World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:12, No:10, 2018

An Automatic Speech Recognition of Conversational Telephone Speech in Malay Language

Authors : M. Draman, S. Z. Muhamad Yassin, M. S. Alias, Z. Lambak, M. I. Zulkifli, S. N. Padhi, K. N. Baharim, F. Maskuriy, A. I. A. Rahim

Abstract : The performance of Malay automatic speech recognition (ASR) system for the call centre environment is presented. The system utilizes Kaldi toolkit as the platform to the entire library and algorithm used in performing the ASR task. The acoustic model implemented in this system uses a deep neural network (DNN) method to model the acoustic signal and the standard (n-gram) model for language modelling. With 80 hours of training data from the call centre recordings, the ASR system can achieve 72% of accuracy that corresponds to 28% of word error rate (WER). The testing was done using 20 hours of audio data. Despite the implementation of DNN, the system shows a low accuracy owing to the varieties of noises, accent and dialect that typically occurs in Malaysian call centre environment. This significant variation of speakers is reflected by the large standard deviation of the average word error rate (WERav) (i.e., ~ 10%). It is observed that the lowest WER (13.8%) was obtained from recording sample with a standard Malay dialect (central Malaysia) of native speaker as compared to 49% of the sample with the highest WER that contains conversation of the speaker that uses non-standard Malay dialect.

Keywords : conversational speech recognition, deep neural network, Malay language, speech recognition
Conference Title : ICNLPSRS 2018 : International Conference on Natural Language Computing, Speech Recognition and Synthesis
Conference Location : Tokyo, Japan
Conference Dates : October 08-09, 2018