An Intelligent Optimization Technique Of Automatic Speech Recognition For Smart Homes

Rabab Hamed M. Aly [1,3,*,†], and Aziza I. Hussein [2,†]

[1] The Higher Institute for Management and Information Technology, Minya, Egypt.
[2] Electrical and Computer Engineering Department, College of Engineering Effat University, Jeddah, Saudi Arabia.
[3] Computer Science Department, Faculty of Computer Science, Nahda University, Beni Suef, Egypt.

Abstract

The creation of new approaches to the design and configuration of smart buildings relies heavily on AI tools and Machine Learning (ML) algorithms, particularly optimization techniques. The widespread use of electronic devices has sparked a strong desire to incorporate the Internet of Things (IoT) into houses, leading to the development of smart homes. As networked gadgets proliferate rapidly, this phenomenon is characterized by rapid proliferation. In smart buildings, smart cities, smart grids, and smart homes, interconnected electronic devices are becoming more popular. The objective of this paper is to enhance the functionality of home automation systems through the performance of speech recognition using the Bat-Salp Swarm Optimization (BSSO). This paper investigates the notion of (BSSO), a data analysis methodology that facilitates the automated construction of analytical models. The implementation of BSSO provides an enhancement to the feature selection process in speech recognition, providing an approximation solution that improves the accuracy of system decisions. The use of the BSSO technique improves the precision of the voice recognition system and also incorporates an Artificial Neural Network (ANN) for the classification part. The findings substantiated the efficacy of the employed methodology.

Keywords

Speech recognition, Smart homes, Salp Optimization Algorithm, and Bat Optimization Algorithm.

1. Introduction

Modern home automation technology is moving us towards the perfect smart home setup. Artificial intelligence (AI) and deep learning are effective in the IoT. In addition, search algorithms rely on metaheuristics, which are essential optimization methods, to solve complex AI problems. Speech recognition and control, data input, voice user interface, and telephone operator automation are the applications. Metaheuristic approaches can mimic physical, biological, or natural principles that exhibit search or optimization traits and predictability [1-5]. Smart buildings, cities, grids, and homes are just a few examples of numerous environments where networked devices are becoming increasingly common. Smart buildings, towns, electricity systems, and residences are using more networked electronics. This paper introduces Bat-Salp Swarm Optimization (BSSO), a data analysis method that automatically creates

* Corresponding author.
† These authors contributed equally.
rabab.cse2010@gmail.com(R. Hamed. M. Aly); azibrahim@effatuniversity.edu.sa (A.I. Hussein).
0000-0002-0383-2078 (R. Hamed. M. Aly); 0000-0001-9610-395X ((A.I. Hussein).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

analytical models. BSSO uses approximation to improve voice recognition feature selection and system decisions. Moreover, the application of the BSSO technique improves the precision of the voice recognition system. The findings substantiated the efficacy of the employed methodologies. The rest of our paper is organized as follows: Section. 2 gives a brief description of related studies. Section.3 illustrates the proposed model. Section.4 shows the experimental results. Finally, Section.5 is the conclusion.

2. LITERATURE REVIEW

Numerous metaheuristic strategies in various applications have been introduced by authors recently. Cross-lingual Voice Conversion is the process of modifying the speaker identification of a voice sample from one speaker to another speaker who speaks a different language from the source speaker. Converters are frequently employed in the provision of a converters, and they displayed the fully integrated power with simulation in different cases. Various methodologies have been devised for voice recognition; however, attaining a high level of precision continues to pose a substantial obstacle [6,7]. In [8], the authors developed a novel deep learning model for voice identification named the Taylor Gradient Descent Political Optimizer (Taylor GDPO). The model that was built exhibited outstanding efficacy, achieving an impressive accuracy rate of 96.93%. In [9], the authors introduced the Kawahara filter as an optimizer for accurately categorizing employees and predicting their performance. In addition, in [10], the authors developed a new method to enhance the accuracy of a certain Automatic Speech Recognition (ASR) engine by automatically adjusting the front-end speech augmentation. The suggested technique allowed users to use Automatic Speech Recognition (ASR) on consumer electronics devices with less difficulty, even when there are changes in the surrounding noise levels. A genetic algorithm (GA) was utilised to determine parameter values for the front-end speech enhancement customised for specific environments. The generated values can be allocated in a dynamic manner to input speech signals by first clustering the surroundings based on noise characteristics. The evaluations showed that our technique surpassed the parameter values determined by a human expert. The authors presented a new algorithm for machine learning in [11]. Initially, Mel-frequency cepstral coefficients were utilised to extract the features from the voice signal datasets. An innovative method was proposed that combines the grey wolf optimizer and the Naïve Bayes machine learning algorithm for the purpose of categorization. According to the results, their proposed classification algorithm shows superior performance compared to existing machine learning approaches. In recent decades, the home automation system has gained considerable popularity, leading to increased comfort and improved quality of life.

