

Research Article

Pengenalan Suara sebagai Pengendali Mobile Robot dengan Metode Adaptive Neuro-Fuzzy Inference System

Voice Recognition as a Mobile Robot Controller with the Adaptive Neuro-Fuzzy Inference System Method

M. Agung Suhendra^{1*}, Timbo Faritcan Parlaungan², Tedi Sumardi¹

¹ Program Studi Fisika, Fakultas Sains, Universitas Mandiri, Subang 41211 Indonesia

² Program Studi Teknik Informatika, Fakultas Teknik, Universitas Mandiri, Subang 41211 Indonesia

*Corresponding Author: agung@universitasm mandiri.ac.id

ARTICLE INFO	ABSTRACT
<p><i>Article history:</i> Submitted : 20 January 2023 Revised : 29 January 2023 Accepted : 2 February 2023 Published : 26 February 2023</p> <hr/> <p>Keywords: Speech Recognition Wavelet Transformation Anfis Simple Robot Ultrasonic Sensor</p>	<p>Voice recognition or speech recognition is a biometric technology that has very wide applications, one of which is for simple robot motion control. There are three stages in this research, namely data acquisition, feature extraction, and data classification. For feature extraction, the wavelet transform method is used which can analyze non-stationary and non-linear signals, while for data classification, the Adaptive Neuro-Fuzzy Inference System (Anfis) method is used. The result of data classification is 92.25% and 7.75% error. So, based on the results of the classification accuracy, the robot can be moved via voice commands and to anticipate the error value, the ultrasonic sensor feature is added to the robot as an alternative control.</p>
<p><i>Keywords:</i> Speech Recognition Transformasi Wavelet Anfis Robot Sederhana Sensor Ultrasonic</p>	<p>ABSTRAK</p> <p>Pengenalan suara atau <i>speech recognition</i> merupakan teknologi biometrika yang memiliki aplikasi yang sangat luas, salah satunya yaitu untuk kendali gerak robot sederhana. Ada tiga tahapan dalam penelitian ini yaitu akuisisi data, ekstraksi ciri dan klasifikasi data. Untuk ekstraksi ciri digunakan metode transformasi wavelet di mana dapat menganalisis sinyal <i>non-stasioner</i> dan <i>non-linear</i> sedangkan untuk klasifikasi data digunakan metode <i>Adaptive Neuro-Fuzzy Inference System</i> (Anfis). Adapun hasil klasifikasi data sebesar 92,25 % dan error 7,75 %. Maka berdasarkan hasil akurasi klasifikasi tersebut, robot dapat digerakkan melalui perintah suara dan untuk mengantisipasi nilai error ditambahkan fitur sensor ultrasonik pada robot sebagai alternatif kontrol.</p>

Pendahuluan

Teknologi pengenalan suara atau biasa disebut *speech recognition* merupakan salah satu teknologi biometrika yang tidak memerlukan biaya besar serta peralatan khusus. Pada dasarnya setiap manusia memiliki sesuatu yang unik/khas yang hanya dimiliki oleh dirinya sendiri. Suara merupakan salah satu dari bagian tubuh manusia yang unik dan dapat dibedakan dengan mudah. *Speech recognition* merupakan suatu proses penggalan informasi karakteristik ucapan dari suara seseorang, kemudian dioperasikan melalui komputer dan mengenali isi ucapan tersebut. banyak disiplin ilmu yang terlibat dalam teknologi ini seperti pemrosesan sinyal, teori informasi, fonetik, linguistik, kecerdasan buatan, dan yang lainnya.

