

Date of publication 2020 00, 0000, date of current version xxxx 00, 2020.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

A Healthcare Monitoring System for the Diagnosis of Heart Disease in the IoMT Cloud Environment Using MSSO-ANFIS

MOHAMMAD AYOUB KHAN, FAHAD ALGARNI

College of Computing and Information Technology, University of Bisha, Bisha, 67714, Saudi Arabia

Corresponding author: Fahad Algarni (e-mail: fahad.algarni@ub.edu.sa).

ABSTRACT The IoT has applications in many areas such as manufacturing, healthcare, and agriculture, to name a few. Recently, wearable devices have become popular with wide applications in the health monitoring system which has stimulated the growth of the Internet of Medical Things (IoMT). The IoMT has an important role to play in reducing the mortality rate by the early detection of disease. The prediction of heart disease is a key issue in the analysis of clinical dataset. The aim of the proposed investigation is to identify the key characteristics of heart disease prediction using machine learning techniques. Many studies have focused on heart disease diagnosis, but the accuracy of the findings is low. Therefore, to improve prediction accuracy, an IoMT framework for the diagnosis of heart disease using modified salp swarm optimization (MSSO) and an adaptive neuro-fuzzy inference system (ANFIS) is proposed. The proposed MSSO-ANFIS improves the search capability using the Levy flight algorithm. The regular learning process in ANFIS is dependent on gradient-based learning and has a tendency to become trapped in local minima. The learning parameters are optimized utilizing MSSO to provide better results for ANFIS. The following information is taken from medical records to predict the risk of heart disease: blood pressure (BP), age, sex, chest pain, cholesterol, blood sugar, etc. The heart condition is identified by classifying the received sensor data using MSSO-ANFIS. A simulation and analysis is conducted to show that MSSA-ANFIS works well in relation to disease prediction. The results of the simulation demonstrate that the MSSO-ANFIS prediction model achieves better accuracy than the other approaches. The proposed MSSO-ANFIS prediction model obtains an accuracy of 99.45 with a precision of 96.54, which is higher than the other approaches.

INDEX TERMS: Internet of Things, heart disease, LCSA, ANFIS, MSSO, Internet of Medical Things

1. INTRODUCTION

A trained practitioner with the adequate experience and expertise can identify and diagnose heart disease. The term heart disease, also referred to as cardiac disease, incorporates various conditions, the symptoms of which include high blood pressure, arrhythmia, stroke, and heart attack. The challenge for healthcare institutions is to provide quality treatment at a reasonable price. An incorrect clinical diagnosis and low-quality treatment may lead to inadequate outcomes. Healthcare institutions may use decision support systems (DSSs) as a tool for cost reduction. Healthcare typically includes a large number of patient records, multiple disease diagnoses, and resource management, etc. The pervasive development of the Internet of Things (IoT) and its use in medical research has improved the effectiveness of remote health monitoring systems [1]. The IoT is an

interconnection of various physical objects to continuously observe physical events [2, 3, 4]. Using IoT technology, a heart disease monitoring system can capture and transmit the patient's physical parameters to a remote healthcare facility center in real time [5]. The heart is the most important organ of the human body. Thus, heart disease is a significant health issue [6]. In the area of clinical research, heart disease prediction is a primary and challenging issue and is based on the observation of several symptoms that include chest pain, chest congestion, blood pressure, shortness of breath, and cold sweats etc. [8, 9]. To assist diagnosis, IoT sensor values are taken as input to predict heart disease.

A decade ago, diseases could only be diagnosed after a thorough physical and clinical examination. Today, a smartwatch can help identify health irregularities, for example, an elderly person's irregular heartbeat. If a health

monitoring device detects an erratic heart rhythm and the person is also in a motionless state, an alarm is activated. Many technologies such as RFID, Bluetooth and body area networks are currently used in healthcare applications [28]. For example, a smartwatch has a temperature sensor, an oximeter, accelerometer and a GPS which measures body temperature, blood oxygen saturation, changes in movement and the wearer's position, respectively. Sensors in clothing monitor ECG signals, EMG activity and blood pressure. The observed data is sent to the cloud for storage and is communicated to stake holders such as doctors, family members and hospitals via cell phone or personal digital assistant. Wearable devices have recently become increasingly popular with wide applications in health monitoring systems which has led to the growth of the Internet of medical things (IoMT) [1]. The IoMT consists of many kinds of medical devices connected to the computer system of a healthcare provider through the Internet. These devices can produce, store, analyze and distribute health data. IoMT products include wearables, remote patient monitors, sensor-enabled beds, infusion pumps and health tracking devices. The goal of IoMT is to increase patient satisfaction and the quality of care given by healthcare providers and insurers.

Machine learning (ML) and data mining techniques reduce computing costs and time. The diagnosis of medical illnesses and interventions to change a patient's lifestyle are one of the applications of ML. The heart disease is usually regarded as one which primarily affects the elderly; however, it is becoming more prevalent across different age groups.

As technology has advanced, medical institutions have significantly changed over the past few years. But those in rural areas have fewer medical services. E-health services are an efficient solution for reducing the risk of heart disease, mortality and clinical test expenses [26, 27]. Sensors which record blood strain, heart-beat and movement patterns are also used to track health.

Remote monitoring systems have become more effective in recent years. Remote monitoring system algorithms have evolved from simple algorithms to ones which are more complicated and informative. They now not only supply patient information such as the amount of sleep the person has had, but can also provide the patient with more detailed data. In recent studies, ML methods have been applied to present more complex knowledge about heart disease. These techniques are able to make predictions, identify anomalies in a patient and classify the data. Classification is the mechanism by which a disease is determined or by which a certain disease is identifiable.

To reduce the risks associated with heart disease, this requires serious attention in the healthcare system. To diagnose heart disease, a machine learning technique can be applied using appropriate classifiers. However, the accuracy of such classifiers has been found to be low. The proposed IoMT framework for the diagnosis of heart disease uses a

modified salp swarm optimization and adaptive neuro-fuzzy inference system to improve the performance of classifier. Several measurements are taken from medical records to assess a patient's risk of heart disease, such as blood pressure (BP), age, sex, chest pain, cholesterol, blood sugar level etc. Heart disease is identified by classifying the received data using MSSO-ANFIS. The contributions of this work are as follows:

- First, features from the pre-processed data are extracted using the metaheuristic crow search algorithm. However, this search strategy does not guarantee convergence due to poor search space exploration. To overcome this issue, a Levy-based crow search algorithm (LCSA) is explored.
- Second, in this framework, predictions are made on heart conditions using an extremely non-linear, complex and dynamic computational processes known as ANFIS. The learning parameters of ANFIS are optimized using MSSO to provide better results.
- The results of the simulation demonstrate that the MSSO-ANFIS prediction model obtains higher accuracy than the other approaches.

The paper is structured as follows: Section II presents survey of the related studies on heart prediction methods. Section III presents the proposed model and algorithms for heart disease prediction. Section IV presents the analysis on the simulation outcome, and Section V concludes the paper.