Several authors, including [12], conducted a comprehensive analysis of contemporary and emerging home automation systems. A mobile application is utilised to oversee and control household appliances through the implementation of various communication tactics. Several research efforts, including [12,13], have examined the operational processes of several wireless communication technologies, such as ZigBee, Wi-Fi, Bluetooth, and GSM. Furthermore, this paper examines several home automation systems, providing a thorough analysis of their benefits and limitations. Furthermore, this paper examines several home automation systems, providing a thorough analysis of their benefits and limitations. The Bat Algorithm is based on the echolocation abilities of microbats. Echolocation is a form of sonar used by microbats to locate prey and identify obstacles or threats in total darkness. Bat algorithms are utilized in diverse fields [14-16]. In [14], the authors implemented a load shifting technique to effectively decrease the total electricity cost. In order to accomplish this objective, they proposed merging the bat algorithm (BA) and crow search algorithm (CSA) to create a hybrid system known as the bat-crow search algorithm (BCSA). An innovative algorithm for reducing the Peak to Average Ratio in Single-Phase H-Bridge Multilevel Inverters is described in [15-18]. Following this description of related works, we will get into the approach and models that are based on some of the related work that has come before.

3. METHODOLOGY

The paper's structure contains three main elements (see Figure 1). The preprocessing stage is followed by the feature selection stage, which combines two primary optimisation techniques: Bat optimisation and Salp Swarm optimisation approach (BSSO). The last part is the classification stage, which utilises an Artificial Neural Network (ANN). The subsequent analysis will focus on the main discussion of each stage illustrated in the block.

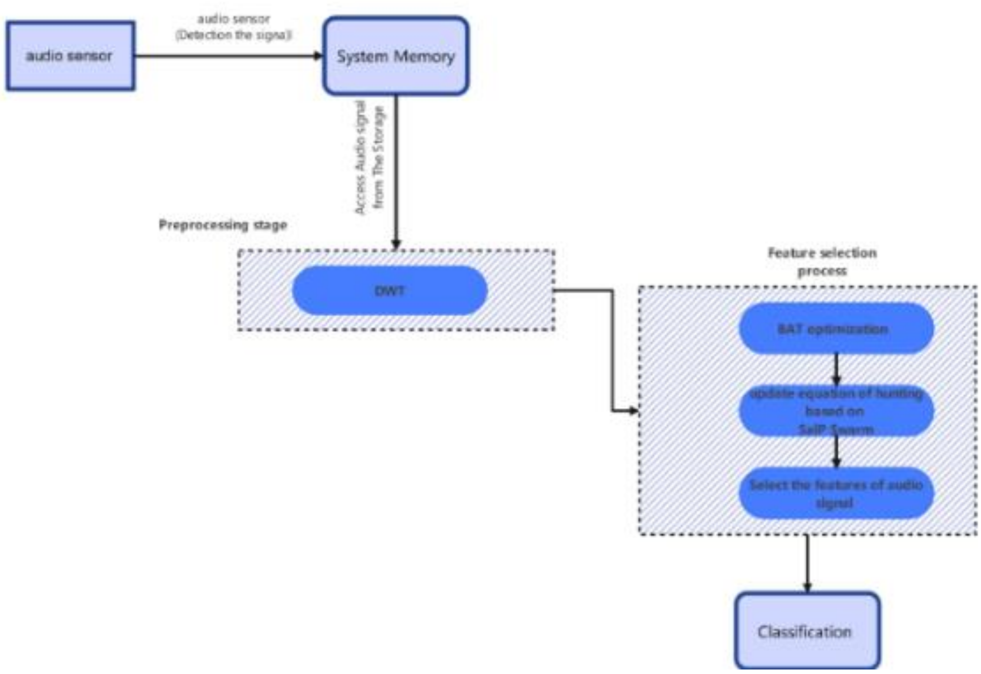


Figure 1: The proposed design of system.

