Research Article

Identification of *Tajweed* Recognition using Wavelet Packet Adaptive Network based on Fuzzy Inference Systems (WPANFIS)

[1*]Ratu Mutiara Siregar, [2]Budy Satria, [3]Andi Prayogi, [4]Muhammad Akbar Syahbana Pane, [5]Elsa Elvira Awal, [6]Yessi Ratna Sari

[1,3,4]Institut Teknologi Sawit Indonesia, Deli Serdang, North Sumatra 20371, Indonesia

[2]Institut Teknologi Mitra Gama, Bengkalis, Riau 28784, Indonesia

[5]Universitas Buana Perjuangan Karawang, Telukjambe Timur, Karawang, West Java 41361, Indonesia

[6]Universitas Lancang Kuning, Jalan Yos Sudarso KM. 8 Rumbai, Pekanbaru, Indonesia

*Corresponding Author: ratu_ms@itsi.ac.id

Citation: R.M.Siregar, et.al., "Identification of Tajweed Recognition using Wavelet Packet Adaptive Network based on Fuzzy Inference Systems (WPANFIS)". *Iota*, **2024**, ISSN 2774-4353, Vol.04, 01. https://doi.org/10.31763/iota.v4i1.703

Academic Editor : Adi, P.D.P

Received : January, 07 2024

Accepted : January, 18 2024

Published : February, 09 2024

Abstract:
This research aims to develop a system capable of processing voice input to recognize Al-Quran reading by recitation of *Tajwid*, using wavelet signal extraction and classification of *Tajwid* rules using ANFIS. The process stages include data acquisition, audio data pre-processing, extraction using wavelet packets, division of training data and test data, and classification. The data obtained were 20 observations from 10 observations carried out in data pre-processing. The wavelet decomposition process produces six main features as ANFIS input variables and 64 rules. Then the data was separated into 17 observations for training data and three for testing data. The test results obtained from the training that had been carried out produced plots that were too fit; in this experiment, the WPANFIS classification model got 100% appropriate classification and SSE values that were the same as the training result, 0.00081225.

Keywords: Al-Qur'an, Wavelet Packet Decomposition, ANFIS, Tajweed, Voice Recognition

1. INTRODUCTION

According to *Tajweed* law, the Al-Qur'an must be read with good and correct pronunciation [1]. All rules and regulations for reading the Al-Qur'an are called the Science of Tajweed. Understanding and study earnestly will give Muslims good and correct abilities to read the Al-Qur'an and make it enjoyable to listen to [2]. Studying the Al-Qur'an requires being with a teacher and extensive training until the trainee reaches the appropriate skills. By looking at the process of learning the Al-Qur'an, participants also need more time to study and practice from meetings with the teacher. This situation can be supported more intensively with a system that can recognize and correct mistakes made by students of the Al-Qur'an when reading it.

Based on research conducted on 20 students taking the *Al-Quran Tilawah* (Reading the *Al-Quran*) level 2 course, there were a number of errors in reading the Al-Quran. The value 29% of errors is in the pronunciation of letters (*makhraj and nature*), 23% of errors

in applying *tarqiq wa tafkhim* (thick-thin), 23% of errors in reading in the rules of *ghunnah* (pronunciation and non-pronunciation), errors in mad as much as 14% of the reading, 8% of errors are in reading letters and their meanings (punctuation marks), and 3% of errors are in *waqf wa ibtida'* (stopping and starting) [3]. In this problem, it is necessary to carry out research related to voice recognition technology. Speech recognition techniques are widely used and focus on spoken words to obtain results through spoken sounds or dialect pronunciation [4].

The proposed method for identifying errors in recitation recognition uses a Wavelet Packet Adaptive Network based on Fuzzy Inference Systems (WPANFIS) [5]. Wavelet and wavelet packet analysis have been proven effective signal-processing techniques for various signal-processing problems and are designed for speech/voice recognition purposes [6]. The ANFIS method has the linguistic capabilities of fuzzy systems and neural networks' learning and parallel computing capabilities [7].