II. LITERATURE SURVEY

Al-Makhadmeh and Tolba [16] presented an IoT-based heart disease recognition system that uses a deep belief neural network model. The data collected were analyzed for missing values. The distribution of the data was analyzed by the authors. The authors also used normalized data using the studentized method. The features were extracted from noise-free data for use by the classifier using deep belief networks and a high-order Boltzmann machine. The authors reported a prediction accuracy of 99.03 which helps to minimize heart disease mortality.

Vivekanandan and Sriman [17] designed a complete model consisting of a modified differential evolution (DE) method, a fuzzy analytical hierarchy process (AHP) and a feedforward neural network (FNN). In order to choose the most important features, the modified differential evolution method was adopted. Furthermore, a reduced set of attributes was fed into an optimized model for a fuzzy AHP with FNN to predict heart disease. The simulation results presented are based on the modified differential evolution approach which has an accuracy of 83.

Uyar and İlhan [18] analyzed cardiac disease using a recurrent fuzzy network-based genetic algorithm. The UCI dataset was applied for the evaluation of the proposed heart disease algorithm. The data processing system was used to collect patient information and a fuzzy technique was used for further research.

TABLE 1: COMPARATIVE ANALYSIS OF THE RELATED WORK

S. No.	Authors	Use of IoT Device	Mobility	Methodology	Approach	Accuracy
1.	Al-Makhadmeh et al. [16]	Yes	No	Cloud based IoT wearable medical device	Higher order Boltzmann (HOB) with deep belief neural network (DBNN)	99.03%
2.	Vivekanandan, T. et al. [17]	No	No	Prediction of heart disease	Modified DE with Fuzzy AHP FFNN	83%
3.	Uyar, Kaan et al. [18]	No	No	Diagnosis of heart disease	GA based trained recurrent fuzzy neural networks (RFNN)	97.78%
4.	Ahmed, Fizar [19]	Yes	No	IoT sensor-based heart rate monitoring	k-nearest neighbors' algorithm	96%
5.	Nazari, Somayeh et al. [20]	No	No	A decision support System for likelihood of heart diseases	Fuzzy Analytic Hierarchy Process	-
6.	Mohan, S. et al. [21]	No	No	Prediction of cardiovascular disease	Hybrid random forest with a linear model	88.7%
7.	Ali, Liaqat [24]	No	No	A predictive diagnostic system for Heart Disease	χ^2 - statistical model and optimally configured DNN.	93.33
8.	Haq, Amin U. et al. [25]	No	No	Hybrid intelligent system framework for the prediction of heart disease	A Hybrid Intelligent System	89
9.	Nourmohammadi. et al. [29]	No	No	Diagnosis heart disease	ICA with meta-heuristic	94.03
10.	Rao, MadhuSudana et al. [30]	No	No	Hybrid Disease Diagnosis	Multiobjective optimization with evolutionary parameters	94.05
11.	Proposed	Yes	Yes	IoMT cloud environment for heart disease prediction	Modified Salp swarm optimization with ANFIS	99.45

Due to the efficient training that was accomplished by applying genetic operators, the recurrent network accurately classified and predicted the results based on the patient's data. The authors reported a recognition rate of 97.78 for a small set of data.

Ahmed et al. [19] developed a heart disease prediction method using the Internet of things framework. The proposed work utilizes a support vector machine (SVM). The proposed system obtained the cloud data from the WEKA framework. The patient information was processed using an SVM to predict heart disease. The authors reported their method achieved 97.53 accuracy for heart disease prediction. With an IoT device, their system collected cardiovascular data such as blood pressure, body temperature and heartbeat etc. This technique has been used to diagnose heart disease correctly. The proposed system recognized cardiac disease in minimal time, but when a large volume of data was used, accuracy was compromised.

Nazari et al. [20] proposed a fuzzy analytic hierarchy process (FAHP) to assess the probability of developing heart disease. The authors calculated weights for various criteria that affect cardiovascular development. The proposed system recommends tests only when a high probability of cardiac disease has been identified. The system is useful for medical practitioners in diagnosing the initial findings before prescribing any costly clinical tests. Hence, this system promises to reduce costs and resource consumption.

Mohan et al. [21] presented a technique which utilizes a random forest (RF) and linear model (LM) named HRFLM to identify important features through the use of ML techniques that improve the predictability of the heart disease. The technique exploits the best features of RF and

LM. The model is implemented using various combinations of features and known methods of classification. The authors reported a prediction accuracy of 88.7.

Ali et al. [24] proposed the χ^2 -DNN model for the prediction of the heart disease. The authors discussed the problems of underfitting and overfitting which impact on prediction. The authors used the χ^2 -statistical model to eliminate irrelevant features. The search is performed using deep neural networks. The authors compared the proposed model with existing models such as artificial neural networks(ANN) and deep neural network (DNN) and reported a prediction accuracy of 93.33%.

Haq et al. [25] proposed a system for prediction of heart disease that has many techniques such as SVM, k-nearest neighbour (K-NN), ANN, decision trees (DT), logistic regression (LR), fuzzy logic (FL), rough set, naïve Bayes (NB), and AdaBoost (AB). The performance of the classifier is evaluated using LASSO with k-fold cross-validation, feature selection, and the Relief and mRMR algorithms. The authors show the impact on accuracy and execution time due to the reduction in the set of features. The maximum accuracy achieved was 89.

Nourmohammadi-Khiarak et al. [29] presented a hybrid algorithm for heart disease diagnosis. The authors used an imperialist competitive algorithm with a metaheuristic technique to optimize feature selection. To classify heart disease, the authors used KNN. The authors evaluated the proposed technique on multiple datasets. The sensitivity of the proposed metaheuristic is high, and the accuracy achieved by this method is 94.03.

Rao et al. [30] proposed a hybrid metaheuristic technique using support vector machines and a multilayer perceptron

for heart disease prediction. The proposed technique has application in e-Healthcare and telemedicine. The authors evaluated the proposed technique on 11 different datasets. The results show the hybridization approach provides a good prediction accuracy across various datasets as compared to other methods in the literature.

A summary of the comparative analysis of the related work is given in Table 1 along with the proposed work. There are several studies [16, 17, 18, 19, 20, 21, 22] on machine learning that have been applied to heart disease prediction and other healthcare applications. However, there are fewer studies [25, 29, 30] which use hybrid metaheuristic techniques. A classification of algorithms used for optimization is shown in Fig. 1.

