Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 16 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i16.51121

PAPER

An Intrusion Detection Model for Internet of Medical Things Using BDA-DAN2 Model

Raid Mohsen Alhazmi(✉)

Computer Science, University of Tabuk, Tabuk, Saudi Arabia

ralhazmi@ut.edu.sa

ABSTRACT

The Internet of Medical Things (IoMT) is a subset of the Internet of Things (IoT) where medical devices communicate with one another to share sensitive data. The integration of medical devices into the IoT has greatly assisted the development of the IoMT. These advancements facilitate effective communication and providing care for patients in the healthcare sector. However, they also face specific security and privacy concerns, such as malware attacks and denial of service (DoS) attacks. To overcome this problem, intrusion detection systems (IDS) are introduced, specifically employing deep learning (DL) methodologies. This study proposes a deep learning-based binary dragonfly algorithm (BDA) with a dynamic architecture for artificial neural networks2 (DAN2) model for implementing a robust and accurate IDS in IoMT. The IDS has the following stages: collection of data, preprocessing, selection of features, and classification. The IoMT dataset is employed to train the model to get improved outcomes. The standard scalar technique is used for the data preprocessing process. The BDA algorithm is used for feature selection (FS) of the preprocessed data. The DAN2 model is implemented to classify the selected data and to improve the classification accuracy. The dataset was further divided for training and testing of the model. The performance of the BDA-DAN2 model is assessed utilizing the evaluation parameters of accuracy, recall, precision, and F1-score. The BDA-DAN2 model demonstrates superior performance with 99.12% accuracy, 99.28% precision, 99.40% recall, and 98.56% F1-score during training, and 98.92% accuracy, 98.50% precision, 98.68% recall, and 97.90% F1-score during testing. Experiments confirmed that the binary dragonfly algorithm with the DAN2 (BDA-DAN2) model has the highest accuracy compared to the existing models.

KEYWORDS

Internet of Things (IoT), Internet of Medical Things (IoMT), deep learning (DL), intrusion detection systems (IDS), binary dragonfly algorithm (BDA), dynamic architecture for artificial neural networks2 (DAN2)

Alhazmi, R.M. (2024). An Intrusion Detection Model for Internet of Medical Things Using BDA-DAN2 Model. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(16), pp. 145–169. https://doi.org/10.3991/ijoe.v20i16.51121

Article submitted 2024-07-14. Revision uploaded 2024-10-12. Final acceptance 2024-10-12.

1 INTRODUCTION

The Internet of Things (IoT) is a network that connects numerous objects to the Internet using data-sensing devices and specific protocols. This enables the sharing and exchange of data, as well as enabling monitoring, tracking, administration, and placement. The healthcare system is the most extensive stage for the deployment of the IoT. The IoT is increasing the capacity to enhance many medical and clinical applications. Examples include remote healthcare monitoring, aged care, other fitness programs, and various chronic conditions. Consequently, a wide range of imaging equipment, diagnostic medical devices, and sensors are considered as intelligent devices that can serve as essential components in the IoT [2]. The medical industry expects significant enhancement in efficiency and quality of care through the many advancements in the IoT, commonly known as the Internet of Medical Things (IoMT).

Internet of Medical Things has emerged as a significant and rapidly expanding technical domain in the healthcare sector. The IoMT, which is also known as IoT in healthcare, links people in numerous roles with different medical devices and objects in the medical scene. It also creates a platform for information sharing where people, things, and information sharing between things are all free to develop. The main characteristics of the IoMT encompass the integration of various devices and systems in the medical domain, along with a complete information platform that enables seamless connectivity [1]. An essential aim of the IoMT is to assure limited human involvement in several health care operations and regular patient visits. To accomplish this, the process involves the utilization of automated sensors and advanced machine intelligence algorithms. The IoMT can also decrease expenses for the patients and enhance the effectiveness of medical professionals [4].

During the COVID-19 pandemic, there has been a growth in the use of telehealth techniques. This is mainly due to the physical distance guidelines that require healthcare practitioners to remotely treat patients using IoMT equipment [1]. In 2015, the United Nations (UN) approved the sustainable development goals (SDGs) to be achieved by 2030 [2]. Ensuring good health and well-being is a key goal in the SDGs. Currently, the IoMT has the capability to achieve the objective of promoting excellent health and well-being [3]. IoMT distinguishes itself from IoT by placing particular emphasis on the interactions between devices and humans within the healthcare sector. Additionally, it includes the evaluation and control of all interconnected entities and the procedures of communication [1]. An IoMT-appropriate architecture consisting of four layers. The IoMT architecture comprises perception, network, transport, and application layers, as depicted in Figure 1.

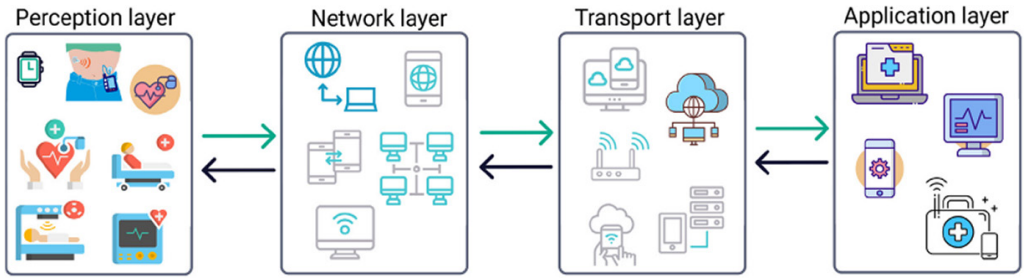

Fig. 1. IoMT architecture

The perception layer comprises medical gadgets, including scanners, wearables, monitors, and sensors, which are utilized for detecting various parameters such as

oxygen levels, temperature, and glucose levels. These gadgets are essential for collecting critical medical data. Furthermore, they function as an interface linking the users to the digital healthcare world. The network layer includes cable, middleware, and wireless technologies that provide the exchange of information between medical equipment, hence establishing the connectivity of all elements within the IoMT environment. Subsequently, the transport layer enables seamless communication from one end to another in order to transmit the gathered physiological information to medical servers for the purposes of storage, data analysis, and processing. The application layer serves as a bridge between the patients and the transport layer, facilitating patient management, medical treatments, and hospital monitoring [6].

However, the privacy and security of IoMT-based healthcare systems have not received enough attention despite their rapid development. Insufficient security in IoMT healthcare systems can lead to affected patient privacy with eavesdropping and late identification of malicious incidents owing to disruptions induced by DoS attacks on IoMT devices. A Distributed DoS (DDoS) attack provides a significant risk to the healthcare system by causing disruption to connected devices and making medical services inaccessible to authorized clients [18]. In 2015, HP Fortify conducted a study that revealed that the top 10 smartwatches at that period all had security problems. These vulnerabilities included insufficient authentication, insecure interfaces, insufficient transfer of data, encryption, privacy problems, and insecure software/firmware [5]. Figure 2 presents the frequently mentioned types of IoMT attacks, such as DoS and DDoS [6].

Fig. 2. Examples of security threats to IoMT [6]

To assure the security and privacy of exchanging data in IoMT platforms, the intrusion detection systems (IDS) act as the best tool for security to counteract a range of cyberattacks. The IDS has the ability to rapidly identify deviations from normal patterns and immediately inform the system to avoid any potential harm. The detection algorithm of IDS plays an important role as it determines the system's capacity to accurately identify various types of cyberattacks while reducing the rate of false alarms. Integrating the IDS with edge computing enables the detection of attacks in close proximity to the data source, resulting in improved effectiveness and efficiency. Intrusion detection emerged as an important development in ensuring the security of IoT networks. Intrusion detection is a system used to handle security interruptions that can occur in any layer of the IoT [25].

Deep learning (DL) and machine learning (ML) can be utilized to create intelligent security provisioning mechanisms in IoT devices. DL algorithms are a specific type of ML algorithm that seek to mimic the learning capabilities of the human brain by processing large volumes of data. A deep learning model typically consists of numerous layers of neurons, where each neuron in each layer produces a non-linear change of its inputs from the previous layer. DL methods are more efficient than typical ML approaches in identifying patterns in large datasets. DL techniques greatly enhance the security of necessary IoMT applications [7].

Although IDS is an exceptionally powerful method for detecting and alleviating cyberattacks, it is not the sole strategy for securing IoMT systems. Alternative options encompass encryption methods for data privacy, blockchain for secure data exchange, multi-factor authentication for access control, and network segmentation to reduce attack surfaces. However, IDS, especially when integrated with DL, are particularly good at identifying real-time anomalies and addressing emerging threats. The motivation behind using IDS in this study is due to its ability to dynamically learn from large datasets, identify patterns, and provide a proactive defense mechanism, especially in environments like IoMT where real-time data is critical. The research aims to illustrate that the integration of IDS with DL, particularly the BDA-DAN2 model, results in enhanced intrusion detection efficacy relative to conventional approaches.

This study introduces a new DL-based IDS model designed to accurately detect harmful intrusions in the IoMT environment. The suggested model will ensure the protection of sensitive medical and healthcare information from malware, attacks, and suspicious threats.

