).
- [25] Moctezuma, L. A.; Molinas, M. EEG-Based Subjects Identification Based on Biometrics of Imagined Speech Using EMD. **Brain Informatics - Lecture Notes in Artificial Intelligence**, v. 11309, pp. 458-467, (2018).
- [26] Jayarathne, I.; Cohen, M.; Amarakeerthi, S. Brainid: Development of an Eeg-based Biometric Authentication System. In *7th IEEE Information Technology, Electronics and Mobile Communication Conference (IEM-CON)*, pp. 1-6, (2016).
- [27] Jayarathne, I.; Cohen, M.; Amarakeerthi, S. Survey of Eeg-based Biometric Authentication. In: *8th IEEE International Conference on Awareness Science and Technology (ICAST)*, pp. 324-329, (2017).
- [28] Marcos, D. P.-B. et. al. Electroencephalogram Subject Identification: a review. *Expert Systems with Applications*, n. 41, pp. 6537-6554, (2014).
- [29] Ruiz-Blondet, M. V.; Jin, Z.; Laszlo, S. Cerebre: a novel method for very high accuracy event-related potential biometric identification. *IEEE Transactions on Information Forensics and Security*, n. 11, pp. 1618-1629, (2016).
- [30] Coretto, G. A. P. et. al. Open Access database of EEG signals recorded during imagined speech. In: *Proceedings of the 12th International*

- Symposium on Medical Information Processing and Analysis*, v. 10160, article 1016002, Tandil, Argentina, (2017).
- [31] Bansal, D.; Mahajan, R. **EEG-Based Brain-Computer Interfaces: Cognitive Analysis and Control Applications**. Academic Press, 2019.
  - [32] Bosi, M.; Goldberg, R.E. **Introduction to Digital Audio Coding and Standards**. Springer, 2003.
  - [33] Wagner, A. et. al. Improved potential quality of intraoperative transcranial motor-evoked potentials by navigated electrode placement compared to the conventional ten-twenty system. *Neurosurgical Review*, v 45, n. 1, pp. 585-593, (2022).
  - [34] Fiore, J. M. **Operational Amplifiers & Linear Integrated Circuits: Theory and Application**, Independently Published, 2016.
  - [35] Dobkin, B.; Williams, J. **Analog Circuit Design: A Tutorial Guide to Applications and Solutions**, Newnes Books, 2011.
  - [36] Deng, L.; O'Shaughnessy, O. *Speech Processing: a dynamic and optimization-oriented approach*. New York: Marcel Dekker, 2003.
  - [37] Carvalho, F.R.; Abe, J.M. **A Paraconsistent Decision-Making Method**. Springer, 2018.
  - [38] Guido, R.C. Paraconsistent Feature Engineering. *IEEE Signal Processing Magazine* (to appear in 2018/2019).
  - [39] Guido, R.C. A Tutorial on Signal Energy and its Applications. *Neurocomputing*, v. 179, pp.264-282, (2016).
  - [40] Guido, R.C. ZCR-aided Neurocomputing: a study with applications. *Knowledge-based Systems*, v. 105, pp. 248-269, (2016).



- [41] Guido, R.C. A Tutorial-review on Entropy-based Handcrafted Feature Extraction for Information Fusion. *Information Fusion*, n.41, pp.161-175, (2018).
- [42] Kaiser, J.F. On a simple algorithm to calculate the energy of a signal. *In: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (IEEE ICASSP)*, pp.381-384, Albuquerque - NM - USA, (1990).
- [43] Guido, R.C. Effectively Interpreting Discrete Wavelet Transformed Signals. *IEEE Signal Processing Magazine*, v. 34, n.3, pp. 89-100, (2017).
- [44] Guido, R.C. Practical and Useful Tips on Discrete Wavelet Transforms. *IEEE Signal Processing Magazine*, v.32, n. 3, pp. 162-166, (2015).
- [45] Kamble, M. R. et. al. Advances in Anti-spoofing: from the perspective of ASVspoof challenges. *APSIPA Transactions on Signal and Information Processing*, n. 9 - E2, pp. 1-18, (2020).
- [46] Patil, A. T. et. al. Improving the Potential of Enhanced Teager Energy Cepstral Coefficients (ETECC) for replay attack detection. *Computer Speech & Language*, n. 72, article 101281, (2022).