

PEMODELAN JARINGAN SYARAF TIRUAN RESILIENT BACKPROPAGATION UNTUK KONVERSI SUARA GITAR KE CORD

Yosi Nurhayati⁽¹⁾
yoosinhayx5@gmail.com

Agus Buono⁽²⁾
pudesha@gmail.com

Abstract

The guitar is a musical instrument that has a chord as a reference tone. It is a fact that is not all human auditory system can distinguish between high and low tones of a musical instrument in good accurate. Then, in this research we develop a voice guitar to cord conversion using resilient backpropagation neural network (RBNN) as to classifier and Mel Frequency Cepstrum Coefficient (MFCC) as feature extraction. We record 345 for each cord (totally we have 8640 recording data with WAV format). Experiments are conducted for some number of cepstral coefficients (13, 26, and 39), with 100 millisecond as time frame and 40% overlapping between successive frame. Total number of hidden neurons in RBNN model in this experiments are 10, 25, 50 and 100. According to the experiment, the maximum accuracy is 96.88% for 52 number of cepstral coefficients and 100 neurons hidden.

Keywords: *Chord, Resilient Backpropagation Neural Network, Mel Frequency cepstral Coefficients (MFCC).*

1. Pendahuluan

Dalam dunia musik dikenal adanya istilah *chord*, *chord* inilah yang akan menjadi acuan pada musik atau lagu. *Chord* merupakan rangkaian nada yang membangun keharmonisan pada musik (Wisnudisastra 2009). Salah satu contoh alaf musik yang menghasilkan bunyi-bunyian dan memiliki *chord* sebagai acuan nada adalah gitar. Seorang pemain musik harus mengerti dan mengenali *chord* apa saja pada lagu yang dimainkannya. Bagi musisi hal ini sangat mudah dilakukan karena sistem pendengaran dalam mengenali dan memainkan nada-nada *chord* pada suatu lagu atau musik sudah terlatih. Lain halnya bagi seorang pemula, hal ini akan sulit dilakukan karena sistem pendengaran dalam mengenali dan memainkan nada-nada *chord* belum terlatih. Solusi untuk menanggulangi masalah tersebut salah satunya yaitu dengan penggunaan komputer sebagai sistem pengenalan *chord* pada alat musik gitar. Hal ini dikarenakan karakteristik komputer yang konsisten terhadap eksekusi perintah yang diberikan dan memudahkan pengguna sistem untuk mendapatkan hasil pengenalan *chord* yang akurat.

Pada penelitian sebelumnya yang dilakukan oleh Wisnudisastra (2010) dengan menggunakan metode *codebook* sebagai *feature matching* dan menggunakan teknik

¹ Departemen Ilmu Komputer, FMIPA, Institut Pertanian Bogor

² Departemen Ilmu Komputer, FMIPA, Institut Pertanian Bogor

MFCC untuk *feature extraction* dalam mengenali sebuah *chord*. Metode dan teknik tersebut menghasilkan rataan akurasi sebesar 91% untuk percobaan menggunakan 13 koefisien *cepstral* dan sebesar 96% untuk percobaan menggunakan 26 koefisien *cepstral*. Penelitian lainnya yang terkait dilakukan oleh Susanto (2007) dengan objek dan metode berbeda yaitu identifikasi pembicara menggunakan Jaringan Syaraf Tiruan (JST) *Resilient Backpropagation* diperoleh rataan akurasi terbaik sebesar 92.8%.