1.1 Problem statement and scope of the research

This study will focus on effective intrusion detection in the IoMT environment. The security and privacy of these IoMT healthcare systems are the most important parameters for the development of IoMT across the world. Traditional IDS may not be suited to manage unique challenges such as cyberattacks. Based on the limitations of traditional IDS systems, which often rely on predefined signatures or rule-based detection, making them less effective against sophisticated or zero-day attacks. Consequently, there is a need for a specific method that combines feature selection (FS) and classification with DL techniques to design an efficient and robust IDS model for IoMT. In contrast, the proposed BDA-DAN2 model addresses these challenges by leveraging DL and FS techniques, allowing it to dynamically adapt to new threats and detect anomalies in real time. Unlike traditional IDS, the BDA-DAN2 model can learn from evolving datasets, improving its accuracy and robustness in detecting complex attack patterns in IoMT environments. This adaptability and precision make it a more advanced and reliable solution compared to existing IDS. This research concentrates on developing and evaluating a combination of FS and classification models for IoMT-IDS. The suggested model aims to identify and detect potential security vulnerabilities in IoT environments that have limited resources, such as unique and emerging assaults.

1.2 Research objectives

The suggested approach includes three sequential steps: first, the dataset pre-processing; second, selecting features; then, applying the selected features for classification.

The research objectives for this work on IoMT intrusion detection and classification using the BDA-DAN2 model can be outlined as follows:

- To develop a new hybrid FS and classification model for IoMT-IDS.
- To enhance the performance and efficiency of IoMT-IDS by implementing the BDA technique for selecting features and the DAN2 classifier for classification.

- To assess the BDA-DAN2 model's performances on the IoMT dataset.
- To assess its precision, accuracy, F1-score, and recall in detecting security threats in IoMT environments.
- To compare the performances of the BDA-DAN2 model with existing approaches and to show the efficiency of the suggested model in enhancing IoMT security.

The work is presented in the following structure: Section 2 reviews the related works on FS and classification-based IoMT-IDS models. Section 3 presents the proposed BDA-DAN2 model, encompassing data pre-processing, BDA-based feature selection, and DAN2-based classifications. Section 4 assesses the BDA-DAN2 model's performance, discusses the findings, and compares it with other models. Lastly, Section 5 presents the conclusions and potential future research suggestions.

2 RELATED WORKS

In this section, the state-of-the-art work related to IDS for the IoMT network implementing ML or DL methodologies and the research gap in the related work is discussed.

Three DL models were introduced in [8] to develop an IDS for the IoMT network. By implementing various strategies, the detection of attacks in the IoMT platform was enhanced by utilizing patient biometrics and network flow statistics. The data preprocessing phase involved the utilization of widely used techniques such as simple imputer and standard scalar. To minimize the effects produced by the attacker, many DL models were introduced, including LinSVM, ConvSVM, and CatEmbedding. With a 100% accuracy rate, these models identify network intrusions better than the latest technologies by applying the integrated features of a network traffic flow meter and data of the patient biometric. Among the three models, ConvSVM and Cat_Embeddings, both DL models, achieved a perfect accuracy of 100%. However, the LinSVM model achieved a slightly lower accuracy of 99.94%.

In [9], a DNN Classifier Model for intrusion detection in IoMT was introduced, which is based on a hybrid PCA-GWO approach. The use of the DNN model in this approach aids in decreasing the quantity of features and samples that are collected for the process of classification. The initial preprocessing of the raw data set involved applying the One-Hot encoding strategy to convert the categorical data into numerical data. Subsequently, the pre-processed data set received dimensionality reduction using PCA as the initial step, followed by GWO as the subsequent step. This method aimed to decrease the number of attributes and identify the most significant ones. There was a 15% increase in accuracy. The findings also demonstrate a 32% reduction in the training time needed for the classification model, specifically designed for the IoMT framework. This methodology facilitates faster notifications to healthcare professionals in the case of an intrusion in their environment.

In [10], a novel genetic algorithm based random forest (RF) technique was designed with the aim of efficiently detecting and analyzing malicious activity and cyberattacks in IoMT devices and their surroundings. This technique was utilized on two actual datasets, namely UNSW_2018_IoT_Botnet and NSL-KDD. The results of the simulations obtained utilizing the WEKA demonstrated that the method outperformed conventional ML algorithms regarding recall, F1-score, precision, and accuracy metrics. The model attained a remarkable accuracy of 99.999% and a precision of 99.9%. In addition, the RF algorithm achieved a recall of 100% and an F-measure of 99.9%. The technique attained a low error rate for MAE and RMSE when applied to

the NSL-KDD. The GA-RF algorithm successfully attained a remarkable accuracy rate of 99.999%, along with perfect precision and recall scores of 100%, when applied to the UNSW_2018_IoT_Botnet.

In [11], a highly efficient anomaly-based IDS (AIDS) was introduced in the IoMT environment. The AIDS system seeks to use both network- and host-based methods to effectively collect log files from IoMT devices and the gateway, along with traffics from IoMT networks, while also considering the computational expense. A test was carried out to simulate the performances of six commonly used ML algorithms for anomaly detection. The techniques examined were decision tree (DT), linear regression (LR), k-nearest neighbor (KNN), support vector machines (SVM), RF, and naïve Bayes (NB). During the preprocessing stage, the numeric values of all features were normalized. The network component of the "TON_IoT Telemetry data set" was utilized for the purpose of training and evaluating the ML algorithms. The evaluation results indicate that the DT, RF, and KNN algorithms were better suited for intrusion detection.

To identify cyberattacks in the IoMT, researchers devised a method that combines ML and DL techniques with the Harris-Hawk-Optimization (HHO) algorithm in a study referenced as [12]. HHO was employed for FS because of its rapid coverage capability. The initial data underwent preprocessing through the utilization of normalization and one-hot encoding techniques to standardize the data. The model was trained and tested using the NSL-KDD dataset. Classification is performed using RF, SVM, bagging, boosting, and RNN (recurrent neural network). The findings indicate that the efficiency was improved when classifiers were integrated with HHO. The findings demonstrate that this model attained a remarkable accuracy of 0.998. The model demonstrated outstanding precision of 0.997. The recall score of both the HHO-RF and HHO-RNN models was 0.987, which was the highest compared to all other models.

The paper [13] suggested SafetyMed, an IDS for the IoMT that utilizes long short-term memory (LSTM) models and convolutional neural networks (CNN). The paper introduces SafetyMed, a system that provides comprehensive protection for the IoMT by effectively mitigating twelve different forms of assaults. SafetyMed is the pioneering IDS designed to safeguard IoMT devices against malevolent picture data and sequence network traffic. This is a novel approach to achieve an optimal balance between the detection rate and false positive rate, resulting in an improved detection capability. The dataset CIC-IDS2017, which is highly notable, was utilized. The performance of the network was evaluated using the CICIDS2018 and CICIDS2019 datasets, which were not utilized during the training phase. The FPR of SafetyMed has an average value of only 0.71%, indicating its high level of reliability. The accuracy of the SafetyMed device was 98.47%.

The study [14] showcases the utilization of a deep RNN (DRNN) and supervised ML models (KNN, ridge classifier, DT, and RF) to create a highly effective IDS in the IoMT platforms. The key objective of this IDS was to accurately classify and predict unforeseen cyber threats. The particle swarm optimization (PSO) algorithm was employed to perform FS on the NSL-KDD dataset. Out of the 40 characteristics available, 21 attributes were selected while considering only one class. The utilization of RF led to an enhancement in the categorization accuracy. The PSO technique was integrated with ML techniques to enhance the precision of the model. The accuracy values for PSO-RF, PSO-DT, PSO-KNN, and PSO-RC are 99.76%, 99.58%, 98.90%, and 97.61%, correspondingly. PSO-RF attains superior precision, accuracy, and MCC, but at the cost of the F1 score.

In [15], a DL framework based on blockchain was proposed. This framework offers two degrees of security and privacy. A decentralized approach was employed

to train the model using a hybrid deep learning architecture consisting of bidirectional LSTM (BiLSTM) and CNN. This approach aimed to minimize latency and reduce processing costs. The framework was trained on all local devices using hybrid DL, namely a combination of BiLSTM and CNN. The BiLSTM is composed of a pair of modules: a feed-forward module and a backward module, which are then concatenated at the end. The freely available datasets based on IoT-ToN from UNSW were utilized. To ensure privacy, it was encoded with a lightweight decryption and encryption technique utilizing homomorphic encryption. Homomorphic encryption enables the execution of addition or multiplication operations on encrypted data. The model was suggested for implementation in cross-domain networks within healthcare systems.

In [16], an improved approach for estimating redundancy in FS for IDS in the IoMT was presented. This method was based on the logistic redundancy coefficient gradual upweighting. Initially, the min-max normalizing technique was utilized to ensure data range fidelity. MIFS, a widely used method for FS, has the ability to efficiently select pertinent features irrespective of the data distribution. This quality makes it well-suited for early detection situations where there was a lack of adequate data attack patterns. The model was employed to measure the redundancy of the features. The data set WUSTL EHMS-2020 was employed. During the training process, the model was modified by taking into account the prediction errors, which were determined by evaluating the loss functions. Following the completion of the training, the model's results were assessed by utilizing the dataset. This involved computing the precision, recall, accuracy, and various pertinent metrics to evaluate the performances of the model. The evaluation showed that the proposed LRGU effectively identified a brief set of important features that performed better than existing strategies. The LRGU surpassed other solutions, demonstrating its effectiveness.