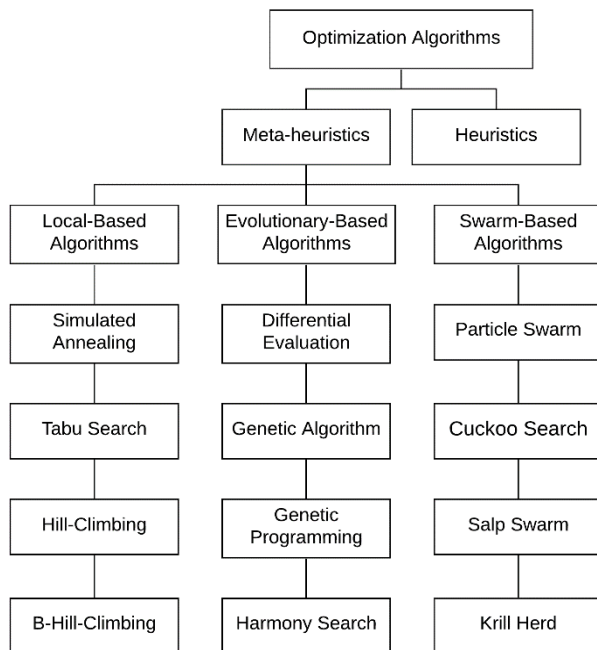


FIGURE 1. Classification of optimization algorithms

Metaheuristics algorithms are usually considered a part of machine learning and soft computing techniques. In metaheuristic algorithms, operation determination is repeatedly carried out until the search process converges or the preset stop conditions are met. The metaheuristic algorithms achieve higher accuracy compared to existing algorithms. The nature inspired salp swarm optimization (SSO) algorithm has showed its efficiency in various fields of engineering since its inception. In this work, we propose a combination of the MSSO and ANFIS techniques. The regular learning process in ANFIS is dependent on gradient-based learning which has a tendency to become trapped in

minima. The SSO is flexible, gradient-free and less prone to the local optima trap. The learning parameters are optimized utilizing MSSO to provide better results for ANFIS. We modified the salp swarm optimization algorithm to improve the search capability using Levy flight. Levy flight has the capability to achieve local optimization and efficiency for searching for a food location by maximizing the search space diversity.

III. PROPOSED METHODOLOGY

With early diagnosis and treatment, the risk of heart disease can be minimized. In this research, we propose an IoMT-based framework for the improved prediction of heart disease. The IoMT device collects patient information on their heart before and after the onset of heart disease. The health parameters of patients can be tracked remotely, constantly, in real-time and then stored and transmitted to the data center, such as the cloud, which significantly increases the efficacy, accessibility and cost efficiency of the healthcare system.

In order to identify the heart condition of the patient, the sensor data collected are passed on to the hospital administration. Training and testing phases are conducted to determine the condition of a patient's heart. The UCI data repository [23] is used to train the data values, followed by preprocessing, feature selection using the Levy crow search algorithm, and classification using MSSO-ANFIS. The final classification results indicate whether the heart condition of the patient is normal or abnormal. Based on these results, the necessary action will be taken by the doctor. The architecture of the proposed framework is presented in Fig. 2. The framework has five components, namely IoMT, network infrastructure, cloud infrastructure, dataset collection and a prediction system. The self-explanatory sub-component presents various functions of the main components.

A. Heart disease Prediction System

After obtaining the patient data, the heart condition of the patient is monitored and identified. The proposed IoMT framework for heart disease prediction performs training and testing to identify the heart condition of the patient. To train the system, the data values from the UCI data repository are used [23]. The data values from the dataset first undergo pre-processing, after which the feature selection process and classification are undertaken. After training, the patient's sensor data is tested by classifying and comparing the training results. Each phase of the training process is explained in the next sections.

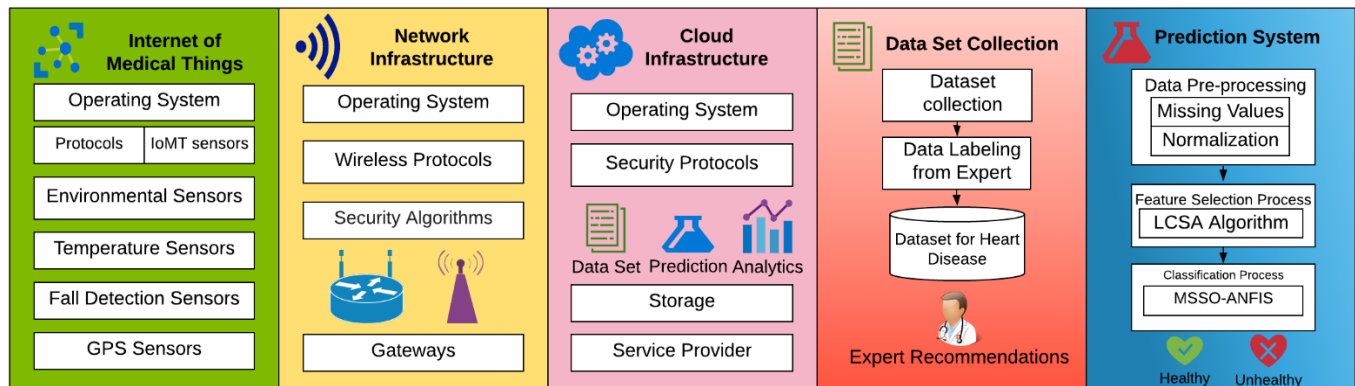


FIGURE 2. Proposed healthcare monitoring system for diagnosis of heart diseases

1) Preprocessing

First, the data values are collected from the UCI dataset, after which these are preprocessed which requires noise removal or the replacement of missing data. The noiseless data helps to effectively detect patterns associated with heart disease. The median studentized residual approach is applied to remove undesirable or noisy data because it investigates the correlation between the data in the dataset. This process of noise removal enhances the recognition process of heart disease.

a) Replacing Missing Values

The first step is to examine the data presented in the dataset and calculate the median for the missing values. The median value is determined by arranging the data into an increasing order, and thereafter the mid value is calculated. Missing and unrelated values are replaced with the aid of the median value.

b) Data Normalization

The data should be standardized within the range 0 to 1 after removing the missing values to reduce the difficulty of evaluating the patterns of heart disease. The residual studentized approach is used as a normalization process based on the calculation of the standard deviation [7]. The normalization process takes place through multiple data distributions and regression analysis for heart disease prediction. Equation (1) shows the regression equation for the normalization of data [7]:

$$P = \beta_0 + \beta_i Q + \epsilon_i, \text{ for } i = 1, 2, \dots, n \quad (1)$$

Every pair of random data samples complies with the above model. In equation (2), the regression model of the data is as follows [7]:

$$P_i = \beta_0 + \beta_i Q + \epsilon_i^* \quad (2)$$

The data samples in equation (2) with the equal variance and error are expressed as ϵ_i . The least squares values are represented by β_0 and β_i respectively. Equations (1) and (2) are used to determine the residual value which is represented by ϵ_i . The sample data and the average values are calculated using the deviation. In equation (3), the average value is estimated as follows [7]:

$$\mu = \frac{\sum_{i=1}^n Q_i}{N} \quad (3)$$

where N represents the frequency of data and Q_i is the input data. The formula for data normalization is given in equations (4) and (5) [7]:

$$n_r = \frac{\epsilon_i^*}{\sigma_i} \quad (4)$$

$$n_r = \frac{Q_i - \mu_i}{\sigma_i} \quad (5)$$

ϵ_i^* - residual value
 σ_i - variance

2) Feature Selection

Feature selection is the process of reducing the input dataset for further processing and analysis to discover the most appropriate part of the information. This process helps improve the performance of the prediction algorithms. The feature selection process also improves the process of modeling. In this work, features from the preprocessed data are extracted using the crow search algorithm [10], defined in the next section.

a) Levy-based CSA

The crow search algorithm (CSA) [10] belongs to the metaheuristic category of algorithms. Inherently, crows are considered to be intelligent birds due to their memory and their aptitude to hide food. This algorithm emulates the intelligence of a group of crows to search and recover additional food [10]. In this algorithm, the crows track other birds to learn where their food is stored. Once the bird leaves its nest, the crow eats or steals the food of the other birds. If another crow or another bird steals the first crow's food, it changes the food's hiding place so that the crow will not be a victim in the future. The crow uses its own knowledge and experience to anticipate the thief's actions to decide the safest place to protect its foods from thieves.

