



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE CIÊNCIAS DA SAÚDE DE PORTO ALEGRE
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS DA INFORMAÇÃO E
GESTÃO EM SAÚDE**

Dieine Estela Bernieri Schiavon de Andrade

**Desenvolvimento de Uma Rede Neural Convolucional para
Classificação de Sinais Acústicos da Deglutição de Indivíduos com Disfagia
Orofaríngea**

Porto Alegre

2024

Dieine Estela Bernieri Schiavon de Andrade

**Desenvolvimento de Uma Rede Neural Convolucional para
Classificação de Sinais Acústicos da Deglutição de Indivíduos com Disfagia
Orofaríngea**

Dissertação no Programa de Mestrado Acadêmico
em Tecnologias da Informação e Gestão em Saúde
da Universidade Federal de Ciências da Saúde de
Porto Alegre.

Orientadora: Profa. Dra. Carla Diniz Lopes Becker

Coorientadora: Profa. Dra. Rafaela Soares Rech

Porto Alegre

2024

Catálogo na Publicação

Andrade, Dieine Estela Bernieri Schiavon de
Desenvolvimento de Uma Rede Neural Convolucional para
Classificação de Sinais Acústicos da Deglutição de
Indivíduos com Disfagia Orofaríngea / Dieine Estela
Bernieri Schiavon de Andrade. -- 2024.
109 p. : il., graf., tab. ; 30 cm.

Dissertação (mestrado) -- Universidade Federal de
Ciências da Saúde de Porto Alegre, Programa de
Pós-Graduação em Tecnologias da Informação e Gestão em
Saúde, 2024.

Orientador(a): Profa. Dra. Carla Diniz Lopes Becker ;
coorientador(a): Profa. Dra. Rafaela Soares Rech.

1. Rede Neural Convolucional. 2. Ausculta Cervical. 3.
Sinais Acústicos. 4. Disfagia. 5. Aprendizado Profundo.
I. Título.

Dieine Estela Bernieri Schiavon de Andrade

**Desenvolvimento de Uma Rede Neural Convolucional para
Classificação de Sinais Acústicos da Deglutição de Indivíduos com Disfagia
Orofaringea**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em
Tecnologias da Informação e Gestão em Saúde da Universidade Federal de
Ciências da Saúde de Porto Alegre como requisito parcial para a obtenção do título
de Mestre em Tecnologias da Informação e Gestão em Saúde.

Orientadora: Profa. Dra. Carla Diniz Lopes Becker

Coorientadora: Profa. Dra. Rafaela Soares Rech

Aprovada em: 18 de outubro de 2024.

BANCA EXAMINADORA:

Prof. Dr. Fernando Neves Hugo
New York University

Prof. Dr. Airton Tetelbom Stein
Universidade Federal de Ciências da Saúde de Porto Alegre

Prof. Dr. Mirko Salomón Alva Sánchez
Universidade Federal de Ciências da Saúde de Porto Alegre

AGRADECIMENTOS

Agradeço ao meu marido, Paulo, e aos meus filhos, Mariana e Eduardo, pelo amor, apoio incondicional, compreensão e incentivo. Por vibrarem a cada conquista e pelas palavras de conforto em momentos mais difíceis.

Às minhas orientadoras, Carla Diniz Lopes Becker e Rafaela Soares Rech por compartilharem suas experiências e conhecimentos, os quais me fizeram crescer como pesquisadora e me guiaram na construção deste trabalho.

A todos os membros do grupo de pesquisa DigDis - Sistema de Diagnóstico da Disfagia apoiado por Inteligência Artificial. A colaboração e o compromisso de cada um foram fundamentais para a realização deste trabalho. Um agradecimento especial à coordenadora, Professora Rafaela, pela disponibilização do acesso ao banco de dados, pela orientação incansável e pelos valiosos conselhos ao longo de todo o processo. Agradeço também a todos os bolsistas e voluntários, que foram parceiros essenciais nesta jornada, sempre prontos para compartilhar ideias e enfrentar desafios juntos.

A todos os participantes da pesquisa, que generosamente se dispuseram a contribuir para o avanço do conhecimento nesta área. A colaboração de cada um foi fundamental para o nosso progresso.

Aos meus colegas e professores pela oportunidade de convívio e aprendizado ao longo do curso. Cada momento compartilhado foi fundamental para minha formação e crescimento.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, p bolsa de estudos que permitiu dedicação exclusiva a este trabalho.

Para concluir, gostaria de expressar minha sincera gratidão aos meus pais, Nelson e Edy, às minhas irmãs, Nina e Nara, e ao meu irmão, João, que sempre torceram por mim e acreditaram nos meus sonhos. Muito obrigada!

