TIME IN PHYSICS
Journal for Theoretical, Instrumentation, Material-Molecular, and Education Physics
Volume 2, Issue 2, Pages 77-84 | e-ISSN: 2986-27601
DOI: 10.11594/timeinphys.2024.v2i2p77-84
ejournal.universitasmandiri.ac.id/index.php/timeinphys

Research Article

Implementation of Adaptive Neural Fuzzy Inference Systems (Anfis) For Speech Recognition Applications In Smart Home Control

Roni Permana[1,*], Mada Sanjaya WS[2], Hasniah Aliah[2]

[1] Department of Primary Teacher Education, Faculty of Teacher Training and Education, Universitas Mandiri, Subang 41211, Indonesia
[2] Department of Physics, Faculty of Science and Technology, UIN Sunan Gunung Djati, Bandung, Indonesia
*Corresponding Author: r.permana@universitasmandiri.ac.id

ARTICLE INFO

Article history:
Received : 30 November 2024
Revised : 1 December 2024
Accepted : 3 December 2024
Available Online : 6 December 2024

Keywords:
MFCC
ANFIS
Signal Processing
Digital Signal Processing
Control Systems

ABSTRACT

Signal Processing is signal processing that is related to the presentation, transformation, and manipulation of signal content and information. Digital Signal Processing is signal processing that is related to the presentation, transformation, and manipulation of signal content and information in digital form. The speech control system is very efficient. Speech signals are signals that change over time at a relatively slow speed. If observed at short intervals (between 5 and 100 miles per second), the practical characteristics are constant, but if observed at longer intervals, the characteristics appear to change according to the sentences spoken. This study uses the signal pattern recognition method with the MFCC and ANFIS methods as learning. The performance results of the entire system obtained an accuracy value with 6 feature extractions in 2 respondents, namely 65% -72.5% and the smarthome control system worked well.

Keywords:
MFCC
ANFIS
Pengolahan Sinyal
Pengolahan Sinyal Digital Sistem
Kedali

ABSTRAK

Pengolahan Sinyal adalah pemrosesan sinyal yang mempunyai kaitan dengan penyajian, perubahan bentuk, dan manipulasi dari isi sinyal dan informasi. Pengolahan Sinyal Digital adalah pemrosesan sinyal yang mempunyai kaitan dengan penyajian, perubahan bentuk, dan manipulasi dari isi sinyal dan informasi dalam bentuk digital. Sistem kendali dengan ucapan sangatlah efisien. Sinyal ucapan merupakan sinyal yang berubah terhadap waktu dengan kecepatan yang relatip lambat. Jika diamati pada selang waktu yang pendek (antara 5 sampai dengan 100 mil perdetik), karakteristik praktis bersifat tetap, tetapi jika diamati pada selang waktu yang lebih panjang karakteristiknya terlihat berubah-ubah sesuai dengan kalimat yang diucapkan. Penelitian ini menggunakan metode pengenalan pola sinyal dengan metode MFCC dan ANFIS sebagai pembelajarannya. Hasil kinerja dari keseluruhan sistem didapat nilai akurasi dengan 6 ekstraksi ciri pada 2 responden yaitu 65%-72,5% dan sistem kendali smarthome bekerja dengan baik.

Introduction

With the advancement of computer technology, intelligent robots began to grow rapidly around the 1950s. Improvements in computer processing power and the reduction in the size of physical components have led to the creation of robots intelligent enough to perform tasks typically carried out by humans.

