BAB I

PENDAHULUAN

I. LATAR BELAKANG

Manusia memiliki akal (otak) dan indera yang merupakan anugerah Tuhan Yang Maha Esa. Akal dan indera berfungsi sebagai alat untuk berpikir dan menerima rangsangan (kepekaan). Indera menerima rangsangan dan diteruskan ke otak (akal) guna mengenali rangsangan yang diterima. Manusia melakukan proses pengenalan dengan cara belajar sejak lahir. Seorang bayi pertama kali belajar mengenali suara dan wajah ibunya, kemudian baru belajar mengenali lingkungan sekitarnya.

Manusia memiliki ciri/perbedaan biologis antara satu dengan yang lainnya, baik bentuk tubuh, suara, mata, dan lain-lain. Dewasa ini, komputer pintar telah memanfaatkan ciri biologis manusia tersebut sebagai alat keamanan (*security*) dengan cara mengenalinya, yang dinamakan *biometric*. *Biometric* yang telah berkembang saat ini diantaranya adalah pengenalan sidik jari, suara, dan retina mata manusia.

Manusia memiliki suara (sinyal suara) yang berbeda-beda. Pengenalan suara dilakukan manusia melalui proses belajar, begitu juga dengan komputer. Berawal dari mesin Von Neumann, pada akhir tahun 1980-an, impian manusia untuk menciptakan *Brain Like Computer* (BLC) atau *Neural Intellegence* mengalami perkembangan pesat. Walaupun jelas tidak mungkin mampu menyamai kecanggihan otak manusia, BLC diharapkan bisa meniru fungsi

This document is Undip Institutional Repository Collection. The author(s) or copyright owner(s) agree that UNDIP-IR may, without changing the content, translate the submission to any medium or format for the purpose of preservation. The author(s) or copyright owner(s) also agree that UNDIP-IR may keep more than one copy of this submission for purpose of security, back-up and preservation:
(http://eprints.undip.ac.id)

kepandaian biologis. Fungsi itu antara lain mengolah informasi dengan menggunakan teknik pelatihan dalam proses belajar, pengenalan pola serta menentukan keputusan. Karena meniru fungsi sistem syaraf tersebut, BLC akhirnya lebih terkenal sebagai *Artificial Neural Network* (ANN) atau Jaringan Saraf Tiruan (JST).

JST telah banyak diaplikasikan untuk analisis data, pengenalan pola, dan fungsi kendali. Salah satu teknik pelatihan pada JST adalah propagasi balik (*Back Propagationi*/BP) yang merupakan bagian dari model *supervised learning* dari JST. BP banyak dimanfaatkan untuk berbagai penelitian pada JST khususnya *signal processing*, karena melalui model *supervised learning* dapat ditentukan kesalahan (*error*) pengenalan sesuai dengan yang diharapkan. Berdasarkan uraian di atas penulis tertarik untuk mengangkat penelitian mengenai pengenalan suara dengan memanfaatkan teknik pelatihan BP dari model *supervised learning* JST guna melatih komputer dalam pengenalannya.

II. PERUMUSAN MASALAH

Suara manusia adalah berupa sinyal-sinyal analog. Padahal komputer hanya bisa melakukan proses pada sinyal digital. Sehingga perlu mengubah sinyal analog ke sinyal digital. Perbedaan sinyal digital suara manusia dilihat dari ciri-cirinya, untuk itu diperlukan suatu analisa sinyal digital, sehingga dapat diekstraksi ciri-ciri (*feature extraction*) sinyal tersebut.

Setelah diketahui ciri-ciri tiap sinyal suara, maka perlu diajarkan (*training*) ke komputer yakni menggunakan propagasi balik (BP) dari model

This document is Undip Institutional Repository Collection. The author(s) or copyright owner(s) agree that UNDIP-IR may, without changing the content, translate the submission to any medium or format for the purpose of preservation. The author(s) or copyright owner(s) also agree that UNDIP-IR may keep more than one copy of this submission for purpose of security, back-up and preservation:
(http://eprints.undip.ac.id)

supervised learning jaringan saraf tiruan. BP adalah suatu algoritma pelatihan (*training algorithm*) yang berbentuk *multilayer* dari JST dengan memanfaatkan *hidden layer* yang ada. Beberapa modifikasi dilakukan guna meningkatkan performasi JST pada beberapa keadaan tertentu, yaitu dengan pembaharuan nilai bobot menggunakan momentum dan mengubah nilai *learning rate* pada saat pelatihan. Permasalahan yang muncul adalah bagaimana menganalisa sumber suara untuk mendapatkan ciri-cirinya, dan dalam pelatihan pada JST berapa jumlah *hiden layer* beserta *node*nya, juga momentum serta nilai *learning rate* yang diperlukan agar pengenalan suara tersebut bisa optimal

III. PEMBATASAN MASALAH

Input yang digunakan terdiri atas sumber suara (orang) yang terbatas. Dengan mengucapkan kalimat yang telah ditentukan sebelumnya.