In [17], a particle swarm optimizer (PSO) deep neural networks (DNN) PSO-DNN was developed to create a highly efficient and precise IDS in the IoMT. The information on sensing devices of IoT was transmitted to the cloud servers utilizing protocols of IoT networks, such as MQTT and AMQP. The data preprocessing phase included the utilization of the standard-scalar approach. To increase the precision and efficiency of the model, the FS method employed the application of PSO. PSO employs a range of DL and ML models to investigate the precise forecasting of IoMT threats. The IoMT dataset was utilized for the purpose of evaluating the model. The ARGUS tool was used to construct the dataset that included the biometric data of the patients as well as network traffic. The DNN outperformed both the LSTM and CNN models. A DNN's accuracy performance slightly improved as its total layers were increased from one to two.

A ML strategy utilizing intelligent intrusion detection was proposed in [18] to effectively counter cyberattacks in IoMT networks. The feature vector is analyzed using two data-driven kernel techniques, namely kernel partial least square (KPLS) and kernel principal component analysis (KPCA), to identify significant features. The kernel extreme learning machine (KELM) was proposed as a classification model to determine if the traffic flows were malicious or benign. The dataset is crucial for evaluating the resilience and efficacy of the detection model. In this technique, a contemporary IoMT dataset called WUSTL-EHMS-2020 is utilized. The KPLS-KELM attains the accuracy of 99.9%, and the model had superior accuracy rates of 99.95%. The KPLS model exhibited a delay of 6.40 seconds compared to other models, and its accuracy rate was 71.8%, which was the poorest performance among all the models.

In [19], a self-tuning LSTM IDS for the IoMT was developed using fuzzy logic (FL). The usage of this strategy is more favorable compared to early stopping of static since it enables the model to dynamically choose the most suitable moment to cease

the real learning progress based on training, rather than relying on a predefined and unchangeable number of epochs. The FL system assesses the performance of the model by taking into account accuracy, the rate of improvement, and validation loss. It then uses this information to adaptively modify the patience parameter. This FST-LSTM technique has superior detection rates across varying amounts of features, hence demonstrating its efficacy in spotting IoMT environment intrusions. The FST-LSTM model typically exhibits a greater detection rate compared to other models, particularly when it comes to feature counts in the mid-range.

This study introduces a novel IDS for a safe big data platform, utilizing Spark and employing multimodal fusion techniques [20]. The model utilizes a decision-based fusion approach that incorporates many processes, including initialization, processing the data, FS, and multimodal classification, to enhance the detection of infiltration. This work introduces the chaotic butterfly optimization (CBO) algorithm to extract the set of features that are most successful. This is followed by a multimodal classification process. Subsequently, the multimodal deep learning classifiers, specifically the CNN, LSTM, and gated recurrent units (GRU), were utilized to enhance the detection rate. In addition, this work employed ideas such as Hadoop MapReduce and Spark to enhance the speed of processing large datasets on a parallel computing platform. Several experiments were conducted utilizing the NSLKDDCup99 dataset repository as a benchmark. This model achieves a high level of performance on the given dataset, with a detection rate of 99.59%, precision of 98.93%, and accuracy of 99.21%.

A system called explainable deep neural network (XMeDNN) was proposed in reference [21] for detecting intrusions in the IoMT. The system underwent training and testing using the dataset WUSTL-EHMS-2020. ARGUS was utilized to extract network-flow features from raw recorded traffic. In order to demonstrate the superiority of DNNs over classical machine learning methods in handling this classification problem, we employ five distinct classifiers for training and testing purposes. The XMeDNN system achieved an average accuracy of 97.578% in 10-fold cross-validation, with an average F1 score of 0.97634. This system exhibited superior performance in comparison to prior works in the same field.

A novel hybrid architecture called “ImmuneNet” [22] was developed to accurately detect and safeguard healthcare data against the most recent intrusion threats using DL techniques. The model utilized correlation-based feature selection (CFS) to enhance efficiency. A statistical FS technique was conducted to identify the important features of the dataset that can enhance the effectiveness of the classifiers. ImmuneNet’s performance was evaluated by comparing it to various other ML algorithms using the Canadian Institute for Cybersecurity’s IDS datasets from 2017 and 2018, as well as the Bell DNS 2021 dataset. These datasets consist of comprehensive real-time data on the newest cyberattacks. ImmuneNet achieved an accuracy value of 99.2% on a CIC Bell DNS 2021 dataset with better class balance. It also achieved accuracies of 99.8% on the CIC IDS 2018 and 99.63% on the 2017 datasets, respectively. The model produced a score of ROC-AUC as 99.19%, much surpassing the scores of the other techniques.

2.1 Research gap analysis

Although there have been notable improvements in the development of IDS for IoMT networks, there are still many research gaps and issues that require attention. Most of the approaches, like ML/DL models, use extensive preprocessing steps,

but often they produce minimum results in accuracy. In addition, models such as GA-RF and HHO-RNN demonstrate high accuracy. However, they frequently choose general datasets commonly used in various models rather than a unique dataset. The gap is addressed by utilizing a unique IoMT-specific dataset rather than relying on general datasets. Although models such as GA-RF and HHO-RNN attain great accuracy, the suggested model was selected due to its advancements in FS through the binary dragonfly algorithm (BDA) and its dynamic architecture (DAN2) for enhanced classification. The research indicates that the BDA-DAN2 model surpasses conventional models, especially in handling IoMT-specific data, indicating an enhancement of established methodologies. Furthermore, the computational cost and complexity of the models such as SafetyMed are too expensive for IoMT networks. Further validation and real-world implementation are necessary to thoroughly evaluate their effectiveness in various healthcare environments.

3 PROPOSED METHODOLOGY

This study proposes a model to improve intrusion detection in IoMT environments by integrating BDA for FS and DAN2 for classification. The model aims to address the security challenges faced in IoMT environments and improve the accuracy of detection by incorporating various strategies. Figure 3 shows the research model IDS framework for the IoMT applications.

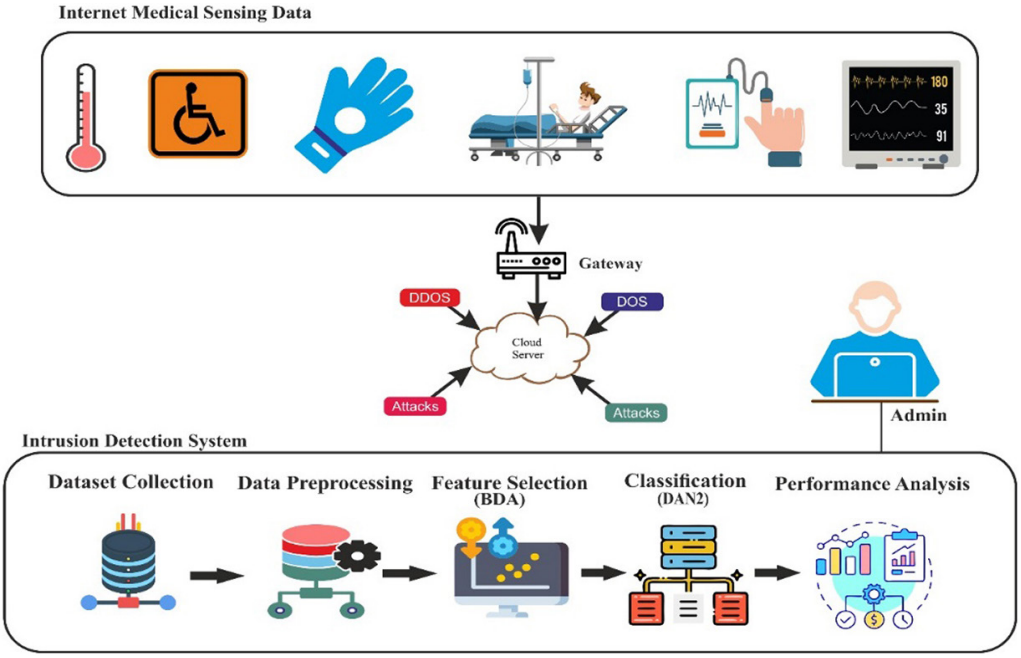

Fig. 3. Framework of the proposed architecture

The framework comprises three main components: the Internet of medical sensing data, the attack layer, and IDS. The initial component of the work includes a range of smart medical gadgets, such as smart gloves, smart monitors, smart wheelchairs, etc., which are linked via the Internet and given specific IP addresses. All these devices transmit confidential patient data at regular intervals, which is then kept in a private cloud. The second element relates to the attack segment or the actions

carried out by unauthorized individuals. In this section, many types of attacks, such as DoS and DDoS, can occur. The intruders refer to attackers from the general public who have an interest in accessing the stored sensitive data in a cloud. The data can be retrieved through other means, such as accessing it while it is stored in the cloud, interrupting it during transmission, or capturing it when sent to a doctor. If the information has been altered before transmitted to the healthcare officials or doctors, false interpretation may be provided regarding the patient's health, potentially resulting in death. In order to address this challenge, the third element is specifically designed to identify any unauthorized access to the network. The proposed methodology primarily focuses on the IDS. The intrusion detection model that has been presented consists of several stages: dataset collection, data pre-processing, FS, and classification. In the pre-processing stage, the normalization technique is incorporated by using the standard scalar technique to improve the model performance. BDA is used for feature selection from the processed data. The classification process is done by using the DAN2 algorithm to minimize the complexity of the model. Using the IoMT dataset, the efficacy of this novel combination model will be evaluated and compared to other existing methods.

