

Universidade de Pernambuco

Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Dissertação de Mestrado

Área: Computação Inteligente

Título: Tecnologia assistiva para pessoas com disartria por meio de Sistema de reconhecimento de fala (ASR) personalizado.

Orientador: Sérgio Murilo Maciel Fernandes – smurilo@ecomp.poli.br

Descrição:

As Tecnologias Assistivas têm como intuito auxiliar as pessoas com quaisquer tipos de limitações a seguir dignamente as suas vidas pela redução das dificuldades apresentadas [1]. As Tecnologias Assistivas são fundamentais para dar ao deficiente a autonomia que ele necessita para realizar as suas tarefas diárias e ser socialmente inclusivo. Problemas relacionados a saúde, especialmente aqueles relacionados aos acidentes vasculares cerebrais (AVC), Doença de Parkinson, Esclerose Lateral Amiotrófica (ELA), Esclerose Múltipla (EM) e Lesão cerebral Traumática, em maior ou menor proporção, provoca nos indivíduos dificuldade no processo relativo a fala, no que se conhece por disartria[2], que os podem levar a uma condição de fragilidade de interação social, inclusive podendo gerar custos financeiros, por vezes proibitivo, às famílias[3]. A disartria, grupo de distúrbios de fala, decorrente de condições neurológicas que enfraquecem ou paralisam os músculos que produzem a fala, afeta a qualidade de vida das pessoas por dificultar a sua comunicação [4]. Pessoas com fala disártrica apresentam características patológicas, como pronúncia descontínua, por vezes explosiva, pausas indevidas e fala lenta, dentre outras, que as diferenciam de uma fala normal dificultando a compreensão das palavras [5]. O reconhecimento de fala automática (ASR), que tem funcionado de modo adequado na conversão de fala saudável, ainda apresenta dificuldades quando esta conversão envolve falas disártricas graves, devido a complexidade da fala disártrica e a falta de volume de dados para treinamento, conforme [6]. Uma maior precisão dos sistemas ASR deve considerar a extração de características dos sinais de áudio em termos de parâmetros, capazes de identificar o conteúdo linguístico, removendo informações irrelevantes e ruídos de fundo agregados a forma de onda, de modo a se obter informações auditivas relevantes. Muitas das técnicas de extração de características estão disponíveis [7], algumas das quais há vários anos, incluindo Codificação preditiva linear (LPC), Coeficientes Cepstrais de Mel-Frequência (MFCCs), Transformadas Wavelet Discretas (DWTs), Análise espectral de potência (FFT), Análise cepstral em escala Mel (MEL), Filtragem de espectros relativos de coeficientes de domínio logarítmico (RASTA), Derivada de primeira ordem (DELTA). Por outro lado, a gradação dos níveis de disartria apresenta implicações significativas para as equipes médicas pois oferece informações fonoaudiológicas essenciais a respeito da capacidade de comunicação das pessoas afetadas, permitindo deste modo a definição de protocolos específicos para melhoria da qualidade de vida das pessoas afetadas. A introdução de técnicas de computação inteligente, por meio de classificação acústica dos níveis de disartria, têm se mostrado promissoras, porém sofrem com a disponibilidade limitada dos dados. O uso de algoritmos de aprendizagem profunda, com base em imagem, possibilita a extração de informações complexas e de alta dimensionalidade, permitindo uma classificação mais acessível de diversos níveis de gravidade [3]. Apesar da disponibilidade de dispositivos de ASR, estes não têm sido adequados às pessoas com disartria por serem de uso geral. Para uma melhor experiência dos usuários há a necessidade de obtenção de palavras comuns de uso pelos indivíduos conforme sistema descrito em [8]. Este sistema tem duas fases: uma dependente de texto, onde as palavras do falante serão gravadas em um banco de dados; e outra independente de texto, que extrai as características das palavras do falante pré-gravadas para identificar o que ela está dizendo. A proposta de dissertação é orientada a indivíduos com disartria de qualquer nível de gravidade. Neste projeto definiremos várias fases: a fase de aquisição das palavras de cada indivíduo; a fase de extração das características das palavras do falante; a fase de treinamento deste base de dados particular a ser criada; a fase da extensão desta base de dados por meio de técnicas de IA; a fase de criação de aplicativo para cada indivíduo, com pequenas adaptações relativas as bases de dados individuais. O sistema traz uma nova proposta para inclusão de pessoas com disartria na sociedade ao enfatizar as práticas específicas de cada pessoa, independente do nível de gravidade da disartria, ao invés de práticas coletivas, de uso geral, complexas e nem sempre de eficiente usabilidade. Este projeto de sistema de reconhecimento de fala disártrica pode ajudar a melhorar a comunicação dessas

peçoas, permitindo a tradução de sua fala distorcida em texto compreensível, especialmente devido ao avanço das tecnologias da informação e comunicação (TIC) e da utilização de algoritmos de inteligência artificial.

Referências Bibliográficas:

- [1] **ASSISTIVE Technology** 2022 <https://www.who.int/news-room/fact-sheets/detail/assistive-technology> Acessado em: 02-06-2024.
- [2] **Dysarthria & Speech: Symptoms, Causes, Treatments** 2018 Cleveland Clinic. [Online]. Disponível em: <https://my.clevelandclinic.org/health/diseases/17653-dysarthria> [Accessed: 01- junho-2024]
- [3] HEEB DESAI, R. et al. **Exploring the Perspectives of Adults Aging With Long-Term Physical Disabilities on Physical Activity: A Qualitative Study** American Journal of Health Promotion, p. 08901171221151123, 2023.
- [4] AL-ALI, A., AL-MAADEED, S., SALEH, M., Kumar, R. et al. **The Detection of Dysarthria Severity Levels Using AI Models: A Review** vol.11, 2023, IEEE-ACCESS DOI 10.1109/ACCESS.2024.3382574
- [5] Qian, Z.; Xiao, K. **A Survey of Automatic Speech Recognition for Dysarthric Speech**. Electronics **2023**, 12, 4278. <https://doi.org/10.3390/electronics12204278>
- [6] SHAHAMIRI, S.R., LAL, V., SHAD, D., **Dysarthric Speech Transformer: A Sequence-to-Sequence Dysarthric Speech Recognition System** IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 31, 2023
- [7] LABIED, M., BELANGOUR, A., **Automatic Speech Recognition Features Extraction Techniques: A Multi-criteria Comparison** (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 8, 2021
- [8] MOHAMED NASIR, A. B., NASIR, N., R., M., SALLEH, F., H., M., **SPEESH: speech-based mobile application for dysarthric speech recognition** International Conference on Applied and Practical Sciences ICAPS (2021) Journal of Physics: Conference Series **1860** (2021) 012003 IOP Publishing doi:10.1088/1742-6596/1860/1/012003