This research aims to develop a system that is capable of processing voice input for the recognition of Al-Quran reading in the form of Tajwid pronunciation, using wavelet signal extraction and classification of Tajweed rules using a Wavelet Packet Adaptive Network Based on Fuzzy Inference Systems. Some of the previous research, specifically on the analysis of each, has provided a sharp analysis of the Wavelet Packet Adaptive Network [14-18] and the application of the ANFIS method [19-22].

2. THEORY

Voice recognition is part of artificial intelligence technology [8]. In this section, the author will provide information related to supporting theories in conducting this research.

2.1 *Tajweed*

The meaning of *Tajweed* is an improvement. In its truest sense, it articulates each letter from its point of articulation and assigns rights and obligations. Reading the *Qur'an* without using tajwid and consistent practice will help you achieve perfect reading. With advances in information technology and artificial intelligence techniques, many Al-Quran applications have been developed for both computer and mobile devices to help the *Tajweed* learning process [9].

2.2 Voice Recognition

Voice recognition is a technology that allows computers or electronic devices to receive and identify voice input from users. This technology is widely used to make things easier for users to control electronic devices, autofill forms, or help transcribe voice into text [10].

2.3 Wavelet Packet Decomposition

Wavelet theory is a relatively recently developed concept. Wavelet was first introduced by Alfred Haar in 1909. Jean Morlet and Alex Grossmann coined the word "wavelet" in the early 1980s, and it comes from the French "*ondelette,*" which means small wave. The word "*onde,*" which means wave, was then translated into English as "wave" and then combined with the original word to form the new word "wavelet." Voice recognition is a technology that allows computers or electronic devices to receive and

identify voice input from users—this technology [11]. One of the practical impacts of this technique is that the resulting outputs are at the same level of decomposition, so they have the same size.

2.4. Adaptive Network based on Fuzzy Inference Systems

The adaptive neuro-fuzzy inference system model is developed or implemented using the fuzzy inference system and neural network [12]. Moreover, in this research, the process of extracting voice characteristics into a system which is then identified using the Adaptive Neuro Fuzzy Inference System (ANFIS), and so that it is then understood by the system and displayed output results in the form of text, percentage numbers, and spoken readings. The first step in the classification process using the ANFIS method is to train the system to get appropriate output results. The system is introduced and then tested by providing several test samples to see the value of the system's accuracy [13].

3. METHOD

The stages carried out in this research were implementing automatic speech recognition, extracting speech signals using wavelet packets, and classifying using the Adaptive Network-based Fuzzy Inference System (ANFIS). The process stages include data acquisition, audio data pre-processing, extraction using wavelet packets, division of training data and test data, and classification. The proposed research method can be seen in Figure 1.