RESUMO

Introdução: A deglutição é um processo contínuo que se inicia na cavidade oral e termina quando o alimento chega ao estômago. Qualquer alteração no trânsito do bolo alimentar, que interfira na segurança e na eficiência durante o trajeto, é considerado de disfagia. A disfagia afeta significativamente a vida de quem sofre dessa condição, causando consequências severas à saúde do indivíduo. O exame padrão ouro para diagnóstico de disfagia é a videofluoroscopia, que é um exame de raio-x dinâmico, que está relacionado ao desconforto e riscos ao indivíduo. A avaliação clínica ainda é permeada de subjetividade e o uso do aprendizado de máquina mostra-se promissor para a qualificação da mesma.

Objetivo: Analisar a acurácia diagnóstica de um modelo baseado em Rede Neural Convolucional que permitirá classificar deglutições de indivíduos com disfagia orofaríngea e indivíduos com deglutição normal ou funcional. **Método:** Estudo transversal de acurácia diagnóstica, realizado a partir de um banco de dados com sinais acústicos de deglutição de indivíduos adultos e idosos, captados por meio da ausculta cervical digital e classificados por um fonoaudiólogo capacitado como deglutição normal e com disfagia orofaríngea. O diagnóstico de disfagia orofaríngea foi confirmado pela videofluoroscopia da deglutição. As gravações de áudio completas, com todos os eventos de deglutição amostrados em uma frequência de 4 kHz, foram transformadas em espectrogramas e escalogramas e, em seguida, em imagens. Uma arquitetura de rede neural convolucional binária foi desenvolvida para classificar deglutições normais e alteradas. Os modelos foram treinados usando um holdout de 3 vias e validação cruzada de 10 vezes. Os espectrogramas e escalogramas foram usados separadamente como entrada para a rede para extração de características. Assim, as frequências presentes nas gravações de áudio foram representadas visualmente ao longo do tempo. Quatro experimentos foram realizados com gravações de áudio de deglutição de saliva, líquido, pastoso e sólido, alternando o tipo de entrada para a rede. Os três primeiros experimentos verificaram o desempenho dos modelos na classificação da deglutição normal ou alterada. O último experimento usou exclusivamente sinais de deglutição com consistência pastosa para verificar o desempenho dos modelos na classificação do

tipo de deglutição, se normal ou anormal e a forma, se livre ou controlada.

Resultados: O modelo com espectrograma obteve melhores resultados no primeiro experimento, com acurácia de 90,54% e recall de 91,89%. No segundo experimento, o modelo com escalograma obteve melhor desempenho, com uma acurácia de 89,44%, recall de 95,92%. Os modelos avaliados no terceiro experimento tiveram a mesma acurácia de 73,68%. No quarto experimento, três modelos obtiveram acurácia e recall de 100%; apenas o modelo com avaliação intragrupo (disfagia) teve desempenho inferior, com acurácia de 70% e recall de 75%. **Conclusão:** Os modelos treinados com espectrogramas e escalogramas tiveram resultados semelhantes em todos os experimentos, sendo promissores para a classificação automática de sinais de deglutição normal e alterada. Os modelos propostos apresentam potencial para uso na prática clínica para uma identificação precoce e acurada da disfagia orofaríngea.

Palavras-chave: Ausculta Cervical. Disfagia. Rede Neural Convolucional.

ABSTRACT

Introduction: Swallowing is a continuous process that begins in the oral cavity and ends when food reaches the stomach. Any change in the transit of the food bolus that interferes with its safety and efficiency during its journey is considered dysphagia. Dysphagia significantly affects the lives of those who suffer from this condition, causing severe consequences to the individual's health. The gold standard exam for diagnosing dysphagia is videofluoroscopy, which is a dynamic x-ray exam that is related to discomfort and risks to the individual. Clinical evaluation is still permeated by subjectivity, and the use of machine learning shows promise for its qualification. **Objective:** To analyze the diagnostic accuracy of a model based on the Convolutional Neural Network that will allow the classification of swallowing of individuals with oropharyngeal dysphagia and individuals with normal or functional swallowing. **Method:** Cross-sectional study of diagnostic accuracy, carried out from a database with acoustic swallowing signals from adults and elderly individuals, captured through digital cervical auscultation and classified by a trained speech-language pathologist as normal swallowing and with oropharyngeal dysphagia. The diagnosis of oropharyngeal dysphagia was confirmed by videofluoroscopy of swallowing. The complete audio recordings, with all swallowing events sampled at a frequency of 4 kHz, were transformed into spectrograms and scalograms and then into images. A binary convolutional neural network architecture was developed to classify tasks for a supervised learner. The models were trained using a 3-way holdout and 10-fold cross-validation. The spectrograms and scalograms were used separately as input to the network for feature extraction. Thus, the frequencies present in the audio recordings were visually represented over time. Four experiments were performed with audio recordings of swallowing saliva, liquid, pasty, and solid, alternating the input type to the network. The first three experiments verified the performance of the models in classifying normal or abnormal swallowing. The last experiment exclusively used swallowing signals with pasty consistency to verify the performance of the models in classifying the type of swallowing, whether normal or abnormal and the form, whether free or controlled. **Results:** The model with the spectrogram obtained better results in the first experiment, with an accuracy

