

Comparative Analysis of Windows for Speech Emotion Recognition using CNN

Felipe L. Teixeira, Salviano F.P. Soares, J.L. Pio Abreu, Paulo M. Oliveira and João P. Teixeira

CeDRI, UTAD, IEETA, Hospital da Universidade de Coimbra & Faculty of Medicine of the University of Coimbra,

INESC-TEC, SusTEC, Instituto Politécnico de Bragança

The paper presents the comparison of accuracy in the Speech Emotion Recognition task using the Hamming and Hanning windows for framing the speech and determining the spectrogram to be used as input of a convolutional neural network. The detection of between 4 and 10 emotional states was tested for both windows. The results show significant differences in accuracy between the two window types and provide valuable insights for the development of more efficient emotional state detection systems.

Keywords: Speech Emotion Recognition · Hamming · CNN