Scopus

Documents

Alghifari, M.F.[a] , Gunawan, T.S.[a b] , Wan Nordin, M.A.B.[a] , Qadri, S.A.A.[a] , Kartiwi, M.[c] , Janin, Z.[d]

On the use of voice activity detection in speech emotion recognition
(2019) *Bulletin of Electrical Engineering and Informatics*, 8 (4), pp. 1324-1332.

DOI: 10.11591/eei.v8i4.1646

[a] Department of Electrical and Computer Engineering, International Islamic University Malaysia, Malaysia
[b] School of Electrical Engineering and Telecommunications, UNSW, Australia
[c] Information Systems Department, International Islamic University Malaysia, Malaysia
[d] Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, 40450, Malaysia

Abstract
Emotion recognition through speech has many potential applications, however the challenge comes from achieving a high emotion recognition while using limited resources or interference such as noise. In this paper we have explored the possibility of improving speech emotion recognition by utilizing the voice activity detection (VAD) concept. The emotional voice data from the Berlin Emotion Database (EMO-DB) and a custom-made database LQ Audio Dataset are firstly preprocessed by VAD before feature extraction. The features are then passed to the deep neural network for classification. In this paper, we have chosen MFCC to be the sole determinant feature. From the results obtained using VAD and without, we have found that the VAD improved the recognition rate of 5 emotions (happy, angry, sad, fear, and neutral) by 3.7% when recognizing clean signals, while the effect of using VAD when training a network with both clean and noisy signals improved our previous results by 50%. © 2019 Institute of Advanced Engineering and Science. All rights reserved.

Author Keywords
Deep neural network; Speech emotion recognition; Voice activity detection

Correspondence Address
Gunawan T.S.; Department of Electrical and Computer Engineering, International Islamic University MalaysiaMalaysia;
email: tsgunawan@iium.edu.my