Journal of Acoustical Engineering Society of Iran, Vol. 8, No. 2, 2021

(Research Article)
Improving the average precision of Persian vowel classification from speech signal by using convolutional neural network

M. Asgari*, N. Akbari
Broadcast Engineering Department, Islamic Republic of Iran Broadcasting University

Received: 2020/02/16, Accepted: 2021/03/02

Abstract

One approach to speech recognition is to model speech based on a number of phonetic units. Because the frequency and temporal characteristics of vowels are more stable than other phonems, it is important to recognize vowels to distinguish speech. In this research, the aim is to present a model using modern methods, such as deep neural network to improve the accuracy of vowel recognition and increase its applications. 30 speakers (15 females and 15 males) read all the combinations of consonants with 6 Persian vowels. After preprocessing, the speech data is segmented into frames containing only the vowels and its spectrogram is extracted. These spectrogram are given as input to the neural network with two hidden layers. Speech of 25 speakers were used for training and speech of 5 speakers were used for testing. The average of accuracy of 6 Persian vowels for the proposed model was 93.17% (total average of vowel detection error is 6.83%). In previous works the average of vowel detection error was 9.7% to 19.6% that in proposed model improved from 2.87% to 12.77%.

Keywords: Persian vowel recognition, Classification, Convolutional neural network, Persian vowel dataset.

pp. 51-59 (In Persian)

* Corresponding author E-mail: m.asgari@iribu.ac.ir

[Downloaded from joasi.ir on 2026-11-11]

[DOR: 20.1001.1.23455748.1399.8.2.3.0]