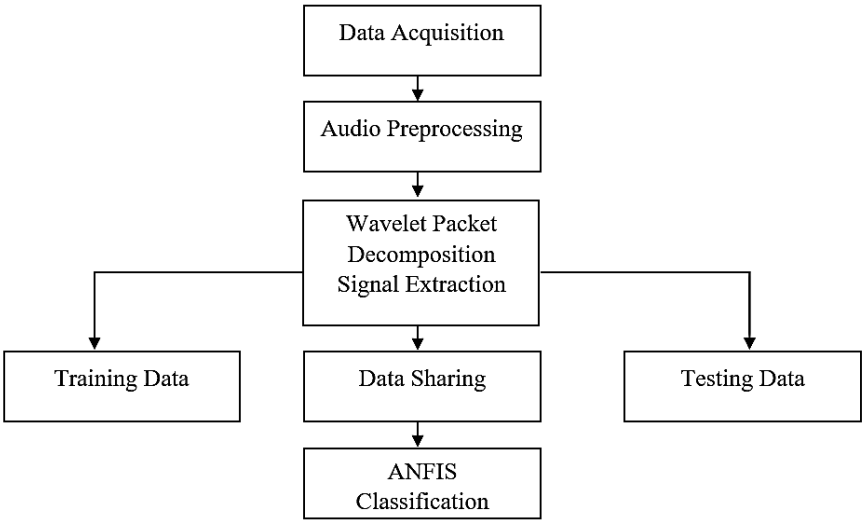

Fig 1. Proposed research method

3.1 Data Acquisition

The data in this research was created through the process of recording the reading of one verse from the Koran using a smartphone from 2 tajwid pronunciations, namely *ghunnah* (*idgham bighunnah*), where *nun mati* meets the letters *ya* with the meaning *fathah* and *ikhfa*, where the pronunciation of ta with the meaning *kashrah tanwin* meets the letters *sya*'. The data source comes from 4 people with the characteristics of 1 male voice (*Quran expert*) and three female voices who are classified as intermediate readers of the Qur'an (at least have undergone *tahsin* level 2 learning) with three repetitions of reading. The audio data is still as long as reading 1 verse, which will only take the pronunciation of

ghunnah and *ikhfa* as a data group. So, the total data obtained a total of 10 observations will be pre-processed into 20 data samples.

3.2 Pre-processing Voice Signals

The sound signal obtained at the data acquisition stage then goes to the pre-processing stage to obtain the legal pronunciation of tajwid *ghunnah* and *ikhfa.* A manual detection method for the beginning and endpoint of the speech signal is used to separate the speech signal. The sampling rate used is 4.8 kHz, and the bit-depth resolution is 16-bit PCM. Data is saved in *.wav format. The sampling rate used is in Table 1.

Table 1. Sampling Rate

No	Pre-processing Signals	
	Parameter	**Value**
1	Sampling Rate	4.8 kHz, 16-bit
2	Data	Isolated Tajweed Pronunciation
3	Speaker	1 male, 3 female
4	Bit Rate	705 kbps
5	Number of Channels	1 (Mono)
6	Audio Formats	PCM *.wav

Eliminate signal noise using signal denoising using wavelets. This denoising stage consists of three stages:

1. Calculate the packet decomposition of a speech signal with level 7 and use order 10 Daubechies wavelets (db10).
2. At each level from 1 to 7, soft thresholding is carried out in the details of the coefficients.
3. Calculate wavelet reconstruction based on the original approach at level 7, the modification of the detailed coefficient level at level 7, and the detailed coefficients approach at levels between 1 and 7.

3.3 Wavelet Packet Decomposition

Speech waveforms are detailed and highly non-stationary. After data pre-processing is carried out, the stage continues with wavelet packet decomposition. Wavelet packet decomposition is applied to the speech signal using Daubechies-10 wavelet packet filter ψ and level 7, in the equation s (1) is speech s_i i is the wavelet coefficient of x where we get 28 = 256 terminal points from A1 to A256 of the input signal.

$$\psi = -\sum_i s_i^2 \log|(s_i^2) \quad (1)$$

Entropy-based criteria describe the nature of information related to accurately representing a given signal. Entropy is a common concept in various fields, especially in signal processing. A method for measuring entropy emerged as an ideal tool for ordering

non-stationary signals. Next calculate the entropy in equation (2) as defined from the waveform at the terminal node signal obtained from the wavelet packet decomposition:

$$E_{(s)} = \frac{\sum i|s_{i\,|P}}{N} \quad (2)$$

E is the wavelet packet entropy in the form of a real number, and s is the signal point terminal (s_i) *i* waveform of the signal points. So, we obtain a new vector matrix with the entropy value of the speech signal.

3.4 WPANFIS Structure

Both artificial neural networks and fuzzy logic are used in this ANFIS architecture. ANFIS consists of an if-then rule learning algorithm and input-output pairs; for ANFIS training, a neural network is used. To simplify things, we assume that FIS has six inputs (x1, x2, x3, x4, x5, x6), namely, maximum value, minimum value, arithmetic mean value, geometric mean value, standard deviation, and variance, each of 256 wavelet packet entropy value, with one output based on the Sugeno fuzzy model. ANFIS architecture and training meters can be seen in Table 2.