The control system is a crucial and integral part of robotic systems. An automatic control system can manage mechanical systems without supervision and repeated data input, as it is equipped with a series of processors to provide automatic control commands. The control system or mobile robot control, in general, still commonly uses remote control.[1]

Voice-based control systems are often an alternative in research, as the control process can be easy and efficient for individuals with physical limitations. However, it is not as simple as it may seem, as the voice recognition process requires machine learning methods to extract sound patterns or voice features by studying the characteristics of previously recorded sounds.[2]

Speech recognition systems focus on extracting features from certain parts of the message information, which contains the spoken text. This spoken text includes linguistic units known as phonemes, representing the keywords of a message or the complete message itself. Meanwhile, speaker identification emphasizes analyzing the voice of the speaker, ensuring that each word spoken by the speaker is distinguishable.[3]

Human voice signals exhibit a very high level of variability. A voice signal produced by different speakers generates distinct sound patterns. One approach to handling this variability is by using the ANFIS pattern recognition method, which operates similarly to a standard FIS (Fuzzy Inference System) but differs in its calculation (algorithm).[4]

Advances in technology within the field of digital signal processing have significantly impacted human life and continue to evolve. Digital signal processing, or digital signal recognition, involves transforming or analyzing information represented as a sequence of discrete numbers.[5]

Research on speech recognition has been extensively conducted. By applying the ANFIS method to cases involving word recognition or specific words in the Indonesian language, the results obtained align with expectations. This research leverages the method for applications in smart home control systems.

Methodology

Data Acquisition

At this stage, a speech recognition application is designed, consisting of two parts: speech recognition and data transmission to the control system, as illustrated in Figure 1. The input data takes the form of voice signals, which are then extracted and classified.

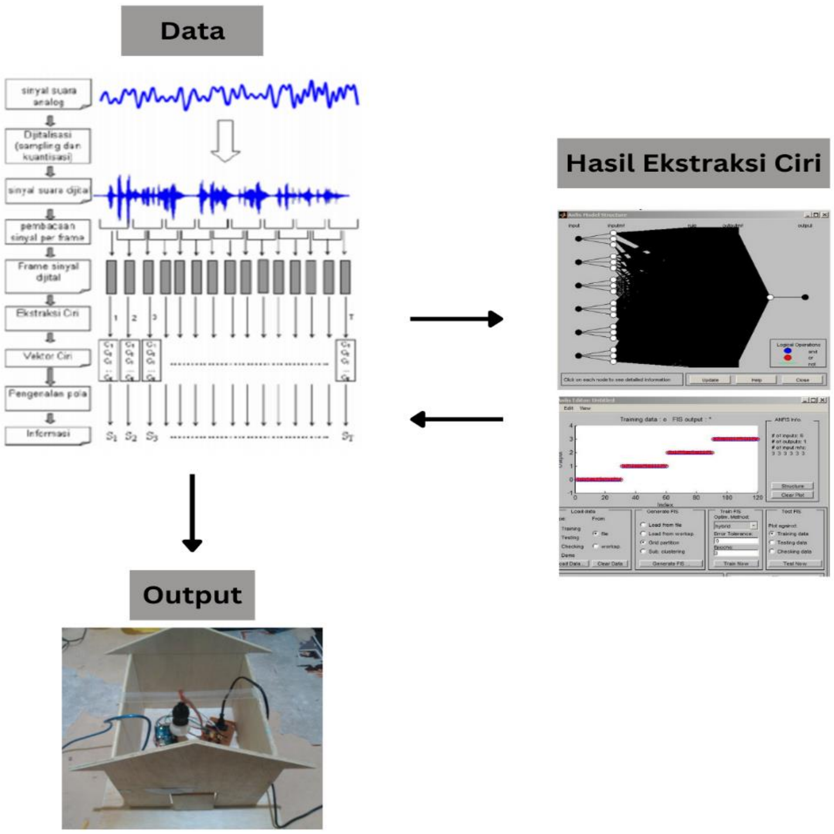

Figure 1. Research Flow Diagram

Feature Extraction

Feature extraction is performed using the MFCC (Mel-Frequency Cepstral Coefficients) method. MFCC is an algorithm that transforms voice characteristics into coefficients representing the voice's features, based on the calculation of frequencies audible to humans or audio sounds. This method is highly effective for applications in any field related to voice processing.[6-8]

The MFCC process involves several stages: Frame Blocking, Windowing, FFT, Mel-Frequency Wrapping, and Cepstrum. The stages of the Mel-Frequency Cepstral Coefficients method are illustrated in Figure 2.

