

RESUMO

GOMES JUNIOR, Sergio Pinto. *Reconhecimento de Emoções em Sinais de Fala Usando Transferência de Aprendizado*. 99 f. Dissertação (Mestrado em Engenharia Eletrônica) - Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, 2019.

A fala tem se tornado um meio de interação entre o ser humano e os computadores cada vez mais importante. Visando tornar essa interação ainda mais natural, pesquisadores têm proposto diferentes sistemas de reconhecimento de emoções na fala. Na área de reconhecimento de emoções em sinais de fala, as redes neurais profundas vêm sendo foco de intensa investigação. Visto isso, neste trabalho foi avaliado o efeito da técnica de transferência de aprendizado e do aumento da base de dados na acurácia de uma rede neural convolucional residual para a predição de emoções, comparando-a com outras técnicas de classificação tais como: a ResNet sem pré-treino, o Modelo de Mistura de Gaussianas e a Rede Neural Probabilística. Para isto, foram utilizadas as amostras das classes Felicidade, Neutra, Raiva e Tristeza contidas nas bases de dados IEMOCAP e EmoDb visando o treino e teste dos sistemas propostos. Nos experimentos com o GMM foi alcançada uma taxa de reconhecimento de 85,77% para a base de dados EmoDb e 66,83% para a IEMOCAP. Já a rede probabilística desenvolvida nesse trabalho conseguiu classificar corretamente 79,64% das amostras de teste da base de dados EmoDb. Nos experimentos com a ResNet, foram gerados os espectrogramas dos sinais de voz para serem utilizados no lugar de imagens. Nesses experimentos foi observado que as técnicas de aumento da base e de transferência de aprendizado contribuem significativamente para um melhor reconhecimento das emoções. Nesse caso, a rede convolucional classificou corretamente 81,26% das amostras.

Palavras-chave: Reconhecimento de emoções; Fala; Redes Neurais Convolucionais; Transferência de Aprendizado.