Using Deep Learning and EMG to predict non-audible speech

Rommel T. Fernandes
Seaver College of Science and Engineering
Loyola Marymount University
Los Angeles, CA 90045
Email: rferna16@lion.lmu.edu

***Abstract*—Many post-stroke victims deal with physiological problems such as speech impediments due to aphasia. With the advancement of Human-Computer Interface (HCI) research, this paper aims to create a project plan based on Silent Speech Interfaces that use Deep Learning and Machine Learning to predict non-audible speech. We will first briefly introduce HCI systems such as Silent Speech Interfaces, and go over how Deep Learning and Machine Learning can be used to predict speech. Next, we will go over our project plan that will investigate Deep Learning and Silent Speech Interfaces and go over the research problem, which will identify the main goal and objectives. Finally, we explain our research plan which will dive into the sub goals required to complete the main objective of building a silent-speech recognition system using Electromyography (EMG) and Deep Learning.**

I. Introduction

Over the past few years, HCI has been an [1]. increasing field of study. Human Computer Interaction can be described a as feedback loop between a human and computing interface. With the increased usage of sensors worn on humans, such as watches, heart rate monitors, and other smart sensors, researchers are trying to extract the information and classify typical human activities. One of the ways HCI is used is for Silent Speech Interfaces (SSI). SSI aims to use signal extracting systems like electromyography (EMG) and electroencephalography (EEG) to convert signals of silent or non-audible speech and use a machine to classify the results. This feedback loop involves feature extraction, model training, and activity inference [1].

For the purpose of this research, non-audible speech can be classified as the inability to verbalize words or sentences through the use of sound in an effective way. SSI systems are not new; they have been proposed in the past. What's new is the computing resources and type of algorithms used to classify speech in SSI systems. In the past, machine learning algorithms such as Decision Tree, Support Vector Machine, Naïve Bayes, and Hidden Markov Models were used as classifiers for speech. Though powerful, they require extensive feature extraction from EMG signals. Typically, with traditional feature extraction, only shallow features can be learned from those approaches, leading to undermined performance. Recently, we have witnessed the incremental development of Deep Learning, which alleviates the issue of feature engineering, since the models extract valuable information through several iterations. Therefore, using EMG signals to classify non-verbal speech does not have to a be a laborious task.

The rest of this paper will discuss the research problem we are trying to achieve. We will define the main goal and a set of sub-goals such that the main goal is achieved when all the sub-goals are met. The paper will then transition into the research plan which will discuss the approach that will be followed to solve the proposed research problem. We will list the necessary resources for the execution of the research plan such as the mathematics, simulation and modeling of the project. Finally, with this project proposal, we can make conclusions to specific methods one can use for furthering research possibilities in SSI systems.

II. Conclusion

The conclusion goes here.

Acknowledgment

The authors would like to thank...

References

[1] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech Recognition with Deep Recurrent Neural Networks," *arXiv:1303.5778 [cs]*, Mar. 2013, arXiv: 1303.5778. [Online]. Available: http://arxiv.org/abs/1303.5778