of 90.54% and a recall of 91.89%. In the second experiment, the better model was with a scalogram that obtained an accuracy of 89.44% and a recall of 95.92%. The models evaluated in the third experiment had the same accuracy of 73.68%. In the fourth experiment, three models obtained accuracy and recall of 100%; only the model with intragroup evaluation (dysphagia) had lower performance, with an accuracy of 70% and recall of 75%. **Conclusion:** The models trained with spectrograms and scalograms had similar results in all experiments, promising for automatically classifying normal and altered swallowing signals. The proposed models can be used in clinical practice for early and accurate identification of oropharyngeal dysphagia.

Keywords: Cervical Auscultation. Dysphagia. Convolutional Neural Network.

SUMÁRIO

1 INTRODUÇÃO.....	13
1.1 TEMA E OBJETIVOS.....	16
1.1.1 Objetivo Geral.....	16
1.1.2 Objetivos Específicos.....	16
1.2 JUSTIFICATIVA DO TRABALHO.....	16
2 REFERENCIAL TEÓRICO.....	18
2.1 DEGLUTIÇÃO NORMAL.....	18
2.2 DEGLUTIÇÃO FUNCIONAL.....	19
2.3 DISFAGIA.....	20
2.4 SONS DA DEGLUTIÇÃO.....	21
2.5 AUSCULTA CERVICAL E AUSCULTA CERVICAL DE ALTA RESOLUÇÃO.....	22
2.6 ESPECTROGRAMA.....	23
2.7 ESCALOGRAMA.....	26
2.7.1 Transformada Wavelet.....	26
2.7.1.1 Transformada Wavelet Contínua.....	28
2.8 APRENDIZADO DE MÁQUINA.....	29
2.9 REDES NEURAIS ARTIFICIAIS.....	30
2.9.1 Redes Neurais Convolucionais.....	31
2.10 FUNÇÕES DE ATIVAÇÃO.....	34
2.10.1. ReLU.....	35
2.10.2. Sigmóide.....	36
2.10.3. Softmax.....	37
2.11 VALIDAÇÃO CRUZADA.....	37
2.12 MÉTRICAS DE DESEMPENHO.....	38
3 REVISÃO DE LITERATURA.....	41
3.1 SELEÇÃO E CARACTERÍSTICAS DOS ESTUDOS SELECIONADOS.....	44
4 MATERIAIS E MÉTODOS.....	51
4.1 AMBIENTE DE DESENVOLVIMENTO.....	51
4.2 BANCO DE DADOS.....	51
4.3 PRÉ-PROCESSAMENTO.....	53
4.3.1 Extração de Features dos Espectrogramas.....	54
4.3.2 Extração de Features dos Escalogramas.....	55
4.4 ARQUITETURA E CONFIGURAÇÃO DOS MODELOS.....	55
5 ARTIGO.....	57
6 RESULTADOS E DISCUSSÃO.....	80
6.1 ANÁLISE COMPARATIVA DOS MODELOS - EXPERIMENTOS 1 E 2.....	80

6.2 ANÁLISE COMPARATIVA DOS MODELOS – EXPERIMENTO 3.....	86
6.3 ANÁLISE COMPARATIVA DOS MODELOS – EXPERIMENTO 4.....	90
7 CONCLUSÃO.....	94
7.1 LIMITAÇÕES.....	95
7.2 OPORTUNIDADES PARA PESQUISAS FUTURAS.....	96
REFERÊNCIAS.....	97
ANEXO A.....	109

REFERÊNCIAS

ABHANG, P. A.; GAWALI, B. W.; MEHROTRA, S. C. **Introduction to EEG-and speech-based emotion recognition**. Academic Press, 2016.

ABOOFAZELI, M; MOUSSAVI, Z. Automated classification of swallowing and breadth sounds. In: **The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. IEEE, 2004. p. 3816-3819.

ADDISON, P. S. **The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance**. CRC press, 2017.

ADKINS, C. et al. Prevalence and Characteristics of Dysphagia Based on a Population-Based Survey. **Clinical Gastroenterology and Hepatology**, v. 18, n. 9, p. 1970-1979.e2, ago. 2020.

