ABSTRACT

The use of online communication had become a normality seeing the increase in active users of online meeting applications during the Covid-19 period. Active users of online meetings consisted of various groups but the accessibility for users with special needs was still not fully met. Deaf people could only use sign language or lip reading as method to understand what the other person was saying. Lip reading was a difficult task to do hence it could be assisted by machines through recognized lip patterns. This research developed a classification model of lip patterns in the speech of Indonesian syllables using the EfficientNetV2-BiLSTM architecture. EfficientNetV2 was used in the feature extraction process to produce 1280 features per image. BiLSTM was used to classify the sequence of lip pattern images from which its features had been extracted. The model had 2 BiLSTM layers with 256 and 128 nodes and a dense layer with 64 nodes. The model was trained using 0,00005 learning rate, 0,2 dropout rate and 128 batches. The model was tested using 2764 test data or 20% of all data. The model evaluation results showed that the model had good performance with an average f1-score of 0.98239 and an accuracy of 0.98513.

Keywords: EfficientNetV2, BiLSTM, pattern recognition, lip motion pattern, lip reading, Indonesia language