3.1 Dataset details

This section provides a detailed discussion of the IoMT dataset, which is used for evaluating the data model. The dataset was generated utilizing a testbed for monitoring real-time health. The test bed consists of the sensor elements connected to the body of a patient, the software defined networks (SDN) controller, and the gateway of networks for displaying the traffic in the networks. The testbed utilizes network traffic and sensor data to detect anomalies and identify attacks. The attack data set was generated by simulating three attacks in the environment. The attacks mentioned are data injection, spoofing attacks, and man-in-the-middle attacks. A man-in-the-middle attack occurs when an attacker penetrates the health monitoring system of the patient and is able to monitor and modify the network traffic in real time. This assault results in losses of patient data integrity and confidentiality within the network. The deliberate alteration of patient health network packets as they pass through the gateway is known as a "data injection attack." This exploit leads to the breach of patient data integrity. The interception of network traffic passed via the gateway is made easier by spoofing attacks. It violates the confidentiality of the patient data. The ARGUS program was used to produce a dataset that comprises network traffic and patients' biometrics. The biometric data comprises peripheral STO2, temperature, pulse rate, ECG ST segment data, diastolic BP, and respiration rate. The records of network traffic flow and associated metrics were collected in order to get the comprehensive network traffic characteristics. The data set comprises a total of 44 features, with 35 of them being network traffic features. The data set outputs were categorized as either normal or attack traffic. The attack was categorized as zero, whereas normal was designated as one.

Table 1 displays the dataset, which consists of 14,272 normal network records and 2,046 attack sample network records. 1400 attack samples were selected at random to achieve a balanced proportion of normal traffic or an attack in the data set. In real-time network operating, instances of attacks were rare, and the ratio of normal traffic to attack traffic was significantly elevated. In order to replicate the actual situation, we determined the number of attack samples in the data set and utilized an imbalanced data set for a performance calculation [17].

Table 1. Dataset description

Raw Data		Random Selection	
Attack	Normal	Attack	Normal
2046	14,272	1400	14,000

3.2 Data preprocessing

Deep learning is a data-driven technique used for decision-making and task performance. The models' performance is negatively impacted by inadequate datasets that lack a comprehensive set of features. The original dataset often contains duplicated data, missing values, abnormalities, and other challenges that impact its overall quality. The pre-processing is an important and vital stage in the training of DL models. This section explains an overview of the techniques employed for pre-processing [10].

This study utilizes popular approaches such as standard scalar for pre-processing. The standard scalar is a widely recognized technique utilized to standardize data. It is generally used as a preprocessing step before applying ML or DL methodologies for normalizing datasets. The purpose of employing feature scaling is to adjust the distribution of values so that the observed mean value is zero and the standard deviation is one. Equation (1) demonstrates the process of feature scaling using the standard scalar method [8]. It addresses the issue of modeling any other condition or non-constant variance that requires normality.

$$x_{feature\ saled} = \frac{x_{feature\ current\ valve} - mean}{standard\ deviation} \tag{1}$$

Where, $x_{feature\ saled}$ represents the feature value that has been scaled as a result of the successful completion of the scaling process. $x_{feature\ current\ valve}$ displays the current feature value [8].

3.3 Feature selection using BDA

Feature selection is a technique used to identify and remove unnecessary and redundant information from data, resulting in a more effective learning process by focusing on the most relevant characteristics. FS is the process of eliminating duplicate and unwanted variables from a dataset to improve the accuracy of detection and decrease the time required for developing a model. In addition to the intricacy of replicas, feature selection may assist in eliminating certain computations [14]. It is necessary to choose appropriate features to enhance accuracy and speed up the prediction process of the system. The FS in this study utilizes the binary dragonfly (DA) algorithm.

Dragonfly algorithm. The DA is a newly introduced algorithm that is based on swarm intelligence. The DA imitates the migrating and hunting strategies of idealized dragonflies. The hunting strategy employed by dragonflies is termed as static swarm feeding, wherein they fly in compact groups within a limited area to locate sources of food. The process of migration is referred to as dynamic swarm (migratory). During this phase, the dragonflies exhibit unidirectional flight in larger groups, facilitating the migration of the swarm. Like previous algorithms inspired by nature, the DA algorithm comprises two phases: exploration, which is influenced

by exploitation, and the behavior of static swarming, which is influenced by the behavior of dynamic swarming [27]. The DA method has demonstrated superior performance in all types of optimization problems, including continuous, multi-objective, single-objective, and discrete issues. It has outperformed more advanced algorithms such as PSO and differential evolution (DE). A binary variant of the differential algorithm dubbed BDA. BDA utilizes a transfer function (TF) for converting a continuous search region into discrete. The ability of Big Data Analytics was assessed on a selection of FS issues, and the results confirmed the satisfactory performance of this method [27].

The BDA is a variant of the DA that operates on binary data. BDA is employed in this study to enhance performance and precision by selecting relevant features. The search space is structured in the form of a hypercube, with each location in the search space being represented by a position vector $y = \{y_1, y_2, ..., y_d\}$. DA was initially introduced to address challenges related to continuous optimization. By combining the present position vectors with the step vectors, the individual position is updated. This method requires modification to effectively address binary optimization difficulties. Utilizing TensorFlow can facilitate the transformation of a continuous optimization method to a binary version.

The locations were transformed from continuous to distinct utilizing two stages with the help of TFs. In the current iteration (t), the input to Equation (2) is the value of *dth* dimensions of the *jth* step vectors (velocity), which is used to calculate the likelihood of transforming that element to either zero or one. Next, assign a value of either 0 or 1 to the element, depending on the result of Equation (3). The utilization of the TF in Equation (2) is founded on a prior suggestion from the research literature.

$$T\left(v_d^j(t)\right) = \left|\left(v_d^j(t)\right) / \sqrt{1+\left(v_d^j(t)\right)^2}\right| \tag{2}$$

The *jth* position vector's element was transformed to 1 or 0 utilizing Eq. (3) by applying the outcome $T\left(v_k^j(t)\right)$ acquired from Equation (2).

$$Y(t+1) = \begin{cases} -Y_t & r < T\left(v_k^j(t)\right) \\ Y_t & r \geq T\left(v_k^j(t)\right) \end{cases} \tag{3}$$

Where function r generates the range within 0 to 1 of random number. When determining whether the Y_t values was flipped, the r value plays a significant influence. If the value of $T\left(v_k^j(t)\right)$ was minimum, then the probability of $Y(t+1)$ switching to a new value will also be low.

The "OF" is an important consideration that must be taken into consideration when developing an optimization issue. The goal of wrapper-based FS techniques is to minimize the count of features while maximizing the accuracy of the learning process. Both of these conflicting objectives should be taken into account when formulating the "OF." The objective function in this study incorporates the minimization of both the classification error rate and selection ratio, as represented by Equation (4).

$$Fitness = \alpha \times ERR(D) + \lambda \times \frac{|F|}{|N|} \tag{4}$$

ERR(D) denotes the rate of classification errors, which is achieved by utilizing the KNN classifier), $|F|$ is the number of selected features and the initial number of

features is denoted by $|N|$. The parameters α and λ are both within the range of [0, 1], with α being the complement of λ. The weights of the classification error rate and selection ratio are represented by these two parameters, respectively [23]. Table 2 represents the data set features and the descriptions of all the features and its types [17].

Table 2. Features in the dataset

Metrics	Descriptions	Type
SrcLoad	Source load (bits per second)	Flow metric (FM)
SrcBytes	Source bytes	FM
DstBytes	flow record's Destination bytes	FM
SrcJitter	Source jitter	FM
DstLoad	Destination load (bits per second)	FM
DstJitter	Destination jitter	FM
TotPkts	Total packets count	FM
Heart_Rate	Heart rate	Biometric (BM)
TotBytes	Total bytes of the packets	FM
Resp_Rate	Respiration rate	BM
DIA	Diastolic BP	BM
Temp	Temperature	BM
SYS	Systolic BP	BM
Pulse_Rate	Pulse rate	BM
ST	ECG ST segment	BM

It has a total of 44 features, of which 35 are network traffic features. The BDA was utilized to choose 15 features out of a total of 44 characteristics in the IoMT dataset. Next, the selected attributes are inputted into the classification algorithm to categorize them for training and testing. This classification technique will improve the efficiency of the model.

3.4 Classification using DAN2

In this study, the classification is done by using the DAN2 model. DAN2 utilizes a distinct structure compared with existing neural network methodologies, such as feedforward backpropagation (FFBP). The fundamental idea of the DAN2 model aims to acquire and store knowledge in all the layers, transmit and adjust this knowledge to subsequent layers, and repeat these processes till the performances of the network are met. The DAN2 architecture was represented in Figure 4. The DAN2 architecture, similar to traditional neural networks, comprises an input layer (IL), hidden layers (HL), and an output layer (OL). The IL of the model receives external data. In DAN2, the number of HL is not predetermined, which sets it apart from standard neural networks. The entities are generated in a sequential and dynamic way till a desired degree of accuracy in performance was attained. Furthermore, the suggested method employs a predetermined quantity of hidden nodes (4) within all the HLs. This model was not random but rather supported by the estimation method.