ALZUBAIDI, L. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. **Journal of Big Data**, v. 8, n. 1, 31 mar. 2021.

AMINI, S. et al. Automated detection of mild cognitive impairment and dementia from voice recordings: A natural language processing approach. **Alzheimer's & Dementia**, 7 jul. 2022.

ANÉAS, G. C. G.; DANTAS, R. O. A videofluoroscopia da deglutição na investigação da disfagia oral e faringea. **GE Jornal Português de Gastreenterologia**, v. 21, n. 1, p. 21–25, jan. 2014.

ARAÚJO, B. C. L. et al. Avaliação clínica e videofluoroscopia no diagnóstico de disfagia na encefalopatia crônica da infância. **Radiologia Brasileira**, v. 47, n. 2, p. 84–88, 2014.

ARVEDSON, J. et al. **American Speech-Language-Hearing Association Roles of Speech-Language Pathologists in Swallowing and Roles of Speech-Language Pathologists in Swallowing and Feeding Disorders:**

Technical Report Dysphagia Document Review and Revision Working Group.

2015. Disponível em:

<https://ugc.futurelearn.com/uploads/files/94/9e/949e14c3-5410-4d16-81a4-255b924646ce/Roles_of_Speech

[Language_Pathologists_in_Swallowing_and_Feeding_Disorders__Technical_Report.pdf](#)>. Acesso em: 18 fev. 2024.

ASLAM, M.; VAEZI, M. F. Dysphagia in the elderly. **Gastroenterology & hepatology**, v. 9, n. 12, p. 784–95, 2013.

ASHA. Adult Dysphagia: Overview. **Asha.org**, 2019. Disponível em:<<https://www.asha.org/practice-portal/clinical-topics/adult-dysphagia/>> Acesso em: 10 fev. 2024.

ASIF, S. et al. An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning. **Multimedia Tools and Applications**, 27 fev. 2023.

ASTRID ST. PIERRE et al. **Terms Used to Describe Pediatric Videofluoroscopic Feeding Studies: A Delphi Survey**. 1 jun. 2012

BADRINARAYANAN, V.; KENDALL, A.; CIPOLLA, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 39, n. 12, p. 2481–2495, 1 dez. 2017.

BAHETI, P. **The Essential Guide to Neural Network Architectures**. 2021. Disponível em:<<https://www.v7labs.com/blog/neural-network-architectures-guide>>. Acesso em: 11 dez. 2023.

BAHETI, P. **12 Types of Neural Networks Activation Functions: How to Choose?** 2022. Disponível em:<<https://www.v7labs.com/blog/neural-networks-activation-functions>>. Acesso em: 11 dez. 2023.

BALAGOPALAN, A. et al. Comparing Pre-trained and Feature-Based Models for Prediction of Alzheimer's Disease Based on Speech. **Frontiers in Aging Neuroscience**, v. 13, 27 abr. 2021.

BALAJI, S. **Binary Image classifier CNN using TensorFlow**. Disponível em:
<<https://medium.com/techiepedia/binary-image-classifier-cnn-using-tensorflow-a3f5d6746697>>.

BANDINI, A.; SMAOUI, S.; STEELE, C. M. Automated pharyngeal phase detection and bolus localization in videofluoroscopic swallowing study: Killing two birds with one stone? **Computer Methods and Programs in Biomedicine**, v. 225, p. 107058, out. 2022.

BAUMGARTEN, A. et al. Dental status, oral prosthesis and chewing ability in an adult and elderly population in southern Brazil. **Clinics**, v. 72, n. 11, p. 681–685, 7 nov. 2017.

BELLO, S. A. et al. Review: Deep Learning on 3D Point Clouds. **Remote Sensing**, v. 12, n. 11, p. 1729, 28 maio 2020.

BORR, C.; HIELSCHER-FASTABEND, M.; LÜCKING, A. Reliability and Validity of Cervical Auscultation. **Dysphagia**, v. 22, n. 3, p. 225–234, 25 abr. 2007.

BUEHRING, B. et al. Tongue Strength Is Associated with Jumping Mechanography Performance and Handgrip Strength but Not with Classic Functional Tests in Older Adults. **Journal of the American Geriatrics Society**, v. 61, n. 3, p. 418–422, 4 fev. 2013.

BUJANG M.A.; DNAN, T.H. Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. **Journal of clinical and diagnostic research: JCDR**. 2016;10(10), YE01–YE06.

BÜLOW, M. Psychogenic dysphagia. **Principles of Deglutition: A Multidisciplinary Text for Swallowing and its Disorders**, p. 771-776, 2013.

