

Kendali Pintu Air Otomatis Berbasis *Speech Recognition* Menggunakan Metode MFCC dan Jaringan Syaraf Tiruan

Imam Fadli

Program Studi Magister Ilmu komputer, Program Pascasarjana Universitas Budi Luhur Jakarta
ifad2011@gmail.com

ABSTRACT

In this research, the speech of command for sluice recorded and then the coefficient of signal speech taken using MFCC algorithm. These coefficients becomes features of speech signal and using as input in neural network to be training data. After training, the testing using data test conducted to see how the algorithm of neural network have good working and how the effectiveness of controlling automatic sluice. There is 5 coefficients of the word of "BUKA" (English: "OPEN") and so is word of "TUTUP" (English: "CLOSE"). That 5 coefficients was taken as features then repeated 15 times to be used as training data in neural network. The total number of command speech is 30. The 20 test data for "BUKA" and 20 for "TUTUP" used to test the system in controlling sluice by speech. After the test we get the result the accuracy for "BUKA" is 75% and 55% for "TUTUP". Testing with giving command using "OPEN" and "CLOSE" give us the result by there is no stability in output of the value of neural network (y). From the neural network training, the output of training is 1.01325 for "BUKA" and 0.0930902 for "TUTUP". With using the other word we get the value of y is changeable inconstantly.

Keywords: control system, microcontroller, automatic sluice, MFCC, Neural Network.

I. PENDAHULUAN

Pengontrolan sistem buka-tutup pintu air (*sluice system*) yang dikendalikan menggunakan kontrol suara (*speech recognition*) biasanya mengandalkan modul *speech recognition (SR Module)* yang hubungkan dengan mikrokontroler. Selain itu, penggunaan *Google Speech* juga menjadi pilihan lain untuk mengendalikan *device* tertentu. Sayangnya, penggunaan *SR Module* memiliki kelemahan yakni *high cost* dan penggunaan *Google Speech* juga mensyaratkan kehadiran internet terus menerus. Oleh karena itu perlu dicarikan metode yang *low cost* sekaligus menghilangkan ketergantungan terhadap teknologi internet. Penggunaan ekstraksi sinyal suara MFCC bersama teknologi Jaringan Syaraf Tiruan dipilih untuk mengatasi hal tersebut.

Penelitian di bidang kontrol berbasis suara (*speech recognition*) telah banyak dilakukan di antaranya : Takiguchi, dkk., (2008) dengan gagasan *a new mobile robot with hands-free speech recognition*, Rambabu, dkk., (2011) menerapkan pemroses kata untuk meningkatkan performa desain *recognized words* untuk digunakan sebagai perintah dalam mengontrol pergerakan robot industri menggunakan pendekatan *Hidden Markov Model (HMM)*, Sementara itu, Blessington, dkk., (2012) melakukan kajian penelitian dalam bidang kontrol robot berbasis suara menggunakan modul *speech recognition* yang dihubungkan dengan mikrokontroler untuk menggerakkan aktuator robot. Gosavi, dkk., (2013) memanfaatkan pengkodean pada Matlab yang dihubungkan dengan rangkaian pada robot.

Sedangkan penelitian di bidang kendali pintu otomatis telah dilaporkan oleh A.O. Oke1, dkk., (2009) dengan mendesain sistem kontrol pintu menggunakan *smart card*, Verma & Tripathi, (2010) mengajukan model sistem penguncian pintu menggunakan teknologi RFID (*Radio Frequency Identification*), Bingol, dkk (2014), membangun *prototype smart-home* di mana di dalamnya salah satunya mengontrol sistem buka-tutup pintu menggunakan teknologi PLC dan Internet, dan Ilkyu Ha (2015), mengajukan sistem pengaman pintu yang diintegrasikan dengan konsep Internet of Things (IoT).

II. TINJAUAN PUSTAKA

A. MFCC (*Mel-Frequency Cepstrum Coefficient*)

MFCC didasarkan atas variasi *bandwidth* kritis terhadap frekuensi pada telinga manusia yang merupakan filter yang bekerja secara linier pada frekuensi rendah dan bekerja secara logaritmik pada frekuensi tinggi (Sanjaya, 2014). Filter ini digunakan untuk menangkap karakteristik fonetik penting dari sinyal ucapan. Untuk meniru kondisi telinga, karakteristik ini digambarkan dalam skala mel-frekuensi, yang merupakan frekuensi linier di bawah 1000 Hz dan frekuensi logaritmik di atas 1000 Hz.