Table 2. ANFIS Architecture

No	ANFIS Architecture	
	Architecture	**Value**
1	Layer	4.8 kHz, 16-bit
2	Input	Isolated Tajweed Pronunciation
3	Output	One male, three female
4	Type MF	705 kbps
5	Rule	

Table 3. Training Parameters

No	Training Parameters	
	Training Parameters	**Value**
1	Learning Rule	Hybrid
2	Sum-Squade Error	0.00081225
3	Epoch Reached To Sse Value	2/100

4. RESULT AND DISCUSSION

4.1 Wavelet Decomposition Process

The data obtained were 20 observations from 10 observations carried out in data pre-processing. The wavelet decomposition process produces six main features as ANFIS input variables and 64 rules. Data was separated into 17 observations for training data and three observations for testing data. The wavelet decomposition process can be seen in Figure 2.

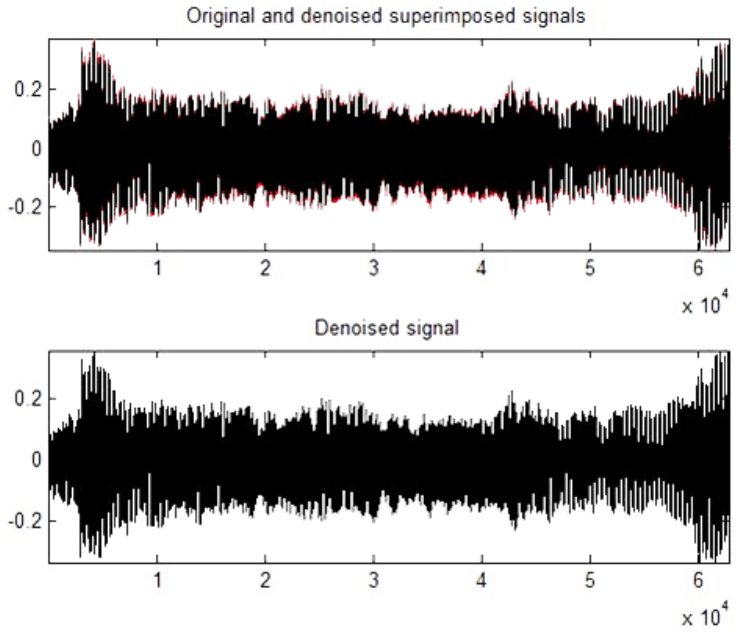

Fig 2. Wavelet decomposition process

4.2 ANFIS Structure

The ANFIS structure can be seen as in Figure 3 from the wavelet decomposition results.

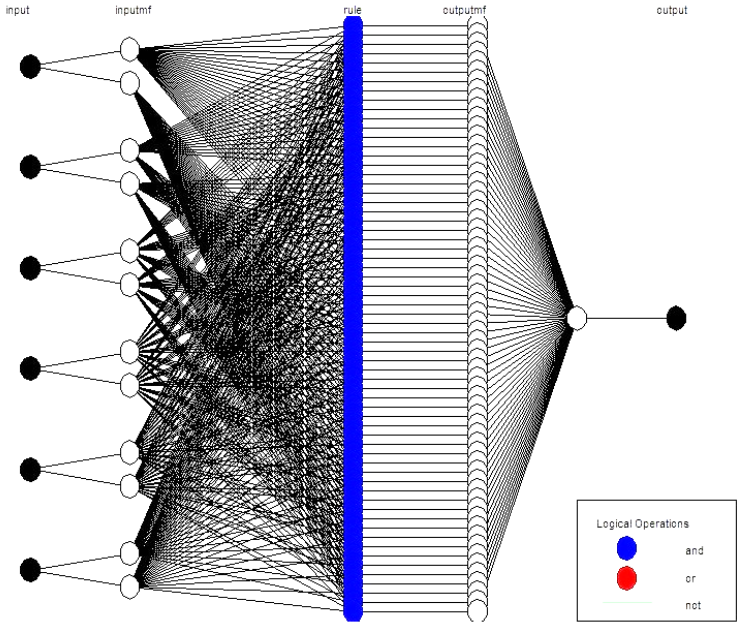

Fig 3. ANFIS Structure

4.3 Training Result

The ANFIS training error result value is 0.00081225; from 100 training epochs, the smallest number of errors (sum squared error) was achieved in the second epoch. The training results can be seen in Figure 4.