BURMAN, P. A Comparative Study of Ordinary Cross-Validation, v-Fold Cross-Validation and the Repeated Learning-Testing Methods. **Biometrika**, v. 76, n. 3, p. 503, set. 1989.

CANTONI, M.M.; OLIVEIRA, B. G. DE; NEVADO, H. M. Introdução à análise acústica da fala com o Praat. **Texto Livre Linguagem e Tecnologia**, v. 15, 22 jul. 2022.

CARDOSO, M. C. A. F.; FONTOURA, E. G. Valor da ausculta cervical em pacientes acometidos por disfagia neurogênica. **Arq Int Otorrinolaringol**, v. 13, n. 4, p. 431-9, 2009.

CHEN, X.; KAMAVUAKO, E. Enhancing Intake Monitoring: Transfer Learning for Audio-Based Detection of Swallowing Events. In: **2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON)**. IEEE, 2024. p. 479-484.

CHILUKURI, P.; ODUFALU, F.; HACHEM, C. Dysphagia. **Missouri Medicine**, v. 115, n. 3, p. 206–210, 2018.

CHLASTA, K.; WOŁK, K. Towards Computer-Based Automated Screening of Dementia Through Spontaneous Speech. **Frontiers in Psychology**, v. 11, 12 fev. 2021.

CHOLLET, F. **Deep Learning with Python**. Manning Publications Co. USA. 2018. ISBN:9781617294433.

CHOLLET, F. **Deep Learning with Python**. Second Edition. Manning Publications Co. USA. 2021. ISBN: 9781617296864.

CIRESAN, D. C. et al. Flexible, high performance convolutional neural networks for image classification. In: **Twenty-second international joint conference on artificial intelligence**. 2011.

CLAVÉ, P. et al. Approaching oropharyngeal dysphagia. **Revista Espanola de Enfermedades Digestivas**, v. 96, n. 2, p. 119-131, 2004.

CLAVÉ, P.; VERDAGUER, A.; ARREOLA, V. Disfagia orofaríngea en el anciano. **Medicina Clínica**, v. 124, n. 19, p. 742–748, maio 2005.

CRARY, M. A.; MANN, G. D. C.; GROHER, M. E. Initial Psychometric Assessment of a Functional Oral Intake Scale for Dysphagia in Stroke Patients. **Archives of Physical Medicine and Rehabilitation**, v. 86, n. 8, p. 1516–1520, ago. 2005.

DE LIMA, P. C., Wavelets: uma introdução. **Matemática Universitária**, v. 33, p. 13-44, 2002.

DOBIE, R. A., VAN HEMEL, S., & National Research Council. **Basics of sound, the ear, and hearing. In Hearing Loss: Determining Eligibility for Social Security Benefits**. National Academies Press (US). 2004. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK207834/>>. Acesso em 10 set. 2023.

DSA, E. **Capítulo 8 - Função de Ativação**. 2022. Disponível em: <<https://www.deeplearningbook.com.br/funcao-de-ativacao/>>. Acesso em: 11 jan. 2024.

EDWARDS, W. History and Symptoms of Esophageal Disease. **Springer eBooks**, p. 103–118, 1 jan. 1974.

EL NAQA, I., MURPHY, M.J. (2015). **What Is Machine Learning?**. In: El Naqa, I., Li, R., Murphy, M. (eds) Machine Learning in Radiation Oncology. Springer, Cham. https://doi.org/10.1007/978-3-319-18305-3_1.

ERTEKIN, C.; AYDOGDU, I. Neurophysiology of swallowing. **Clinical Neurophysiology**, v. 114, n. 12, p. 2226–2244, dez. 2003.

FERRUCCI, J. L. et al. Sons da deglutição na prática fonoaudiológica: análise crítica da literatura. **Einstein (São Paulo)**, v. 11, n. 4, p. 535–539, dez. 2013.

FURKIM, A. M.; SACCO, A. B. DE F. Eficácia da fonoterapia em disfagia neurogênica usando a escala funcional de ingestão por via oral (FOIS) como marcador. **Revista CEFAC**, v. 10, n. 4, p. 503–512, dez. 2008.

GONG, Y., QIAO, J., HUANG, Y., ZHANG, Q., DOU, Z. **Mealcoach: Contact Microphone-Based Meal Supervision For Post-Stroke Dysphagia Patients**. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 1-5). IEEE. Disponível em: <<https://ieeexplore.ieee.org/document/10340832>>. Acesso em: 11 fev. 2024.

HAULCY, R.; GLASS, J. Classifying Alzheimer's Disease Using Audio and Text-Based Representations of Speech. **Frontiers in Psychology**, v. 11, 15 jan. 2021.

HIRAN, K.K. et al. **Machine Learning: Master Supervised and Unsupervised Learning Algorithms with Real Examples (English Edition)**. BPB Publications, 2021.