Biasanya frekuensi pencuplikan yang digunakan di atas 10000 Hz agar dapat meminimalkan efek *aliasing* pada konversi *analog-digital*. Algoritma MFCC diawali dengan *frame blocking* sinyal suara, *windowing*, FFT, *Mel-Frequency Wrapping*, dan Cepstrum. Hasil akhir dari proses ini adalah *cepstrum coeficient* dalam bentuk *mel-frequency* yang merepresentasikan ciri sinyal suara.

B. Jaringan Syaraf Tiruan *Backpropagation*

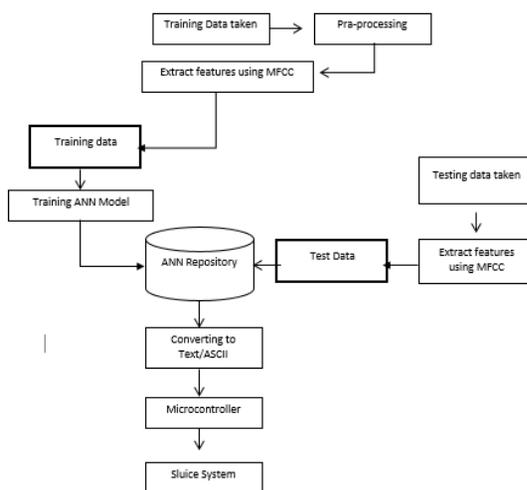
Istilah “propagasi balik” atau “penyiaran kembali” diambil dari cara kerja jaringan ini yaitu bahwa gradien *error* unit-unit tersembunyi diturunkan dari penyiaran kembali *error-error* yang diasosiasikan dengan unit-unit *output*. Hal ini karena nilai target untuk unit-unit tersembunyi tidak diberikan. Metode ini

menurunkan gradien untuk meminimalkan penjumlahan *error* kuadrat (MSE, *Minimum Squared Error*) *output* jaringan. Nama lain dari propagasi balik adalah **aturan delta yang digeneralisasi** (*generalized delta rule*) (Diyah, 2006).

Jaringan syaraf tiruan propagasi balik terdiri dari banyak lapisan (*multi layer*) yang terdiri dari: lapisan *input*, lapisan tersembunyi, dan lapisan *output*.

III. METODE PENELITIAN

Penelitian ini fokus pada penggunaan *speech recognition* untuk mengontrol sistem pintu air otomatis menggunakan metode ekstraksi fitur sinyal suara MFCC dan Jaringan Syaraf Tiruan (JST).



Gambar 1. Metode Penelitian

Gambar 1 menunjukkan metode penelitian yang dimaksud. Berdasarkan gambar 5 dapat dijelaskan langkah-langkahnya. Pertama, untuk mengumpulkan data ucapan perintah direkam melalui *speaker built in* dalam ruang tertutup untuk mengurangi derau. Suara dengan derau rendah akan dipilih menjadi *input* dalam pelatihan di jaringan syaraf tiruan. Setelah direkam, sinyal suara akan langsung diekstrak fiturnya sebagai ciri sinyal suara dengan mengekstrak 5 buah koefisien untuk kata “BUKA” dan 5 koefisien untuk kata

“TUTUP” kemudian diulang dengan perulangan 5 kali, 10 kali, dan 15 kali. Semua jenis perulangan dilakukan uji coba *training* dalam JST. Kami ingin melihat perulangan yang mana yang memiliki performa terbaik dalam JST. Setelah proses *training* selesai dan didapatkan hasil pengujian terbaik, maka dipilih pengujian tersebut menjadi *repository* model jaringan syaraf tiruan untuk mengonfirmasi data uji.

Data uji diambil tanpa melalui pra-proses dan langsung diekstrak koefisiennya dan hasilnya dikonfirmasi atau dicocokkan dengan model JST yang telah dibuat. Keluarannya (Y) diasosiasikan dengan data target yakni 1 untuk kata buka dan 0 untuk kata tutup. Jika Y mendekati 1 maka sistem akan mengubahnya menjadi teks/kode ASCII dan mengirimkannya ke mikrokontroler untuk membuka motor *servo* dan sebaliknya jika mendekati 0 maka *servo* akan menutup.

