



CLONAGEM DA VOZ HUMANA POR SÍNTESE DE VOZ COM O USO DE INTELIGÊNCIA ARTIFICIAL

ALESSANDER ALVES NOVAES; LUIZ ALBERTO PINTO

Introdução: Este estudo foca na síntese de fala usando Algoritmo Genético para imitar vozes, visando reproduzir vozes naturais ou sintéticas. Utilizando o modelo Tacotron 2 com Espectrograma Mel como entrada, o sistema gera fala de alta qualidade em tempo real, alinhando texto com fala para produzir vozes expressivas em diferentes idiomas. Apesar da escassez de estudos em português brasileiro, este projeto desenvolve um sistema avançado para essa língua. **Objetivos:** O objetivo é criar vozes sintéticas por inteligência artificial, úteis não apenas em publicidade e cinema, mas também para auxiliar pacientes com problemas de fala como paralisia facial ou ELA. Sistemas de IA convertem texto em fala, permitindo que pacientes expressem desejos. Tecnologias inovadoras criam vozes personalizadas, facilitando uma comunicação íntima e reconhecível. Essas soluções integram-se a tecnologias avançadas como interfaces cérebro-computador, possibilitando comunicação por meio do pensamento, promovendo interações inovadoras. **Metodologia:** Para este estudo, foram usadas 178.873 vozes masculinas brasileiras, com 5.490 frases distintas. A avaliação do modelo seguiu uma divisão estratégica dos dados: 70% para treinamento, 20% para testar a eficácia e 10% para validação. O Tacotron, baseado em redes neurais profundas, foi fundamental na criação das vozes sintéticas. Este modelo TTS é crucial para a produção de vozes sintéticas, usando o espectrograma Mel como entrada, capturando detalhes sonoros para precisão na geração de voz sintética. **Resultados:** Os resultados revelaram um valor de similaridade de cosseno de 0,9047 entre os Mel Spectrograms. Essa alta similaridade sugere uma relação substancial entre os elementos avaliados, embora não sejam idênticos. **Conclusão:** Demonstra-se que a síntese de voz com o Tacotron possui potencial significativo na geração de vozes sintéticas realistas e de alta qualidade. Essa abordagem baseada em redes neurais profundas captura nuances da fala humana, abrindo possibilidades em assistentes virtuais e soluções de comunicação para pacientes com dificuldades vocais.

Palavras-chave: **FALA SINTÉTICA; MODELO TACOTRON2; REDES NEURAIS; ALGORITMO GENÉTICO; SÍNTESE DE VOZ**