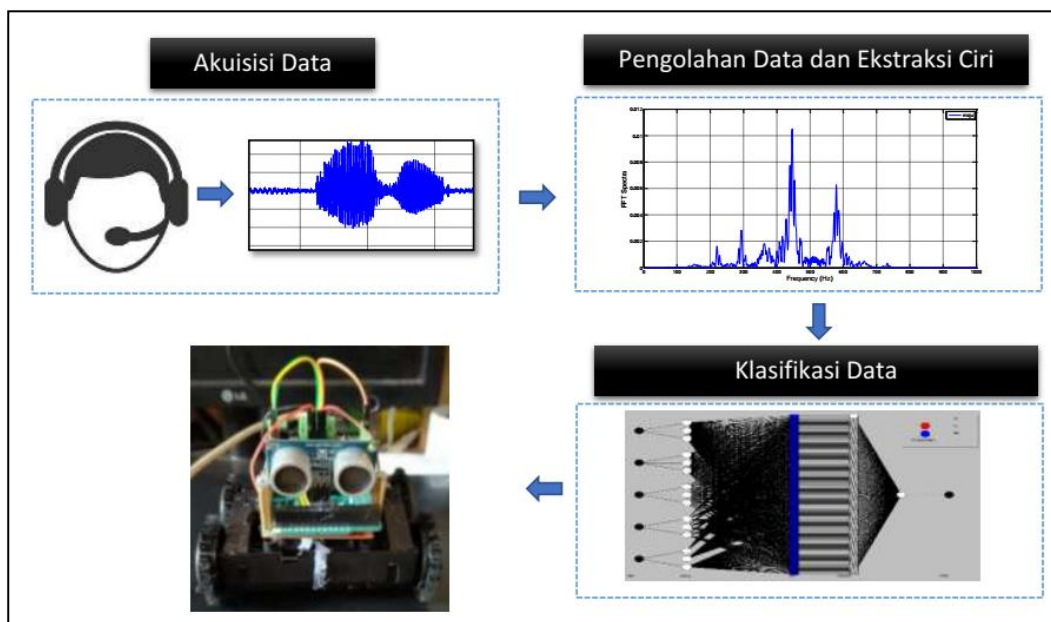
Adapun aplikasinya sangat luas diantaranya, digunakan untuk keamanan,¹ kendali robot dan aplikasi-aplikasi lainnya.²

Pada penelitian ini, penulis akan menjelaskan penelitian terkait implementasi teknologi pengenalan suara untuk kendali robot sederhana dengan gerakan maju, mundur, kanan dan kiri dengan menambahkan sensor *ultrasonic* sebagai kendali alternatif bila terjadi kesalahan dalam hasil perintah dari pengolahan data suara yang diproses. Secara prinsip, pengolahan data suara ini ada beberapa tahapan yaitu akuisisi data suara, ekstraksi ciri sinyal suara, pengenalan pola suara dan klasifikasi data. Untuk ekstraksi ciri sinyal suara digunakan metode transformasi wavelet di mana akan membuang sinyal *noise* dan mendapatkan karakteristik ciri suara. Terakhir, data yang didapat akan diklasifikasikan dengan menggunakan *Adaptive Neuro-Fuzzy Inference System* (Anfis).

Metodologi

Akuisisi Data

Pada penelitian ini, sinyal suara didapat dari rekaman seseorang dengan mengucapkan empat kata yaitu; maju, mundur, kanan dan kiri. Perekaman ini dilakukan sebanyak 10 kali pengucapan pada tiap-tiap katanya dengan lama waktu pengucapan selama 2 detik dengan frekuensi 1000 Hz. Untuk memperjelas hasil rekamannya kita menggunakan *headset* pada saat perekaman. Adapun alur proses penelitiannya terlihat pada Gambar 1.



Gambar 1. Skema penelitian

Pengolahan Sinyal dan Ekstraksi Ciri

Data hasil rekaman masih terdapat *noise* sehingga harus dilakukan filter dengan menggunakan metode wavelet. Secara teori, wavelet ini merupakan transformasi matematika yang digunakan untuk menganalisis sinyal bergerak. Sinyal bergerak ini dianalisis untuk diperoleh informasi spektrum frekuensi dan waktunya secara bersamaan. Selain itu, Teori wavelet menyajikan kerangka kerja gabungan untuk sejumlah teknik yang telah dikembangkan secara independen untuk pemrosesan sinyal berganda. Khususnya, transformasi wavelet menjadi perhatian untuk analisis sinyal *non-stasioner* dan *non-linear* serta menganalisis sinyal

dalam domain waktu, dekomposisi, dan sinyal kompresi. Untuk mengidentifikasi karakteristik sinyal suara diperlukan transformasi wavelet yang didalamnya terdapat fungsi wavelet sebagaimana terlihat pada persamaan berikut.³⁻⁵