Data uji diambil tanpa melalui pra-proses dan langsung diekstrak koefisiennya dan hasilnya dikonfirmasi atau dicocokkan dengan model JST yang telah dibuat. Keluarannya (Y) diasosiasikan dengan data target yakni 1 untuk kata buka dan 0 untuk kata tutup. Jika Y mendekati 1 maka sistem akan mengubahnya menjadi teks/kode ASCII dan mengirimkannya ke mikrokontroler untuk membuka motor *servo* dan sebaliknya jika mendekati 0 maka *servo* akan menutup.

A. Data Latih

Data latih terdiri dari 5 koefisien yang merepresentasikan fitur sinyal suara. Semua nilai koefisien ini dihasilkan dari metode MFCC kemudian dilakukan pengulangan dengan 5, 10, dan 15 kali pengulangan. Kami menggunakan kata “BUKA” dan kata “TUTUP”.

B. Data Target

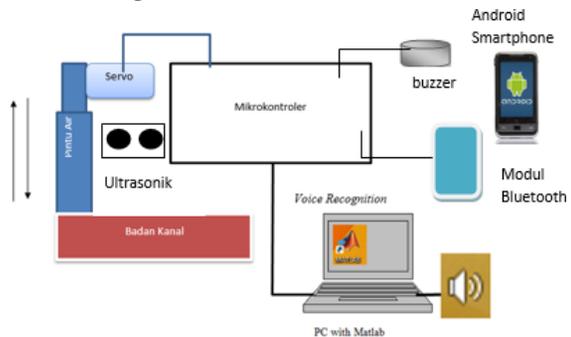
Data target digunakan karena sistem *training* dalam JST menggunakan *supervised learning*. Di sini, kami menggunakan 1 sebagai target kata “BUKA” dan 0 sebagai target kata “TUTUP”.

C. Data Uji

Data uji diambil bagaimana JST memberikan sebuah respons berdasarkan data latih yang telah

digunakan. Di sini, kami menggunakan kata “BUKA” dan “TUTUP”. Kami juga menggunakan data uji selain dua kata tersebut.

D. Rancangan Sistem



Gambar 2. Rancangan Sistem

IV. HASIL DAN PEMBAHASAN

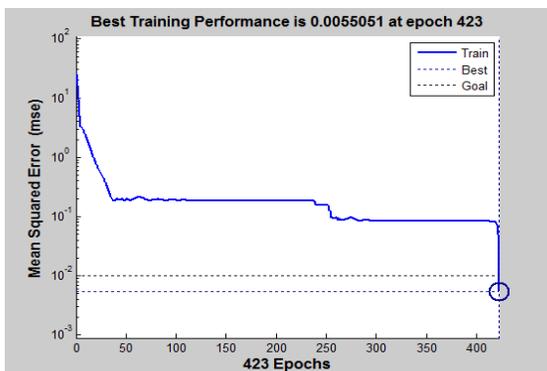
Dengan memvariasikan berbagai jumlah pengulangan data *input* pada penggunaan metode MFCC maka dihasilkan berbagai model pelatihan pada JST. Berbagai model ini akan dipilih model pelatihan yang memiliki performa pengujian terbaik.

A. Pelatihan dengan Data Latih.

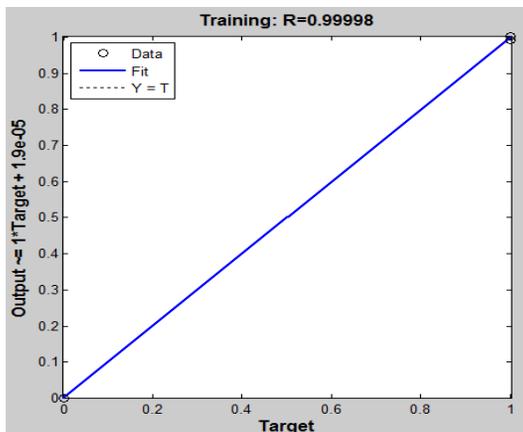
Setelah dilakukan pelatihan dengan 5, 10, dan 15 perulangan, kita dapatkan perbandingan performa JST seperti pada Gambar 3.