At all the HLs, the network was trained by simultaneously analyzing all findings in the set of training, with the objective of reducing a designated training accuracy metric, such as the MSE value or Figure 4 demonstrates that every HL was made of four nodes. The starting node was the constant or bias (i.e., 1) input node, also called the C node. The "current accumulated knowledge element" (CAKE node) is a second node that integrates function from the earlier training step. The 3rd and 4th nodes indicate the rest of the nonlinear components of the current step, which is determined by a transfer function that combines and normalizes the weighted variables of input. CURNOLE nodes indicate the present residual nonlinear component.

Figure 4 depicts the input node as I, the constant nodes as C, the G_k and H_k nodes as the CURNOLE nodes, and E_k as CAKE nodes (CN). The last CN shows the variable that is influenced by other variables, often known as the output. At each layer, the input for producing the next output value (E_q). Consists of the previous four nodes: (C, M_q, H_q, *and* E_{q-1}) arcs parameters resulting in the output nodes. The variables (α_q, b_q, c_q, d_q) reflect the weights assigned to all inputs in the output calculations for the subsequent layers. The μ_q, which connects the CURNOLE nodes, was utilized as an argument for CURNOLE nodes. It represents the related contributions of all input vectors to the last values of output at all layers. The training method starts with the specialized layers in which the CN gathers the linear aspect of the input. Therefore, the content of the input was obtained by taking the input variable's weighted sum and the constant input nodes. These weights can be simply obtained using conventional LR. Once the target accuracy level was attained, the relationship was linear and the training procedure was terminated. For difficulties in classifying data, this process was substituted by another approach. Nonlinear relations require the addition of more hidden layers. In all the subsequent layers, the CAKE node's input was calculated by taking a linear combination (weighted sum) of the CAKE, CURNOLE, and C nodes from the previous layer. During the training process, the CN effectively retains and transfers a sufficient amount of knowledge acquired from previous layers. The procedure guarantees that the performance or acquired information up to this point is refined and enhanced without being forfeited. This characteristic of DAN2 incorporates the process of retaining knowledge into this model. The DAN2 method guarantees a reduction in residual error and a monotonic rise in accumulated knowledge during network training.

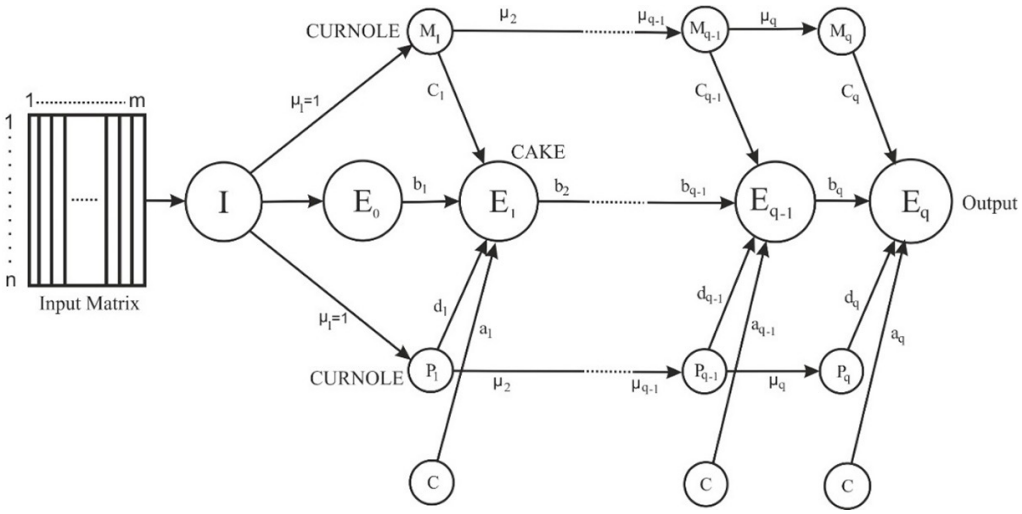

Fig. 4. DAN2 network architecture

The training procedure involves the development of partitions that might include both linear and nonlinear components among classes. The initial CN utilizes ordinary least square (OLS) or various computationally common techniques to capture the linear aspect of the input data. Subsequently, the algorithm modifies the input to represent the process nonlinearity in successive rounds. The model employs the vector projection methodology to execute data transformations. The transformations involve generating the reference vector $R = \{r_j\, j = 1, 2, ..., m\}$, where the overall characteristics is m in the observation records. Each record is then projected in this vector for normalizing the data. The process of normalizing establishes the measurement of the angle, α_i within the records i and R as reference vectors. The model utilizes a set of α_i's for training the networks and modifies their values in all the iterations. The process of normalizing may be mathematically expressed using the trigonometric function *Cosine* $(\mu_q \alpha_q + \theta_q)$. In every HL k of the architecture, this alteration on the value of output is measured by varying $(\mu_q \alpha_q + \theta_q)$. The angle alteration $(\mu_q \alpha_q + \theta_q)$ was similar to rotating μ_q and shifting θ_q the reference vectors, hence altering the effect of estimated input vector and its contributions to the iteration outputs. Two nonlinear factors μ_q and θ_q are used in the *Cosine* $(\mu_q \alpha_q + \theta_q)$. If the cosine function is expanded into the form *A Cosine* $(\mu_q \alpha_i)$ + *B Sine* $(\mu_q \alpha_i)$. This form of function is the TF in the model. This version is shown by the two CURNOLE nodes in Figure 4. For a particular HL k, if the preceding layers' captured *Cosine* $(\mu_q \alpha_i + \theta_q)$ terms do not sufficiently represent the nonlinear process behavior, a newer layer with extra nodes set was created automatically, with a new *Cosine* $(\mu_q \alpha_i + \theta_q)$ term. This method resembles how the Fourier series integrates additional phases to enhance how close it gets to a function. As a result, the number of layers in the DAN2 system changes based on how complicated the process is and how precise the results need to be. This model's output can be written as the sum of the constant CURNOLE nodes and CAKE. The mathematical formula for the relationship at iteration (layer) q is shown by Eq. (5):

$$E_q(X_i) = a_q + b_q E_{q-1}(X_i) + c_q M_q(X_i) + d_q H_q(X_i) \tag{5}$$

Where, X_i indicates the n independent input samples, $E_q(X_i)$ indicates the output values at layer q, $M_q(X_i)$ = *Cosine* $(\mu_q \alpha_q)$ and $P_q(\mu_q \alpha_i)$ = *Sine* $(\mu_q \alpha_i)$ indicates the transferred nonlinear variables, and α_q, b_q, c_q, d_q, and μ_q are parameters at iteration q. During the training process, the linear part is first picked up using OLS. The training stops when the amount of accuracy that was wanted is reached. Alternatively, the model adds more layers for showing the nonlinear part of the phase by decreasing a total error measure called $SSE_q = \sum_i [E_q(X_i) - E\hat{}(X_i)]^2$. Putting $E_q(X_i)$ into Eq. (5) gives:

$$SSE_q = \sum_i [a_q + b_q E_{q-1}(X_i) + c_q Cos(\mu_q \alpha_i) + d_q Sin(\mu_q \alpha_i) - E\hat{}(X_i)]^2 \tag{6}$$

where observed output values are $E\hat{}(X_i)$. Minimizing Equation (6) needs five parameters for estimation. This equation was linear in the set of parameters A_q, where $A_q = \{\alpha_q, b_q, c_q, d_q\}$ and nonlinear in parameter μ_q. It is stated that a number of nonlinear optimization methods can be used to find the nonlinear parameter μ_q. We also show that this method leads to steady increases in knowledge gained at each layer, a decrease in overall error, and better network training.

The method developed by DAN2 has a great level of scalability. The algorithm consists of a sequence of layers that were created dynamically and automatically.

At all the layers, the observation vectors were transformed by projecting them onto the reference vector, resulting in the creation of the angle α_i, as previously mentioned. During the optimization and training process, just converted data is used to calculate Eq. (6)'s five parameters. Thus, irrespective of the initial data size collection, the model just requires to calculate the four linear parameters set, $A_q = \{\alpha_q, b_q, c_q, d_q\}$, and one nonlinear parameter μ_q at all the iterations. The angle, α_i, at which it is projected is then calculated. The training process exclusively utilizes the α_i values and solely requires the computation of the five parameters. Consequently, the initial issue of dealing with an enormous number of features is resolved by transforming it into a model with only five parameters per layer. This approach allows for the problem size to be scaled and showcases the model's ability to handle larger datasets. The scalability of DAN2 sets it apart from standard artificial neural networks [28].

The initial linear layer:

$$E_0(X) = a_0 + \sum_j b_{0j} x_{ij} \tag{7}$$

The CN of subsequent hidden layers at iteration k:

$$E_q(X_i) = a_q + b_q E_{q-1}(X_i) + c_q M_q(X_i) + d_q P_q(X_i) \tag{8}$$

input and transfer function at iteration of CURNOLE node's q (q = 1, 2, ... Q; where Q was the number of HL or maximum sequential iterations was defined as:

A random set of m constant was specified, which represents the "reference" vectors Z (default $r_j = 1$ for every j = 1, 2, ..., q, m).