HONDA, T. et al. Characterization of Swallowing Sound: Preliminary Investigation of Normal Subjects. **PLOS ONE**, v. 11, n. 12, p. e0168187, 13 dez. 2016.

HOLT, P.R. **Management and evaluation of dysphagia**. Supplement to Annals of Long-Term Care. 2004;2-7.

JO, A.H.; KWAK, K.-C. Speech Emotion Recognition Based on Two-Stream Deep Learning Model Using Korean Audio Information. **Applied Sciences**. 2023, 13, 2167. <https://doi.org/10.3390/app13042167>

JONES, B. **Normal and abnormal swallowing: imaging in diagnosis and therapy**. Springer Science & Business Media, 2003.

JOTZ, G.P. **Disfagia - Abordagem Clínica e Cirúrgica - Criança, Adulto e Idoso**. SP: Grupo GEN, 2016. E-book. ISBN 9788595155022. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595155022/>. Acesso em: 15 fev. 2024.

KHALIFA, Y. et al. Autonomous Swallow Segment Extraction Using Deep Learning in Neck-Sensor Vibratory Signals From Patients With Dysphagia. **IEEE Journal of Biomedical and Health Informatics**, v. 27, n. 2, p. 956–967, 1 fev. 2023.

KAPOOR, A. et al. **Deep Learning with TensorFlow and Keras: Build and deploy supervised, unsupervised, deep, and reinforcement learning models**. Packt Publishing Ltd, 2022. Disponível em: <<https://github.com/PacktPublishing/Deep-Learning-with-TensorFlow-and-Keras-3rd-edition>>. Acesso em: 11 dez. 2023.

KIM, J. K. et al. Deep Learning Analysis to Automatically Detect the Presence of Penetration or Aspiration in Videofluoroscopic Swallowing Study. **Journal of Korean Medical Science**, v. 37, n. 6, 2022.

KIMURA, S. et al. Novel Approach Combining Shallow Learning and Ensemble Learning for the Automated Detection of Swallowing Sounds in a Clinical Database. **Sensors**, v. 24, n. 10, p. 3057, 2024.

LAGUARTA, J.; SUBIRANA, B. Longitudinal Speech Biomarkers for Automated Alzheimer's Detection. **Frontiers in Computer Science**, v. 3, 8 abr. 2021.

LECUN, Y. et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, 1998.

LESLIE, P.; MCHANWELL, S. The physiology of swallowing. **Scott-Brown's Otolaryngology: Head and Neck Surgery**, 2008.

LI, F. et al. Classification of Heart Sounds Using Convolutional Neural Network. **Applied Sciences**, v. 10, n. 11, p. 3956, 6 jun. 2020.

LOGEMANN, J. A. Dysphagia: Evaluation and Treatment. **Folia Phoniatica et Logopaedica**, v. 47, n. 3, p. 140–164, 1995.

LIDA, Y. et al. Detection of aspiration from images of a videofluoroscopic swallowing study adopting deep learning. **Oral Radiology**, v. 39, n. 3, p. 553–562, 8 fev. 2023.

MAHAJAN, P.; BATHS, V. Acoustic and Language Based Deep Learning Approaches for Alzheimer's Dementia Detection from Spontaneous Speech. **Frontiers in Aging Neuroscience**, v. 13, 5 fev. 2021.

MAHDI, H. et al. Tuning In: Analysis of Audio Classifier Performance in Clinical Settings with Limited Data. **arXiv preprint arXiv:2402.10100**, 2024.

MATSUO, H.; SAKUMA, K. Pathophysiology of Cachexia and Characteristics of Dysphagia in Chronic Diseases. **Asia-Pacific Journal of Oncology Nursing**, p. 100120, jul. 2022.

MATSUO, K.; PALMER, J. B. Anatomy and Physiology of Feeding and Swallowing: Normal and Abnormal. **Physical Medicine and Rehabilitation Clinics of North America**, v. 19, n. 4, p. 691–707, nov. 2008.

MCFEE, B. et al. librosa: Audio and music signal analysis in python. In: **Proceedings of the 14th python in science conference**. 2015. p. 18-25
Disponível em:
<https://www.researchgate.net/publication/328777063_librosa_Audio_and_Music_Signal_Analysis_in_Python>.

MILLER, A. J. Overview of Deglutition and Digestion. **Principles of Deglutition**, p. 3–17, 25 jul. 2012.

Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. **Guia alimentar para a população brasileira** / Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Atenção Básica. – 2. ed., 1. reimpr. – Brasília : Ministério da Saúde, 2014. 156 p. : il. ISBN 978-85-334-2176-9. 2014

MORITA, K. et al. Factors related to masticatory performance in healthy elderly individuals. **Journal of Prosthodontic Research**, v. 62, n. 4, p. 432–435, out. 2018.