Dari Gambar 3 proses *training* berhenti pada *epoch* ke 423 dari 5000 *epoch*. Ini berarti *training* bekerja dengan baik karena berhenti sebelum *epoch* maksimum. Nilai MSE didapatkan 0,0055050.

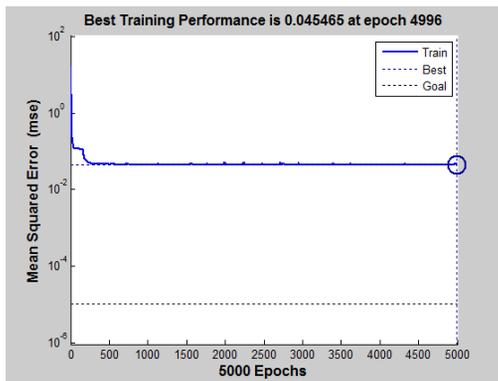
Proses pengujian BP-ANN menggunakan data 5 ciri menghasilkan nilai korelasi R (Gambar 4) yang cukup tinggi yakni 0,99998. Nilai akurasi pengujian dapat dilakukan dengan mengakarkan nilai R sehingga diperoleh 0,9998 atau 99,9% (sangat akurat).



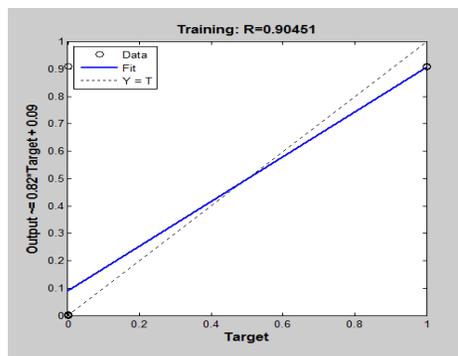
Gambar 3. Best Training Performance



Gambar 4. Nilai Korelasi R

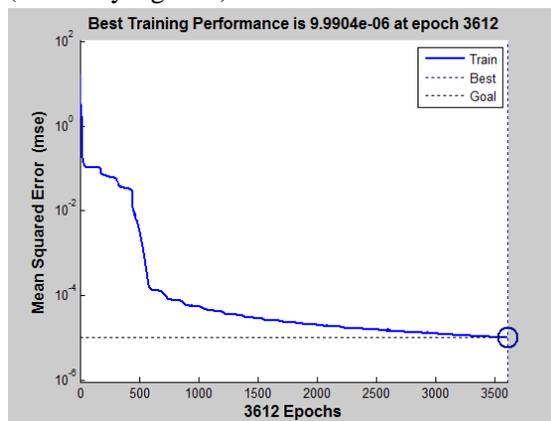


Gambar 5. Pengujian dengan 10 kali pengulangan



Gambar 6. Nilai Korelasi R pada pengujian 10 kali pengulangan

Dari grafik di atas dapat dibaca bahwa pelatihan pada jaringan syaraf tiruan menghasilkan performa terbaik yakni berhenti pada epoch 4996 atau mendekati maksimum epoch yakni 5000. Selanjutnya untuk pembacaan regresi, nampak grafik regresi pada gambar di bawah ini di mana nilai output terjadi pada $Y = 0,82 * T + 0,09$. Pada Gambar 6 terlihat bahwa selama masa *training output* yang dihasilkan oleh program memiliki korelasi yang cukup kuat dengan target. Seperti yang terlihat, koefisien korelasi yang dihasilkan yaitu sebesar 0.90451 (korelasi yang baik).

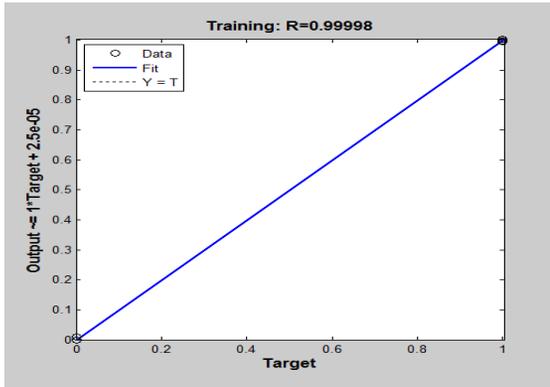


Gambar 7. Performa pelatihan dengan 15 pengulangan

Pada pengujian ke-3 dengan pola 15 kali pengulangan, terjadi penurunan nilai error yang baik yakni menjadi $9,99 \times 10^{-6}$. Nilai MSE ini sangat kecil sehingga sangat baik untuk dijadikan referensi proses pengujian data uji. Untuk nilai

korelasi juga menghasilkan nilai R yang sangat baik yakni 0,9998 seperti yang ditunjukkan oleh Gambar 8.

