

ABSTRACT

Dialect is a variation of the language used by a group of people, sometimes in a particular region. It plays an essential role in automatic speech recognition (ASR). In general, ASR gives high accuracy for a dialect-specific case, but it obtains a low accuracy for a multi-dialect application, such as for the Indonesian language that has hundreds of dialects. In this research, a system to recognize various dialects in Indonesia is developed. First, an utterance is pre-processed using both normalization and framing. Second, its features are then extracted using the Mel frequency cepstrum coefficients (MFCC), which is one of the feature extraction methods for the best acoustic signals. Finally, a deep recurrent neural network (DRNN) is used to learn and classify dialect characteristics. Evaluation of the dataset of five major dialects in Indonesia shows that the greater the Epoch and Bath Size, the greater the accuracy produced by the DRNN. However, accuracy is not directly proportional to the value of both parameters. The Epoch of 30 and Batch Size of 30 are the optimum parameters that yield the highest accuracy of 87.0% for the training set. Evaluation of the testing set shows that it gives an accuracy of 85.4% for the unseen dialects.

Kata Kunci: *deep recurrent neural network, dialect recognition, Indonesian, Mel frequency cepstrum coefficients, speech recognition.*