For certain search configurations, the crow search algorithm has shown its capability to provide the optimal solution. However, this search strategy does not guarantee convergence due to the poor exploration of the search space [10]. The search strategy of the crow algorithm poses major challenges when faced with extremely multimodal formulations. To overcome such difficulties, the proposed method uses a Levy-based CSA. In LCSA, Levy flight behavior is used to perform random movements [11, 12, 14].

The step size of the Levy flights is governed by a hard-tailed distribution of probability, typically referred to as the Levy distribution. This distribution is more effective than the uniform random distribution in the quest for space exploration.

P_n is the population which is randomly initialized with N crows. The search algorithm assumes Mx_{it} as the maximum count for iteration. The position of the e^{th} crow in the d^{th} dimensional search space in the k^{th} iteration is calculated as follows by equation (6):

$$C_{e,k} = [c_{e,k}^1, c_{e,k}^2, \dots, c_{e,k}^n],$$

$$e = 1, 2, \dots, P_n; k = 1, 2, \dots, Mx_{it} \quad (6)$$

Every crow is supposed to be able to remember the most visited location denoted as $B_{e,k}$ to hide the food before the start of the next iteration, which is calculated by equation (7) as follows:

$$B_{e,k} = [b_{e,k}^1, b_{e,k}^2, \dots, b_{e,k}^n] \quad (7)$$

To discover the place where the food was hidden, crow c chases crow e while crow e is unaware that crow c is following it. In this way, the objective of crow c is satisfied. Assume that crow e is aware of the existence of crow c , crow e uses a random path in order to safeguard the food. This action is represented by the random movement in the crow search algorithm [10]. In the proposed method, the random movement is done using the Levy flight method. The awareness probability (AP) determines the type of behavior that is regarded in each crow e . R_e is a randomly sampled number that is distributed uniformly between 0 and 1. Equation (8) presents a Levy walk based on the Levy distribution as follows [11, 12, 13]:

$$R_e = \text{Levy} \sim u = t^{-\lambda}, (1 < \lambda \leq 3) \quad (8)$$

The operation of the movement is expressed through following model as equation (9) [10]:

$$C_{e,k+1} = \begin{cases} C_{e,k} + R_e \cdot Fl_{e,k} \cdot (B_{c,k} - C_{e,k}) & R_e \geq AP \\ \text{move to random place,} & \text{otherwise} \end{cases} \quad (9)$$

The parameter flight length $Fl_{e,k}$ represents the amplitude of the movement from crow $C_{e,k}$ in the best possible position $B_{c,k}$ of crow c . R_e denotes the random number that has a distribution in the range [0, 1]. When the position of the crows changes, the positions are evaluated. Equation (10) is used to update the memory vector [10] as follows:

$$B_{e,k+1} = \begin{cases} C_{e,k+1}, & \text{if } O(C_{e,k+1}) \text{ is better than } f(B_{e,k}) \\ B_{e,k}, & \text{otherwise} \end{cases} \quad (10)$$

The objective function is denoted by $O(.)$ which is to be minimized.

3) Classification Using ANFIS

In this framework, heart condition predictions are made using an extremely non-linear, complex and dynamic computational process known as the adaptive neuronal fuzzy inference system, an effective combination of ANN and fuzzy logic [14]. Evidence shows that the ANFIS can predict

most of the applications by applying the proper number of rules. The ANFIS is a widely used algorithm, however it is computationally intensive and complex. The large number of inputs has an impact on the number of rules and tunable parameters which may grow exponentially. Furthermore, the regular learning process in ANFIS is dependent on gradient-based learning which has a tendency to become trapped in local minima [14]. Equation (11) and (12) contain the fundamental rules of ANFIS as follows [14].

$$\text{Rule 1: If } F_1 \text{ is } B_i \text{ and } F_2 \text{ is } Z_i, \text{ then} \\ \text{Rules}_i = sF_i + t_iF_{i+1} + u_i \quad (11)$$

$$\text{Rule 2: If } F_1 \text{ is } B_{i+1} \text{ and } F_2 \text{ is } Z_{i+1}, \text{ then} \\ \text{Rules}_{i+1} = s_{i+1}F_i + t_{i+1}F_{i+1} + u_{i+1} \quad (12)$$

where B_i, Z_i, B_{i+1} , and Z_{i+1} specify the fuzzy sets. F_i and F_{i+1} represent the different optimized feature values obtained from the previous step. The parameters $s_i, t_i, u_i, s_{i+1}, t_{i+1}$, and u_{i+1} are the predicted design parameters in the training process. These specified parameters are optimized utilizing MSSO to provide a better result. Fuzzy set parameter optimization using MSSO is elucidated briefly in the next section. The layers in ANFIS are explained as follows [14].

Layer 1: The first layer is named fuzzification. The fuzzification layer decides the input values and their membership function (MF), which are computed as follows using equation (13):

$$L_{1,i} = \mu_{f_i}(F_i) \quad (13)$$

where:

F_i – input to node i

μ_{f_i} – membership function

The output produced by every node denotes the degree of membership which is expressed by the input of MF. In this framework, the Gaussian kernel is used as the membership function which is expressed in equation (14) as follows:

$$\mu_{f_i} = \exp\left(-\frac{\|s_i - t_i\|^2}{2u_i^2}\right) \quad (14)$$

where s_i, t_i , and u_i are the MF parameters which are responsible for the curve of the membership function. The membership function parameters (s_i, t_i , and u_i) are called premise parameters.

Layer 2: This is a rule layer that generates the strengths for the rules. The product of all the input signals constitutes the output for this layer $L_{2,i}$ as expressed in equation (15).

$$L_{2,i} = H_i = \mu_{f_i}(F_i) \times \mu_{f_i}(F_{i+1}) \quad (15)$$

Layer 3: This layer is responsible for the normalization of firing strength. Node i computes \bar{H}_i , the normalization as shown in equation (16).

$$L_{3,i} = \bar{H}_i = \frac{H_i}{H_1 + H_2 + H_3 + H_4 + H_5 + H_6}, i = 1, 2, \dots, 6 \quad (16)$$

Layer 4: The normalized values are given to the fourth layer (consequence parameter set). The node function of the adaptive node in layer 4 is expressed as in equation (17).

$$L_{4,i} = \bar{H}_i \cdot \text{Rules}_i \quad (17)$$

where

\bar{H}_i – previous layer normalized firing strengths
 $Rules_i$ – system rules

Layer 5: The defuzzificated values in layer four are transferred to layer five to generate the final output. The total output is computed as the sum of all the input signals. In this layer, the circle node is labeled Σ as shown in equation (18):

$$L_{5,i} = \sum_i H_i Rules_i = \frac{\sum_i H_i Rules_i}{\sum_i H_i} \quad (18)$$

a) Parameter Optimization Using MSSO

The parameters of the ANFIS are optimized using the modified salp swarm optimization algorithm. Salp has a jellyfish type structure with a transparent body and tissue. They live in deep oceans and move by water force to find food, and they are organized into swarms called salp chains [15].