Dari ketiga pengujian di atas maka disimpulkan bahwa pengujian terbaik terjadi pada pengujian ke-3 yakni dengan pola 15 kali pengulangan. Dengan demikian, maka pengujian ke-3 ini akan diambil sebagai *repository* dalam instrumen dan nantinya akan menjadi acuan dalam menguji data uji (Tabel 1).



Gambar 8. Nilai R dengan 15 kali perulangan

Table 1. Perbandingan hasil pelatihan

Pengujian Ke-	Jumlah Pengulangan	Nilai Error	Akurasi
I	5	0,005505	99,8%
II	10	0,45455	95,1%
III	15	$9,9 \times 10^{-6}$	99,9%

Akurasi dihasilkan dari R^2 . Berdasarkan perbandingan seperti pada Tabel 1, kami memilih pengujian ke-3 dengan 15 kali pengulangan sebagai model JST. Hasil ini akan menguji data uji selanjutnya.

B. Pengujian dengan Data Uji

1) Pengujian dengan kata “BUKA” dan “TUTUP”.

Pengujian dengan 20 kata BUKA dan 20 kata TUTUP memberikan hasil 75% nilai akurasi untuk kata BUKA dan 55% untuk kata TUTUP.

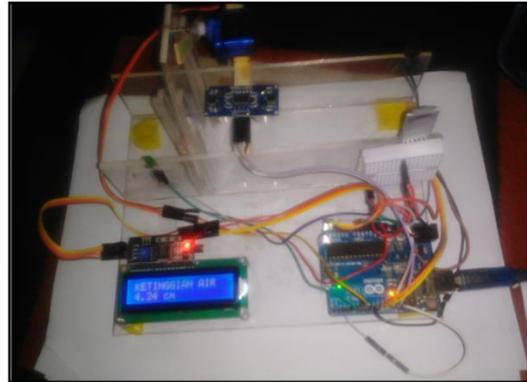
2) Pengujian dengan kata lain

Pengujian dengan kata “OPEN” dan “CLOSE” memberikan nilai *output* JST (Y) menjadi tidak stabil. Pada JST, *output* Y adalah

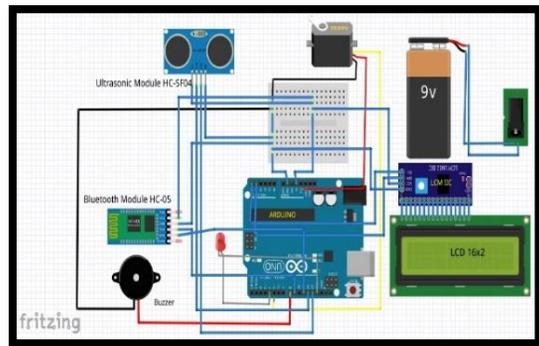
1,01325 untuk kata BUKA dan 0,0930902 untuk kata TUTUP. Dengan menggunakan kata lain didapatkan nilai Y berubah-ubah/tidak konstan untuk masing-masing kata tersebut.

3) Pengujian Fungsionalitas Sistem

Tujuan dari pengujian ini adalah untuk melihat sejauh mana komponen masing-masing sistem bekerja dengan baik. Setelah mendesain dan membuat sistem dengan melakukan *coding* dalam lingkungan pengembangan Matlab dan Arduino IDE, sistem bekerja dengan baik. System yang dibangun dapat dilihat pada Gambar 9 dan 10.



Gambar 9. Prototipe Sistem



Gambar 10. Rangkaian elektronik sistem

V. KESIMPULAN

Prototype system pintu air dapat dibangun dengan menggunakan komponen utama motor *servo* dan mikrokontroler. Perintah dengan metode *Speech Recognition* MFCC dan Jaringan Syaraf Tiruan memiliki kemampuan untuk

melatih ciri suara dengan akurasi yang baik. Untuk penelitian selanjutnya, penggunaan metode *Speech Recognition* yang lain dengan pra-proses yang baik dapat menjadi rekomendasi untuk meminimalkan derau sinyal suara sehingga akurasi menjadi meningkat.

