

## Research Article

# Internet of Medical Things (IoMT) and Reflective Belief Design-Based Big Data Analytics with Convolution Neural Network-Metaheuristic Optimization Procedure (CNN-MOP)

**A. Sampathkumar** <sup>1</sup>, **Miretab Tesfayohani** <sup>2</sup>, **Shishir Kumar Shandilya** <sup>3</sup>, **S. B. Goyal**,<sup>4</sup>  
**Sajjad Shaukat Jamal**,<sup>5</sup> **Piyush Kumar Shukla**,<sup>6</sup> **Pradeep Bedi** <sup>7</sup>, and **Meshal Albeedan**<sup>8</sup>

<sup>1</sup>Department of Applied Cybernetics, Faculty of Science, University of Hradec Kralove, Hradec Kralove, Czech Republic

<sup>2</sup>Department of Information Technology, Dambi Dollo University, Dambi Dollo, Ethiopia

<sup>3</sup>School of Computing Science and Engineering, VIT Bhopal University, Bhopal, India

<sup>4</sup>City University, Petaling Jaya 46100, Malaysia

<sup>5</sup>Department of Mathematics, College of Science, King Khalid University, Abha, Saudi Arabia

<sup>6</sup>Computer Science and Engineering Department, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, (Technological University of Madhya Pradesh), Bhopal-462033, Madhya Pradesh, India

<sup>7</sup>Galgotias University, Greater Noida, Uttar Pradesh, India

<sup>8</sup>Department of Computer Science, Faculty of Engineering and Technology, Liverpool John Moores University (LJMU), Liverpool L3 3AF, UK

Correspondence should be addressed to Miretab Tesfayohani; miremsc2011@gmail.com

Received 10 November 2021; Accepted 29 January 2022; Published 18 March 2022

Academic Editor: Anastasios D. Doulamis

Copyright © 2022 A. Sampathkumar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent times, the Internet of Medical Things (IoMT) is a new loomed technology, which has been deliberated as a promising technology designed for various and broadly connected networks. In an intelligent healthcare system, the framework of IoMT observes the health circumstances of the patients dynamically and responds to backings their needs, which helps detect the symptoms of critical rare body conditions based on the data collected. Metaheuristic algorithms have proven effective, robust, and efficient in deciphering real-world optimization, clustering, forecasting, classification, and other engineering problems. The emergence of extraordinary, very large-scale data being generated from various sources such as the web, sensors, and social media has led the world to the era of big data. Big data poses a new contest to metaheuristic algorithms. So, this research work presents the metaheuristic optimization algorithm for big data analysis in the IoMT using gravitational search optimization algorithm (GSOA) and reflective belief network with convolutional neural networks (DBN-CNNs). Here the data optimization has been carried out using GSOA for the collected input data. The input data were collected for the diabetes prediction with cardiac risk prediction based on the damage in blood vessels and cardiac nerves. Collected data have been classified to predict abnormal and normal diabetes range, and based on this range, the risk for a cardiac attack has been predicted using SVM. The performance analysis is made to reveal that GSOA-DBN\_CNN performs well in predicting diseases. The simulation results illustrate that the GSOA-DBN\_CNN model used for prediction improves accuracy, precision, recall, F1-score, and PSNR.

## 1. Introduction

In the modern world, a better healthcare system is the main challenge for the growing population of the world. The vision of the Internet of Medical Things (IoMT) is to provide a better and more pervasive health monitoring system. The IoMT is the integration of medical devices through Wi-Fi and permits device-to-device (D2D) communication. In recent days, the most challenging issue is the time needed for web services. Three-dimensional (3D) video can be downloaded at sporadic intervals by keeping in mind the latest technological trends. The collected voluminous data with less delay are obtained for accurate data measurement. It will increase the device resource allocation ability and offers quicker speed for the heterogeneous networks. The IoMT comprises various heterogeneous networks, for instance, Wi-Fi, Bluetooth, ZigBee, and other cellular platforms. The D2D communication is the central part of the IoMT platform with high efficiency and reliability. The main traits of an intelligent healthcare system are to offer less delay and high throughput and reliability, which are very important for an effective and accurate diagnosis and consultation. The critical time analysis is the key parameter to be considered for emergency healthcare applications. The highly reliable and delay-tolerant communication and transmission of data was achieved through IoT-driven wearable devices [1].

Metaheuristic optimization approaches are employed in the partitioning clustering methods with which a dataset is partitioned as groups according to the specific measure considered as a fitness function [2]. This function has a greater impact on the nature of forming these groups. When an appropriate fitness function is selected, the process of partitioning is converted as an optimization problem. Here, partitioning is performed by minimizing the distance or maximizing the similarity between the patterns, or the frequency is optimized in N-dimensional space. These techniques are generally employed in various research works as they are able to clustering large-scale datasets, like signal/image processing for segmenting images, analyzing homogeneous users to classify groups, generating precise hidden equalizers, organizing humans effectively in robotics based on their actions, matching aftershocks in seismology from background conditions, obtaining high-dimensional data report, mining web text and recognizing image patterns in computer science domain, managing portfolio in control studies, classifying diseases in medical anthropology using the medical records of patients and investigations, distributing sensor nodes in a wireless sensor network for improving lifetime, and grouping publications in library based on the contents [3].

Nowadays, big data is extensively used for analyzing the huge data for the predictors, business people, and researchers to estimate the predictions with much accuracy than the traditional analysis. Big data is structured with the five dimension maps such as velocity, value, volume, veracity, and variety. Now the researchers are working to handle scalability and the high dimension of the databases with high processing needs. Volume defines the size of the data, and velocity defines the arrival of data streams

continuously from which most valuable information is gathered [4].

Moreover, in big data, the throughput, connectivity, and speed of computing have been enhanced for digital devices that improve the retrieval, progression, and fabrication of the data. Authenticity controls the information standard from different places. Variety designates the communicative path of information between different places, for example, the primary data incorporating the conventional structure. The example of data source consists of both structured traditional and relational data, along with that it also comprises of semi-structured, quasi-structured, and unstructured data such as sensor data, audio, video, text, and graph. In healthcare, big data helps in predicting epidemics, curing diseases, improving the survival rate, and avoiding unnecessary deaths. As the population in the world is increasing and the life span of individuals is long, there occur rapid changes in providing treatment, and many decisions are made based on these changes. With big data, patients take the right decision at the right time. With the data report of the patient, "proactive care" required for individuals is identified or the change required for them is analyzed so that degradation in their health can be avoided [5].

However, machine learning, as well as statistical approaches, undergoes a few necessary changes to maintain and follow some specific constraints while estimating high-dimensional data; moreover, existing issues are solved when input variables are reduced before data mining approaches are applied. Thus, there are two ways techniques are employed for dimensionality reduction in machine learning. While exploring the redundant input data or in feature selection, the essential variables are taken from the primary dataset. In the dimensionality reduction process, data are eliminated forming a new dataset with few input variables, where each column holds the combined input variables providing the same information as that of the input variables. In statistical modeling, this process is termed exploration [6]. The existing works are related to the data optimization with machine learning techniques, which does not improve the predictive accuracy with their optimization level.

The contribution of the work is as follows:

- (i) To present the metaheuristic optimization algorithm for big data analysis in the IoMT using gravitational search optimization algorithm (GSOA) and reflective belief network with convolutional neural networks (DBN-CNNs).
- (ii) To perform data optimization using GSOA for the collected input data. The input data were collected for the diabetes prediction with cardiac risk prediction based on the damage in blood vessels and cardiac nerves.

Big data in healthcare is being used to predict epidemics, cure disease, improve quality of life, and avoid preventable deaths. With the world's population increasing and everyone living longer, the models of treatment delivery are rapidly changing, and data are driving many of the decisions behind those changes. Big data can help patients make the right

decision in a timely manner. From patient data, analytics can be applied to identify individuals who need “proactive care” or need a change in their lifestyle to avoid health condition degradation. For example, patients in early stages of some diseases (e.g., heart failure often caused by some risk factors such as hypertension or diabetes) should be able to benefit from preventive care, thanks to big data.

In Section 2, the related research works are presented. Section 3 shows the proposed model for data optimization and diabetes data classification. Evaluation criteria are discussed in Section 4. The conclusion is finally presented in Section 5.