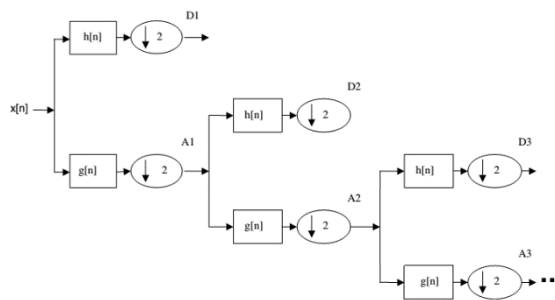
$$\psi_{a,b} = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \tag{1}$$

Di mana $a, b \in \mathbb{R}$, $a > 0$, and \mathbb{R} adalah *wavelet space*. salah satu fitur dari transformasi wavelet ini adalah transformasi diskrit wavelet atau *Discrete wavelet transform* (DWT) di mana fungsinya adalah untuk menganalisa sinyal terdapat 2 koefisien yaitu aproksimasi dan dekomposisi sebagaimana secara teori dapat dijelaskan melalui persamaan matematika.

$$a_{(i)} = x(k) * \varphi_{i,l}(k), \tag{2}$$

$$d_{(i)} = x(k) * \psi_{i,l}(k), \tag{3}$$

Di mana $a_{(i)}(l)$ and $d_{(i)}(l)$ adalah koefisien aproksimasi dan dekomposisi. Untuk lebih jelasnya terlihat pada *flowchart* yang ada pada Gambar 2.⁴



Gambar 2. Struktur transformasi diskrit wavelet.

Klasifikasi Data

Metode Anfis selain digunakan sebagai tahapan untuk klasifikasi, tapi juga dapat digunakan untuk aplikasi optimasi, pemodelan, prediksi, dan deteksi sinyal. Secara teori, Anfis ini memiliki kesamaan dengan *fuzzy*, akan tetapi ada kelebihan pada metode ini karena memiliki algoritma pembelajaran terhadap sekumpulan data serta dapat beradaptasi menyesuaikan pada sekumpulan data tersebut.

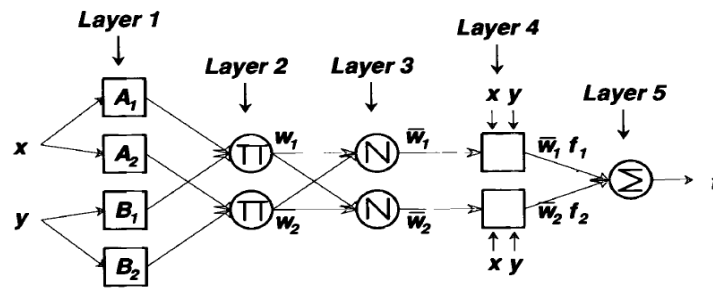
Pada Gambar 3 terlihat skema arsitektur Anfis itu ada dua input x_1, x_2 dan satu output y . Ada dua aturan pada basis aturan model sugeno dengan persamaan matematika sebagai berikut.

$$\text{If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } B_1 \text{ then } y_1 = c_{11}x_1 + c_{12}x_2 + c_{10} \tag{4}$$

$$\text{If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } B_1 \text{ then } y_2 = c_{21}x_1 + c_{22}x_2 + c_{20} \tag{5}$$

Jika α predikat untuk aturan ke dua aturan adalah w_1 dan w_2 , maka dapat dihitung rata-rata terbobot.⁶⁻⁷

