World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:13, No:01, 2019

Emotion-Convolutional Neural Network for Perceiving Stress from Audio Signals: A Brain Chemistry Approach

Authors : Anup Anand Deshmukh, Catherine Soladie, Renaud Seguier

Abstract : Emotion plays a key role in many applications like healthcare, to gather patients' emotional behavior. Unlike typical ASR (Automated Speech Recognition) problems which focus on 'what was said', it is equally important to understand 'how it was said.' There are certain emotions which are given more importance due to their effectiveness in understanding human feelings. In this paper, we propose an approach that models human stress from audio signals. The research challenge in speech emotion detection is finding the appropriate set of acoustic features corresponding to an emotion. Another difficulty lies in defining the very meaning of emotion and being able to categorize it in a precise manner. Supervised Machine Learning models, including state of the art Deep Learning classification methods, rely on the availability of clean and labelled data. One of the problems in affective computation is the limited amount of annotated data. The existing labelled emotions datasets are highly subjective to the perception of the annotator. We address the first issue of feature selection by exploiting the use of traditional MFCC (Mel-Frequency Cepstral Coefficients) features in Convolutional Neural Network. Our proposed Emo-CNN (Emotion-CNN) architecture treats speech representations in a manner similar to how CNN's treat images in a vision problem. Our experiments show that Emo-CNN consistently and significantly outperforms the popular existing methods over multiple datasets. It achieves 90.2% categorical accuracy on the Emo-DB dataset. We claim that Emo-CNN is robust to speaker variations and environmental distortions. The proposed approach achieves 85.5% speaker-dependant categorical accuracy for SAVEE (Surrey Audio-Visual Expressed Emotion) dataset, beating the existing CNN based approach by 10.2%. To tackle the second problem of subjectivity in stress labels, we use Lovheim's cube, which is a 3-dimensional projection of emotions. Monoamine neurotransmitters are a type of chemical messengers in the brain that transmits signals on perceiving emotions. The cube aims at explaining the relationship between these neurotransmitters and the positions of emotions in 3D space. The learnt emotion representations from the Emo-CNN are mapped to the cube using three component PCA (Principal Component Analysis) which is then used to model human stress. This proposed approach not only circumvents the need for labelled stress data but also complies with the psychological theory of emotions given by Lovheim's cube. We believe that this work is the first step towards creating a connection between Artificial Intelligence and the chemistry of human emotions.

Keywords : deep learning, brain chemistry, emotion perception, Lovheim's cube
Conference Title : ICSERAC 2019 : International Conference on Speech Emotion Recognition and Affective Computing
Conference Location : Paris, France
Conference Dates : January 24-25, 2019