## 2. Related Works

The healthcare system involves machine learning (ML) approaches in the fields such as diagnosing, predicting, and surveillance. It is believed by the health monitoring agents that by using ML techniques, life can be saved [7]. Using SVM, diabetic nephropathy type 2 patients can be detected [8]. Moreover, earlier detection of abnormalities can be achieved using a decision tree-based approach when integrated with genotype and clinical data of several type 2 diabetic patients. Data can be classified gender-wise using support vector machine (SVM), Naïve Bayes (NB), decision tree (DT), and random forest (RF) approaches to produce satisfactory results. In [9], ML approaches are embedded with data mining profiling for gaining knowledge from the larger collection of information. Classification approaches like SVM, NB, RF, and logistic regression (LR) are used. In [10], for cardiovascular and cerebrovascular events, an automated prediction model was developed using heart rate variability (HRV) analysis. This was suitable for patients with 55 years and more who have higher risks. 10-fold cross-validation based on HRV features with data mining approaches such as NB, DT, RF, SVM, AdaboostM1 (AB), and multilayer perceptron neural network (MLPNN) were used for prediction. Reference [11] focused on antidiabetic drug failure and developed a prediction model to maintain an exponential increase in diabetic type 2 patients. SVM was used for training large-scale medical datasets. In [12], the risk of diabetic neuropathy was examined. When diabetes has been spread in the entire body, the nervous system is affected leading to cardiac arrest. Heart rate variability (HRV) was estimated using a multiscale Allan vector and the features of ECG helped in automatic detection using ML techniques. A graph-based machine learning system (GBMLS) was introduced to diagnose diabetic neuropathy effectively. In [13], a healthy and asthma patient was comparatively examined with alternative devices to obtain the feature vector of asthma patient with the guidelines of GINA. RF, AB-RF, and MLPNN were used to develop an ML prediction model. Specificity, sensitivity, and accuracy were the parameters considered. Leave-one-out (LOO) validation methods were used to train and test the dataset to eradicate the overfitting problem. In [14], a surveillance system was introduced to effectively monitor the impacts of dengue hemorrhagic fever (DHF) and the rate of *Aedes aegypti* mosquito infection using SVM, which depends on the

climate conditions and geographical area. In [15], a prediction system was designed for detecting influenza-like illness (ILI) at the earlier stages. Using NB and SVM classifiers not only produced better results but logistics regressions (LR) and sequential minimal optimization (SMO) also are suitable. In [16], an automatic coronary artery disease (CAD) diagnosis was introduced where this disease caused cardiac arrests. The diagnosis was made using tunable Q-wavelet transform (TQWT), and the heart rate signals were observed and monitored using raw ECG (electrocardiogram). With threefold cross-validation, least square support vector machine (LS-SVM) approach was used for classification. In [17], heart disease was predicted and analyzed using BagMOOV ensemble model with a multi-purpose-weighted voting approach. Reference [18] studied the mood disorder which is a psychological behavior of humans and investigated a psychiatry solution using ML techniques. Three ML approaches, namely, SVM, least average shrinkage and selection operator (LASSO), and relevance vector machine (RVM) were developed using MATLAB to predict the possibilities of a suicide attempt. In [19], risks, diagnosis, and prediction of breast cancer were examined with four ML approaches, namely, SVM, DT, NB, and k-nearest neighbor (KNN). In [20], the classification of UCI’s disease datasets was studied in which SVM integrated with endocrine-based particle swarm optimization (EPSO) and artificial bee colony (ABC) was employed. In [21], it was revealed that SVM with fruit fly optimization algorithm (FOA) was used in several medical datasets, namely, Wisconsin Parkinson’s dataset, breast cancer dataset, thyroid disease diagnosis, and Pima Indian diabetes dataset. In [22], a model for diabetes prediction utilizing machine learning approaches was predicted. The supervised machine learning algorithms used for the prediction model are decision tree, Naïve Bayes, artificial neural network, and logistic regression. Performance evaluation is performed in this method by employing parameters such as accuracy, recall, precision, and F1-score. In [23], the data-driven SmartWork system’s AI component was described, comprising the predictive models which are personalized and decision support tools. In these subsystems, long-term predictive models and data mining techniques are implemented for providing probabilistic prediction of particular risk indicators motivated towards decision-making and T2DM intervention, among other chronic conditions.

Based on the comparison discussed, it is not enough for both optimizations to improving the accuracy. So, this research aims to propose metaheuristic algorithm for data optimization of diabetic data and the predictive analysis for cardiac attack risk prediction. Here, the IoMT module is used for data collection and uses a gravitational algorithm for data optimization, and then the data classification is done using DBN-based CNN. Finally, the predictive analysis will be done using SVM on the basis of image predictive analysis.

## 3. System Model

This section discusses the proposed metaheuristic algorithms based on data optimization for diabetes data, which

leads to the predictive analysis of cardiac attack risk based on blood vessels and cardiac nerves and cardiac nerve damages. The dataset, obtained from public healthcare, contains more than 100,000 records comprising 55 attributes. Few among them are age, gender, race, number of procedures, number of medications, number of diagnoses, and readmissions.

The data have been initially collected using the IoMT module, and these data have been clustered for improving the optimization of data to be processed. Here we use metaheuristic algorithms for data optimization in which gravitational search optimization algorithm has been used. Then using these optimized data, the diabetic data have been classified for identifying the abnormal range. Here we establish a deep belief network (DBN), where the classification is carried out using CNN. By this classification, the normal range and abnormal range of diabetes have been classified. The normal diabetes range has been updated to the hospital database, and for the abnormal diabetes range, the cardiac nerve and blood vessel damages have been analyzed using SVM-based image predictive analysis. The proposed diabetic data analysis has been given in Figure 1. In the IOMT module, a dataset for diabetes is collected and the data from the dataset are clustered in the data clustering phase, and the optimization of data is done by metaheuristic data optimization by using GSOA. The optimized data are classified by the classification algorithm DBN-based CNN. From the classification results, the normal range of diabetes is sent to the hospital database, whereas an abnormal range of diabetes is sent to the analysis of predicting cardiac attack risk, which will be done by various parameters.

**3.1. Metaheuristic-Based Gravitational Search Optimization Algorithm (GSOA).** Gravitational search optimization algorithm (GSOA) is a stochastic population-based metaheuristic approach that was developed based on Newton's laws of gravity and motion [24]. Originally, the basic GSOA model was developed to find a solution for continuous optimization problems. A set of agents/objects were introduced in the search space within dimension to determine an optimal solution where the principle of Newton's laws was followed. Here, the position of every agent describes a candidate solution  $X_i$ , which is a vector in the search space. An agent whose performance is higher obtains more gravitational mass as heavier objects gain more attraction radius. In the life span of GSOA,  $X_i$  is adjusted successively by an agent with the positions of best agents in KGSOA adapting Newton's laws. To explain in detail, a system with  $s$  agents is assumed, where the position of the agent is given by

$$X_i = (x_i^1, \dots, x_i^d, \dots, x_i^n); \quad i = 1, 2, \dots, s, \quad (1)$$

where  $x_i^d$  represents the position of agent  $i$  in dimension  $d$  with search space dimension  $n$ . For every agent, the gravitational mass after estimating the current data fitness is computed as given in the following equations:

$$q_i(t) = \frac{\text{fit}_i(t) - \text{worst}(t)}{\text{best}(t) - \text{worst}(t)}, \quad (2)$$

$$M_i(t) = \frac{q_i(t)}{\sum_{j=1}^s a_j(t)}, \quad (3)$$

where  $M_i(t)$  and  $\text{fit}_i(t)$  are the gravitational mass and fitness value of  $i^{\text{th}}$  agent, respectively, at time  $t$ . The  $\text{best}(t)$  and  $\text{worst}(t)$  are given in the following equations:

$$\text{best}(t) = \min_{j \in (1, \dots, s)} \text{fit}_j(t), \quad (4)$$

$$\text{worst}(t) = \max_{j \in (1, \dots, s)} \text{fit}_j(t), \quad (5)$$

Agent acceleration is estimated by adding the forces of every agent present in the set of KGSA based on the gravitational law, which is determined by using equation (6) and agent acceleration which is estimated by the motional law is expressed in equation (7):

$$F_i^d(t) = \sum_{j \in K_{\text{best}}, j \neq i} \text{rand}_j G(t) \frac{M(t)M_i(t)}{R_{ij}(t) + z} \quad (6)$$

$$\cdot (x_j^d(t) - x_i^d(t)),$$

$$\begin{aligned} a_i^d(t) &= \frac{F_i^d(t)}{M_i(t)} \\ &= \sum_{j \in K_{\text{best}}, j \neq i} \text{rand}_j G(t) \frac{M(t)M_i(t)}{R_{ij}(t) + z} (x_j^d(t) - x_i^d(t)), \end{aligned} \quad (7)$$

where  $r$  and  $j$  are random numbers distributed evenly ranging between  $[0, 1]$ ,