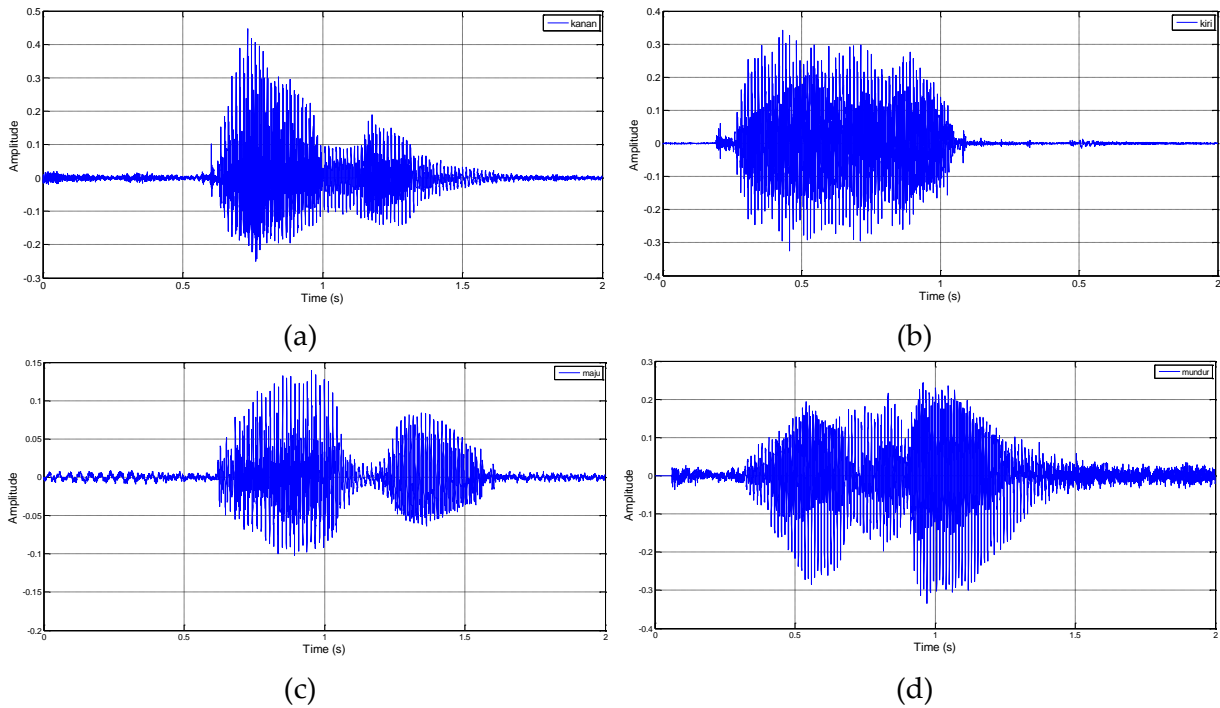
$$y = \frac{w_1 y_1 + w_2 y_2}{w_1 + w_2} = \bar{w}_1 y_1 + \bar{w}_2 y_2 \tag{6}$$



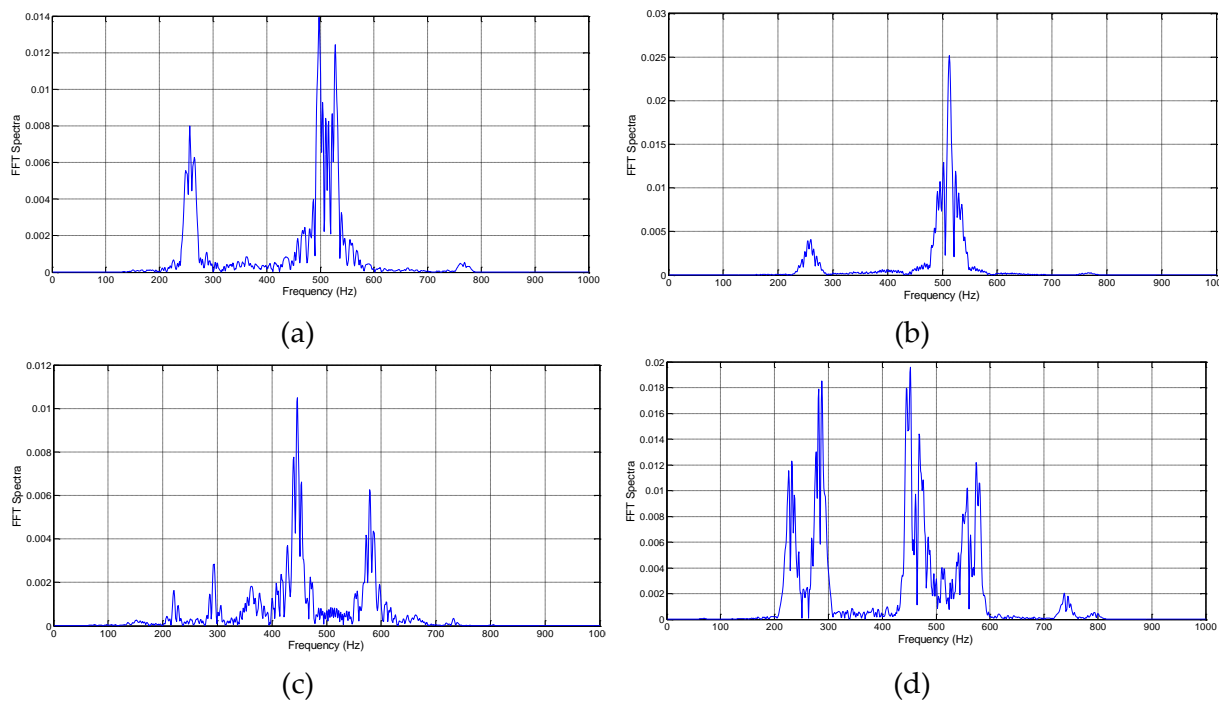
Gambar 3. Contoh struktur ANFIS Roger Jang