The salp population consists of two classes, namely leaders and followers. The front of the line is the leader, while the other salp are known as followers. We modified the salp swarm optimization algorithm to improve its search capability using Levy flight. Levy flight has the capability to achieve local optimization and efficiency in the search for a food location by maximizing the search space diversity. Furthermore, this can enhance salp swarm optimization to achieve effective optimization outcomes. To achieve this objective, we modified the salp leader position update as follows [15]:

$$V_j^1 = \eta W_j + r_1 \left((u_j - l_j)r_2 + l_j \right) * Levy \quad r_3 \geq 0 \quad (19)$$

$$V_j^1 = \eta W_j - r_1 \left((u_j - l_j)r_2 + l_j \right) * Levy \quad r_3 < 0 \quad (20)$$

where:

V_j^1 – position of the leader
 W_j – food source position
 u_j, l_j – upper and lower bound
 r_1, r_2, r_3 – random numbers

The position of the leader salp is updated using equation (19) and (20) [15]. In this proposed methodology, the food source position is multiplied by inertia weight η , which is denoted as the MSSO. Inertia weight $\eta \in [0,1]$, the MSSO algorithm's new parameter, speeds up convergence during the search. It also combines the ability to exploit and explore to escape from numerous local solutions in feature selection tasks and to accurately assess the solution. The r_1 balances the exploitation and exploration and is calculated using equation (21) [15]:

$$r_1 = 2e^{-\left(\frac{4e_i}{MI_x}\right)^2} \quad (21)$$

where:

e_i – current iteration
 MI_x – max. iteration

This controls the upcoming position and step size. Equation (22) is used to update the position of the follower salp as follows [15]:

$$V_j^i = \frac{1}{2} \omega t^2 + \delta_0 t, \quad i \geq 2 \quad (22)$$

where:

V_j^i – position of the i^{th} salp follower
 t – time
 δ_0 – initial speed

Variable ω is calculated in equation (23) as follows [15]:

$$\omega = \frac{\delta_{final}}{\delta_0} \quad (23)$$

Since the time in optimization is iterated, the difference is 1. Assuming $\delta_0 = 0$ leads to equation (24) [15]:

$$V_j^i = \frac{1}{2} (V_j^i + \eta V_j^{i-1}), \quad i \geq 2 \quad (24)$$

The inertia weight value is also multiplied in the update procedure of the follower position. In this way, the parameters are optimized so they compensate for the shift in frequency.

Time Complexity: The time complexity of the MSSO-ANFIS can be derived by combining the complexity of the MSSO and ANFIS. The complexity of the SSO algorithm [15] with t as number of iterations, d -dimension, n as number of solutions, and COF as the cost of the objective is calculated using equation (25)[15].

$$O(t(d * n + COF * n)) \quad (25)$$

The complexity of the ANFIS algorithm can be computed with reference to equations (11)-(18). The complexity of the major parts of the ANFIS can be categorised as $O(1), O(n), O(n^2)$. Rules such as $Rules_i = sF_i + t_i F_{i+1} + u_i$ have a running time $O(1)$ while $L_{1,i} = \mu_{f_i}(F_i)$ has a running time of $O(n)$. The complexity of the $\mu_{f_i} = \exp\left(-\frac{\|s_i - t_i\|^2}{2u_i^2}\right)$ can be expressed as $O(n^2)$.

The time complexity of ANFIS can be obtained by adding all three categories $O(1) + O(n) + O(n^2)$. The overall running time of MSSO-ANFIS can be computed as shown in equation (26)

$$O(t(d * n + COF * n) + O(1) + O(n) + O(n^2)) \approx O(n^2) \quad (26)$$

Limitations: The proposed MSSO has the advantage of being simple, with a rapid search speed, and ease of hybridization compared to the other optimization algorithms. While the new version is developed for hybridization between the two algorithms, MSSO and ANFIS have developed an exploration and exploitation ability but suffer from the problem of slow and premature convergence. Also, the probability distribution keeps on varying by generation.

IV. RESULTS AND DISCUSSION

Heart disease prediction using the proposed MSSO-ANFIS is verified using the JAVA and MATLAB platform. The performance of the MSSO-ANFIS methods is analyzed with the existing methods regarding several performance measures, as explained in the next section. Two datasets (Hungarian and Framingham) are used in the simulation [23].

The datasets consist of around seventy-six attributes, but in this simulation, we use thirteen important attributes as shown in Table 2.

TABLE 2: DATASET ATTRIBUTES

Attributes	Description
#AGE	Young, medium, old, very old
#SEX	1-male; 0-female
#CPT	Type of chest pain: 1-typical angina 2-atypical angina 3-non-anginal pain 4-asymptomatic
#RBP	resting blood pressure (Low, Medium, High, Very High)
#SCH	serum cholesterol in mg/dl (Low, Medium, High, Very High)
#FBS	blood sugar (fasting) > 120mg/dl (1-Yes; 0-No)
#RES	electrocardiographic results(resting) 0-normal 1-having ST-T wave abnormality 2-left ventricular hypertrophy
#MHR	Max. heart rate (Low, Medium, High, Very High)
#ELA	exercise induced angina (1-Yes; 0-No)
#OPK	Depression induced by exercise relative to rest (Low, Risk, Terrible)
#PES	the slope of the peak exercise ST segment 1-upsloping 2-flat 3-downsloping
#VCA	major vessels (0-3)
#THA	Thallium scan (3-normal; 6-fixed defect; 7-reversible defect)

A. Performance Analysis of LCSA

The performance of the presented LCSA technique for optimal feature selection is compared with existing CSA with respect to the fitness value against the number of iterations. A graphical representation of the analysis of the performance of the LCSA and CSA techniques is given in Fig. 3.