- (i)  $\varepsilon$ , a small value, helps to get rid of division by zero error when  $R_{ij}(t)$  is zero.
- (ii)  $R_{ij}(t)$  represents the Euclidean distance of agents  $i$  and  $j$ , defined as  $\|X_i(t) - X_j(t)\|_2$ .
- (iii)  $K_{\text{best}}$  indicates the set of first agents in KGSOA having best fitness value and higher gravitational mass, where KGSOA is the time function which is initially assigned a  $K_{\text{initial}}$  value and gets decreased with time.
- (iv)  $G(t)$  is the gravitational constant which initially takes  $G_{\text{initial}}$  value and decreases with time till  $G_{\text{end}}$  is reached as given in the following equation:

$$G(t) = G(G_{\text{initial}}, G_{\text{end}}, t). \quad (8)$$

Then, agent velocity and next position are estimated using the following equations:

$$v_i^d(t+1) = \text{rand}_i * v_i^d(t) + a_i^d(t), \quad (9)$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1). \quad (10)$$

**3.2. Establishing DBN Architecture.** Generally, when compared with a general neural network, deep neural network (DNN) is superior. In DNN, the input layer is separated

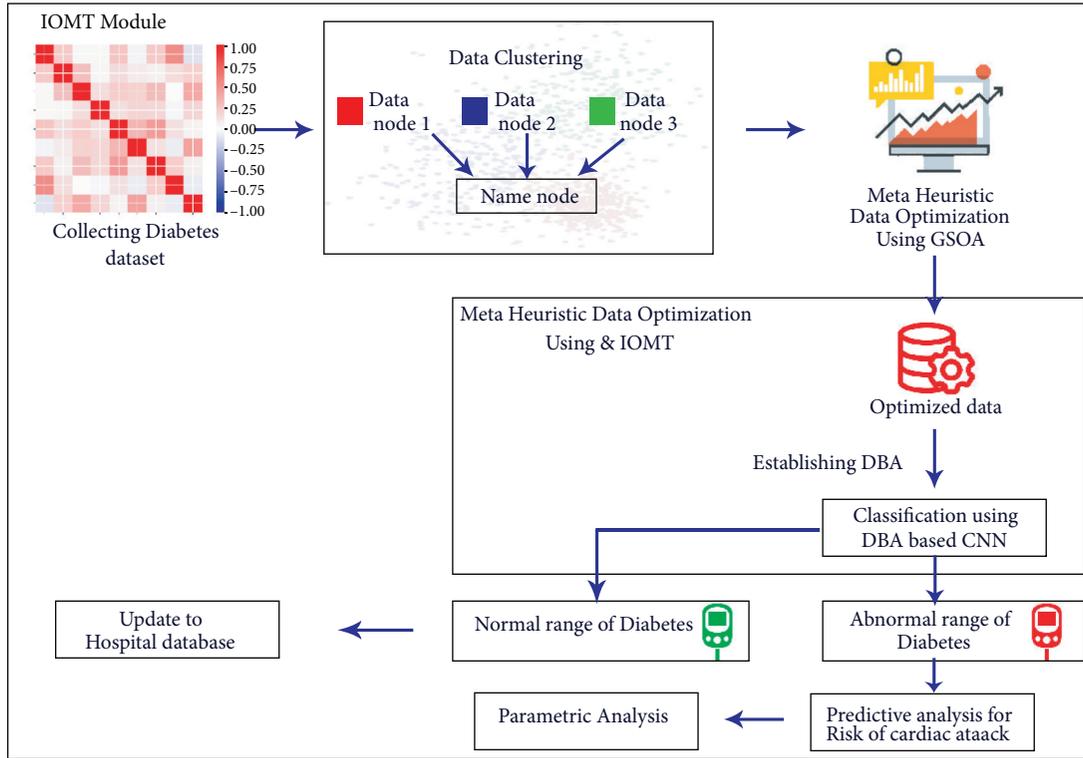


FIGURE 1: Proposed design for metaheuristic algorithms for data optimization with the IoMT.

from the output layer by several hidden layers. The method of training the networks differs. Particularly, in DNN, the unsupervised learning method is used for adjusting the weights in hidden layers, and the network is capable of identifying the optimal features from the input provided. Hence, DNN is flexible and enables high-order modeling of the nonlinear complex relationship between input provided to the network and output produced. The advantage of using DNN for learning the features and classifying data is proved with various pattern recognition applications like vision, speech, and natural language processing [24–26]. The outcomes observed have motivated the researchers to develop automated pattern recognition systems using DL methods. In spite of some challenges, developing a suitable training model for DNN is still a challenging one.

With the presence of several hidden layers in DNNs and numerous parameters, training has to be done with utmost care. In the process of training, the feature detection layers are randomly initialized. In DNN, a series of generative models, namely, a single visible input and a hidden layer are considered for initializing weights and are trained without taking discriminative information into account. At last, the standard back propagation approach is used to train DNN discriminatively. From the investigation, it is observed that the standard gradient-based random initialization approach used for initializing weights of the network produces very little performance in DNN containing more than two layers. As DNNs with several parameters and numerous hidden layers increase the computational complexity, training is much slower and even gets stuck with local minima providing unexpected results. Here, the parameters are initialized during

unsupervised pretraining, so that the process of optimization ends with local minima of the cost function. The architecture of DBN with RBM is given Figure 2 [24].

In RBM, the energy function  $E$  taking parameters  $v$  and  $h$  representing a pair of visible and hidden vectors, respectively, has the general form with weight matrix  $W$  as in

$$E(v, h) = -a^T v - b^T h - v^T W h, \quad (11)$$

where  $a$  and  $b$  indicate the bias weights for visible and hidden units, respectively. With  $v$  and  $h$  in terms of  $E$ , the probability distribution  $P$  is given by

$$P(v, h) = \frac{1}{Z} e^{-E(v, h)}. \quad (12)$$

Here the normalizing constant  $Z$  is given by

$$Z = \sum_{v, h} e^{-E(v, h)}. \quad (13)$$

Moreover, the probability of  $v$  over hidden units is the sum of above given equations and is given by

$$P(v) = \frac{1}{Z} \sum_h e^{-E(v, h)}. \quad (14)$$

Log-likelihood difference of training data in terms of  $W$  is estimated as

$$\begin{aligned} \sum_{n=1}^{n=N} &= \frac{\partial \log P(v^n)}{\partial w_{ij}} \\ &= v_i h_{j\text{data}} - v_i h_{j\text{model}}. \end{aligned} \quad (15)$$