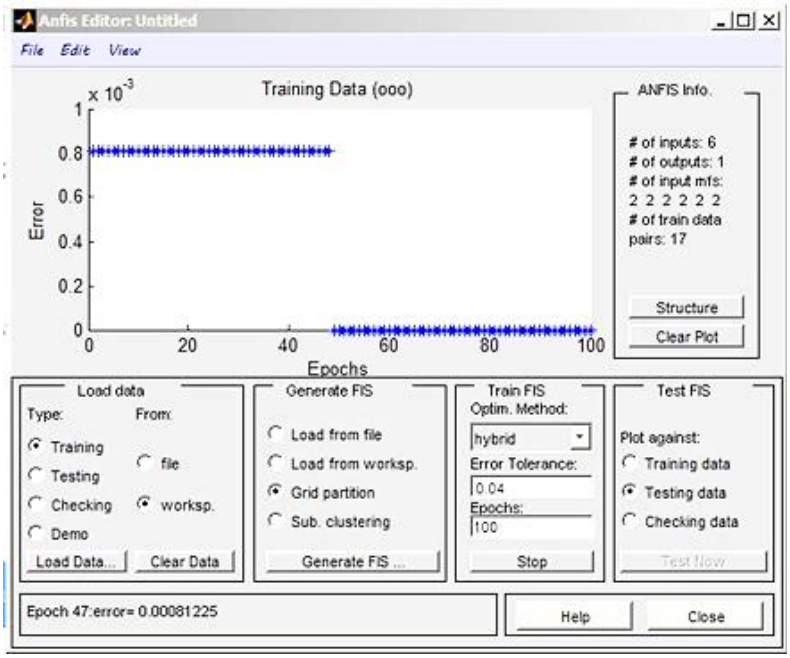

Fig 4. Training Result

4.4 Testing Result

The testing results on the training produce plots that are too fit, so in this experiment, the WPANFIS classification model gets 100% appropriate classification and the same SSE value as the training results, 0.00081225. The test results can be seen in Figure 5.

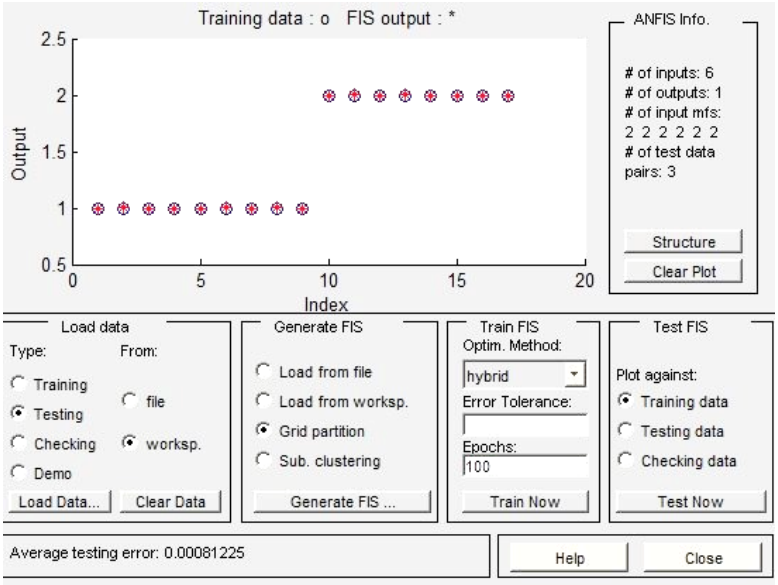

Fig 5. Training Result 2

5. CONCLUSION AND SUGGESTION

This research developed an automatic speech recognition (ASR) system using pattern recognition and speech recognition techniques from 4 speakers who read the Qur'an. Moreover, using wavelet packages to obtain entropy from speech signals, *Tajweed* pronunciation extraction was developed. Meanwhile, the classification of Tajweed rules using ANFIS produces a classification model that is too precise. WPANFIS recognized the legal recitation of ghunnah and ikhfa readings quite well with the 20 observation data obtained in this research. Processing audio signals, both music and voice, requires sophisticated systems to speed up sound data processing; this is a weakness in research, so little data can be obtained. The existence of a corpus and dataset on the pronunciation of Tajweed phonemes will really help work in the field of research on the introduction of reading and pronunciation of the Qur'an more widely in the future, which is possible with the support of an expert system.