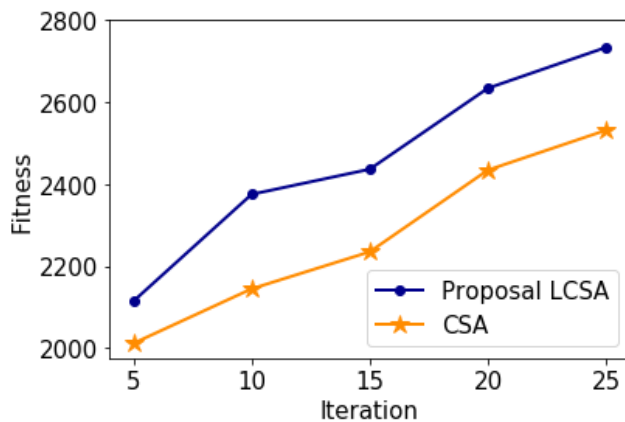


FIGURE 3. Fitness analysis of the proposed LCSA with existing CSA. Fig. 3 evaluates the performance of LCSA and CSA in terms of fitness value (FV). For a different number of iteration values, such as 5, 10, 15, 20, and 25, the existing CSA has FVs of 2011.35, 2144.325, 2234.637, 2434.127, and 2531.357, respectively, whereas the proposed LCSA has FVs of 2114.356, 2374.956, 2436.245, 2633.897, and 2733.587, respectively. Hence, this proves that the proposed LCSA technique has a higher FV than CSA.

B. Classification Performance

The template for the confusion matrix to measure the performance of the proposed classifier is given in Fig. 4 which shows the information on the actual and predicted classifications.

Fig. 4 shows the classification accuracy of the MSSO-ANFIS system for predicting the status of a heart patient.

		Predicted class		
		Heart Disease	No Heart Disease	
Actual class	Positive	TP	FN	<i>Sensitivity</i>
	Negative	FP	TN	<i>Specificity</i>
		<i>Precision</i>	<i>NPV</i>	<i>Accuracy</i>

FIGURE 4. Confusion matrix

The results obtained from the proposed algorithm are benchmarked based on several parameters, namely sensitivity, specificity, accuracy, precision, and F1-score which employ the terms true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The standard formula is used to determine the values of sensitivity, specificity, accuracy, precision, and F1-score.

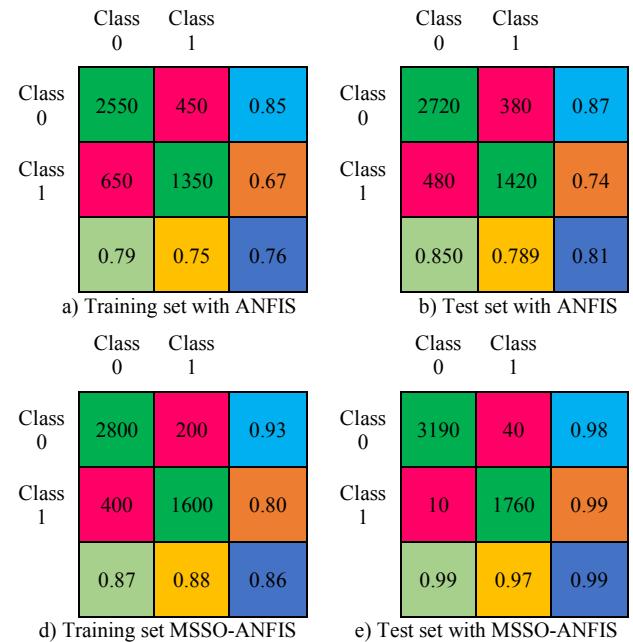


FIGURE 5. Confusion matrix for the dataset

Fig. 5(a) and (b) illustrates the outcomes obtained from the training set and test set respectively using the conventional ANFIS algorithm. In Fig. 5(a), it can be observed that of the 3200 heart disease patients (Class 0), 2550 were correctly classified who have the heart disease while 650 were misclassified. However, of the 1800 patients without heart disease (Class 1), 1350 were correctly classified while 450

were misclassified. In Fig. 5(b), which is obtained from the test set, it can be seen that of the 3200 heart disease patients, 2720 were correctly classified who have the heart disease while 480 were misclassified. Furthermore, of the 1800 patients without heart disease, 1420 were correctly classified while 380 were misclassified. The classification results obtained from the training set and test set using the proposed MSSO-ANFIS algorithm are presented in Fig. 5(d) and (e), respectively. From Fig. 4(d), this can be analysed that of the 3200 heart disease patients (Class 0), 2800 were correctly classified as having a heart disease while 400 were misclassified.

In Fig. 5(e), which is obtained from the test set, it can be seen that of the 3200 heart disease patients, 3190 were correctly classified who has the heart disease while 10 were misclassified. Furthermore, of the 1800 patients without heart disease, 1760 were correctly classified while 40 were misclassified.

The overall performance of the proposed MSSO-ANFIS algorithm has an average accuracy of 99.45% (Fig. 5e) for patients in both classes (Class 0 and Class 1) compared with 81.2% accuracy for the conventional ANFIS algorithm (Fig. 5b). The obtained results show that the proposed MSSO-ANFIS algorithm is superior than existing algorithms.

C. ROC plot analysis

The ROC curve is developed by plotting the false positive rate (FPR) versus true positive rate (TPR). The values of area under curve (AUC) are in the range of 0.5 to 1, where 0.5 indicates a bad classifier while values closer to 1 indicate a good classifier.

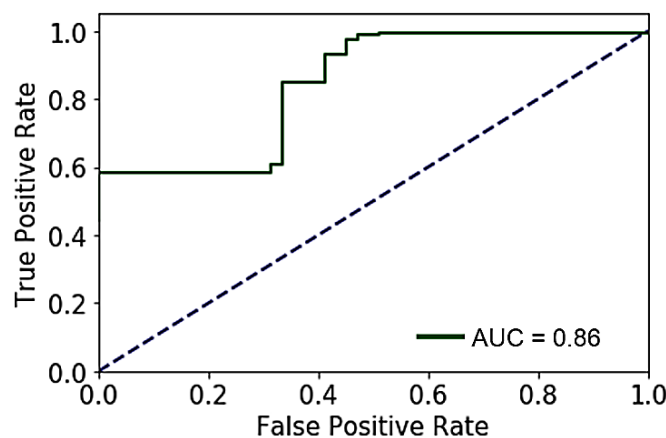


FIGURE 6. ROC for the training dataset using MSSO-ANFIS

Fig. 6 presents the ROC plotted for the results obtained by the proposed MSSO-ANFIS method on the training dataset. The obtained AUC is 0.86 which is close to 1.

We compare the value of AUC with some of the existing works. Table 3 shows that the proposed technique has better AUC values compared with the existing works [24, 25].

TABLE 3: COMPARATIVE ANALYSIS OF AUC

Authors	Methods	AUC
Liaquat Ali et al. [24]	χ^2 -DNN	0.95
Haq, A. et al [25]	Logistics regression	.88
Proposed	MSSO-ANFIS	0.99

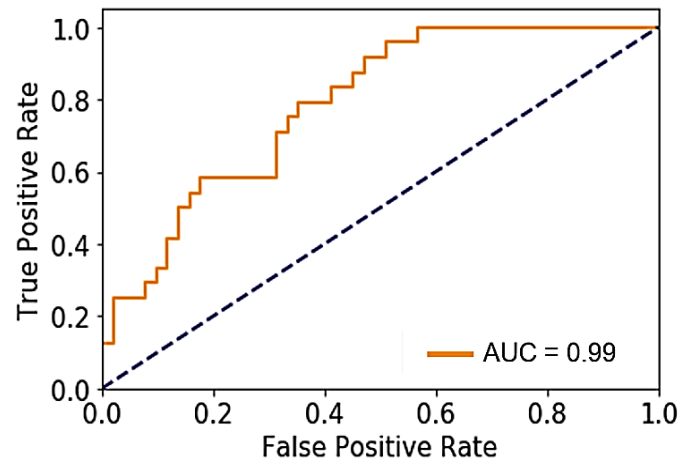


FIGURE 7. ROC for test dataset using MSSO-ANFIS

Fig. 7 plots the ROC for the test dataset after applying MSSO-ANFIS. It can be observed that the test data cover a wider area between the upper-left corner and diagonal of the curves than the area of the training dataset. The AUC obtained for the test dataset is 0.99 which is closer than the AUC obtained from the training dataset.