3.1. Preprocessing:

The first block pre-processes. The Discrete Wavelet Transform (DWT) converted the signal from time to frequency after having filtered it. The Discrete Wavelet Transform (DWT) at its maximal decomposition level produces a signal with precise frequency resolution [20]. The MATLAB digital signal processing program is used to apply three stages of Discrete Wavelet

Transform (DWT). The DWT divides signals into approximation and detailed sub bands. Most noise affects the high-frequency sub band, while vital information is in the low-frequency sub band. This level uses Daubechies wavelet [20-21].

3.2. Feature selection

The major feature section approach is bat optimisation. The Bat Algorithm (BA) mimics bat hunting. BAs operate on this idea. The bats in the swarm are thought to travel randomly at a certain velocity Vi and position Xi.

They search for prey using a steady frequency f, variable wavelength λ, and a specific loudness $A0$. By manipulating pulse wavelength and emission rate $r \in [0, 1]$, they may determine their proximity to the target. Loudness can range from a high positive number $A0$ to a constant minimum value Amin. Frequency f falls within [$f$min, $f$max], similar to wavelength range [λmin, λmax] [22]. The algorithm's implementation focuses on three illustrated principles:

1) Bats employ echolocation to perceive distance and possess the ability to differentiate between prey and obstacles.

2) To search for prey, bats fly randomly with the velocity vi at position xi with varying frequencies fi (from a minimum frequency to a maximum frequency) or with varying wavelengths and loudness's A0. Automatically adjusting wavelength (or frequency) and rate r of pulse emission can also be accomplished depending on target proximity.

3) In the third step, assume that the loudness varies from the largest positive value A0 to the minimum constant value Amin. As a result of the above rules, the position vector of the bat represents a solution in the search space. The algorithm randomly initializes the bats since the global optimal position vector is not known a priori [23].

The bat optimization techniques will apply based on the following equation:

$$x_{i,j} = x_{lb} + (x_{ub} - x_{lb})rand \tag{1}$$

$$f_i = f_{min} + (f_{max} - f_{min})\beta, \tag{2}$$

$$v_i^t = v_i^{t-1} + \left(x_i^t - x_*\right)f_i, \tag{3}$$

$$x_i^t = x_i^{t-1} + v_i^t, \tag{4}$$

$$x_{new} = x_{old} + \varepsilon A^t, \tag{5}$$

$$v_i^t = w_i(t)v_i^{t-1} + \left(x_i^t - x_*\right)f_i, \tag{6}$$

$$w_i(t) = w_{max} - (w_{max} - w_{min})arctan\left(\frac{4t}{N\_gen}\right) \tag{7}$$

Where i th bat, i = 1, 2,3,. . . , n; j = 1, 2,3,. . . , d; n signifies the population size, d signifies the dimension of the search space. x_{lb} and x_{ub} are lower, upper bounds for j th dimension, rand is a random number between [0,1], and β is a random number between [0, 1]. x_* signifies the current global best solution, and the bat updates the velocity, position according to the change of frequency f , ε ∈ [−1, 1], w_{max} and w_{min} respectively represent the maximum and minimum weight, t represents the current number of iterations, and N_gen characterizes the maximum number of iterations i, At correspond to the average loudness in current iteration and Loudness

Ai which is a vector of values for all bats and also Rate r_i of pulse emission which is a vector for all bats controlling the diversification of bat algorithm [22].

However, the SSO is a bio-inspired algorithm that mimics Salp's natural foraging movements. By mimicking Salps' hunting and foraging techniques, the SSO hopes to efficiently handle complex optimization problems [15,23]. Equation (4) is revised in this study using the SSO hunting equation (9). Keep in mind that the value of the variable c1 changes over time according to equation (8), where L is the maximum number of iterations and l is the current iteration value. The BBSO approach, a novel approach for feature selection in acoustic signal analysis, and the bat optimization algorithm are both greatly improved by this version.

$$C_1 = 2e^{-(\frac{4l}{L})^2} \tag{8}$$

$$x_j^i = c_1 \times rand(x_j^i + x_j^{i-1}) \tag{9}$$

In addition, in the classification phase, we employ hybrid learning utilizing the ANN algorithm, which combines the descent and least-squares methods to determine the parameters [24].