MURADELI, J. ssqueezepy GitHub repository. 2022.

NAJMI, A. H et al., The continuous wavelet transform and variable resolution time-frequency analysis. **Johns Hopkins apl technical digest**, v. 18, n. 1, p. 134-140, 1997.

NAKAMORI, M. et al. Swallowing sound evaluation using an electronic stethoscope and artificial intelligence analysis for patients with amyotrophic lateral sclerosis. **Frontiers in neurology**, v. 14, p. 1212024, 2023.

NICOSIA, M. A. et al. Age Effects on the Temporal Evolution of Isometric and Swallowing Pressure. **The Journals of Gerontology Series A: Biological Sciences and Medical Sciences**, v. 55, n. 11, p. M634–M640, 1 nov. 2000.

O'NEIL, K. H. et al. The Dysphagia Outcome and Severity Scale. **Dysphagia**, v. 14, n. 3, p. 139–145, maio 1999.

OPPENHEIM, A. V. Speech spectrograms using the fast Fourier transform. **IEEE Spectrum**, v. 7, n. 8, p. 57–62, ago. 1970.

PADOVANI, A. R. et al. Dysphagia Risk Evaluation Protocol. **Revista da Sociedade Brasileira de Fonoaudiologia**, v. 12, n. 3, p. 199–205, 2007.

PANEBIANCO, M. et al. Dysphagia in neurological diseases: a literature review. **Neurological Sciences**, v. 41, p. 3067-3073, 2020.

PAYASI, S. **Understanding Confusion Matrix**. 2022. Disponível em: <<https://www.c-sharpcorner.com/article/understanding-confusion-matrix/>>. Acesso em: 10 fev. 2024.

PRIYADARSHINEE, P. et al. Alzheimer's Dementia Speech (Audio vs. Text): Multi-Modal Machine Learning at High vs. Low Resolution. **Applied sciences**, v. 13, n. 7, p. 4244–4244, 27 mar. 2023.

RABINER, L.; SCHAFER, R. **Introduction to Digital Speech Processing, Foundations and Trends R in Signal Processing**, vol 1, no 1–2, pp 1–194, 2007. ISBN: 978-1-60198-070-0. Disponível em: <https://research.iaun.ac.ir/pd/mahmoodian/pdfs/UploadFile_2643.pdf>. Acesso em 10 nov. 2023.

RECH, R.S. **Perspectivas Epidemiológicas Sobre a Disfagia Orofaríngea em Idosos Independentes da Comunidade**. Tese (Doutorado em Epidemiologia). 2020. Universidade Federal do Rio Grande do Sul.

REDDY, C. S., PARK, E., LEE J. T. **Comparative Analysis of Deep Learning Architectures for Penetration and Aspiration Detection in Videofluoroscopic Swallowing Studies**. 2023. *IEEE Access*. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/10251495>>. Acesso em: 11 fev. 2024.

RENGIFO, J. P. B. et al. Caracterización de sonidos deglutorios adquiridos mediante auscultación cervical en sujetos sanos y con disfagia. **Revista EIA**, v. 19, n. 38, p. 3831 pp. 1–12, 7 jun. 2022.

ROBERT, J. [jiaaro/pydub@ GitHub](https://github.com/jiaaro/pydub). 2022.

ROSENBEK, J.C. et al. A penetration-aspiration scale. **Dysphagia**, v. 11, p. 93-98, 1996.

SAAB, R. et al. Machine-learning assisted swallowing assessment: a deep learning-based quality improvement tool to screen for post-stroke dysphagia. **Frontiers in Neuroscience**, v. 17, p. 1302132, 2023.

SABRY, A. et al. Automatic Estimation of Laryngeal Vestibule Closure Duration Using High-Resolution Cervical Auscultation Signals. **Perspectives of the ASHA Special Interest Groups**, v. 5, n. 6, p. 1647–1656, 17 dez. 2020.

SAKAI, K. et al. Relationship between tongue strength, lip strength, and nutrition-related sarcopenia in older rehabilitation inpatients: a cross-sectional study. **Clinical Interventions in Aging**, v. Volume 12, p. 1207–1214, ago. 2017.

SHARMA, S; SHARMA, S; ATHAIYA, A. Activations Functions in Neural Networks. **International Journal of Engineering Applied Sciences and Technology**, 2020 Vol. 4, Issue 12, ISSN No. 2455-2143, Pages 310-316 Published Online April 2020 in IJEAST (<http://www.ijeast.com>) 310.