D. Comparative analysis of MSSO-ANFIS

Here, the performance of the proposed MSSO-ANFIS is analysed to determine its ability to identify the heart condition of patients. The proposed MSSO-ANFIS is compared with existing techniques, such as HOBDBNN [16], GA-RFNN [18], HRFLM [21], ANN-FuzzyAHP [22], χ^2 -DNN [24], logistics regression [25], ICA with meta-heuristic [29] and hybrid intelligent systems [30] in terms of sensitivity, specificity, precision, F1-score, accuracy, and classification error. These measures are computed based on the standard formula. Table 4 illustrates the evaluation results of the proposed and existing techniques. The existing HOBDBNN, GA-RFNN, HRFLM, ANN-FuzzyAHP, χ^2 -DNN, logistics regression, ICA with meta-heuristic and hybrid intelligent systems methods give sensitivity values of 96.43, 97.25, 92.8, 91.1, 85.36, 77, 96.27, and 93.49 respectively. Of these, the existing RFNN obtains the highest value for sensitivity, but the proposed MSSO-ANFIS gives a sensitivity value of 97.89, which is higher than the other approaches.

Similarly, in relation to the accuracy of heart disease prediction, the proposed MSSO-ANFIS obtains the highest level of accuracy (99.45) compared with the existing HOBDBNN (99.03), GA-RFNN (96.43), HRFLM (88.4), ANN-FuzzyAHP (91.10), χ^2 -DNN(93.33), logistics regression (89), ICA with meta-heuristic (94.03) and hybrid intelligent systems (94.05). Likewise, with respect to the remaining measures, the proposed MSSO-ANFIS gives the highest values. In addition, in Table 4, the MSSO-ANFIS technique is compared with the existing techniques for classification error.

TABLE 4: PERFORMANCE EVALUATION METRICS.

Authors	Methods	Performance Metrics (%)					
		Sensitivity	Specificity	Precision	F1-Score	Accuracy	Error
Al-Makhadmeh et al. [16]	HOBDBNN	96.43	97.76	95.89	98.495	99.03	0.97
Uyar, K. et al. [18]	GA-RFNN	97.25	95.8	94.64	95.93	96.43	3.57
Mohan, S. et al. [21]	HRFLM	92.8	82.6	90.1	90	88.4	11.6
Samuel, O. W. et al. [22]	ANN-FuzzyAHP	91.1	84.0	93.12	94.82	91.10	8.9
Ali, Liaqat et al. [24]	χ^2 -DNN	85.36	100	-	-	93.33	6.67
Haq, A. et al [25]	Logistics regression	77	98			89	11
Nourmohammadi et al. [29]	ICA with meta-heuristic	96.27	90.36	-	-	94.03	5.97
Rao, MadhuSudana et al.[30]	hybrid intelligent systems	93.49	94.77	94.07	94.05	94.05	5.95
Proposed model	MSSO-ANFIS	97.89	97.88	96.54	98.79	99.45	0.55

Each existing classifier's performance is assessed individually, and the results are recorded for further study. The proposed MSSO-ANFIS achieves the lowest error (0.55) when compared with the existing HOBDBNN (0.97), GA-RFNN (3.57), HRFLM (11.6), ANN-FuzzyAHP (8.9), χ^2 -DNN (6.67), logistics regression (11), ICA with meta-heuristic (5.97) and hybrid intelligent systems (5.95). Therefore, we can state that the MSSO-ANFIS is superior than the other classification techniques in terms of achieving and maintaining a heart patient monitoring system.

V. CONCLUSION AND FUTURE WORK

Healthcare monitoring and prediction systems play an important role in saving many lives, particularly when patients are remotely located. In this work, an IoMT-based healthcare monitoring system for heart disease prediction using MSSO-ANFIS was proposed. The proposed LCSA for feature selection achieved the highest fitness values for all iterations. The proposed MSSO-ANFIS technique gives higher values for precision, recall, F1-score, and accuracy and the lowest values for classification error when compared with the existing HOBDBNN, GA-RFNN, HRFLM, ANN-FuzzyAHP, χ^2 -DNN, logistics regression, ICA with meta-heuristic and hybrid intelligent systems methods. From these results, we can state that the proposed MSSO-ANFIS works well in accurately detecting the heart condition of patients and continuously monitoring patients' heart conditions. Based on the identified heart condition, the doctor can immediately give further treatment if necessary. In addition, a further study will be carried out with other feature-selection and optimization techniques to improve the effectiveness of predictive classifiers for the diagnosis of cardiovascular disease. Furthermore, the research proposed will be realized using the wearable technologies and available products in the market.

ACKNOWLEDGMENT

The authors would like to express their sincere thanks to the University of Bisha, Bisha, Kingdom of Saudi Arabia for the facilities and support provided during the research.

REFERENCES

- [1] Ning, Zhaolong & Dong, Peiran & Wang, Xiaojie & Hu, Xiping & Guo, Liang & Hu, Bin & Guo, Yi & Qiu, Tie & Kwok, Ricky. (2020). Mobile Edge Computing Enabled 5G Health Monitoring for Internet of Medical Things: A Decentralized Game Theoretic Approach, IEEE Journal on Selected Areas in Communications, 2020
- [2] K. Wei, L. Zhang, Y. Guo and X. Jiang, "Health Monitoring Based on Internet of Medical Things: Architecture, Enabling Technologies, and Applications," in IEEE Access, vol. 8, pp. 27468-27478, 2020, doi: 10.1109/ACCESS.2020.2971654.
- [3] Ahmadi, H., Arji, G., Shahmoradi, L. et al. The application of internet of things in healthcare: a systematic literature review and classification. Univ Access Inf Soc 18, 837–869 (2019). <https://doi.org/10.1007/s10209-018-0618-4>.
- [4] Md. Zia Uddin et al., A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare, Information Fusion, Vol. 55, 2020, pp. 105-115, doi:10.1016/j.inffus.2019.08.004.
- [5] Joyia, Gulraiz & Liaqat, Rao & Farooq, Aftab & Rehman, Saad. (2017). Internet of medical things (IOMT): Applications, benefits and future challenges in healthcare domain. Journal of Communications. 12. 240-247. 10.12720/jcm.12.4.240-247.
- [6] J. Vijayashree and P.H. Sultana, "A machine learning framework for feature selection in heart disease classification using improved particle swarm optimization with support vector machine classifier," Programming and Computer Software, vol. 44, no. 6, pp. 388–397, 2018.
- [7] Sarkar SK, Midi H, Rana MS. Detection of outliers and influential observations in binary logistic regression: An empirical study. Journal of Applied Statistics 2011;11:26-35
- [8] N. Prabakaran and R. Kannadasan, "Prediction of Cardiac Disease Based on Patient's Symptoms," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, 2018, pp. 794-799, doi: 10.1109/ICICCT.2018.8473271.
- [9] P. Sudeshna, S. Bhanumathi and M. R. A. Hamlin, "Identifying symptoms and treatment for heart disease from biomedical literature using text data mining," 2017 International Conference on Computation of Power, Energy Information and Communication (ICCPEIC), Melmaruvathur, 2017, pp. 170-174, doi: 10.1109/ICCPEIC.2017.8290359.
- [10] Alireza Askarzadeh, A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm, Computers & Structures, Vol. 169, 2016, Pages 1-12, <https://doi.org/10.1016/j.compstruc.2016.03.001>.
- [11] Anis Farhan Kamaruzaman et al., Levy Flight Algorithm for Optimization Problems – A Literature Review, Applied Mechanics and Materials Vol. 421 (2013) pp 496-501, doi:10.4028/www.scientific.net/AMM.421.496