Hasil dan Diskusi

Hasil rekaman suara terlihat pada Gambar 4. adalah salah satu data dari 10 data pengulangan pada setiap suara. Pada grafik tersebut telah dilakukan normalisasi data terlebih dahulu. Secara umum, terlihat perbedaan pola sinyal yang berbeda dari setiap ucapan suara yang berbeda-beda yaitu kanan, kiri, maju dan mundur. Selanjutnya dilakukan tahapan ekstraksi ciri dengan menggunakan transformasi wavelet untuk membuang sinyal *noise*, juga untuk mengubah sinyal dari domain waktu ke domain frekuensi. Hal ini dilakukan untuk melihat frekuensi spektrum yang dihasilkan dari setiap sinyal suara. Adapun hasil transformasi wavelet terlihat pada Gambar 5. Di mana pada tahap ekstraksi ciri pun mempunyai frekuensi spektrum yang berbeda-beda. Seperti pada Gambar 5.a. dan 5.b antara kanan dan kiri pola spektrumnya mempunyai kemiripan akan tetapi pada rentang nilai 200-300 Hz spektrum yang kanan lebih tinggi.



Gambar 4. Grafik sinyal suara (a) kanan, (b) kiri, (c) maju dan (d) mundur.



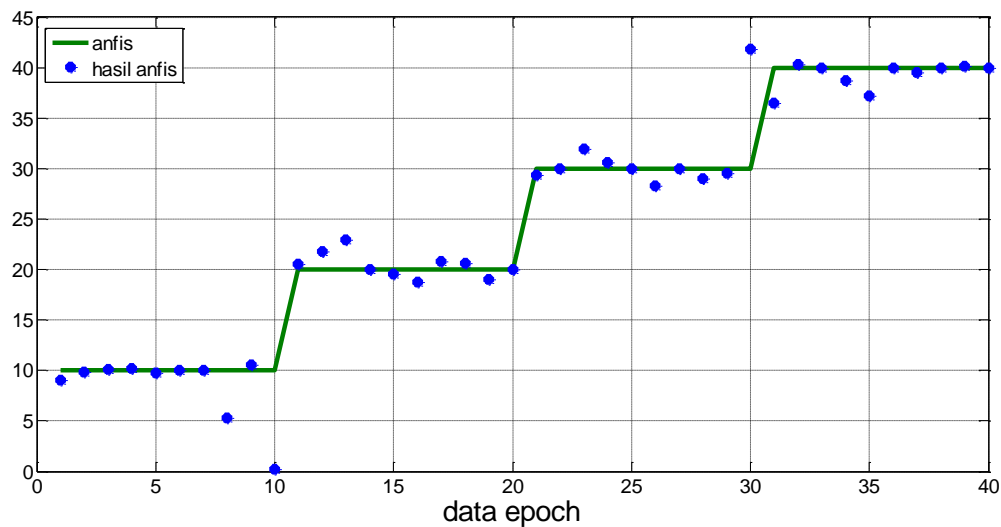
Gambar 5. Grafik frekuensi spektrum suara (a) kanan, (b) kiri, (c) maju dan (d) mundur.