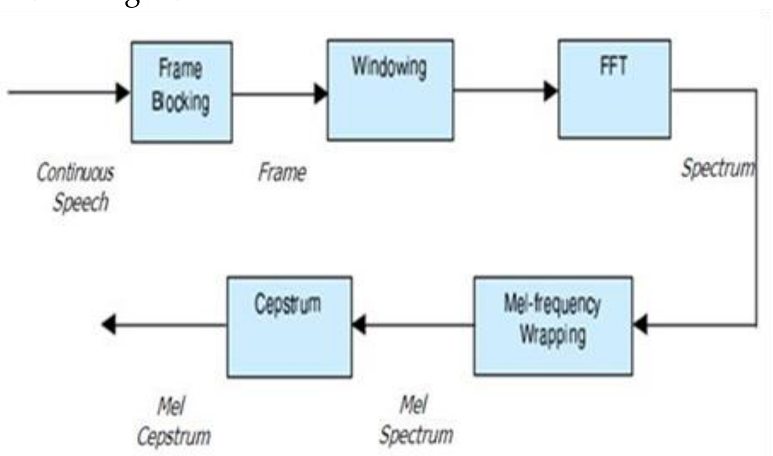

Figure 2. Diagram of the Extraction Process Using the MFCC Method

Data Classification

Neuro-fuzzy is a combination of two systems: fuzzy logic systems and artificial neural networks. A neuro-fuzzy system is based on a fuzzy inference system trained using learning algorithms derived from artificial neural networks. Consequently, the neuro-fuzzy system inherits all the advantages of both fuzzy inference systems and neural networks.

Due to its learning capability, the neuro-fuzzy system is often referred to as ANFIS (Adaptive Neural Fuzzy Inference Systems). The initialization process of ANFIS follows a well-known structure, as illustrated in Figure 3, which represents the ANFIS structure.

.

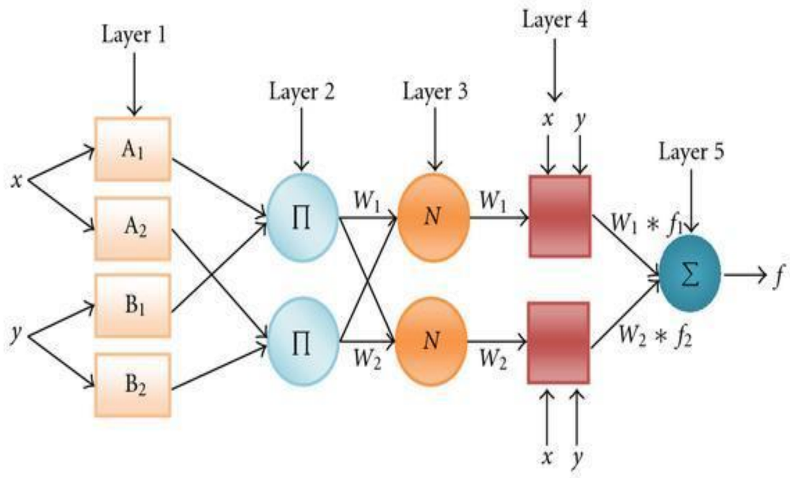

Figure 3. ANFIS Structure

Design

The design of the speech recognition and electrical system is implemented as an interconnected system controlled by voice. System testing aims to evaluate the accuracy of speech recognition and measure the response time required to interact with the smart home system. In this study, MATLAB software is used for data processing and program development, which will serve as a learning platform for the smart home control system application.

The accuracy of speech recognition is calculated by comparing the number of correct outputs generated by the system with the total amount of data. The accuracy percentage is calculated using the following formula:

$$\text{Accuracy} = \frac{\sum \text{Correct Words}}{\sum \text{Words Tested}} \times 100\%$$

Results and Discussion

Speech Recognition System Testing

Testing the accuracy of the smart home control system requires thorough testing and analysis. The testing is conducted in two stages: system testing and real-time testing. System testing is carried out to achieve optimal results.

