

O USO DA INTELIGÊNCIA ARTIFICIAL COMO RECURSO DECODIFICADOR DE VOCALIZAÇÕES

FRASSON, Isabela
SACILOTTI, Adani

isabela.frasson@fatec.sp.gov.br
adani.sacilotti@fatec.sp.gov.br

Fatec Jundiaí – Deputado Ary Fossen
Fatec Jundiaí – Deputado Ary Fossen

1. INTRODUÇÃO

A vocalização não verbal pode ser, por vezes, a única forma possível de comunicação oral de uma pessoa, vocalização essa que pode ter suas intenções claras ou não, e pode ser compreendida com certo grau de dificuldade ou completamente incompreendida, dependendo do grau de familiaridade do receptor (quem interpreta) com o comunicador (quem vocaliza). Faz-se necessário, então, um interpretador linguístico que não se baseie em palavras, fonemas ou sílabas de um idioma já decodificado, mas que seja capaz de interpretar as vocalizações particulares do comunicador dentro do seu próprio contexto e apresentá-las em um código de saída mais comumente compreendido, como a língua materna dos receptores, por exemplo, a fim de uma comunicação mais efetiva.

Esse processo de “tradução” envolve o uso de um decodificador que seja baseado na própria comunicação de origem - a linguagem pessoal e única do comunicador. A decodificação da linguagem é realizada por meio da captura, interpretação e catalogação das vocalizações, que incluem grunhidos, gemidos, gritos e sons em geral, consideradas em métricas como frequência, amplitude e pausas. Essas catalogações são personalizadas de acordo com as necessidades de cada usuário e identificadas em tempo real por meio de um aplicativo para smartphone.

2. METODOLOGIA

Para cumprir os objetivos desta pesquisa, foi empregada a metodologia de pesquisa bibliográfica. Isso permitiu examinar o estado da arte, as potencialidades e os desafios do uso da Inteligência Artificial, mais especificamente das técnicas de *Machine Learning* na construção de modelos capazes de identificar padrões na comunicação não verbal.

3. RESULTADOS E DISCUSSÕES

É entendida por comunicação não verbal toda aquela à parte do discurso da fala. Nota-se que existem várias formas de comunicação não verbal. Nesta pesquisa, será abordada somente a comunicação não verbal vocal, com pouca relevância dada à expressão corporal, uma vez que o treinamento da IA é baseado apenas em um banco de dados sonoros.

O impedimento da fala e o uso de vocalizações não-verbais como grunhidos, gritos e gemidos, podem ser decorrentes de diversas situações, como diagnósticos de transtorno do espectro autista (TEA), paralisia cerebral, distúrbios genéticos ou congênitos, síndromes, atrasos globais, acidentes, entre outros fatores. De qualquer forma, todos indivíduos comunicam suas vontades e necessidades com expressividade, fazendo uso das habilidades que possuem e ferramentas que estejam ao seu alcance.

Pensando, então, na necessidade da comunicação por vias não verbais, no caso da ausência ou impossibilidade da comunicação não verbal em conjunto com a comunicação verbal, surge a questão: quais ferramentas existem atualmente que facilitam essa comunicação, ou que se propõem a fazer essa ponte entre comunicador e receptor, e de quais tecnologias elas se utilizam?

É explorada, então, a Comunicação Aumentativa e Alternativa (CAA), um conjunto de métodos e procedimentos desenvolvidos para aprimorar a comunicação, podendo substituir ou complementar a fala e/ou a escrita.

Uma das tecnologias exploradas, e onde reside o foco deste estudo, é o *Real-World Communicative and Affective Nonverbal Vocalizations* (ReCANVo [Comunicações Não Verbais Expressivas do Mundo Real, em tradução livre]): um sistema pioneiro na área de tecnologia assistiva, que criou novas metodologias de aquisição de dados e técnicas de análise pós processamento dos sinais, utilizando uma base de dados sonoros própria construída através de mais de 7000 amostras. O processamento dos dados se dá a partir do desenvolvimento de um modelo de aprendizado de máquina personalizado para cada usuário, baseado em sua própria base de dados [1].

Os modelos de aprendizado de máquina foram capazes de classificar com precisão as vocalizações não verbais em diferentes categorias, permitindo a identificação da intenção expressiva ou comunicativa por trás das vocalizações. Os resultados também indicam a eficácia dos modelos personalizados, sugerindo que os mesmos foram bem-sucedidos na interpretação das características únicas de expressividade e dos padrões das vocalizações não verbais de cada comunicador.

4. CONCLUSÕES

Este trabalho evidencia um marco significativo no uso da Inteligência Artificial para interpretar vocalizações não verbais, destacando-se pela sua abordagem inovadora e aplicação prática e também promove a inclusão de pessoas com diferentes capacidades de expressão vocal. A criação de um corpus de conhecimento específico e a expansão do banco de dados para abranger uma gama mais ampla de vocalizações são passos cruciais para a universalidade e eficácia do sistema. Este projeto, portanto, não apenas contribui para o campo tecnológico, mas também enfatiza a importância de práticas éticas e responsáveis no desenvolvimento de tecnologias que impactam diretamente a vida das pessoas.

5. REFERÊNCIAS

[1] NARAIN, J. et al. Conference: ICMI '20: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION, out. 2020).

AGRADECIMENTOS

À instituição Fatec Jundiaí – Deputado Ary Fossen, pela disponibilidade de uso dos recursos tecnológicos para a pesquisa e corpo docente excepcional; à Professora Orientadora deste projeto, Adani, pela ampla disponibilidade, certo direcionamento, feedback detalhado e acolhimento desde o início da graduação; ao professor José Madureira, pelo auxílio na compreensão da tecnologia de Inteligência Artificial e reuniões inspiradoras, e ao coordenador do curso de Análise e Desenvolvimento de Sistemas, Peter Jandl, pelas excelentes conversas e orientações acadêmicas valiosas.