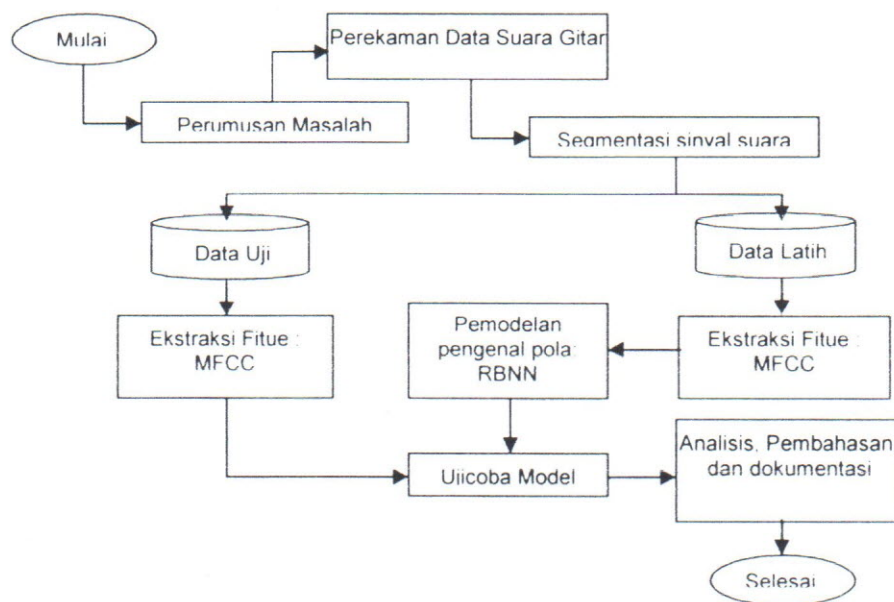
Oleh karena itu, pada penelitian ini akan dibangun sistem dengan menggunakan metode Jaringan Syaraf Tiruan (JST) *Resilient Backpropagation* dan menggunakan teknik MFCC sebagai ekstraksi cirinya. Dalam pengenalan suara, MFCC dapat merepresentasikan sinyal lebih baik dibandingkan dengan *Linear Prediction Cepstrum Coefficient* (LPCC) dan teknik lainnya (Buono 2009). Metode LPCC memiliki keunggulan dalam hal waktu komputasi yang lebih sederhana, namun tingkat akurasinya tidak sebaik MFCC. Pada *JST Resilient Backpropagation* terdapat proses pelatihan yang berulang-ulang dan memiliki tambahan faktor naik juga faktor turun. Dengan kelebihan tersebut, penelitian ini diharapkan bisa mendapatkan nilai akurasi yang lebih baik dari penelitian sebelumnya.

Selanjutnya, paper ini disajikan dengan susunan sebagai berikut: bagian 2 mendeskripsikan metode percobaan, data yang dipergunakan, serta pemrosesan yang dilakukan. Bagian berikutnya difokuskan pada diskusi mengenai hasil dan pembahasan. Akhirnya, pada bagian 4 akan disajikan kesimpulan dari penelitian ini.

2. Metode Penelitian, Data, dan Pemrosesan

2.1 Metode Penelitian

Penelitian ini dilaksanakan mengikuti tahapan seperti disajikan pada Gambar 1.



Gambar 1. Blok Diagram Penelitian

2.2 Data Penelitian

Penelitian ini menggunakan data hasil perekaman suara gitar yang telah dikonversi ke bentuk digital berformat WAV. Data suara gitar merupakan dua campuran dari 12 *chord* mayor dan 12 *chord* minor (Tabel 1). Perekaman suara gitar dilakukan ditempat yang hening untuk meminimalkan adanya *noise*. Perekaman dilakukan 15 kali untuk setiap dua campuran *chord* sehingga dihasilkan 8640 data suara gitar.

Suara gitar direkam menggunakan ukuran waktu tiga detik dengan *sampling rate* sebesar 11.000 Hz dan *bit rate* sebesar 16 bit. Data suara gitar berupa campuran *chord* C ke *chord* C#, *chord* C ke *chord* Cm dan seterusnya masing-masing sebanyak 15 kali. Untuk setiap campuran digunakan sepuluh kali perekaman untuk data latih dan lima kali perekaman untuk data uji, sehingga terdapat 5760 data latih dan 2880 data uji.

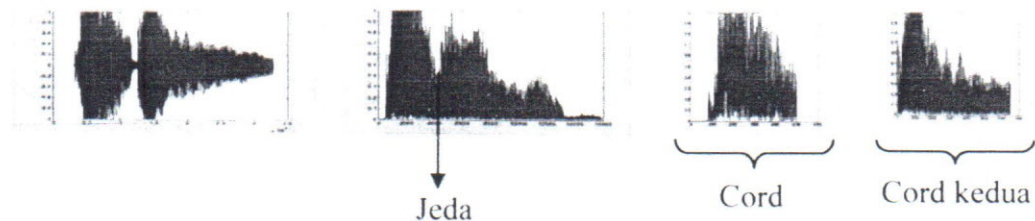
Tabel 1.