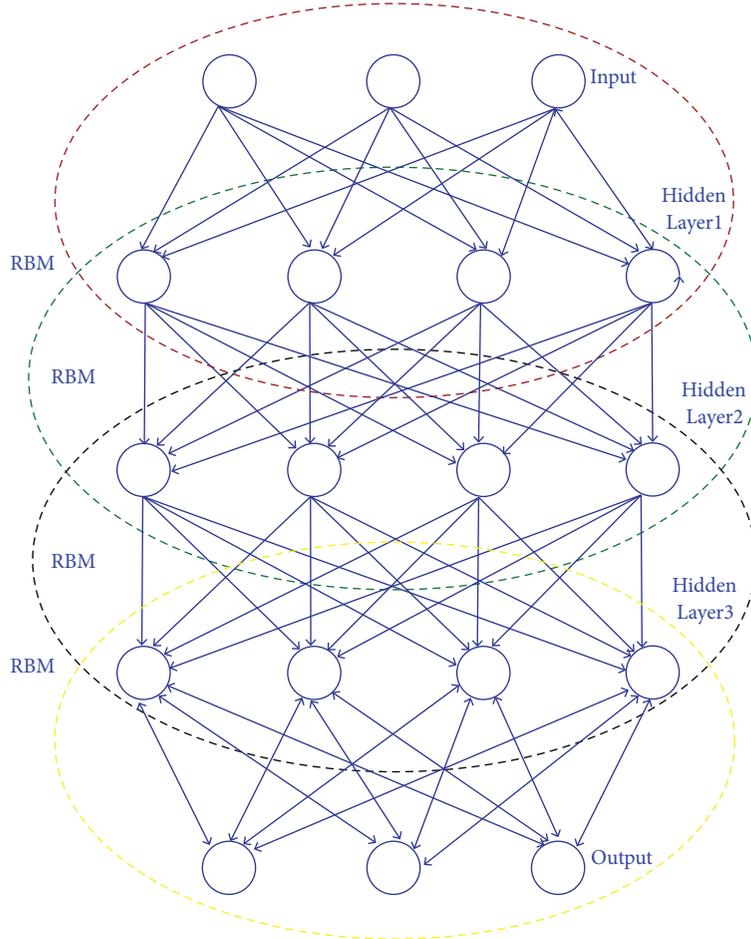


FIGURE 2: Architecture of the DBN with RBM.

where  $v_i h_{j\text{data}}$  and  $v_i h_{j\text{model}}$  represent the values expected for data and distribution model, respectively. For log-likelihood-based training data, network weights are computed using the learning rate  $\varepsilon$  as in

$$W_{ij} = \varepsilon (\langle v_i h_{j\text{data}} \rangle - \langle v_i h_{j\text{model}} \rangle) (v_i h_{j\text{data}}). \quad (16)$$

As neurons are not connected at either the hidden or the visible layers, it is possible to obtain unbiased samples from  $v_i h_{j\text{data}}$ . Further, the activation of hidden or visible units is conditionally independent for given  $h$  and  $c$ , respectively. For given  $v$ , conditional property is described in:

$$P(h|v) = \prod_j P(h_j|v), \quad (17)$$

where  $h_j \in \{0, 1\}$  and the probability of  $h_j = 1$  is given in

$$P(h_j = 1|v) = \sigma \left( b_j + \sum_i v_i W_{ij} \right). \quad (18)$$

Here the logistic function  $\sigma$  is specified as in

$$\sigma(x) = (1 + e^{-x})^{-1}. \quad (19)$$

Likewise, when  $v_i = 1$ , the conditional property is estimated by

$$P(v_i = 1|v) = \sigma \left( a_i + \sum_j W_{ij} h_j \right). \quad (20)$$

Generally, with  $\langle v_i, h_j \rangle$ , the unbiased sampling is not straightforward; however, it is applicable for reconstructing the first sampling of  $v$  from  $h$  and then Gibbs sampling is used for multiple iterations. With this Gibbs sampling, every unit of the hidden and visible layers is updated in parallel. At last, with  $\langle v_i, h_j \rangle$ , the proper sampling is computed by multiplying the expected and updated values of  $h$  and  $v$ . RBM weights can be used for initializing feedforward neural networks. Among several supervised and unsupervised approaches, this work focuses on DBN as it is commonly used for classifying diabetic data.

Convolutional neural network (CNN), which is deep learning knowledge-based neural network model, is a well-described and widely used technique for classifying images. It incorporates linear convolutional layer (conv), a fully connected layer (FC), this model consists of nonlinear function with an activation function above the linear function. This nonlinear function impinges on every component of input and pooling layer and minimizes the size of the final outcomes. Multiple perceptrons are used to analyze the image inputs, and it is trained with learnable weights and

bias value, which is used in several parts of images to segregate pixel values. One main advantage of using CNN is it uses a local spatial domain for the input images; in addition, it also shares a few sharable parameters and a lesser number of weights. This technique is predominately more efficient than other models due to less computation complexity and less usage of memory. The CNN architecture is shown in Figure 3 [25].

**Convolutional layer:** basically, the image taken as input is resized to  $3 \times 224 \times 224$ , which is the standard size for the CNN model. The resized image undergoes a stack of multiple layers as convolutional layers of various receptive fields. In the convolution layer, the basic operation is a convolution where several sequential mathematical operations are performed on convoluting sliding kernel matrix over the input matrix by which feature data are extracted and then those data are mapped to the successive layers. At each co-ordinate, an element-wise matrix multiplication carries out, and the outcomes are combined to obtain a feature matrix. Convolution recurrent is a distinguishable kind of linear model that assists in various platforms such as image processing, statistics, and physics. Convolution is estimated by more than one axis. Where, two-dimensional  $I$  is the input image and  $K$  kernel filter, respectively, then the convoluted image is calculated by

$$S(i, j) = \sum_m \sum_n I(m, n)k(i - m, j - n). \quad (21)$$

**Pooling layer:** the pooling layer, termed also as pool, is the successive layer of the convolutional layer whose main function is to minimize the spatial domain representation thereby reducing the network computations performed. Generally, in CNN, the pooling kernel is  $2 \times 2$  in size with a stride of 2.

**Fully connected layer:** this FC layer is replicated in CNN due to convolution.  $n1 \times n2$  is its usual size, where  $n1$  and  $n2$  are the size of the input tensor ( $7 \times 7 \times 512$ ) and output tensor, respectively, which is normally an integer.

**Dropout:** this layer is abbreviated as ‘‘Drop.’’ Usually, it is employed for eliminating the input overfits; the basic role is to improve the conjecture of deep learning methods. Normally, it assigns the weights to the connected network nodes.

**Softmax:** normally represented as ‘‘ $\sigma$ ,’’ is a deep learning model followed by many layers and the convolutional layer is followed by a ReLU layer which determines the nonlinearity of CNN and improves it.

The convolution layers present within both the pooling layers possess the equal channel number, kernel size, and stride. Actually, collecting two  $3 \times 3$  convolution layers and three  $3 \times 3$  convolution kernels is equal to a single  $5 \times 5$  and  $7 \times 7$  convolution layers, respectively. Stacking 2 or 3 small convolution kernels works much quicker than a single huge convolution kernel. Moreover, parameter numbers have

been minimized. ReLU layers, which are inserted among undersized convolution layers, are really useful.

The input data and their corresponding map images are  $S = (S^{(1)} \dots S^{(N)})$  and  $M = (M^{(1)} \dots M^{(N)})$ , respectively. The major objective is to design a model which maps  $S$  to  $M$  with the help of some training data. This is modeled as a probabilistic approach by learning the model of distribution over labels which is represented by

$$P(n(M, i, w_m)|n(S, i, w_s)), \quad (22)$$

where  $n(I, i, w)$  is a patch with  $w * w$  size for image  $I$ , focused on pixel  $i$ . Here,  $W_S W_S$  is preferred to be higher so that more contextual information can be extracted. Its functional form  $f$  is given by

$$\begin{aligned} f_t(s) &= \sigma(a_1(s)) \\ &= P\left(m_1 = \frac{1}{s_1}\right), \end{aligned} \quad (23)$$

where  $a_i$  and  $f_i$  represent the sum of input for the  $i^{\text{th}}$  output and significance of  $i^{\text{th}}$  output component, respectively.  $(x)$ , a logistic utility, is expressed by

$$(x) = \frac{1}{1 + \exp(-x)}. \quad (24)$$

CNN alongside softmax output unit is used for multi-class marking. The softmax output is a vector of size  $L$ , which demonstrates the conveyance more than potential marks for pixel  $i$ . Along these lines for multiclass marking, if the path from pixel  $i$  to output unit  $l$  is thought of, the recomposed condition is given in

$$\begin{aligned} f_{il}(s) &= \frac{\exp(a_{il}(S))}{z} \\ &= P(m_i = l|s), \end{aligned} \quad (25)$$

where  $f_{il}(s)$  is the prediction probability, where pixel  $I$  is mapped to label  $j$ . The advantages of the proposed method are summarized as follows:

- (i) First, CNN possibly handles a huge amount of labelled data from various domains.
- (ii) Second, it is faster when paralleled with graphics processing unit (GPU). Hence this is also extended for a greater number of pixels. The kernel size has been minimized to simulate the training data, and the kernel size minimization is performed by the computational learning process of the proposed method.
- (iii) Every patch in training data has been given by initiative sigma. Due to a large number of training patches, optimization becomes complicated. This can be done using a binary classifier that uses minimum patches. Few of the hyper parameters have been altered to some extent. The analysis over sensitivity has been defined by hyper parameters for them to be tuned with higher accuracy.