- [12] Haklı H and Uğuz H. A novel particle swarm optimization algorithm with Levy flight. *Applied Soft Computing*. 2014, Vol. 23, pp. 333–345, doi.org/10.1016/j.asoc.2014.06.034.
- [13] Yang XS. Firefly algorithm, Levy flights and global optimization. In: Bramer M., Ellis R., Petridis M. (Eds.), *Research and Development in Intelligent Systems XXVI*: Springer, London; 2010. p. 209–218.
- [14] J. -. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, May–June 1993, doi: 10.1109/21.256541.
- [15] Seyedali Mirjalili, Amir H. Gandomi, Seyedeh Zahra Mirjalili, Shahrzad Saremi, Hossam Faris, Seyed Mohammad Mirjalili, Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems, *Advances in Engineering Software*, Volume 114, 2017, Pages 163–191, <https://doi.org/10.1016/j.advengsoft.2017.07.002>.
- [16] Z. Al-Makhadmeh and A. Tolba, "Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: A classification approach," *Measurement*, vol. 147, pp. 106815, 2019.
- [17] T. Vivekanandan and S. N. Iyengar, "Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease," *Computers in Biology and Medicine*, vol. 90, pp. 125–136, 2017.
- [18] Kaan Uyar, Ahmet İlhan, Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks, *Procedia Computer Science*, Vol. 120, 2017, pp. 588–593, doi.org/10.1016/j.procs.2017.11.283.
- [19] F. Ahmed, "An Internet of Things (IoT) application for predicting the quantity of future heart attack patients," *International Journal of Computer Applications*, vol. 164, no. 6, 2017, <http://doi.org/10.5120/ijca2017913773>
- [20] S. Nazari, M. Fallah, H. Kazemipour, and A. Salehipour, "A fuzzy inference-fuzzy analytic hierarchy process-based clinical decision support system for diagnosis of heart diseases," *Expert Systems with Applications*, vol. 95, pp. 261–271, 2018, doi.org/10.1016/j.eswa.2017.11.001
- [21] S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in *IEEE Access*, vol. 7, pp. 81542–81554, 2019, doi: 10.1109/ACCESS.2019.2923707.
- [22] O. W. Samuel, G. M. Asogbon, A. K. Sangaiah, P. Fang, and G. Li, "An integrated decision support system based on ANN and Fuzzy AHP for heart failure risk prediction," *Expert Systems with Applications*, vol. 68, pp. 163–172, 2017, doi.org/10.1016/j.eswa.2016.10.020.
- [23] Kaggle open dataset, URL: <https://www.kaggle.com/datasets>, accessed on Apr 15, 2020, 3:30AM.
- [24] Liaqat Ali, Atiqur Rahman, Aurangzeb Khan, Mingyi Zhou, Ashir Javed, and Javed Ali Khan, "An Automated Diagnostic System for Heart Disease Prediction Based on χ^2 Statistical Model and Optimally Configured Deep Neural Network", *IEEE Access*, vol. 7, pp. 34938–34945, 2019, DOI: 10.1109/ACCESS.2019.2904800
- [25] Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., & Sun, R.. A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. *Mobile Information Systems*, 2018, doi.org/10.1155/2018/3860146
- [26] M. Liu and Y. Kim, "Classification of Heart Diseases Based On ECG Signals Using Long Short-Term Memory," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, 2018, pp. 2707–2710. doi: 10.1109/EMBC.2018.8512761
- [27] Z. Wang, J. Chen, Z. Zheng and J. Wang, "Sample expansion-oriented LDBN heart disease risk forecast model," 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC), Limassol, 2018, pp. 1327–1332., doi: 10.1109/IWCMC.2018.8450376
- [28] Sara Amendola, Rossella Lodato, Sabina Manzari, Cecilia Occhiuzzi, and Gaetano Marrocco, RFID Technology for IoT-Based Personal Healthcare in Smart Spaces, *IEEE INTERNET OF THINGS JOURNAL*, Vol. 1, no. 2, pp. 144–152, 2014
- [29] Nourmohammadi-Khiarak, J., Feizi-Derakhshi, M., Behrouzi, K. et al. New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection. *Health Technol. Vo.* 10, pp. 667–678, 2020, <https://doi.org/10.1007/s12553-019-00396-3>
- [30] MadhuSudana Rao Nalluri, Kannan K., Manisha M., Diptendu Sinha Roy, Hybrid Disease Diagnosis Using Multiobjective Optimization with Evolutionary Parameter Optimization, Volume 2017, Article ID 5907264, 27 pages, doi:10.1155/2017/5907264



Mohammad Ayoub Khan (Senior Member, IEEE) received a Ph. D (Computer Engineering) from Jamia Millia Islamia, New Delhi, India and a Master of Technology (Computer Science and Engineering) from Guru Gobind Singh Indraprastha, New Delhi, India. Presently, he is an Associate Professor at the University of Bisha, Saudi Arabia with interests in the Internet of Things, RFID, wireless sensor networks, ad hoc networks, smart cities, industrial IoT, signal processing, NFC, routing in network-on-chip, real time and embedded systems. He has more than 14 years of experience in his research areas. He has published many research papers and books in reputed journals and international IEEE conferences. He contributes to the research community by undertaking various volunteer activities in the capacity of editor for many journals and as a conference chair.



Fahad Algarni, Dean of the College of Computing and Information Technology at the University of Bisha, Saudi Arabia. He received a PhD from Clayton School of Information Technology, Monash University, Melbourne, Australia. He was awarded a bachelor degree from the Department of Computer Science at King Abdul-Aziz University with First Class Honors. He also graduated with a Master of Information Technology (Computer Networks) from La Trobe University, Melbourne, Australia. His research interests include wireless sensor networks, cloud computing, systems' reliability, IoT, and cybersecurity.