SREENIVAS, A.; MAHESHWARI, M.; JAIN, S.; CHOUDHARY, S.; VADIVU, G. Indian Sign Language Communicator Using Convolutional Neural Network. **International Journal of Advanced Science and Technology** 29. 11015-11031, 2020.

SUBRAMANI, S. et al. **SegNet-Based Deep Representation Learning for Dysphagia Classification**. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/9747732/figures#figures>>. Acesso em: 20 dez. 2022.

Stanford Otolaryngology — Head & Neck Surgery (2023). Disponível em: <<https://med.stanford.edu/ohns/education.html>>. Acesso em: 15 fev. 2024.

SUTTON, R.S.; BARTO, A.G. **Reinforcement learning: An introduction**. Second Edition. MIT press, 2018.

TAKAHASHI, K.; GROHER, M.; MICHI, K. Methodology for detecting swallowing sounds. **Dysphagia**, v. 9, n. 1, 1994.

THOMA, M. **CMGLEE. English: Receiver Operating Characteristic (ROC) curve with False Positive Rate and True Positive Rate. The diagonal shows the performance of a random classifier. Three example classifiers (blue, orange, green) are shown. Drawn by CMG Lee based on <http://commons.wikimedia.org/wiki/File:roc-draft-xkcd-style.svg> . 2018.** Disponível em: <https://commons.wikimedia.org/wiki/File:Roc_curve.svg#filelinks>. Acesso em: 11 jan. 2024.

TRIPATHI, A. M., & PANDEY, O. J. **Divide and distill: new outlooks on knowledge distillation for environmental sound classification**. IEEE Xplore. Transactions on Audio, Speech, and Language Processing, 31, 1100-1113, 2023.

WALTON, J.; SILVA, P. Physiology of swallowing. **Surgery (Oxford)**, v. 36, n. 10, p. 529–534, out. 2018.

WILKINS, T. et al. The prevalence of dysphagia in primary care patients: a HamesNet Research Network study. **Journal of the American Board of Family Medicine: JABFM**, v. 20, n. 2, p. 144–150, 2007.

YANG, X. et al. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. **Chemical Reviews**, v. 119, n. 18, p. 10520–10594, 11 jul. 2019.

YANI, M.; BUDHI IRAWAN, S., Si., M.T.; CASI SETININGSIH, S. T., M.T. Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry's Nail. **Journal of Physics: Conference Series**, v. 1201, p. 012052, maio 2019.

YUAN, L.; LIU, Y.; FENG, H.-M. Parkinson disease prediction using machine learning-based features from speech signal. **Service Oriented Computing and Applications**, 27 jun. 2023.

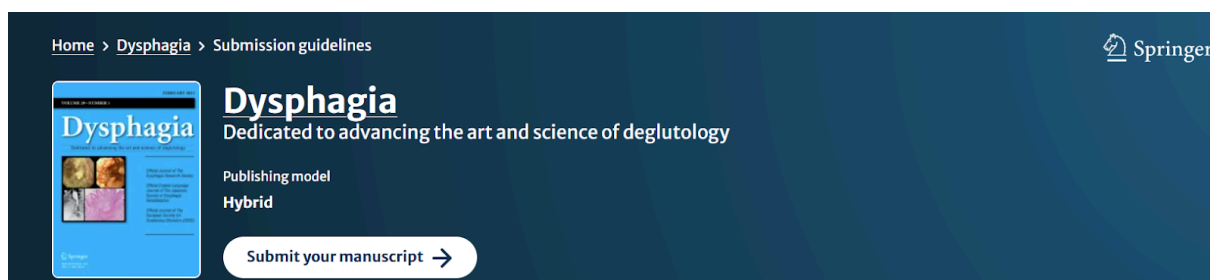
ZHANG, S. et al. Deep Learning-based Multimodal Emotion Recognition from Audio, Visual, and Text Modalities: A Systematic Review of Recent Advancements and Future Prospects. **Expert Systems with Applications**, p. 121692, 18 set. 2023.

ANEXO A

GUIA DE PUBLICAÇÃO DE ARTIGOS PARA AUTORES - DYSPHAGIA

O artigo científico redigido a partir desta dissertação foi submetido para publicação na revista Dysphagia - Fator de Impacto de 2.2 (2023).

As orientações para os autores e os passos para encaminhamento do manuscrito constam no material a seguir.



[Submission guidelines](#)

[Contents](#)

[Instructions for Authors](#)

[Types of Papers](#)

[Editorial procedure](#)

[Manuscript Submission](#)

[Title Page](#)

[Text](#)

[Scientific style](#)

[References](#)

[Tables](#)

[Artwork and Illustrations Guidelines](#)

[Supplementary Information \(SI\)](#)