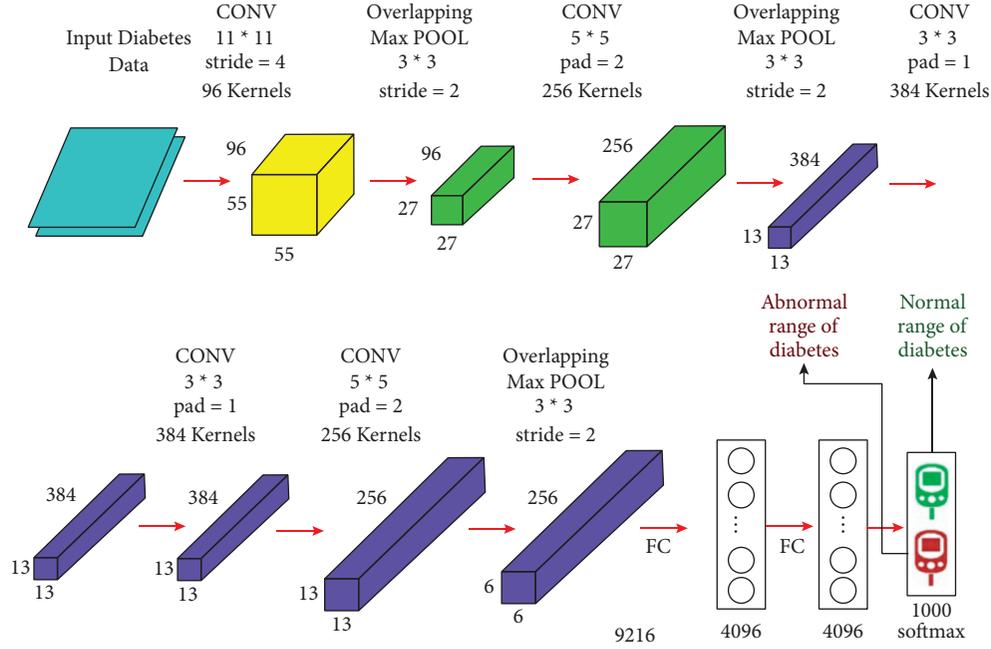


FIGURE 3: Architecture of the convolutional neural network.

The output from the softmax layer will be classified to be normal and abnormal ranges of diabetes, and the abnormal range of diabetes will be carried out for the predictive analysis of cardiac attack based on the cardiac nerve and blood vessel damage. This has been analyzed by using SVM-based image predictive analysis.

**3.3. SVM-Based Image Predictive Analysis.** SVM method is employed in classification and regression. With methods based on SVM, data points with identical properties are grouped together depicting the space. With linear SVM, for the data provided, p-dimensional vector is contemplated and divided by utmost p-1 planes called hyperplanes, which are employed to divide the space for classification and regression. The mathematical form of SVM is given. The line equation is outlined by

$$a_1 = a_2x + b. \quad (26)$$

In (26), “x” stands for the line’s slope and “b” stands for intersect as given in

$$a_1 - a_2x + b = 0. \quad (27)$$

Let  $a' = (a_1, a_2)^T$  and  $z' = (x, -1)$ . Thus,

$$z' \cdot a' = 0. \quad (28)$$

The above equation obtained from two-dimensional vectors is known as the hyper lane equation.  $a' = (a_1, a_2)^T$  is mentioned as  $z'$  as given in

$$z' = \frac{a_1}{\|a\|} + \frac{a_2}{\|a\|}. \quad (29)$$

Here,

$$\|a\| = \sqrt{a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2}. \quad (30)$$

It is known that

$$\cos(g_1) = \frac{a_1}{\|a\|}, \quad (31)$$

$$\cos(g_2) = \frac{a_2}{\|a\|}.$$

Moreover, this can be expressed as equations (32)–(35):

$$z' = (\cos \cos(\|_1), \cos \cos(\|_2)), \quad (32)$$

$$z' \cdot a = \|z\| \|a\| \cos(\theta), \quad (33)$$

$$\theta = \theta_1 - \theta_2,$$

$$\begin{aligned} \cos(\theta) &= \cos(\theta_1 - \theta_2) \\ &= \cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2), \end{aligned} \quad (34)$$

$$= \frac{z'_1 a_1}{z' \|a\|} + \frac{z'_2 a_2}{z' \|a\|} \frac{z'_1 a_1 + z'_2 a_2}{z'},$$

$$z' \cdot a' = \sum_{i=1}^n z'_i a_i. \quad (35)$$

For n-dimensional vectors, dot product is computed as follows: consider  $f = y(z \cdot a + b)$  when  $\text{sign}(f) > 0$ , then it is a correct categorization, and for  $\text{sign}(f) < 0$ , it is incorrect. If dataset  $D$  is provided, then on a training dataset,  $f$  is calculated by

$$f_i = y_i(z' \cdot a + b). \quad (36)$$

Functional margin ( $F$ ) on the dataset is estimated as given in

$$F = \min_{i=1\dots m} f_i. \quad (37)$$

Through the contrast between the hyperplanes, hyperplane having greatest  $F$  is chosen. For selecting optimal hyperplane, optimal  $z$  and  $b$  values have to be found. Lagrangian function  $L$  is given by the following equations:

$$L(z', b, \alpha) = \frac{1}{2} z' \cdot z' - \sum_{i=1}^m \alpha_i [y_i : (z' \cdot a_i + b) - 1], \quad (38)$$

$${}_b L(z', b, \alpha) = - \sum_{i=1}^m \alpha_i y_i = 0. \quad (39)$$

Thus  $z'$  is calculated as follows:

$$z' = \sum_{i=1}^m \alpha_i y_i a_i, \quad (40)$$

$$\sum_{i=1}^m \alpha_i y_i = 0.$$

In the following equation, replacing  $L$ ,

$$z'(\alpha, b) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j a_i a_j. \quad (41)$$

Thus, the following equation is obtained:

$$\max_{\alpha} = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j a_i a_j. \quad (42)$$

When a point lies over the hyperplane, it is categorized as +1 class describing those cardiac risks are identified while under the hyperplane, it is -1 class describing that no cardiac risks are identified.

## 4. Performance Analysis

For the performance evaluation of the proposed method, MATLAB software is used for implementation. The efficiency of the proposed method is evaluated with a few parametric measures like accuracy, precision, recall, and F1-score. By using the U-matrix error rate of topographical error, then quantization error has been estimated. The dataset chosen for estimation is the diabetic dataset. To test the performance of the model, data are randomly selected from the dataset.

**4.1. Quantization Error.** During training, the required amount of computation level is reduced. Moreover, it estimates the performance level by applying stochastic quantization evaluating the value of the gradients. It regulates the standard distance between every input vector of the node and its winner.

$$Q_e = \frac{1}{N} \sum_{j=1}^N \|x_j - w_j\|. \quad (43)$$

In the above equation,  $w_j$  indicates the weight vector of winner for input  $x_j$ . Thus, the expected  $Q_e$  has to be small.

**4.2. Neuron Utilization  $N_U$ .** This specifies the ratio of the winner neurons of a single input or more inputs in vector map.

$$N_U = \frac{1}{nm} \sum_{i=1}^{mm} u_i. \quad (44)$$

Here by (45), if neuron  $i$  is the winner, then  $u_i = 1$  or  $u_i = 0$ . Therefore,  $N_U$  closer to 1.0 is more expected. The error calculation is shown in Figures 4(a) and 4(b)

**4.3. U-Matrix.** The U-matrix represents the cluster structure of the map illustrating the distances of the adjacent neurons. Output classes of the winning neurons of the U-matrix and winning classes with class names are shown in Figures 5(a) and 5(b)

**4.4. Accuracy.** This gives the ratio of the instances classified correctly which is estimated by

$$\text{accuracy rate} = \frac{\text{true positive} + \text{true negative}}{\text{total instances}} * 100. \quad (45)$$

**4.5. Precision.** This is the measure which reveals the ratio of data transmitted in the network with intrusion. This parameter estimates the correctness and quality of classification which is determined by

$$\text{precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}. \quad (46)$$

**4.6. Recall.** This metric presents the proportion of real positives that are correctly predicted positives given by

$$\text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}. \quad (47)$$

**4.7. F1-Score.** This is generally the mean value of precision and recall. Moreover, statistical measure is used in F1-score to calculate the performance rate of individual classifier of FN and FP as given in (49). Precision judges the accuracy, while recall detects sample instances with respect to faulty or nonfaulty attributes.

$$F1 \text{ score} = \frac{2X \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (48)$$

Table 1 shows some of the observations from the instances of diabetic datasets, which are then classified. The performance measures of various techniques of KNN, NN, and ANN are considered for estimating the efficiency of the proposed GSOA-DBN\_CNN technique. Table 1 presents the comparative analysis of accuracy, precision, recall, and F1-score, which are represented in percentage.