For all the input records X_i, calculate the scalar products:

$$Z \times X = \sum_j r_j x_{ij} \tag{9}$$

The length of the vector R and record vector were calculated

$$X_i : \|Z\| = \sqrt{\sum_j r_j^2} \; ; \|X_i\| = \sqrt{\sum_j x_{ij}^2} \tag{10}$$

Normalize $Z \times X$ to compute

$$(Z \times X)_N = (Z \times X_i) = (Z \times X_i) / \left(\|Z\| \times \|X_i\|\right) \tag{11}$$

Recall that:

$$(Z \times X_i)_N = (Z \times X_i) = \left(\|Z\| \times \|X_i\|\right) \times \cos(angle\ (Z, X_i)) \tag{12}$$

thus,

$$cos(angle\ (Z, X_i)) = (Z \times X_i) / \left(\|Z\| \times \|X_i\|\right) = (Z \times X_i)_N \tag{13}$$

For i = 1, 2, ..., n; compute

$$angle\ (Z, X_i) = arccos\ (Z \times X_i)_N = \alpha_i \tag{14}$$

The nonlinear transferred component was calculated using:

$$M_q(X_i) = \cos(\mu_q \times \alpha_i),\ P_q(X_i) = \sin(\mu_q \times \alpha_i), \tag{15}$$

Here, constant multiplier for iteration k was μ_k. Replacing $M_q(X_i)$ and $P_q(X_i)$ in Eq. 15 will result as follows.

$$E_q(X_i) = a_q + b_q E_{q-1}(X_i) + c_q \cos(\mu_q \times \alpha_i) + d_q \sin(\mu_q \times \alpha_i) \tag{16}$$

Normalization of data in DAN2 could be denoted by the trigonometry functions cos $(\mu_q \times \alpha_i + \theta)$. To reduce the resulting total error, vector Z is rotated and shifted at every layer. If the process of training the model is terminated in advance, it is referred to as under-training or under-fitting of the network. An insufficiently trained model frequently exhibits elevated values of sum of squared errors for whether one or testing and training data. Under training frequently arises once there is an inadequate amount of data available for the model to be properly fitted. DAN2 uses $\mathcal{E}_1 = (SSE_q - SSE_{q-1}) / SSE_q \leq \mathcal{E}_1^*$ for assessing the absence or existence of the model's under-training. Over-training, also known as over-fitting, was a prevalent issue in neural network modelling. An over-fitted (over-trained) neural network model occurs when it performs well on the data it was trained on but performs poorly on new, unknown information. To mitigate the issue of overfitting, it is advisable to partition the available data in-sample into separate validation and training data. At all iterations, q, (q > 1), they measure MSE values for testing (MSEV) and training (MSET) sets and they use $\mathcal{E}_2 = \frac{\left|MSE_T - MSE_V\right|}{MSE_T} \leq \mathcal{E}_2^*$ to guard against over-fitting [29].

4 EXPERIMENTAL ANALYSIS

4.1 Experimental setup

The experiments are carried out with MATLAB installed on a laptop running Windows 10 with an Intel(R) Core i5-9750H CPU and 8GB of RAM. The IoMT dataset was utilized in the performance evaluation of this study. 25% of the data was utilized to test the research methodology, while the remaining 75% was applied to train it. This dataset is publicly available and can be downloaded from the website: https://www.cse.wustl.edu/~jain /ehms/index.html 15.

4.2 Evaluation metrics

The performance efficiency of the BDA-DAN2 model is assessed using the following metrics:

Accuracy quantifies the ratio of accurate classifications out of the overall data.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

Recall is the proportion of properly classified data to the total data analyzed.

$$Recall = \frac{TP}{TP + FN} \tag{18}$$

Precision is determined by comparing the total accurate and incorrect classifications.

$$Precision = \frac{TP}{TP + FP} \tag{19}$$

F1-score is the recall and precision harmonic mean and provides a balanced evaluation of the performance of a model.

$$F1score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{20}$$

True positive (TP) refers to the total instances of a specific attack class properly classified. True negative (TN) represents the total instances correctly designated as not belonging to the class. False positive (FP) is the total number of instances inaccurately classified as belonging to a particular attack class when they do not. False negative (FN) is the total number of instances incorrectly classified as not pertaining to a specific attack class.

4.3 Performance evaluation

The proposed model is trained and tested on the IoMT dataset. The performance of the model was evaluated by the evaluation metrics of accuracy, precision, recall, and F1 score. The data from the dataset are divided into testing and training. As shown in Table 3, the proposed BDA-DAN2 model produces excellent results in various metrics.

Table 3. Research model's training and testing performances

Parameters	Training Performance	Testing Performance
Accuracy	99.12	98.92
Precision	99.28	98.50
Recall	99.40	98.68
F1-score	98.56	97.90

The research model's performances were calculated using the training and testing sets of data that had been split from the dataset, as shown in Table 3. In comparison to the testing set, the research model performed better in the training set. Accuracy can be defined as the proportion of total predictions that were correct according to the classification model out of the total number of predictions. The model achieved 99.12% accuracy in training and 98.92% accuracy in testing. The evaluation of classification methods can be done using accuracy. Accuracy in testing and training differs by 0.2%. An estimation of precision can be obtained by dividing the count of accurately identified positive samples by the sum of correctly identified and incorrectly identified positive samples. The precision indicates the accuracy with which the model predicts whether a sample is positive or negative. The research model's precision was 99.28% during training and 98.50% during testing, with a 0.78% variance.

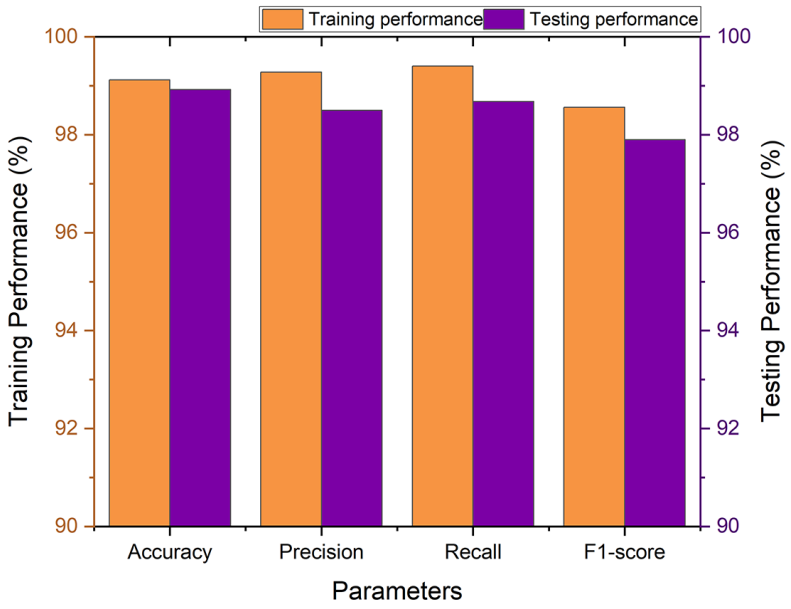

Fig. 5. Graphical plot of research model's training and test performance

Recall is defined as the proportion of TPs vs. TPs and FNs. It is also termed as sensitivity and detection rate. The proposed model achieves a recall score of 99.40% during training and 98.68% during testing, with a 0.72% variance. The F1 score, which can also be referred to as the F1 score, is a score that determines how accurate a model is on a dataset. It is utilized in the evaluation of classification systems, namely in the categorization of samples as either positive or negative. F1-score is a metric that considers recall (sensitivity) in addition to precision. Its definition is the harmonic mean of the model's precision and recall. The research model attained a 98.56% F1 score in training and 97.90% in testing; the variation between them was 0.66%. Figure 5 indicates the graphical chart for the testing and training sets performances of the research model.

The testing performance of the research model and existing models that are part of the literature study are compared and shown in Table 4.

Table 4. Comparison of performance analysis

Model	Accuracy	Precision	Recall	F1-Score
KNN [11]	98.87	97.52	95.66	96.58
SVM [11]	97.85	93.75	93.33	93.54
Safety-Med [13]	97.63	94.87	97	97.73
PSO-DNN [17]	96	–	–	–
FST-LSTM [19]	96.7	–	–	96.6
XMeDNN [21]	97.57	–	–	0.976
PSO-Adaboost [24]	98.5	98.30	96.6	–
DBN- botnet [25]	97.93	96.21	98.54	0.97
PCA-MLP [26]	96.39	–	–	–
Proposed model BDA-DAN2	98.92	98.50	98.68	97.90

Figure 6 represents the accuracy comparison. The research model has an accuracy of 98.92%, which is 0.05% to 2.53% higher than the other models. It shows that the BDA-DAN2 model is the most effective model for accurate classification in the IoMT environment among the existing models. Other high-performing models include KNN with 98.87% accuracy and PSO-Adaboost with 98.5%. The lower accuracy is shown in PSO-DNN with 96% and PCA-MLP with 96.39. The proposed model has the highest accuracy, indicating superior performance in accurately classifying IoMT data compared to existing models.