Berdasarkan pada Gambar 1, setelah melakukan tahapan ekstraksi ciri maka dilakukan tahap klasifikasi data menggunakan anfis, di mana sebagai input data yang diklasifikasikan adalah frekuensi spektrum dari setiap sinyal yang kemudian akan dijadikan sebagai data latih untuk pengujian hasil klasifikasi. Hasil pengujian klasifikasi terlihat pada Tabel 1. Di mana untuk data suara mundur memiliki hasil klasifikasi terbesar yaitu 97 % dan terendah untuk data suara kanan 83 %. Adapun hasil klasifikasinya secara rinci terlihat pula pada Gambar 6. Di mana garis warna hijau menjelaskan tentang nilai klasifikasinya dan warna biru menjelaskan hasil klasifikasinya. Secara keseluruhan, nilai akurasi klasifikasi anfis adalah 92,25 % dan nilai error sebesar 7,75 %. Hal ini disebabkan tidak memperhatikan saat tahapan akuisisi data atau saat pengambilan data suara juga ruangan yang dipakai saat pengambilan data.

Tabel 1.

Hasil pengujian klasifikasi data menggunakan Anfis

Suara	Jumlah Pengujian	Keberhasilan (%)
Kanan	10	83 %
Kiri	10	95 %
Maju	10	94 %
Mundur	10	97 %
Rata-rata		92,25 %



Gambar 6. Hasil pengujian klasifikasi anfis.

Simpulan

Penerapan metode transformasi wavelet sebagai fitur untuk mengekstraksi ciri dan metode anfis untuk mengklasifikasikan data suara dengan empat variasi suara yaitu kanan, kiri, maju dan mundur bisa digunakan untuk aplikasi pengenalan suara atau *speech recognition* yang efektif untuk menggerakkan kendali robot sederhana dengan hasil akurasi klasifikasi 92,25 %. Untuk menutupi nilai kesalahan atau error sebesar 7,75 % digunakan sensor ultrasonik sebagai alternatif kontrol.

Pendanaan

Penelitian ini tidak menerima pendanaan dari sumber manapun.

Ucapan Terima Kasih

Kami ucapkan terima kasih kepada program studi fisika, program studi informatika dan fakultas sains yang mendukung dalam penulisan jurnal ini.

Konflik Kepentingan

Penulis dalam penulisan jurnal ini menyatakan tidak ada konflik kepentingan.

Kontribusi Penulis

Muhamad Agung Suhendra: Konsep dan Metodologi, Analisis data. **Timbo Faritcan:** Penulisan - *review*. **Tedi Sumardi:** Penulisan – *review*.

Daftar Pustaka

1. Kabari LG, Chigoziri MB. Speech recognition using MATLAB and cross-correlation technique. *Eur J Eng Res Sci.* 2019;4(8):1-3. doi:10.24018/ejers.2019.4.8.1437
2. Deuerlein C, Langer M, Seßner J, Heß P, Franke J. Human-robot-interaction using cloud-based speech recognition systems. In: *Procedia CIRP.* Vol 97. Elsevier B.V.; 2020:130-135. doi:10.1016/j.procir.2020.05.214
3. Procházka A, Kukul J, Vyšata O. Wavelet transform use for feature extraction and EEG signal segments classification. In: *2008 3rd International Symposium on Communications, Control, and Signal Processing, ISCCSP 2008.* ; 2008:719-722. doi:10.1109/ISCCSP.2008.4537317

4. Samar VJ, Bopardikar A, Rao R, Swartz K. Wavelet analysis of neuroelectric waveforms: A conceptual tutorial. *Brain Lang.* 1999;66(1):7-60. doi:10.1006/brln.1998.2024
5. Kusumandari DE, Suhendra MA, Rizqyawan MI, et al. Effect of Sundanese music on daytime sleep quality based on EEG signal. *AIP Conf Proc.* 2019;2193(December). doi:10.1063/1.5139394
6. Turnip A, Suhendra MA, Mada Sanjaya WS. Brain-controlled wheelchair based EEG-SSVEP signals classified by nonlinear adaptive filter. *IEEE Int Conf Rehabil Robot.* 2015;2015-September:905-908. doi:10.1109/ICORR.2015.7281318
7. Turnip A, Amri MF, Suhendra MA, Kusumandari DE. Lie detection based EEG-P300 signal classified by ANFIS method. *J Telecommun Electron Comput Eng.* 2017;9(1-5):107-110.