6. ACKNOWLEDGMENTS

We want to thank all the authors involved in this research who helped with the research directly or indirectly, both in thought and time and the campus of Institut Teknologi Sawit Indonesia for funding assistance.

AUTHOR CONTRIBUTIONS

All Authors are responsible for building Conceptualization, Methodology, analysis, investigation, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision of project administration, funding acquisition, and have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. S. M. Abdou et al., "Computer aided pronunciation learning system using speech recognition techniques," in INTERSPEECH 2006 - ICSLP, Ninth International Conference on Spoken Language Processing, 2006, pp. 17–21. doi: 10.21437/Interspeech.2006-287.
2. N. F. Laily and S. Maesurah, "Strategi Peningkatan Kemampuan Dan Pemahaman Siswa Tpq Atas Pelafalan Makhorijul Huruf Dan Ilmu Tajwid Di Desa Baureno, Jatirejo, Mojokerto," Al-Din J. Dakwah dan Sos. Keagamaan, vol. 7, no. 2, pp. 12–26, 2021, doi: 10.35673/ajdsk.v7i2.2365.
3. S. S. Bin Hassan and M. A. Bin Zailaini, "Analysis of Tajweed Errors in Quranic Recitation," Procedia - Soc. Behav. Sci., vol. 103, no. September, pp. 136–145, 2013, doi: 10.1016/j.sbspro.2013.10.318.
4. S. Qiu, "Construction of English Speech Recognition Model by Fusing CNN and Random Deep Factorization TDNN," ACM Trans. Asian Low-Resource Lang. Inf. process., no. May, pp. 1–19, 2023, doi: 10.1145/3597456.
5. R. R. Mutra, D. Mallikarjuna Reddy, J. Srinivas, D. Sachin, and K. Babu Rao, "Signal-based parameter and fault identification in roller bearings using adaptive neuro-fuzzy inference systems," J. Brazilian Soc. Mech. Sci. Eng., vol. 45, no. 1, pp. 3–26, 2023, doi: 10.1007/s40430-022-03954-5.