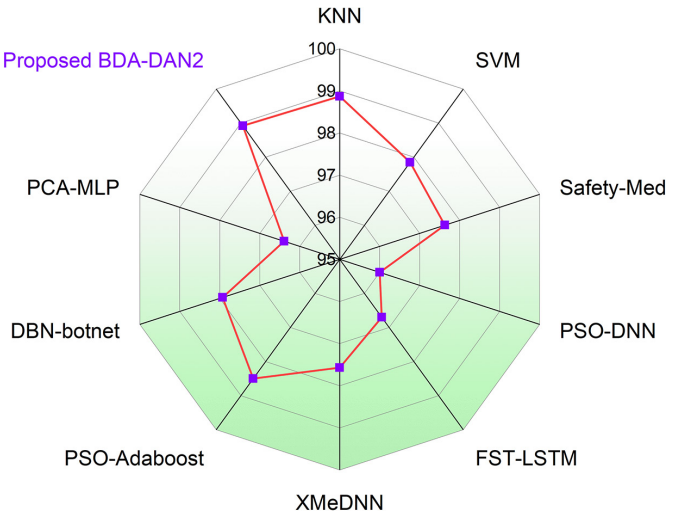

Fig. 6. Graphical plot of accuracy comparison

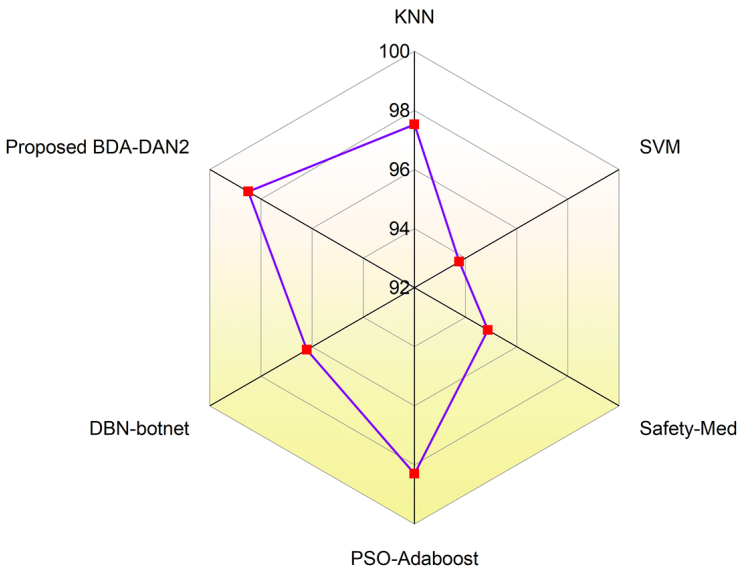

Fig. 7. Graphical plot of precision comparison

The graphical figure for the precision analysis comparison is shown in Figure 7. Among the models, the proposed BDA-DAN2 model achieves the highest precision value of 98.50%, better than the compared models by 0.2% to 4.75%. PSO-Adaboost follows closely with a precision of 98.30%, defining strong performance. KNN also performs well with a precision of 97.52%. DBN-botnet and Safety-Med achieve lower but still good precisions of 96.21% and 94.87%, respectively. SVM has a precision of

93.75%, showing good but less optimal performance. Overall, BDA-DAN2 achieves high precision, showing its effectiveness in correctly identifying positive instances in the IoMT environment.

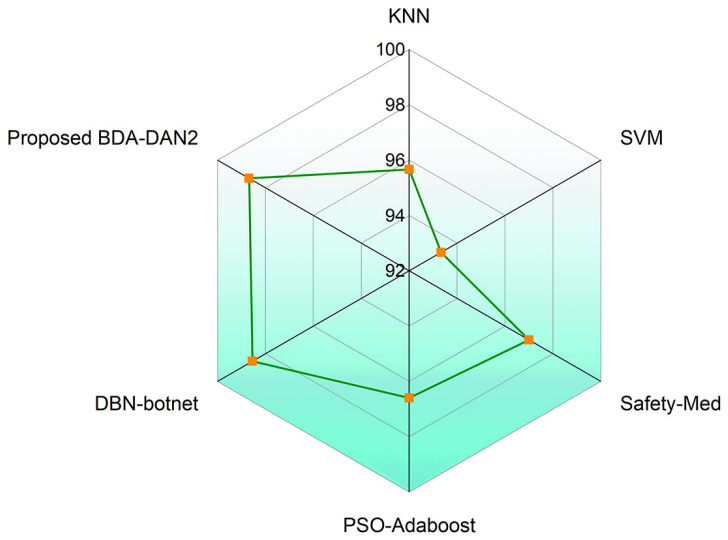

Fig. 8. Graphical plot of recall comparison

The graph for the recall analysis comparison is displayed in Figure 8. The research model's recall was 98.68%, which was 0.14% better than the DBN-botnet model. The proposed model BDA-DAN2 leads with the highest recall of 98.68%, showcasing its superior capability. Safety-Med [13] also performs excellently with a recall of 97%, close to the proposed model. SVM produces the lowest recall of 93.33%, indicating its weaker performance. DBN-botnet achieves recall as 98.54%, close to the BDA-DAN2 model. Overall, the BDA-DAN2 model's recall score is the best and highest compared to the existing models.

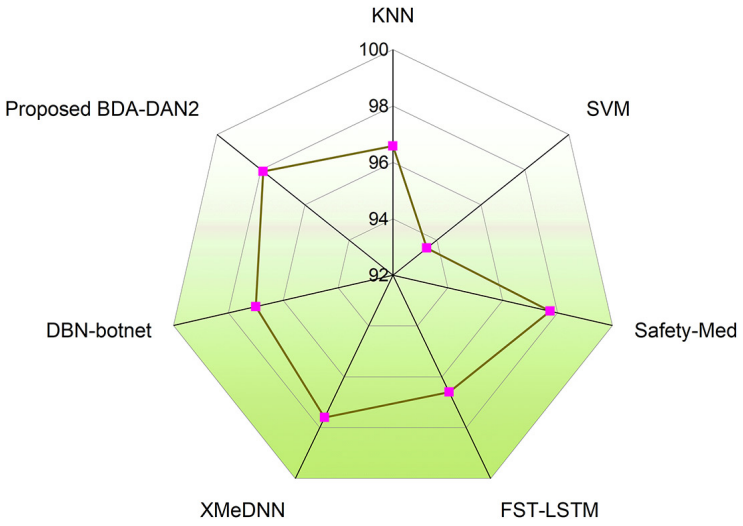

Fig. 9. Graphical plot of F1-score comparison

The graph for the F1 score analysis comparison is displayed in Figure 9. The research model's F1 score value was 97.90%, better than the compared models by 0.17% to 4.36%. Safety-Med [13] also shows a strong F1 score of 97.73%, indicating

its high reliability. XMeDNN and DBN-botnet achieve the high FI score closest to the BDA-DAN2 model. SVM was performed with a lower F1 score of 93.54%, denoting a less optimal balance between recall and precision compared to the other models. As a result, the BDA-DAN2 model produces the highest F1-score value, highlighting its effective performance in the IoMT environment.

The proposed BDA-DAN2 model outperforms the other models in the table by achieving the highest accuracy (98.92%), precision (98.50%), and recall (98.68%), along with a strong F1 score (97.90%). Compared to KNN (98.87% accuracy, 96.58 F1 score) and SVM (97.85% accuracy, 93.54 F1 score), the BDA-DAN2 model delivers better performance across all metrics. It also surpasses models such as PSO-Adaboost, which has a slightly lower accuracy (98.5%) and an incomplete evaluation of precision and F1 score, and safety-med (97.63% accuracy), which, while strong in recall, falls short in precision. The consistent, high-performance metrics of the BDA-DAN2 model demonstrate its superior ability to detect intrusions in IoMT systems compared to existing methods.

Experimental evaluations of the BDA-DAN2 model were conducted using multiple metrics, namely accuracy, precision, recall, and F1 score. The research model includes three main steps. At first, the pre-processing of data was performed using the standard scalar technique to remove the redundant values. Then the processed features from the dataset are fed to the feature selection process using the BDA algorithm. Initially, the dataset contained 44 features; after the feature selection process, 15 features were selected from the IoMT dataset. The data set was split into training and evaluation sets. The training data was employed for training the DAN2 algorithm, while the evaluation data was employed to compute its performances. The main advantage of the proposed model is that it produces excellent accuracy results. The efficiency of the model is improved by the classification algorithm used. Finally, the BDA-DAN2 model achieves the highest accuracy, recall, precision, and F1 score compared to the existing models.

However, the research model has achieved the best performance; the model suffers from several limitations. The dataset size is limited; this may lead to overfitting challenges. The classification performance of the suggested model is relatively low, even when compared to existing models that are considered high. Furthermore, binary classification is used in the BDA-DAN2 model.

In contrast, future research may implement the multiclass classification system for better performance. To overcome the overfitting issue, the size of the dataset can be increased. These directions could help guide further research in the field of IDS for the Internet of Medical Things.

5 CONCLUSION

This research proposed a BDA-DAN2-based IDS for the IoMT environment. The intrusion detection model has a series of workflows, including dataset collection, pre-processing of data, selection of features, and classification. Data collected from the dataset are preprocessed by the standard scalar technique. After preprocessing the data, feature selection is done by the BDA algorithm. The IoMT dataset was used in this research to evaluate the model. To increase the accuracy of the system, DAN2 algorithm is used for the classification of features. The data set was divided for testing and training of the model. BDA-DAN2 model performance was assessed by parameters such as accuracy, precision, F1 score and recall. The research has

obtained 98.92% accuracy, 98.50% precision, 98.68% recall, and 97.90% F1-score. The BDA-DAN2 model's accuracy, precision, F1score, and Recall values were compared with those of the KNN, SVM, Safety-Med, PSO-DNN, FST-LSTM, XMeDNN, PSO-Adaboost, DBN-botnet and PCA-MLP models determined in the literature review. In conclusion, the proposed BDA-DAN2 intrusion detection system model for IoMT environments has been validated and has yielded excellent results. The research methodology is most appropriate for the IoMT application, where smart medical devices connect with one another through a network. Future studies could increase the accuracy of the proposed model by increasing the data size and implementing multi-class classification techniques.