The proposed technique aims to enhance the hardware of the system on a chip by optimising its performance to achieve maximum efficiency. This solution will involve implementing a smart design in the future.

In the following the main algorithm of this paper as follows:

<u>Part 1 : preprocessing an Feature extraction</u> part **Start Loop**

1) Preprocessing of audio signal.
2) define objective function f(x), x = (x1, . . . , xd) T, and Set the initial value of population size n, α, γ, and N.gen
3) Initialize pulse rates ri and loudness Ai and population based on (1)
4) Evaluate and find x∗ where x∗ ∈ {1, 2, . . . , n}
 	
while t ≤ N_gen
for i = 1 to n
5) Adjust frequency (Equation (2))
If i≥2
6) Calculate (8) based on current iteration.
7) Select a solution among the best solutions (best feature of audio signal.
8) Generate a local solution around selected best (5).
end if
9) Evaluate objective function
 	
if (rand < Ai & f(xi))
update x_i^j based on (9)
end if
end for
 	
10) Rank the bats and find the current best x∗
 	
t = t + 1
.end while
Return best selected features
<u>Part 2: Classification stage :</u>
 	
- To Train ANN
Start Loop
1) Select type of optimization (Hybrid)
2) Choose the number of epochs.
3) Compute the error
Err= (actual value-Estimated value)/ (actual value)*100% %% using for estimation of prediction
End Loop
Display control signal

4. RESULTS AND DISCUSSIONS:

Our main goal is to accurately identify audio signal qualities for automation systems. The goal is to master and use BSSO. The system extracted features from the audio signal and used those elements to provide instructions. However, adding an ANN signal categorization component enhanced the system. Both the feature extraction and classification methods had over 94% accuracy. The BSSO feature extraction method [24] employed 100–500 iterations (Figure 2). The model is being conducted using a Windows 11 personal computer with a 2.5GHz Core i7 CPU, 16GB of RAM, and MATLAB 2023a software. The cumulative duration for the system to complete all assignments is three minutes. Table 1 illustrates the enhancement in the system's accuracy achieved by the Salp algorithm compared to the BAT techniques and the other techniques based on the same data of this paper from audio signals.

