CHAPTER I

INTRODUCTION

1.1 Research Background

Dysarthria is one of a neurological disorder that causes people lose their ability to articulate properly (Dardjowidjojo, 2008). According to Yayasan Stroke Indonesia (2012), Dysarthria is one of the Stroke's symptoms, which occur in 15% of the sufferer. There are about 500.000 Stroke survivors, annually, in Indonesia. Hence, there will be, at least, 75000 Dysarthric speakers annually, caused by stroke. In case of the serious amount of Dysarthric people, supporting their life is greatly important. One of many ways is by resolving the gap between Dysarthric speakers and normal speakers, to establish better communication in daily life through Automatic Speech Recognition (ASR).

Recognizing speech for a human is easy, but not for a computer. By adopting the way a human brain works using Artificial Neural Network, a computer can do what human do, in this case, recognizing speech. The foundation of speech recognition is speech classification, where a network is trained to learn the pattern of each class of speech. Therefore, this research will be focused on classification for Dysarthric speech.

Neural Network has been vastly developed over the world and novel methods are emerging nowadays. In fact, training network is about finding the optimal parameters to mitigate errors and maximize the likelihood. Many scholars had conducted research to find the best method to train various type of datasets.

Convolutional Neural Network (CNN) is one of modern method and it has been a success in various recent works related to image recognition. Abdel-Hamid et al (2014) proposed CNN architecture for speech recognition and Adam (2016) proved the satisfying performance of CNN in his research about Speech

Recognition using Deep Learning. CNN outputs probabilities of data to particular classes. In spite of the fact that CNN satisfies speech classification, it still has a problem in handling outliers.

Tang (2013) proved a better performance of CNN using Support Vector Machine (SVM). In line with Hasegawa (2006), on his research about Dysarthric Speech using HMM and SVM, proved that SVM has better performance in handling major deletion of consonants, which matches the case of Dysarthric speech. SVM provides a smarter way to deal with outliers by maximizing the margin of each class and it will result in robust classification. Hence, this research aims to provide vigorous classification result using a combination of CNN and SVM.

1.2 Research Problem

The problems that will be solved in research are :

1. How to implement CNN and SVM for Dysarthric Speech Recognition?
2. How robust is classification result using CNN and SVM in Dysarthric Speech Classification?

1.3 Research Scope

The scopes of this research are :

1. Building Speaker-Dependent Isolated-word Recognition focused on 10 digit words.
2. Using UA Speech database as a secondary dataset to train the model.
3. Focused on severe to moderate Dysarthric speakers.
4. CNN will be used for pre-training data and SVM will be used for classification.
5. Using Radial Basis Function (RBF) kernel for SVM, since it is powerful for nonlinearly separable data (Xu et al, 2001).

1.4 Research Purposes

The purposes of this research are :

1. To build Speaker-Dependent Dysarthric speech classification using CNN and SVM.
2. To know how powerful is the method used for classification method for Dysarthric speech with CNN architecture.

1.5 Research Objectives

This research proposes to bring support for the development of Automatic Speech Recognition for Dysarthric speech.