Figure 6 graphically represents the comparison for various methods in terms of accuracy. It is observed that the proposed GSOA-DBN\_CNN technique produced more

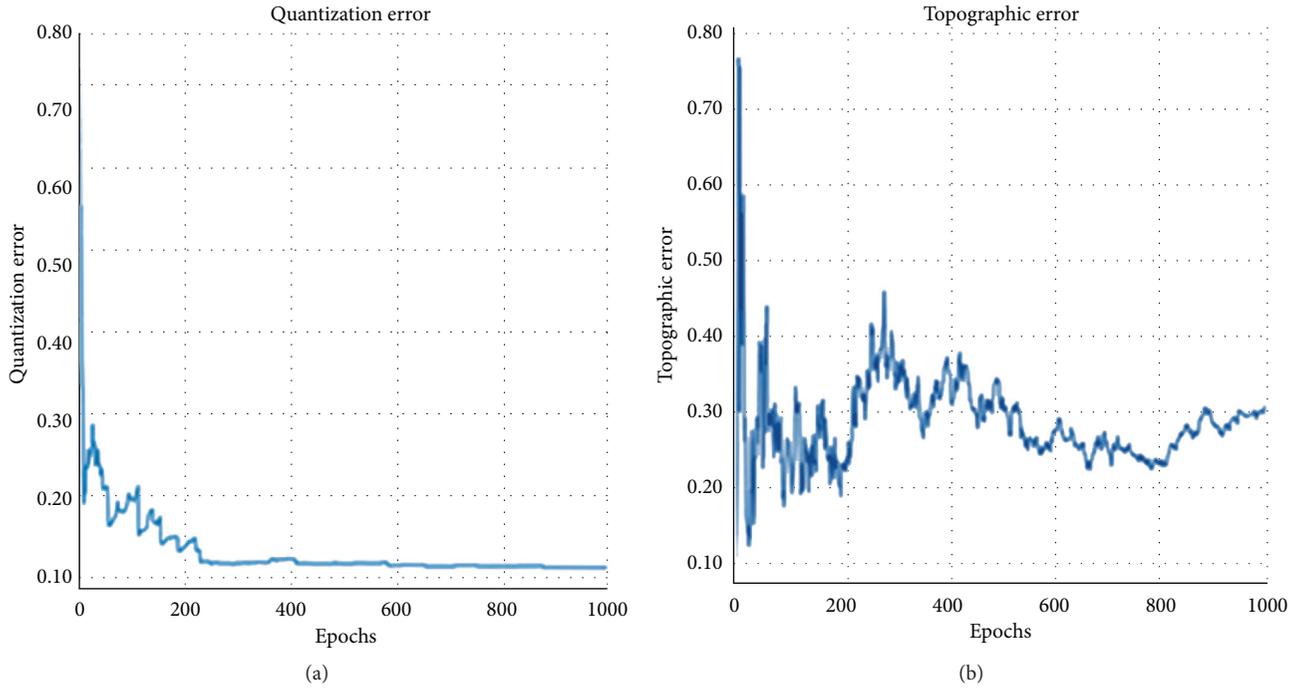


FIGURE 4: (a) Quantization error calculation and (b) topographic error calculation.

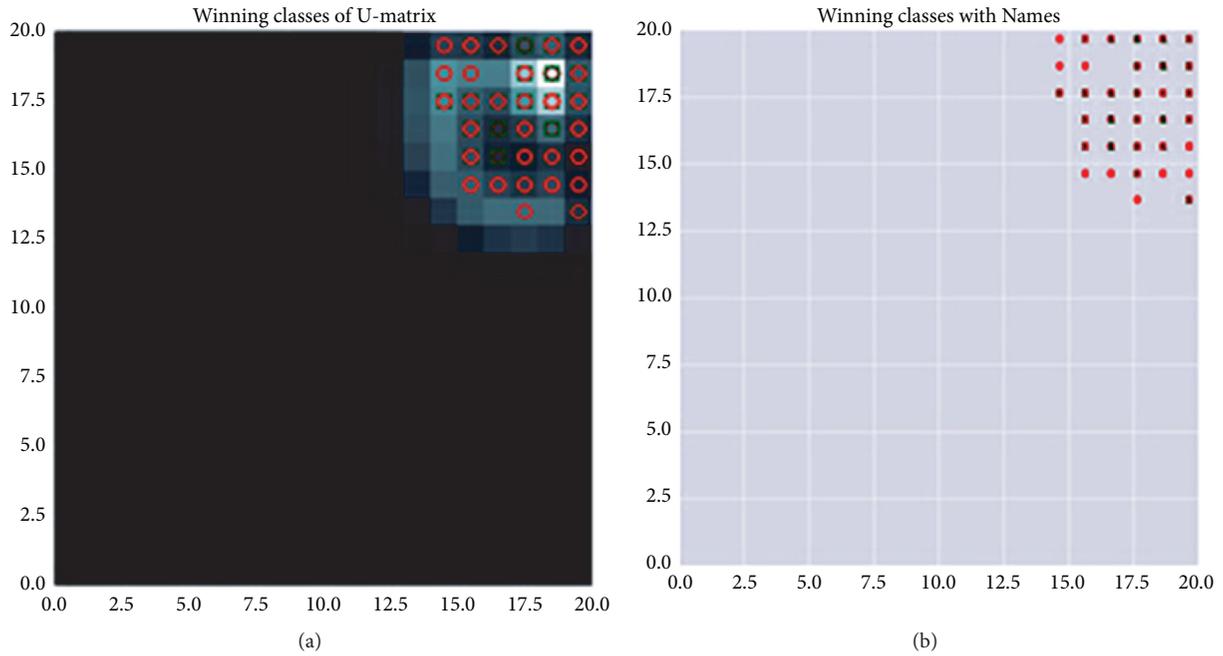


FIGURE 5: (a) Output classes of the winning neurons of the U-matrix and (b) winning neurons with class names.

TABLE 1: Comparison of performance.

Parameters	KNN	NN	ANN	GSOA-DBN_CNN
Accuracy	94	96	97	98
Precision	89	92	94	96
Recall	87	88	91	94
F1-score	82	86	88	92

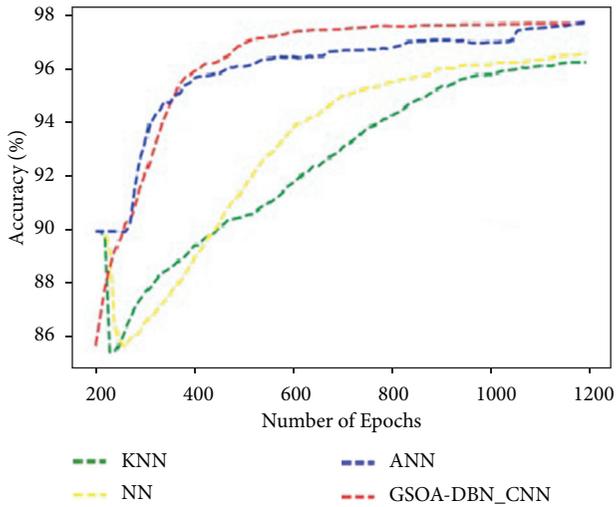


FIGURE 6: Comparison of accuracy.