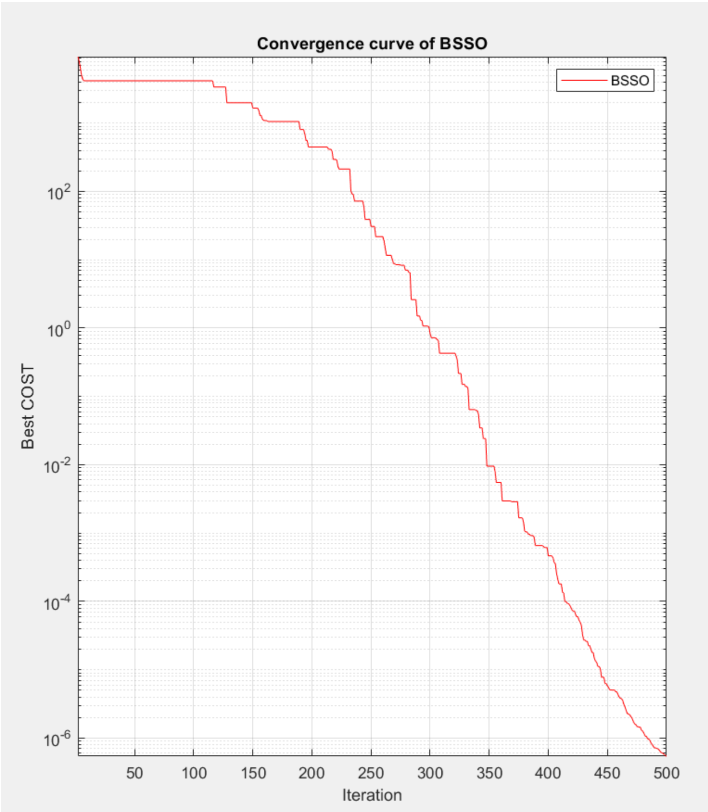


Figure 2: Convergence curve of BSSO

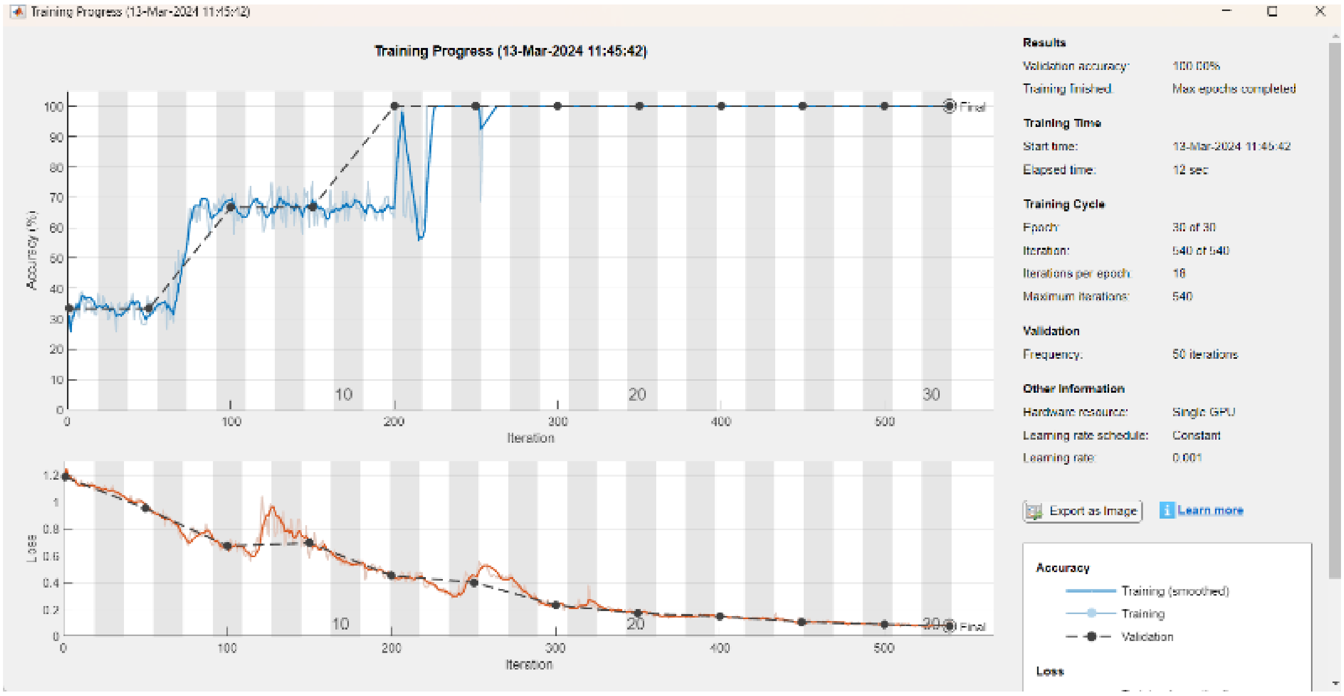
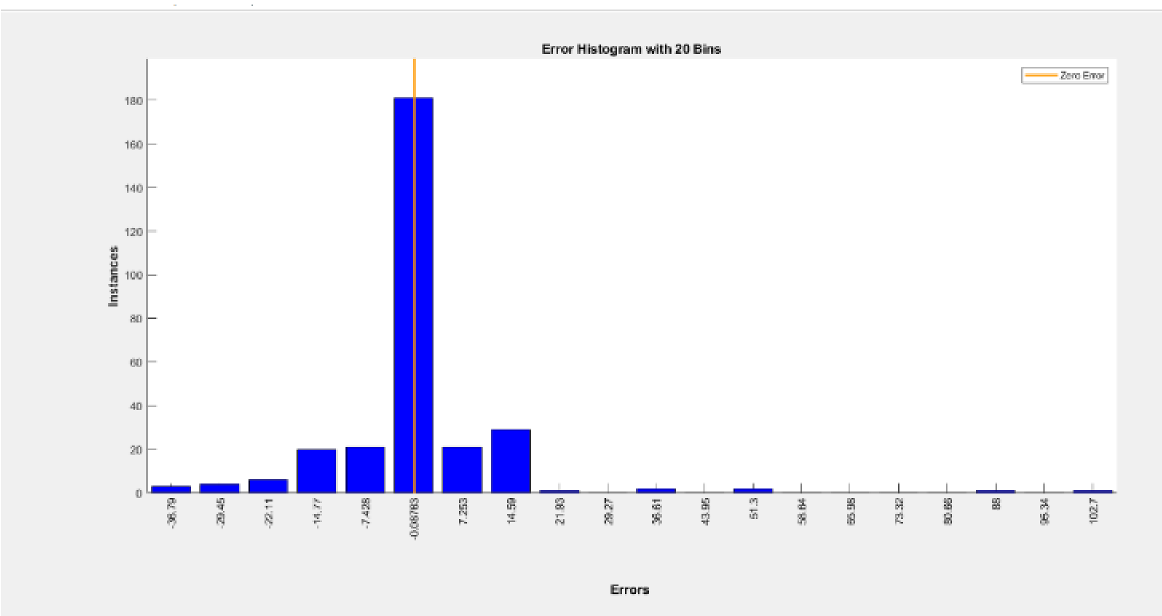


