LipVision: A Deep Learning Approach

International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

Volume 179
Number 8

Year of Publication: 2017

Authors:
Parth Khetarpal, Riaz Moradian, Shayan Sadar, Sunny Doulhani, Salma Pathan

10.5120/ijca2017916029

Abstract

Lip-Reading is the task of interpreting what an individual is saying by analysing his/her mouth patterns while the individual is talking. The paper is conducting a survey on the previously done work on Lip-Reading. It will be discussing the different classifiers used, their efficiency and the end accuracy obtained. Lip-Reading can be used in a myriad of fields such as medical, communication and gaming.

The proposed system will use the GRID corpus dataset in which the videos are recorded from 33 speakers. OpenCV and dlib will be used for face and mouth detection. Then the mouth ROI will be used with the iBug tool to annotate facial landmarks. The architecture consists of Convolutional Neural Networks which will be created and trained in Tensorflow (Open Source Software Library), which are then passed through Connectionist Temporal Classification. It will then be using saliency visualisation technique to interpret and match the learned behaviour and generate text.
References

7. GRID Corpus Dataset http://spandh.dcs.shef.ac.uk/gridcorpus/

Index Terms

Computer Science  Information Sciences

Keywords

Computer Vision, Deep Learning, Pattern Recognition.