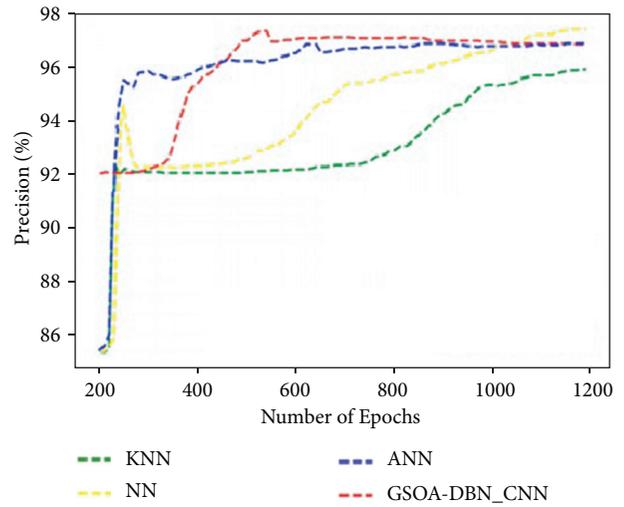


FIGURE 7: Comparison of precision.

accuracy than the existing techniques. Accuracy has been improved for proposed technique as 98%.

Figure 7 graphically represents the comparison for various methods in terms of precision. It is proved that the proposed GSOA-DBN\_CNN technique produced higher precision than the existing techniques. By comparing precision analysis with existing technique, the proposed technique has been improved as 96%.

Figure 8 graphically represents the comparison for various methods in terms of recall. It is noticed that the proposed GSOA-DBN\_CNN technique produced maximum recall than the existing techniques. Recall has been improved as 94% for proposed technique when compared with existing technique.

Figure 9 graphically represents the comparison for various methods in terms of F1-score. It is discovered that the proposed GSOA-DBN\_CNN technique achieves improved F1-score as 92% than the existing techniques.

PSNR graph comparison for proposed GSOA and existing ACO, WOA, GA, and HA is shown in Figure 10.

Figure 10 graphically represents the comparison for various methods in terms of PSNR comparison for data optimization. It is inferred that the proposed GSOA-DBN\_CNN technique achieves improved 0.2 PSNR than the existing techniques.

The comparative results are described in the Table 2, and the overall analyses show that the proposed GSOA-DBN\_CNN achieved accuracy higher than 2% of ANN, 3% of NN, and 4% of KNN. The precision achieved is more than 2% of ANN, 3% of NN, and 6% of KNN. Recall obtained is higher than 2% of ANN, 3% of NN, and 4% of KNN, similarly the F1-score achieved is 2% of ANN, 3% of NN, and 3% of KNN. The proposed GSOA-DBN\_CNN technique achieves improved 0.2 PSNR than the existing AWO, WOA, and GA techniques which is shown in Figure 11.

From the experimental results, it is observed that the proposed GSOA-DBN\_CNN technique achieves accuracy of 98%, 96% of precision, recall of 94%, and PSNR of 0.2 higher than the existing ACO, WOA, and GA techniques. Although

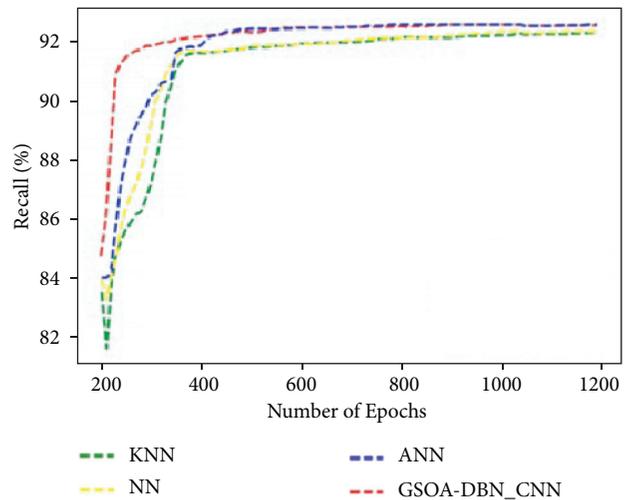


FIGURE 8: Comparison of recall.

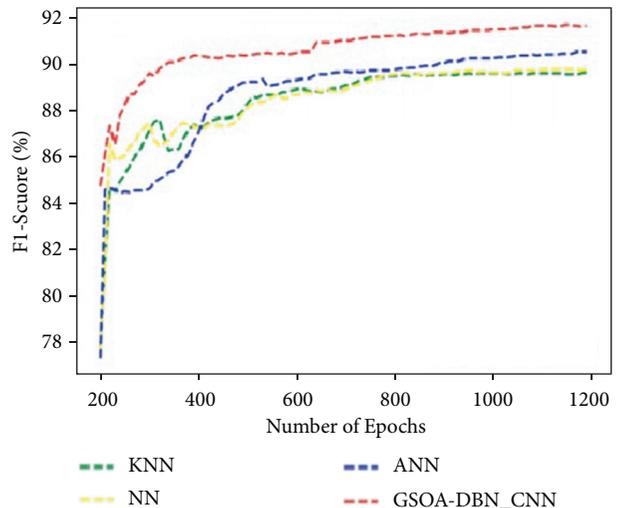


FIGURE 9: Comparison of F1-score.

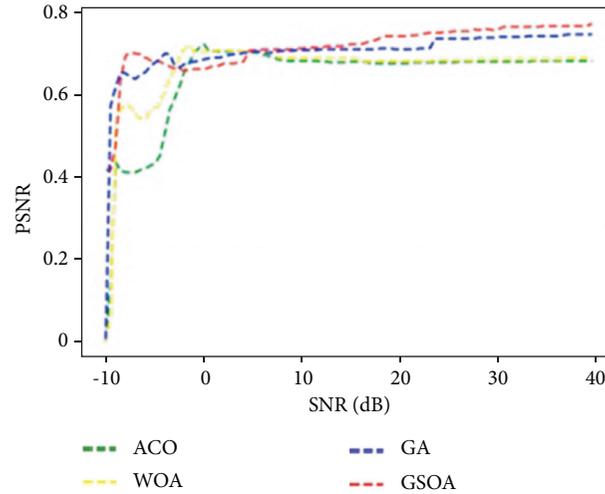


FIGURE 10: PSNR comparison for data optimization.

TABLE 2: Overall result comparison.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
KNN	94	92	90	88
NN	95	95	91	88
ANN	96	96	92	90
GSOA-DBN_CNN	98	98	94	92

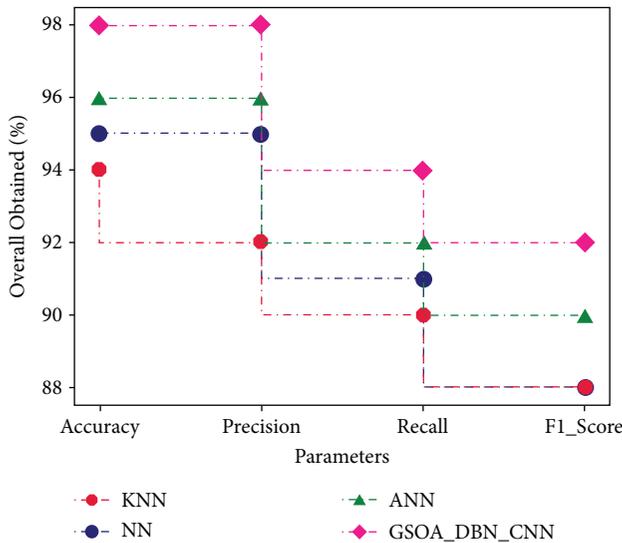


FIGURE 11: Overall comparison.

the proposed GSOA-DBN\_CNN model achieves better results, it lacks efficiency in some areas, which is to be improved.