6 REFERENCES

[1] C. Huang, J. Wan, S. Wang, and Y. Zang, "Internet of medical things: A systematic review," *Neurocomputing*, vol. 557, p. 126719, 2023. https://doi.org/10.1016/j.neucom.2023.126719

[2] S. A. Wagan, J. Ko, I. F. Sidiqui, M. Atique, D. R. Shi, and N. M. F. Quereshi, "Internet of medical thing and trending converged technologies: A comprehensive review on real-time application," *Journal of King Saud University-Computers and Information Science*, vol. 34, no. 10, pp. 9228–9251, 2022. https://doi.org/10.1016/j.jksuci.2022.09.005

[3] S. Razdan and S. Sharma, "Internet of medical things (IoMT): Overview, emerging technologies, and cases studies," *IETE Technical Review*, vol. 39, no. 4, pp. 775–788, 2022. https://doi.org/10.1080/02564602.2021.1927863

[4] R. Hireche, H. Mansori, and A. S. K. Pathaan, "Security and privacy managements in Internet of Medical Thing (IoMT): A synthesis," *Journal of Cybersecurity and Privacy*, vol. 2, no. 3, pp. 640–661, 2022. https://doi.org/10.3390/jcp2030033

[5] Y. Sun, F. P. W. Lou, and B. Loi, "Security and privacy for the Internet of Medical Thing enabled healthcare systems: A survey," *IEEE Access*, vol. 7, pp. 183339–183355, 2019. https://doi.org/10.1109/ACCESS.2019.2960617

[6] M. L. H. Jaimes, A. M. Cruz, K. A. R. Gutierez, and C. F. Uribbe, "Artificial intelligences for IoMT security: A review of intrusions detections system, attack, dataset and Cloud-Fog-Edge architecture," *Internet of Things*, vol. 23, p. 100887, 2023. https://doi.org/10.1016/j.iot.2023.100887

[7] R. Pakrooh, A. Jabari, and C. Fang, "Deep learning-assisted security and privacy provisioning in the Internet of Medical Thing System: A survey on recent advance," *IEEE Access*, vol. 12, pp. 40610–40621, 2024. https://doi.org/10.1109/ACCESS.2024.3377561

[8] P. G. Shambharkar and N. Sarma, "Artificial intelligences driven intrusions detections frameworks for the Internet of Medical Things," *Research Square*, 2023. https://doi.org/10.21203/rs.3.rs-2634004/v1

[9] R. M. Swarna Priya *et al.*, "An effective features engineering for DNNs using hybrid PCA-GWO for intrusions detections in IoMT architectures," *Computer Communication*, vol. 160, pp. 139–149, 2020. https://doi.org/10.1016/j.comcom.2020.05.048

[10] M. Norouzi, Z. G. Aydın, O. C. Tuma, M. Y. Yagcci, M. A. Aydin, and A. Sori, "A hybrid genetic algorithms-based random forests model for intrusions detections approach in Internets of Medical Things," *Applied Sciences*, vol. 13, no. 20, p. 11145, 2023. https://doi.org/10.3390/app132011145

[11] G. Zachos, I. Esop, G. Mantaas, K. Porfyrkis, J. C. Ribero, and J. Rodriguez, "An anomaly-based intrusions detections systems for internet of medical things network," *Electronics*, vol. 10, no. 21, p. 2562, 2021. https://doi.org/10.3390/electronics10212562

[12] S. Abbas, G. A. Sampeedro, M. Abisaado, A. Almador, I. Youasaf, and S. P. Hong, "Harris-Hawks-Optimizations-Based deep recurrent neural networks for securing the internets of medical things," *Electronics*, vol. 12, no. 12, p. 2612, 2023. https://doi.org/10.3390/electronics12122612

[13] N. Faruqui *et al.*, "SafetyMed: A novel IoMT intrusions detections systems using CNN-LSTM hybridizations," *Electronics*, vol. 12, no. 17, p. 3541, 2023. https://doi.org/10.3390/electronics12173541

[14] Y. K. Saheed and M. O. Arawolo, "Efficient cyberattacks detections on the internet of medical things-smart environments based on deep recurrent neural networks and machine learning algorithm," *IEEE Access*, vol. 9, pp. 161546–161554, 2021. https://doi.org/10.1109/ACCESS.2021.3128837

[15] M. A. Almaiah, A. Ali, F. Hajej, M. F. Paasha, and M. A. Alhali, "A lightweight hybrid deep learning privacy preserving model for FC-based industrial Internet of Medical Things," *Sensors*, vol. 22, no. 6, p. 2112, 2023. https://doi.org/10.3390/s22062112

[16] M. Alalhareth and S. C. Hong, "An improved mutual information features selections technique for intrusions detections system in the Internet of Medical Things," *Sensors*, vol. 23, no. 10, p. 4971, 2023. https://doi.org/10.3390/s23104971

[17] R. Chaganti, A. Morade, V. Ravi, N. Vemaprala, A. Duai, and B. Bushan, "A particles swarm optimizations and deep learning approach for intrusions detections systems in Internet of Medical Things," *Sustainability*, vol. 14, no. 19, p. 12828, 2022. https://doi.org/10.3390/su141912828

[18] O. Taouali *et al.*, "Intelligent intrusions detections systems for the Internet of Medical Thing based on data-driven technique," *Computer System Sciences & Engineering*, vol. 47, no. 2, pp. 1594–1609, 2023. https://doi.org/10.32604/csse.2023.039984

[19] M. Alalhareth and S. C. Hong, "An adaptive intrusions detections system in the Internet of Medical Thing using fuzzy-based learning," *Sensors*, vol. 23, no. 22, p. 9247, 2023. https://doi.org/10.3390/s23229247

[20] P. T. Nguyen, V. D. B. Huyh, K. D. Vo, P. T. Phann, M. Elhoaseny, and D. N. Lei, "Deep learning based optimal multimodal fusions frameworks for intrusions detections systems for healthcare data," *Computer, Materials & Continua*, vol. 66, no. 3, pp. 2555–2571, 2021. https://doi.org/10.32604/cmc.2021.012941

[21] M. M. Alani, A. Masatan, and A. Miri, "XMeDNN: An explainable deep neural networks system for intrusions detections in Internet of Medical Things," in *Proceedings of the 9th International Conference on Information Systems Security and Privacy (ICISSP)*, 2023, pp. 144–151. https://doi.org/10.5220/0011749200003405

[22] M. A. Kumar, D. Samiaya, P. M. D. R. Vincent, K. Srinivaasan, C. Y. Chang, and H. Ganesha, "A hybrid framework for intrusions detections in healthcare system using deep learning," *Frontier in Public Health*, vol. 9, p. 824898, 2022. https://doi.org/10.3389/fpubh.2021.824898

[23] A. I. Hammouri, M. Mafaraja, M. A. A. Betar, M. A. Awadalah, and I. A. Doush, "An improved dragonfly algorithm for features selections," *Knowledge-Based System*, vol. 203, p. 106131, 2020. https://doi.org/10.1016/j.knosys.2020.106131

[24] Z. Sun, G. Ann, Y. Yang, and Y. Li, "Optimized machine learning enabled intrusions detections 2 systems for Internet of Medical Things," *Franklin Open*, vol. 6, p. 100056, 2024. https://doi.org/10.1016/j.fraope.2023.100056

[25] S. Manimurugan, S. A. Mutari, M. M. Aborokabah, N. Chilamakurti, S. Ganesan, and R. Pataan, "Effective attacks detections in Internet of Medical Things smart environments using a deep beliefs neural network," *IEEE Access*, vol. 8, pp. 77396–77404, 2020. https://doi.org/10.1109/ACCESS.2020.2986013

[26] A. Judith, G. J. W. Katherine, and S. Silaas, "Efficient deep learning-based cyber-attacks detections for Internet of Medical Things device," *Engineering Proceedings*, vol. 59, no. 1, p. 139, 2023. https://doi.org/10.3390/engproc2023059139

[27] M. Mafarja *et al.*, "Binary dragonfly optimizations for features selection using time-varying transfer function," *Knowledge-Based Systems*, vol. 161, pp. 185–204, 2018. https://doi.org/10.1016/j.knosys.2018.08.003

[28] M. Ghiassi, D. Zimabra, and S. Lee, "Targeted twitter sentiments analysis for brand using supervised features engineering and the dynamic architectures for artificial neural network," *Journal of Managements Information System*, vol. 33, no. 4, pp. 1034–1058, 2016. https://doi.org/10.1080/07421222.2016.1267526

[29] E. Guresen, G. Kayakutulu, and T. U. Dam, "Using artificial neural networks model in stock markets index predictions," *Expert System with Application*, vol. 38, no. 8, pp. 10389–10397, 2011. https://doi.org/10.1016/j.eswa.2011.02.068

7 AUTHOR

Raid Mohsen Alhazmi had received PhD in Information technology from Towson University, USA at 2013. Currently, he is an Associate Professor at the Computer and information Technology College, Department of Computer Science, University of Tabuk. He is interested in software engineering, machine learning, deep learning, and big data analytics (E-mail: ralhazmi@ut.edu.sa).