

INTISARI

Ucapan manusia merupakan salah satu gelombang yang sangat unik. Jenis gelombang ucapan manusia beraneka ragam. Berbagai penelitian menemukan inovasi terhadap komputer yang mampu mengenali ucapan manusia. Hal inilah yang memberikan daya tarik penulis untuk membuat sistem pengenalan ucapan. Penulis berfokus kepada pengenalan ucapan angka 0-9 dalam bahasa Indonesia.

Penelitian ini memproses ucapan manusia yang telah terekam. Perekaman dilakukan menggunakan mikrofon. Data perekaman yang sudah dibuat dalam format *.wav* akan dibagi menjadi 2 yaitu untuk data *training* dan data testing. Pada data *training* maupun data testing memiliki jumlah ucapan yang sama sebanyak 200 data. Data tersebut terdiri atas 20 kali perekaman. Masing-masing pengucapan dilakukan sebanyak 10 kali untuk setiap angka 0 hingga 9 dari 2 orang yang berbeda. Proses yang dilakukan untuk sistem pengenalan ucapan angka meliputi *pre-processing*, ekstraksi ciri MFCC, klasifikasi jaringan syaraf tiruan *backpropagation* dan penentuan hasil ucapan pengenalan. Jaringan syaraf tiruan *backpropagation* dilatih dengan algoritma pelatihan *gradient descent*.

Percobaan dilakukan dengan memvariasi nilai terhadap beberapa parameter jaringan syaraf tiruan diantaranya *learning rate*, jumlah *neuron hidden layer*, dan *epoch*. Hasil pengujian tidak langsung terletak pada variasi *learning rate* 0,07 dengan jumlah *neuron hidden layer* 10 dan *epoch* 500. Hasil pengujian secara *real time* mendapat rata-rata tingkat pengenalan sebesar 86%.

Kata kunci : Pengenalan Ucapan, *Mel Frequency Cepstral Coefficients* (MFCC), Jaringan Syaraf Tiruan *Backpropagation*.

ABSTRACT

Human speech is a very unique wave. Types of human speech waves are various. Many studies have found innovations in computers that are able to recognize human speech. This is what attracts the writer to create a speech recognition system. The author focuses on the recognition of the numbers 0-9 in Indonesian.

This research processes recorded human speech. Recording is done using a microphone. The recording data that has been made in the .wav format will be divided into 2, for training data and testing data. The training data and testing data have the same amount with the speech amount, which are 200 data. The data consists of 20 recordings. Each pronunciation is done 10 times for each number 0 to 9 from 2 different people. The process carried out for the speech recognition system includes pre-processing, MFCC feature extraction, backpropagation neural network classification and determination of speech recognition results. The backpropagation neural network is trained with a gradient descent training algorithm.

Experiments were carried out by varying the values for several parameters of the artificial neural network including learning rate, number of hidden layer neurons, and epoch. The indirect test results lie in the variation of the learning rate 0,07 with the number of hidden layer neurons 10 and epoch 500. The real time test results get an average recognition rate of 86%.

Keywords: Speech Recognition, Mel Frequency Cepstral Coefficients (MFCC), Backpropagation Neural Network.