Chord Mayor dan Minor sebagai kelas dalam penelitian ini

Cord Dasar	Mayor	Minor
C	C	Cm
C#	C#	C#m
D	D	Dm
D#	D#	D#m
E	E	Em
F	F	Fm
F#	F#	F#m
G	G	Gm
G#	G#	G#m
A	A	Am
A#	A#	A#m
B	B	Bm

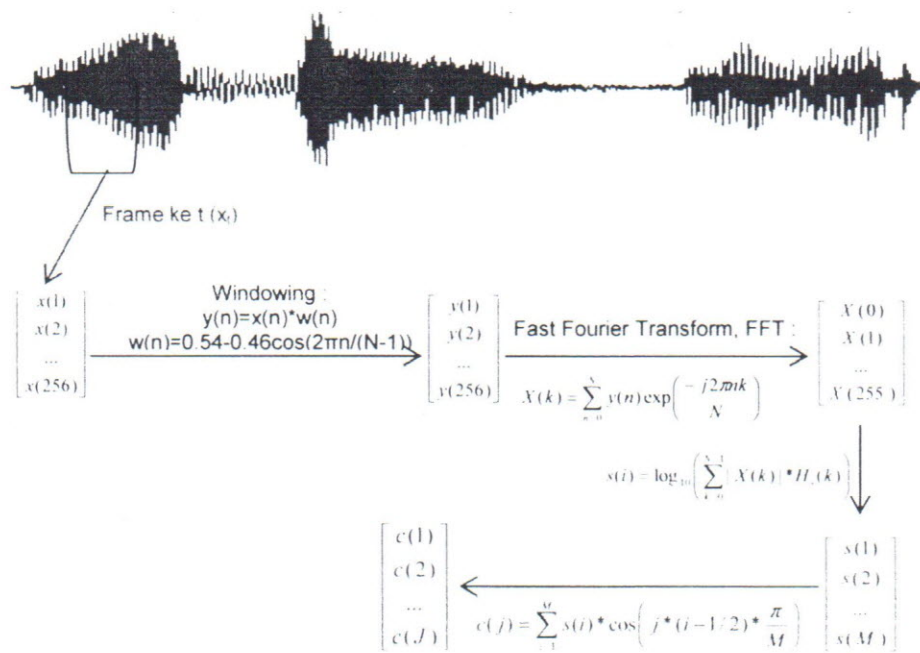
2.3 Pemrosesan Penelitian

Selanjutnya pada setiap data tersebut dilakukan segmentasi sinyal secara manual seperti diperlihatkan pada Gambar 2. Pada tahap ini dilakukan proses segmentasi manual untuk deteksi jeda pada setiap transisi dua *chord*. Data kontinyu yang panjang pada bagian jeda akan dipotong. Data tersebut akan dibuat batas kritis untuk membedakan jeda dan bukan jeda.



Gambar 2. Proses segmentasi dari sinyal suara yang terdiri 2 *chord*

Pada setiap sinyal yang sudah disegmentasi, selanjutnya dilakukan ekstraksi ciri. Ekstraksi ciri merupakan proses untuk menentukan suatu nilai atau vektor yang dapat dipergunakan sebagai pencari objek atau individu (Buono 2009). Pada Penelitian ini, teknik ekstraksi ciri yang dipergunakan adalah MFCC. Tahapan MFCC adalah seperti disajikan pada Gambar 3. Seperti pada Gambar 3 sinyal suara dibaca dari *frame* ke *frame* dengan lebar tertentu yang saling tumpang tindih (*overlap*), kemudian dilakukan *windowing* pada setiap *frame*. Tahap berikutnya, dilakukan transformasi *Fourier* untuk mengubah suara dari domain waktu ke domain frekuensi. Nilai hasil transformasi *Fourier* ini dihitung *spectrum mel* menggunakan sejumlah filter. Koefisien MFCC merupakan hasil transformasi *Cosinus* dari *spectrum mel* tersebut dan dipilih K koefisien. Transformasi *Cosinus* berfungsi untuk mengembalikan domain, dari domain frekuensi ke domain waktu (Ganchev, 2005).



Gambar 3. Tahapan Ekstraksi Sinyal Suara dengan Teknik MFCC (dimodifikasi dari Nilsson dan Ejnarsson, 2002)

Ciri setiap sinyal hasil segmentasi adalah sebuah vektor koefisien MFCC yang merupakan rata-rata dari T vektor koefisien MFCC dari sinyal tersebut. Tahapan selanjutnya adalah memilah data vektor ciri ini menjadi data latih dan data uji untuk mengevaluasi kinerja model RBNN dalam melakukan konversi suara gitar ke cord.

