

Phone calls event identification system for deaf users

Stefano Artuso (BA, audio engineer, developer @ Padius)

Background

We are seeing technology taking quite all fields of our reality. Speech recognition is one of the main trend that is finding more and more space in our life. Ironically, there are simple obstacles that are not really defeated. One set of this problem is related to the deaf people and their difficulties to make a simple emergency call. Padius using text-to-speech and speech-to-text technologies is trying to overcome this problems and help deaf people to make a call with their smartphone. The next big step is to help deaf people during meeting call to associate the right speaker to the corresponding speech, a problem called diarization, useful in different scenarios like forensic.

Problems



- How can a deaf user could have the same user experience of a hearing one? How can we try to:
 - Understand which audio event is flowing (a “hold-on” music, a human operator answering, a pre-recorded welcome message...)
 - Which speaker is talking and if is the same of a previous speaker that in the same call already have said something
- How can we optimize the work of speech-to-text engine?

Objectives

- Develop a system that could classify audio events using deep learning techniques in real time
- Improve this system for the recognition of a real human voice or a recorded one
- Extend this system to a diarization system that can satisfy the problem of "who spoke when?"