The training data consists of 40 voice samples per respondent, tested with two male respondents. Each respondent's voice samples include 10 for "Nyala" (On), 10 for "Mati" (Off), 10 for "Buka" (Open), and 10 for "Tutup" (Close), as shown in Figure 4.

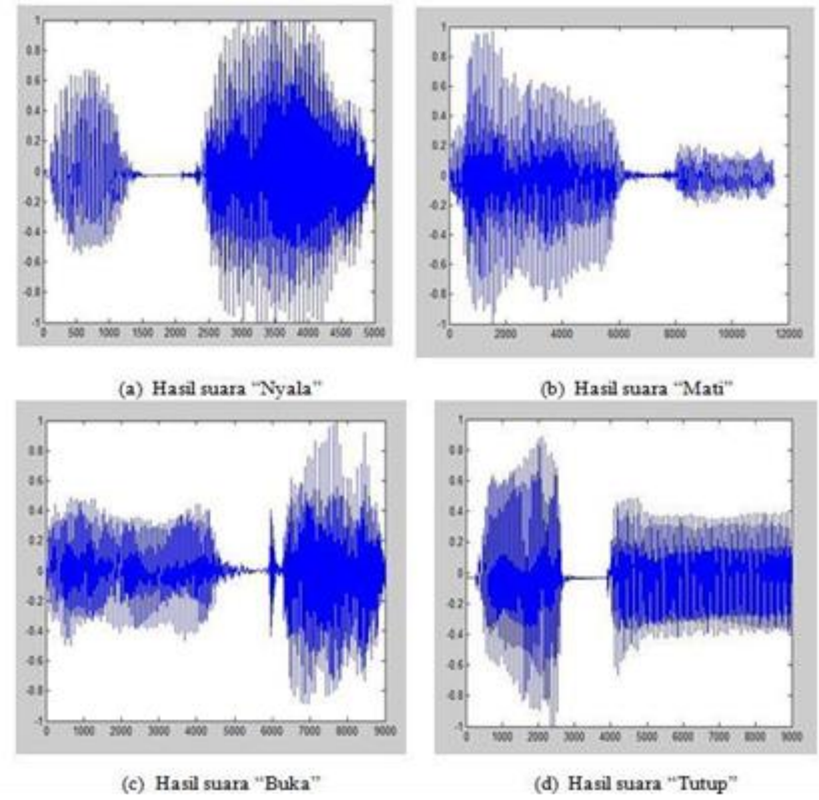

Figure 4. Detected Voice Results for "Nyala," "Mati," "Buka," and "Tutup"