6. E. Avci and Z. H. Akpolat, “Speech recognition using a wavelet packet adaptive network based fuzzy inference system,” Expert Syst. Appl., vol. 31, no. 3, pp. 495–503, 2006, doi: 10.1016/j.eswa.2005.09.058.
7. V. Sreelekha and A. Prince, “ANFIS-Based Fault Distance Locator With Active Power Differential-Based Faulty Line Identification Algorithm for Shunt and Series Compensated Transmission Line Using WAMS,” IEEE Access, vol. 11, no. July, pp. 91500–91510, 2023, doi: 10.1109/ACCESS.2023.3307466.
8. B. Satria, “Prediksi Volume Penggunaan Air PDAM Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 2, no. 3, pp. 674–684, 2018, doi: 10.29207/resti.v2i3.575.
9. F. Ahmad, S. Z. Yahya, Z. Saad, and A. R. Ahmad, “Tajweed Classification Using Artificial Neural Network,” in 2018 International Conference on Smart Applications, Communications and Networking, SmartNets 2018, IEEE, 2018. doi: 10.1109/SMARTNETS.2018.8707394.
10. I. M. Nurrizqy, B. H. Prasetio, and R. M. P. Rekyan, “Infrared-Based Device Control System Using Speech Recognition With Spectrogram and Convolutional Neural Network Based On Microcontroller,” J. Teknol. Inf. dan Ilmu Komput., vol. 10, no. 5, pp. 955–962, 2023, doi: 10.25126/jtiik.2023106909.
11. M. Ohyver and H. Pudjihastuti, “Pemodelan Tingkat Penghunian Kamar Hotel di Kendari dengan Transformasi Wavelet Kontinu dan Partial Least Squares,” ComTech Comput. Math. Eng. Appl., vol. 5, no. 2, p. 1178, 2014, doi: 10.21512/comtech.v5i2.2435.
12. S. Rahmah, W. Witanti, and P. N. Sabrina, “Prediksi Penjualan Obat Menggunakan Metode Adaptive Neuro-Fuzzy Inference System (ANFIS),” JIMP - J. Inform. Merdeka Pasuruan, vol. 7, no. 3, p. 109, 2023, doi: 10.51213/jimp.v7i3.733.
13. F. M. Fauzi, L. Hayat, and D. N. Kusuma Hardani, “Pengenalan Dialek Bahasa Daerah di Pulau Jawa menggunakan Metode Mel-Frequency Cepstral Coefficients dan Adaptive Network-based Fuzzy Inference System,” J. Ris. Rekayasa Elektro, vol. 4, no. 2, p. 39, 2023, doi: 10.30595/jrre.v4i2.11526.
14. S. C. S, A. C. V, A. G and B. M, "RPCA and Wavelet packet Decomposition based Network Traffic Anomaly Detection," 2022 IEEE 19th India Council International Conference (INDICON), Kochi, India, 2022, pp. 1-5, doi: 10.1109/INDICON56171.2022.10039837.
15. M. Li and Y. Wang, "Rolling Bearing Fault Classification Utilizing Adaptive Density Peaks Search Clustering Based on Wavelet Packet Transform," 2021 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD), Nanjing, China, 2021, pp. 1-5, doi: 10.1109/ICSMD53520.2021.9670877.
16. P. -C. Wu, P. -F. Li, Z. -T. Wu and J. -W. Hung, "The Study of Improving the Adaptive FullSubNet+ Speech Enhancement Framework with Selective Wavelet Packet Decomposition Sub-Band Features," 2023 9th International Conference on Applied System Innovation (ICASI), Chiba, Japan, 2023, pp. 130-132, doi: 10.1109/ICASI57738.2023.10179539.
17. M. A. Riyadi, I. Setiawan and A. Amir, "EEG Multiclass Signal Classification Based on Subtractive Clustering-ANFIS and Wavelet Packet Decomposition," 2021 International Conference on Electrical and Information Technology (IEIT), Malang, Indonesia, 2021, pp. 81-86, doi: 10.1109/IEIT53149.2021.9587407.
18. W. Shi and Y. Qin, "Inverter Fault Diagnosis Algorithm Based on IGWO-ADAM-BP Neural Network," 2022 4th International Conference on Applied Machine Learning (ICAML), Changsha, China, 2022, pp. 365-371, doi: 10.1109/ICAML57167.2022.00075.
19. A. D. Firoozabadi et al., "Estimating the Number of Speakers by Novel Zig-Zag Nested Microphone Array Based on Wavelet Packet and Adaptive GCC Method," 2022 8th International Conference on Signal Processing and Communication (ICSC), Noida, India, 2022, pp. 358-363, doi: 10.1109/ICSC56524.2022.10009025.
20. M. Ghanbari and W. Kinsner, "Designing a Neural Network and a Genetic- Algorithm-Based Adaptive Wavelet for Internet Traffic Containing DDoS Attacks," 2021 IEEE 20th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), Banff, AB, Canada, 2021, pp. 173-180, doi: 10.1109/ICCICC53683.2021.9811314.

21. A. Kenge, M. S. Panse, S. Choudhari, H. Pethkar and R. Pinto, "Wavelet Transform Based Compression of Raw SAR Data," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-5, doi: 10.1109/ICCCNT51525.2021.9579774.
22. G. Li, D. Cui, J. Li, Z. Han, Q. Jiang and W. Wang, "Research of Gearbox Failure Diagnosis Based On Nerve Network," 2023 2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS), Bristol, United Kingdom, 2023, pp. 452-457, doi: 10.1109/AIARS59518.2023.00097.