Pencuplikan sinyal suara dilakukan dengan laju pencuplikan 8000 Hz, 8 bit. Setelah itu dilakukan proses normalisasi terhadap sinyal hasil pencuplikan. Pada proses ini data yang diambil sebanyak 5760 buah data. Data yang diambil berdasarkan asumsi bahwa pengucapan satu kata membutuhkan 0,72 detik. Sinyal suara berbeda terhadap perubahan waktu (tidak tetap), maka proses analisa sinyal dilakukan dengan membagi sinyal tersebut ke dalam *frame-frame* yang sempit atau pendek (~30 ms) sehingga sinyal bisa dikatakan bersifat tetap. Proses *windowing* menggunakan *Hamming window*. Orde LPC diambil sebesar 10, berdasarkan rumus pemilihan orde LPC. Metode yang digunakan untuk analisa LPC adalah metode *Levinson-Durbin*.

This document is Undip Institutional Repository Collection. The author(s) or copyright owner(s) agree that UNDIP-IR may, without changing the content, translate the submission to any medium or format for the purpose of preservation. The author(s) or copyright owner(s) also agree that UNDIP-IR may keep more than one copy of this submission for purpose of security, back-up and preservation:
(http://eprints.undip.ac.id)

Proses mengubah sinyal domain waktu menjadi domain frekuensi dilakukan dengan menggunakan *Fast Fourier Transform* (FFT) *radix* 2 *Decimation in Frequency* (DIF) 512 *point*. Keluaran dari proses ini ditentukan sebanyak 32 data, yaitu rata-rata dari 256 *point* FFT (FFT bersifat simetris) per 8 blok.

Proses pelatihan dan pengenalan suara dilakukan menggunakan jaringan saraf tiruan propagasi balik dengan satu *hidden layer*. Inisialisasi nilai bobot dan nilai bias pada jaringan menggunakan inisialisasi *Nguyen-Widrow* dengan mengambil nilai secara acak antara – 0,5 dan 0,5. Proses pelatihan dilakukan dengan menggunakan target dari sinyal masukan awal (suara penulis sendiri) sampai mencapai *error* referensi sebesar 0,0001. Untuk meningkatkan perfomansi jaringan dilakukan beberapa modifikasi, yaitu pembaharuan nilai bobot dengan momentum dan mengubah nilai *learning rate* pada saat pelatihan.

IV. TUJUAN PENULISAN

Penulisan tugas akhir ini bertujuan menyusun perangkat lunak atau program agar komputer dapat mengenali sumber suara (orang) dalam mengucapkan suatu kalimat tertentu. Dengan perangkat yang ada, dilakukan penelitian untuk mengetahui parameter-parameter jaringan saraf tiruan propagasi balik yang optimal bagi pengenalan tersebut. Dan beberapa modifikasi dilakukan terhadap jaringan yang bertujuan untuk meningkatkan kinerja jaringan.

This document is Undip Institutional Repository Collection. The author(s) or copyright owner(s) agree that UNDIP-IR may, without changing the content, translate the submission to any medium or format for the purpose of preservation. The author(s) or copyright owner(s) also agree that UNDIP-IR may keep more than one copy of this submission for purpose of security, back-up and preservation:
(http://eprints.undip.ac.id)

V. SISTEMATIKA PENULISAN

Pada tugas akhir ini terdiri dari tiga bab ditambah kesimpulan dan lampiran program, dengan perincian sebagai berikut :

Bab I adalah pendahuluan berisi latar belakang, perumusan masalah, pembatasan masalah, tujuan penulisan, dan sistematika penulisan.

Bab II adalah teori penunjang, yang berisi dasar-dasar teori tentang pencuplikan sinyal analog, metode *Linear Predictive Coding* (LPC) yang menguraikan mengenai dasar LPC, analisa LPC, dan metode autokorelasi. Kemudian dilanjutkan dengan teori transformasi Fourier dan jaringan saraf tiruan propagasi balik

Bab III adalah pembahasan pengenalan suara manusia, yang berisi tentang proses pencuplikan sinyal, langkah-langkah analisa sinyal dengan metode LPC, *fast fourier transform* (FFT), implementasi jaringan saraf tiruan propagasi balik, dan contoh kasus pengenalan suara manusia.

This document is Undip Institutional Repository Collection. The author(s) or copyright owner(s) agree that UNDIP-IR may, without changing the content, translate the submission to any medium or format for the purpose of preservation. The author(s) or copyright owner(s) also agree that UNDIP-IR may keep more than one copy of this submission for purpose of security, back-up and preservation:
(http://eprints.undip.ac.id)