## 5. Conclusion

The efficiency of the metaheuristic optimization algorithms is proved by solving several issues related to text clustering. But, the trapping of local optima is possible as the focus is on global search rather than on local search, that is exploration instead of exploitation. In this study, a metaheuristic algorithm based on data optimization is proposed for diabetes

data, which leads to the predictive analysis of cardiac attack risk based on blood vessels and cardiac nerves and cardiac nerve damages. For data optimization, GSOA is used since those metaheuristic optimization algorithms are robustly feasible. Then, a deep belief architecture (DBA) is established, where the classification is carried out using CNN. By this classification, the normal range and abnormal range of diabetes have been classified. The normal diabetes range has been updated to the hospital database, and for the abnormal diabetes range, the cardiac nerve and blood vessel damages have been analyzed using SVM-based image predictive analysis. Further, the method has also been compared with ACO, WOA, GA, and HA. The results reveal that the proposed GSOA-DBN\_CNN technique is better in terms of every parameter considered for comparison against the existing techniques. The accuracy achieved by the proposed technique is 98%, precision attained is 96%, the recall has been improved up to 94%, and F1-score obtained by 92% when compared with existing KNN, ANN, and NN. The comparative results are described as follows: the proposed GSOA-DBN\_CNN achieved accuracy higher than 2% of ANN, 3% of NN, and 4% of KNN. The precision achieved is more than 2% of ANN, 3% of NN, and 6% of KNN. Recall obtained is higher than 2% of ANN, 3% of NN, and 4% of KNN, similarly, the F1-score achieved is 2% of ANN, 3% of NN, and 3% of KNN. The proposed GSOA-DBN\_CNN technique achieves an improved 0.2 PSNR than the existing AWO, WOA, and GA techniques. The data optimization has been improved when compared with existing optimization techniques in terms of PSNR. In the future, hybrid deep learning methods can be utilized to further improve the efficiency of the model.

## Data Availability

The data used to support the findings of this study are included within the article.

## Disclosure

This study was performed as a part of the employment of the authors.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

The authors extend their gratitude to the deanship of scientific research at King Khalid University for funding this work through a research group program under Grant no. R.G.P. 1/399/42.

## References

- [1] G. U. Maheswari, R. Sujatha, V. Mareeswari, and E. P. Ephzibah, "The role of metaheuristic algorithms in healthcare," in *Machine Learning for Healthcare*, Chapman and Hall/CRC, Boca Raton, FL, USA, 2020.
- [2] H. Firdaus, S. I. Hassan, and H. Kaur, "A comparative survey of machine learning and meta-heuristic optimization algorithms for sustainable and smart healthcare," *African Journal of Comput. ICT Ref. Format*, vol. 11, no. 4, pp. 1–17, 2018.
- [3] S. Murugan, S. Jeyalakshmi, B. Mahalakshmi, G. Suseendran, T. N. Jabeen, and R. Manikandan, "Comparison of ACO and PSO algorithm using energy consumption and load balancing in emerging MANET and VANET infrastructure," *Journal of Critical Reviews*, vol. 7, no. 9, 2020.
- [4] A. Abugabah, A. A. AlZubi, F. Al-Obeidat, A. Alarifi, and A. Alwadain, "Data mining techniques for analyzing healthcare conditions of urban space-person lung using metaheuristic optimized neural networks," *Cluster Computing*, vol. 23, no. 3, pp. 1781–1794, 2020.
- [5] A. Saha, C. Chowdhury, M. Jana, and S. Biswas, "IoT sensor data analysis and fusion applying machine learning and metaheuristic approaches," *Enabling AI Applications in Data Science*, pp. 441–469, 2021.
- [6] P. Suganya and C. P. Sumathi, "A novel metaheuristic data mining algorithm for the detection and classification of Parkinson disease," *Indian Journal of Science and Technology*, vol. 8, no. 14, p. 1, 2015.
- [7] A. Doulamis, N. Doulamis, A. Angeli et al., "A non-invasive photonics-based device for monitoring of diabetic foot ulcers: architectural/sensorial components & technical specifications," *Inventions*, vol. 6, no. 2, p. 27, 2021.
- [8] J. Li, L. S. Liu, S. Fong et al., "Adaptive swarm balancing algorithms for rare-event prediction in imbalanced healthcare data," *PloS one*, vol. 12, no. 7, Article ID e0180830, 2017.
- [9] S. Chatterjee, J. Byun, K. Dutta, R. U. Pedersen, A. Pottathil, and H. Xie, "Designing an Internet-of-Things (IoT) and sensor-based in-home monitoring system for assisting diabetes patients: iterative learning from two case studies," *European Journal of Information Systems*, vol. 27, no. 6, pp. 670–685, 2018.
- [10] S. Fong, S. Deb, X.-S. Yang, and J. Li, "Feature selection in life science classification: metaheuristic swarm search," *IT Professional*, vol. 16, no. 4, pp. 24–29, 2014.
- [11] N. D. Hoang, "Image processing-based pitting corrosion detection using metaheuristic optimized multilevel image thresholding and machine-learning approaches," *Mathematical Problems in Engineering*, vol. 2020, Article ID 6765274, 19 pages, 2020.
- [12] I. M. El-Hasnony, S. I. Barakat, and R. R. Mostafa, "Optimized ANFIS model using hybrid metaheuristic algorithms for Parkinson's disease prediction in IoT e," *IEEE Access*, vol. 8, pp. 119252–119270, 2020.
- [13] R. Mukherjee and U. M. Diwekar, "Multi-objective optimization of the TEG dehydration process for BTEX emission mitigation using machine-learning and metaheuristic algorithms," *ACS Sustainable Chemistry & Engineering*, vol. 9, no. 3, pp. 1213–1228, 2021.
- [14] R. Singh, "Nature inspired based meta-heuristic techniques for global applications," *The International Journal of Computer Applications & Information Technology*, vol. 12, no. 1, pp. 303–309, 2020.
- [15] A. Tayal, A. Solanki, and S. P. Singh, "Integrated frame work for identifying sustainable manufacturing layouts based on big data, machine learning, meta-heuristic and data envelopment analysis," *Sustainable Cities and Society*, vol. 62, Article ID 102383, 2020.
- [16] J. Mallick, S. Alqadhi, S. Talukdar et al., "Risk assessment of resources exposed to rainfall induced landslide with the development of GIS and RS based ensemble metaheuristic machine learning algorithms," *Sustainability*, vol. 13, no. 2, p. 457, 2021.
- [17] S. Kaul and Y. Kumar, "Nature-Inspired metaheuristic algorithms for constraint handling: challenges, issues, and research perspective," in *Constraint Handling in Metaheuristics and Applications*, Springer, New York, NY USA, 2021.
- [18] T. M. Le, T. M. Vo, T. N. Pham, and S. V. T. Dao, "A novel wrapper-based feature selection for early diabetes prediction enhanced with a metaheuristic," *IEEE Access*, vol. 9, pp. 7869–7884, 2020.
- [19] V. Mazaheri and H. Khodadadi, "Heart arrhythmia diagnosis based on the combination of morphological, frequency and nonlinear features of ECG signals and metaheuristic feature selection algorithm," *Expert Systems with Applications*, vol. 161, Article ID 113697, 2020.
- [20] A. Sampathkumar, S. Murugan, M. Sivaram, V. Sharma, K. Venkatachalam, and M. Kalimuthu, "Advanced energy management system for smart city application using the IoT," *Internet of Things in Smart Technologies for Sustainable Urban Development*, pp. 185–194, 2020.
- [21] V. Manikandan, K. Gowsic, T. Prince, R. Umamaheswari, B. F. Ibrahim, and A. Sampathkumar, "DRCNN-IDS approach for intelligent intrusion detection system," in *Proceedings of the 2020 International Conference on Computing and Information Technology (ICCIT-1441)*, pp. 1–4, IEEE, Tabuk City, Saudi Arabia, September 2020.
- [22] A. Sharma, K. Guleria, and N. Goyal, "Prediction of diabetes disease using machine learning model," in *Proceedings of the International Conference on Communication, Computing and Electronics Systems*, pp. 683–692, Springer, Coimbatore, India, October 2021.
- [23] N. Fazakis, O. Kocsis, E. Dritsas, S. Alexiou, N. Fakotakis, and K. Moustakas, "Machine learning tools for long-term type 2 diabetes risk prediction," *IEEE Access*, vol. 9, pp. 103737–103757, 2021.

- [24] L. Abualigah, A. H. Gandomi, M. A. Elaziz et al., “Advances in meta-heuristic optimization algorithms in big data text clustering,” *Electronics*, vol. 10, no. 2, p. 101, 2021.
- [25] A. A. Sayed, M. M. Abdallah, A. M. Zaki, and A. A. Ahmed, “Big data analysis using a metaheuristic algorithm: twitter as case study,” in *Proceedings of the 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE)*, pp. 20–26, IEEE, Aswan, Egypt, February 2020.
- [26] Y. Zhang and F. Liu, “An improved deep belief network prediction model based on knowledge transfer,” *Future Internet*, vol. 12, no. 11, p. 188, 2020.