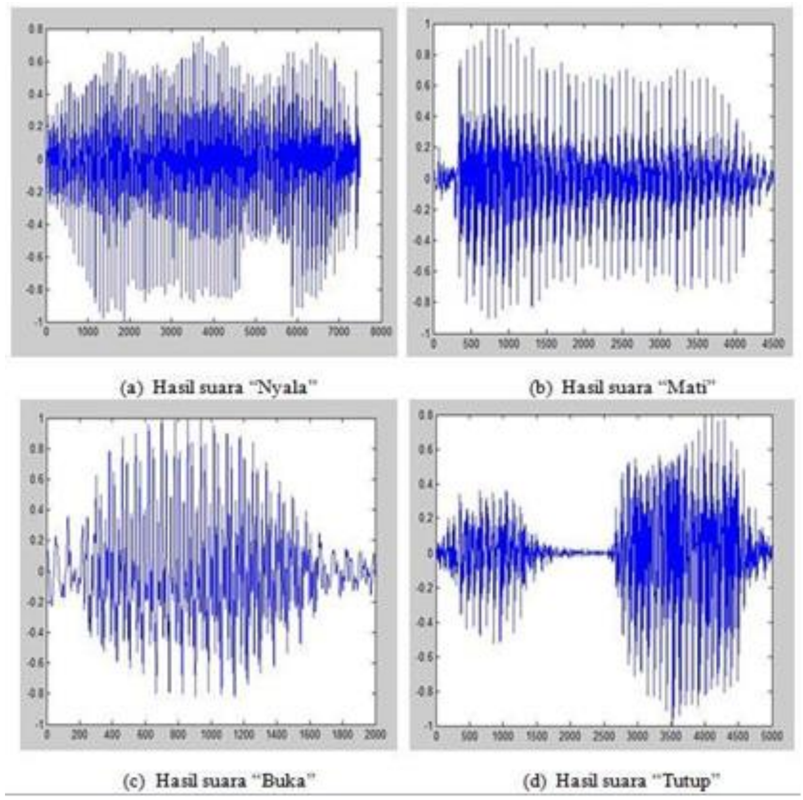

Figure 5. Undetected Voice Results for "Nyala," "Mati," "Buka," and "Tutup"

The two images above show the results of voice signal detection that will later be processed by the smart home control system. In Figure 4, the detected voice signals are read by the smart home control system. In Figure 5, the undetected voice signals are shown, where the system fails to recognize the signals. The difference in signal readings is clearly visible. The failure to detect the voice signals is caused by changes in the signal frequency and the presence of noise that affects the voice signal, resulting in either the system failing to detect the signal or reading it incorrectly.

This system testing is based on 40 training data points and six feature extractions from two respondents. The results are presented in Table 1 and Table 2.

Table 1. Testing Results with 6 Features for Respondent 1

Command Word	Correct	Incorrect	Accuracy %
Nyala	10	0	100%
Mati	7	3	70%
Buka	6	4	60%
Tutup	6	4	60%
Average Accuracy			**72.5%**

Table 2. Testing Results with 6 Features for Respondent 2

Command Word	Correct	Incorrect	Accuracy %
Nyala	6	4	60%
Mati	7	3	70%
Buka	7	3	70%
Tutup	6	4	60%
Average Accuracy			**65%**

From the testing results, the two respondents achieved accuracy rates with average accuracies of 72% and 65%, respectively. The second respondent showed the lowest accuracy with an average of 65%. This result indicates that the system's accuracy is not ideal, likely due to differences in intonation for each word spoken, differences in the respondents' vocal cords, lower frequency values for each respondent, and noise disturbances from the environment, considering that the testing was conducted in a room with high noise levels.

Smart Home Response Testing

The testing of the smart home response significantly impacts the success of this research. The smart home response testing serves as the output of the training data results.