Figure 3: Training of classification based on ANN.

Every signal the smart home sensor unit emits is classified to manage equipment. Categorization uses ANN to classify. Figure 3 shows the ANN classification method and results. Figure 4 displays the classification model managed the smart system utilising audio signals with 98% accuracy and an error histogram. This model classifies audio files using ANN and feature extraction. On audio samples from various sensors, the approach works well.



The algorithm is characterised by its high speed and ability to extract practical values for feature extraction. Additionally, it enhances the model of speech values. Furthermore, the model has shortcomings in terms of the time required for data collection, particularly for historical data collection.

Table 1
Comparing BOA and BSSO results.

Method	Iteration	Training
BOA	200	80%
	500	98%
BSSO(this work)	200	95-98%
	500	99%
ACO-TSP [25]	200	88.9%
	500	95%
NGN[26]	200	88%
	500	96.8%

Figure 4: Errors histogram with 20 Bins of ANN.

5. CONCLUSION

AI technologies and Machine Learning (ML) algorithms, specifically optimization techniques, are crucial for developing innovative methods for designing and configuring smart buildings. The Digital device use has increased interest in connecting the Internet of Things (IoT) into houses, creating smart homes. Networked devices are becoming more prevalent in many situations such as smart buildings, cities, grids, and homes. There has been a rise in interest in the concept of constructing smart homes by IoT into homes because of the use of digital devices. It involves the fast spread of connected devices. Connected gadgets are becoming more common in smart buildings, towns, grids, and families. Bat Salp Swarm Optimization is used to integrate voice recognition with home automation systems to improve performance. This paper examines (BSSO), a data analysis method that generates analytical models automatically. Decisions made by voice recognition systems are enhanced by BSSO's feature selection approximation solution. The BSSO method enhances voice recognition accuracy and adds the ANN for classification. The findings supported the methods. The method is distinguished by its rapidity and capacity to get useful values for feature extraction. Furthermore, it improves the framework of speech values. Moreover, the model has deficiencies in terms of the duration needed for data collection, especially for the acquisition of historical data.

6. Declaration on Generative AI

During the preparation of this work, the author(s) used QuillBot program and wordtune for rephrasing and improving clarity of certain paragraphs, as well as Grammarly for grammar and spelling checks. All content generated or suggested by these tools was critically reviewed and edited by the authors. The author(s) affirm full responsibility for the accuracy, originality, and integrity of the final manuscript.

References

[1] F.M, Iván, P.F. Lamas, and T.M. F. Caramés. "Design, Implementation and Practical Evaluation of a Voice Recognition Based IoT Home Automation System for Low-Resource Languages and Resource-Constrained Edge IoT Devices: a System for Galician and Mobile Opportunistic Scenarios." IEEE Access, Vol.11, 2023, pp. 63623 − 63649.

[2] D.S, "Hyperparameter Optimization of CNN Using Genetic Algorithm for Speech Command Recognition." Advances in Data-driven Computing and Intelligent Systems, vol.2., 2023, pp. 123-135.

[3] J., L., and P., "Review on the Application Areas of Decision-Making Algorithms in Smart Homes." Augmenting Human Intellect. IOS Press,2023, pp.74-92.

[4] S., and M.: "Machine Learning-Integrated IoT-Based Smart Home Energy Management System." Handbook of Research on Deep Learning Techniques for Cloud-Based Industrial IoT, 2023, pp.219-235.