REFERENSI

- Blessington, Praveen., et al., (201),. *Accoustic Controlled Robotic Vehicle*, International journal of Emerging Technology and Advanced Engineering, Vol. 2, Issue 4, April 2012.
- Boerker, Kunal., Gaikwad, Rohan., (2014), *Wireless Controlled Surveillance Robot*, Journal of IJARCSMS Vol. 2 Issue 2, February, 2014.
- Budiharto, Widodo., Suhartono, Derwin., (2014), *Artificial Intelligence: Konsep dan Penerapannya*. Penerbit Andi, Yogyakarta.
- Gosavi, Shradda D., Khot, Uday Pandit., Shah, Seema., (2013), *Speech Recognition for Robotic Control*, Journal of Engineering Research and Applications.
- Gaikwad, Santosh K., Gawali, Bharti W. Yannawar, Pravin., (2010), *A Review on Speech Recognition Technique*, International Journal of Computer Application. Vol. 10 No. 3, November, 2010
- Heriyanto. (2015), *Analisis Deteksi Huruf Hijaiyah Melalui Voice Recognition Menggunakan Kombinasi Energy*, Jurnal Telematika, UPN "Veteran" Vol. 12 No 01, Januari 2015, pp 11-12. Yogyakarta
- Jain, Rachna. Saxena, S.K., (2011), *Voice Automated Mobile Robot*, International Journal of Computer Application. Vol. 16 No. 2, February, 2011.
- Kannan, K., Selvakumar, J., (2015), *Arduino Based Voice Controlled Robot*. International Research Journal of Engineering and Technology (IRJET). Volume: 02 Issue: 01, Mar-2015.
- Kumar, Anup. Chauhaan, Ranjeeta., (2014), *Voice Controlled Robot*, International Journal of Innovative Research in Technology (IJIRT). Vol 1 Issue 11
- Pushpavathi, M., Giribabu, P., Krishna, K. Mohan, (2015), *Motorized PC Control: Pick and Place Robot Using Wireless Technology*, Journal of IRJET. Vol. 02 Issue 05, August 2015.
- Puspitaningrum, Diyah. (2006), *Jaringan Saraf Tiruan*. Penerbit Andi. Yogyakarta.
- Rambabu, D., Raju, R. Naga., Venkatesh, B., (2011), *Speech Recognition of Industrial Robot*, International Journal of Computational Mathematical Ideas, Vol. 3, No. 02, pp 92-98.
- Ronando, Elsen., Irawan, M. Isa., (2012), *Pengenalan Ucapan Kata sebagai pengendali Gerakan Robot lengan Secara real-Time dengan Metode LPC-Neuro Fuzzy*, Jurnal Sains dan Seni ITS Vol. 1 No. 1, Sept, 2012.
- Saini, Preeti., Kaur, Parneet., (2013), *Automatic Speech Recognition: A Review*, International Journal of Engineering and Technology, Vol. 4 Issue 2 – 2013.
- Sanjaya W.S, Mada., Salleh, Zabidin, (2014), *Implementasi Pengenalan Pola Suara Menggunakan MFCC dan ANFIS Sebagai Kontrol Lampu Otomatis*, Al-Hazen Journal of Physics. Vol. 1 No. 1, Desember, 2014.
- Setiawan, Angga., Hidayatno, Achmad., Isnanto, R. Rizal., (2011). *Aplikasi Pengenalan Ucapan dengan Ekstraksi Mel-Frequency Coefficient (MFCC) melalui Jaringan Syaraf Tiruan (JST) Learning Vector Quantization (LVQ) Untuk Mengoperasikan Komputer*, Jurnal Transmisi UNDIP Vol. 13 No. 3, 2011.
- Swamy, Suma., Ramakrishnan, K.V. (2013), *An Efficient Speech Recognition System*, Computer Science & Engineering: International Journal (CSEIJ). Vol. 3 No. 4, August, 2013.
- Takeguchi, Tetsuya., et al., (2008), *Human-Robot Interface Using System Request Utterance Detection Based on Accoustic Features*, International Journal of Hybrid Information Technology Vol. 1 No. 3, July, 2008