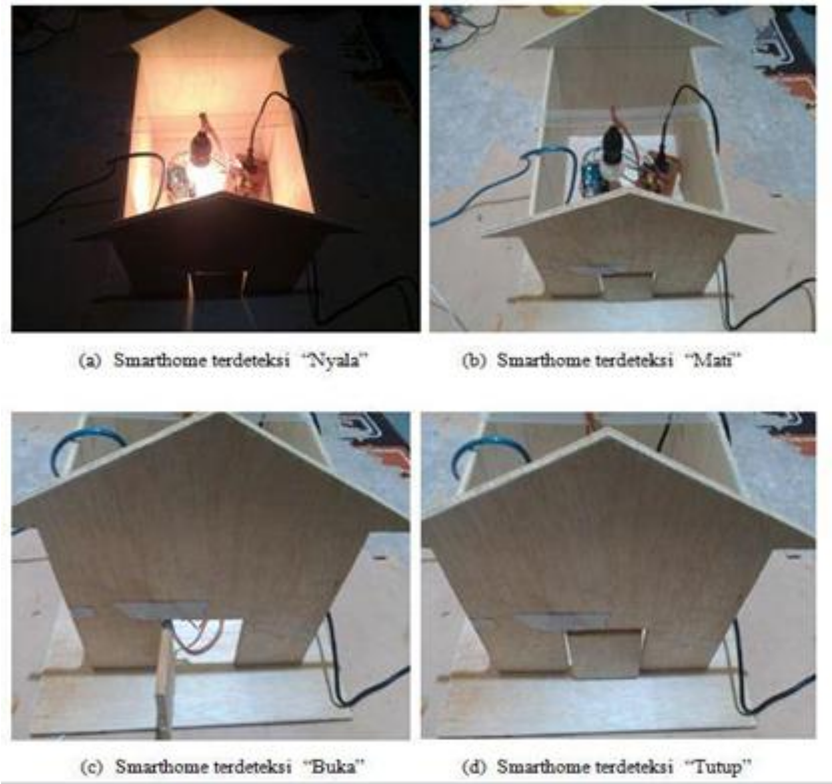

Figure 6. Results of Smart Home Device Testing

Based on the overall system test results, which produced output responses in the smart home system, high precision is required to achieve good accuracy. In this research, the device response is sometimes accurate for each word, providing the correct response, but at other times, it gives less accurate responses, resulting in suboptimal outcomes. The lower accuracy is due to the program, which serves as the brain of the system, being imperfect or not functioning optimally.

Conclusion

Based on the results of the research above, it can be concluded that the testing of the smart home system shows good performance, fulfilling commands sent in real-time from speech to the smart home control system. By using feature extraction with the MFCC method for speech recognition, the system can accept input values from previously created training data with two respondents, achieving an average accuracy range of 65% to 72.5%. The inaccuracy in the data collection process is due to the MFCC method's speech recognition being less effective, along with a looping process in the program that affects the accuracy, thereby impacting the output of the smart home control system. Voice tone or pitch must be the same and consistent during data collection to ensure accurate results.

Acknowledgment

I would like to express my sincere gratitude to my advisor, family, and all those who supported me throughout the research on the implementation of Adaptive Neural Fuzzy Inference Systems (ANFIS) for speech recognition applications in smart home control. Your guidance and encouragement have been invaluable to the success of this project.

Conflicts of Interest

The author declares that there are no conflicts of interest related to the research or the findings presented in this study.

Author Contribution

Roni Permana: Research idea, research planning, data analysis and manuscript writing.
Mada Sanjaya, W.S.: Research advisor.
Hasniah Aliah: Research advisor.

References

1. Cropley, D. H., Kaufman, J. C., & Cropley, A. (2011). Measuring creativity for innovation management. Journal of Technology Management and Innovation, 6(3), 13-30.
2. De Breu, C. K. W., Nijstad, B. A., Bechtoldt, M. N., & Baas, M. (2011). Group creativity and innovation: A motivated information processing perspective. Psychology of Aesthetics, Creativity and the Arts, 5(1), 81-89.
3. Dionne, S. D. (2008). Social influence, creativity and innovation: boundaries, brackets and non-linearity. In M. D. Mumford, S. T. Hunter, & K. E. Bedell-Avers (Eds.), Multi-Level issues in creativity and innovation: Research in multi-level issues (pp.63-73). Amsterdam: JAI Press.
4. Feuer, A. (2011). Developing foreign language skills, competence and identity through a collaborative creative writing project. Language, Culture and Curriculum, 24(2), 125-139.
5. Ghonsooly, B., & Showqi, S., (2012). The effects of foreign language learning on creativity. English Language Teaching, 5(4), 161-167.
6. Luk, J. (2013). Bilingual language plan and local creativity in Hong Kong. International Journal of Multilingualism, 10(3), 236-250.
7. Romero, M., Hyvönen, P., & Barbera E. (2012). Creativity in collaborative learning across the life span. Scientific Research, 3(4), 422-429.
8. Schulz, B. (2011). Syntactic creativity in second language English: wh-scope marking in Japanese-English interlanguage. Second Language Research, 27(2), 313-341.