[5] R., R., M., and R. :"Genetic Algorithm-Based Optimization for Speech Processing Applications." Optimization Techniques in Engineering: Advances and Applications, 2023, pp.215-23.

[6] K., M.,and H.J.: "Low cost smart home automation via Microsoft speech recognition." International Journal of Engineering & Computer Science ,vol.13.3,2013, pp. 6-11

[7] D., and Q.Y.: "Neural network-based speech fuzzy enhancement algorithm for smart home interaction." Journal of Computational Methods in Sciences and Engineering Preprint, 2023, pp.1-12.

[8] M., R.: Speech recognition using Taylor-gradient Descent political optimization based Deep residual network. Computer Speech & Language, vol.78, 2023, pp.101442.

[9] R.HA, and K.R., "Grasshopper KUWAHARA and Gradient Boosting Tree for Optimal Features Classifications." Computers, Materials & Continua, vol. 72.2 ,2022.

[10] R.A, K.H and A.H.: "Design and Optimization of PID Controller based on Metaheuristic algorithms for Hybrid Robots." 2023 20th Learning and Technology Conference (L&T). 26-26 January 2023, pp.85-90.

[11] R., S.: "Automatic speech emotion detection using hybrid of gray wolf optimizer and naïve Bayes." International Journal of Speech Technology, 2021,pp.1-8.

[12] A., M.; R., and A. , "An overview of home automation systems". In: 2016 2nd international conference on robotics and artificial intelligence (ICRAI). IEEE, 1-2 November 2016, pp. 27-31.

[13] K., J., "Energy impacts of smart home technologies." Report A1801,2018. URL: https://www.aceee.org/sites/default/files/publications/researchreports/a1801.pdf.

[14] P., J., N., M., S.M., I., A., Y.,and A.. A Hybrid Bat-Crow Search Algorithm Based Home Energy Management in Smart Grid. In: Barolli, L., Javaid, N., Ikeda, M., Takizawa, M. Complex, Intelligent, and Software Intensive Systems. CISIS 2018. Advances in Intelligent Systems and Computing, Springer, Cham, 19 June 2018,vol 772.
[15] M., S., "Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems." Advances in engineering software, vol.114 ,2017, pp.163-191.
[16] KH., S.,. Energy efficient scheduling of smart home. In: Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications. Springer International Publishing, vol.33, 2019, p. 67-79.
[17] N., B.. "Improved salp swarm algorithm with mutation schemes for solving global optimization and engineering problems." Engineering with Computers, pp.1-23, 2022.
[18] B.Chaabane, S., "Improved salp swarm optimization algorithm: Application in feature weighting for blind modulation identification." Electronics ,pp.10.16 ,2021.
[19] V., H., and M.: "Improving speech recognition using bionic wavelet features." AIMS Electronics and Electrical Engineering, vol.4.2, 2020, pp. 200-215.
[20] T., A. M., A. M., and S. Ch. "Naive Bayes-guided bat algorithm for feature selection." The Scientific World Journal, vol.2013.
[21] Li., Y.:. "An improved bat algorithm based on lévy flights and adjustment factors." Symmetry, vol.11.7,2019,pp.925.
[22] A., S.. "Classification Model for Diabetes Mellitus Diagnosis based on K-Means Clustering Algorithm Optimized with Bat Algorithm." IJACSA) International Journal of Advanced Computer Science and Applications, 2023, pp.14.1.
[23] D., M.. "Dry bean cultivars classification using deep CNN features and Salp swarm algorithm based extreme learning machine." Computers and Electronics in Agriculture,2023,pp.204.
[24] G., A., G.B, and P.V. Ch.. "Electrocardiogram signal classification using VGGNet: a neural network based classification model." International Journal of Information Technology, vol.15.1,2023, pp. 119-128.
[25] R.H.M., K.H.R, and H.FA. "Brain tumors diagnosis and prediction based on applying the learning metaheuristic optimization techniques of particle swarm, ant colony and bee colony." *Procedia Computer Science*, vol.163, 2019, pp.165-179.
[26] R.H.M., K.H.R, and H.FA, and Hesham FA Hamed. "Brain cancer diagnosis and prediction based on neural gas network and adaptive neuro fuzzy." *Procedia Computer Science*, vol.163 ,2019, pp.518-526.