Tahap selanjutnya adalah percobaan menggunakan software Matlab dengan RBNN untuk pengenalan cord dengan menggunakan parameter seperti disajikan pada Tabel 2.

Tabel 2.

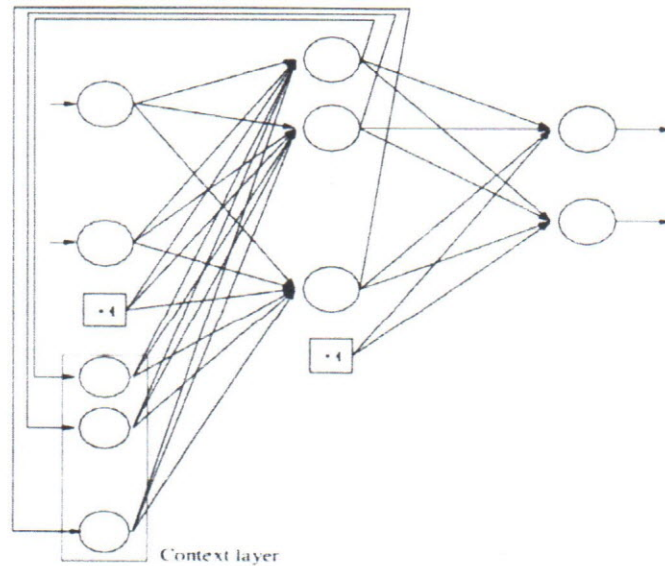
Karakteristik Model Neural Network yang dicobakan

Karakteristik	Spesifikasi
Arsitektur	1 hidden layer
Neuron input	13,26,52 (koefisien hasil ekstraksi ciri)
Neuron output	24 (banyaknya chrod dasar)
Fungsi aktivasi	Sigmoid biner
Inisialisasi Bobot	Nguyen widrow
Hidden Neuron	10, 25, 50, 100
Toleransi Galat	10^{-4}
Learning Rate	0.01
Faktor Naik	1.2
Faktor Turun	0.5
delta0	0.1
Deltamax	50

Struktur dari jaringan resilient neural network adalah seperti disajikan pada Gambar 4 (Engelbreth, 2007). Dalam hal ini ada konteks layer (*context layer*) yang merupakan hal sama dengan hidden layer.

3. Hasil dan Pembahasan

Tabel 3 menunjukkan perbandingan akurasi penggunaan koefisien *cepstral* 13, 26, dan 52 terhadap nilai *hidden neuron* 10, 25, 50, dan 100. Percobaan dengan koefisien *cepstral* 13 dengan nilai *hidden neuron* 10, 25, 50, dan 100, rata-rata akurasi terendah berada pada pemakaian *hidden neuron* 10 yaitu sebesar 47.63% dan tertinggi pada pemakaian *hidden neuron* 100 yaitu sebesar 67.77%.



Gambar 4. Arsitektur Jaringan Syaraf Resilient (Engelbreth, 2007).

Tabel 3.
Perbandingan nilai akurasi koefisien cepstral 13, 26, 52 dengan hidden neuron 10, 25, 50, 100

Koefisien cepstral	Hidden neuron			
	10	25	50	100
13	47.63%	55.65%	62.43%	67.77%
26	74.20%	84.06%	90.10%	93.36%
52	81.31%	90.45%	94.37%	96.88%

Tabel 3 di atas juga menunjukkan bahwa untuk percobaan dengan koefisien cepstral 26, dan 52 dengan nilai hidden neuron 10, 25, 50, dan 100, rata-rata akurasi terendah berada pada pemakaian hidden neuron 10 yaitu sebesar 74.20% dan 81.31%, sedangkan tertinggi pada pemakaian hidden neuron 100 yaitu sebesar 93.36% dan 96.88%.

Pada percobaan ini, nilai hidden neuron berbanding lurus dengan akurasi yang dihasilkan sistem. Pada Tabel 4 penggunaan koefisien cepstral 13, 26 maupun 52, akurasi terendah sama-sama dihasilkan dengan nilai hidden neuron 10 dan akurasi tertinggi dihasilkan dengan nilai hidden neuron 100. Pada percobaan ini nilai hidden neuron berpengaruh terhadap kenaikan akurasi dan juga mempengaruhi waktu pelatihan yang lama. Meskipun nilai hidden neuron tinggi tetapi belum pasti menjamin nilai akurasinya akan tinggi. Jika digunakan nilai hidden neuron lebih dari 100, hal tersebut akan mempengaruhi nilai akurasi tetapi tidak ada peningkatan nilai akurasi yang signifikan. Sebaliknya jika digunakan nilai hidden neuron lebih kecil dari 100, maka nilai akurasinya kecil tetapi waktu pelatihan akan lebih cepat. Gambar 5 merepresentasikan data dari Tabel 5 agar terlihat secara jelas perbandingannya.

