World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:14, No:06, 2020

# Chinese Sentence Level Lip Recognition

**Authors :** Peng Wang, Tigang Jiang

**Abstract :** The computer based lip reading method of different languages cannot be universal. At present, for the research of Chinese lip reading, whether the work on data sets or recognition algorithms, is far from mature. In this paper, we study the Chinese lipreading method based on machine learning, and propose a Chinese Sentence-level lip-reading network (CNLipNet) model which consists of spatio-temporal convolutional neural network(CNN), recurrent neural network(RNN) and Connectionist Temporal Classification (CTC) loss function. This model can map variable-length sequence of video frames to Chinese Pinyin sequence and is trained end-to-end. More over, We create CNLRS, a Chinese Lipreading Dataset, which contains 5948 samples and can be shared through github. The evaluation of CNLipNet on this dataset yielded a 41% word correct rate and a 70.6% character correct rate. This evaluation result is far superior to the professional human lip readers, indicating that CNLipNet performs well in lipreading.

**Keywords :** lipreading, machine learning, spatio-temporal, convolutional neural network, recurrent neural network
**Conference Title :** ICCV 2020 : International Conference on Computational Vision
**Conference Location :** Venice, Italy
**Conference Dates :** June 22-23, 2020