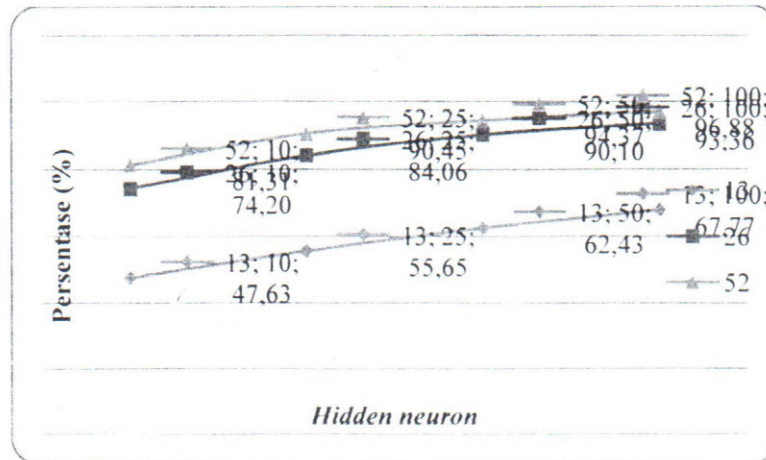
Gambar 5 menggambarkan kecenderungan peningkatan jumlah koefisien cepstral dan hidden neuron terhadap tingkat akurasi ketika nilai overlap dan time frame tetap, yaitu 0,4 dan 100 ms. Tingkat akurasi maksimum yang diperoleh adalah sebesar 96.88% yaitu pada saat nilai koefisien cepstral 52 dan jumlah hidden neuron sebanyak 100, sedangkan tingkat akurasi minimum yang diperoleh adalah sebesar 47.63% ketika nilai koefisien cepstral 13 dan jumlah hidden neuron sebanyak 10.

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat diambil kesimpulan sebagai berikut:

1. Penelitian ini telah berhasil mengembangkan implementasi metode jaringan syaraf tiruan *resilient backpropagation* dan teknik ekstraksi ciri MFCC dalam konversi suara gitar ke cord.
2. Akurasi tertinggi yang dihasilkan sebesar 96,88% dengan penggunaan *time frame* 100ms, *overlap* 40% , koefisien cepstral 52, dan hidden neuron sebanyak 100.
3. Penelitian ini mampu mengidentifikasi campuran nada gitar menggunakan model jaringan syaraf tiruan *resilient backpropagation* dengan penggunaan 52 koefisien

cepstral menghasilkan akurasi yang lebih tinggi dibandingkan dengan 13 dan 26 koefisien *cepstral*.



Gambar 5
Grafik perbandingan akurasi antara koefisien cepstral dengan hidden neuron

Daftar Pustaka

- Buono A. (2009). Representasi Nilai HOS dan Model MFCC Sebagai Ekstraksi Ciri Pada Sistem Identifikasi Pembicara di Lingkungan Ber-Noise Menggunakan HMM.[Disertasi]. Depok: Program Pascasarjana, Universitas Indonesia.
- Engelbrecht, A.P. (2007). Computational Intelligence: An Introduction. John Wiley & Sons, Ltd.
- Ganchev T. (2009). Speaker recognition [disertasi]. Patras (GR): University of Patras.
- M. Nilsson dan M. Ejarsson. (2002). *Speech Recognition using Hidden Markov Model: Performance Evaluation in Noisy Environment*. Master Thesis, Departement of Telecommunications and Signal Processing, Blekinge Institute of Technology
- Susanto, N. (2007). Pengembangan Model Jaringan Syaraf tiruan Resilient Backpropagation Untuk Identifikasi Pembicara Dengan Praproses MFCC [skripsi] Bogor: Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, 2007.
- Wisnudastra E, Buono A. (2010). Pengenalan *Chord* pada Alat Musik Gitar Menggunakan *CodeBook* dengan Teknik Ekstraksi Ciri MFCC. *Jurnal Ilmiah Ilmu Komputer*, Vol 14 No. 1, Mei 2010 : 16 – 21.