IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 September 2023, accepted 26 October 2023, date of publication 31 October 2023, date of current version 3 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3328909

RESEARCH ARTICLE

A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks

MANA SALEH AL RESHAN[1], SAMINA AMIN[2], MUHAMMAD ALI ZEB[2], ADEL SULAIMAN[3], HANI ALSHAHRANI[3], AND ASADULLAH SHAIKH[1], (Senior Member, IEEE)

[1]Department Information Systems, College of Computer Science and Information Systems, Najran University, Najran 61441, Saudi Arabia
[2]Institute of Computing, Kohat University of Science and Technology, Kohat 26000, Pakistan
[3]Department Computer Science, College of Computer Science and Information Systems, Najran University, Najran 61441, Saudi Arabia

Corresponding author: Samina Amin (kustsameena@gmail.com)

This work was supported by the Deanship of Scientific Research at Najran University through the Research Priorities and Najran Research Funding Program under Grant NU/NRP/SERC/12/35.

ABSTRACT Heart Disease (HD) is recognized as the leading cause of worldwide mortality by the World Health Organization (WHO), resulting in the loss of approximately 17.9 million lives each year. HD prediction is found to be a challenging issue that can provide a computerized estimate of the level of HD so that additional action can be simplified. Early detection and accurate prediction of HD play a critical role in providing timely medical interventions and improving patient outcomes. Thus, HD prediction has expected massive attention worldwide in healthcare environments. Deep Learning (DL) based systems played a significant role in various disease prediction and diagnosis with good efficiency. To this end, the main contribution of this paper is to design a robust HD prediction system using Hybrid Deep Neural Networks (HDNNs) involves combining multiple neural network architectures to extract and learn relevant features from the input data. The HDNN is employed to apply its feature learning capabilities and non-linear technology to capture complex patterns and relationships in HD datasets, leading to enhanced prediction accuracy. For this, three DL models, namely Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and a new HDNN model combining both CNN and LSTM along with additional Dense layers are proposed, to develop the hybrid HD prediction architecture. The proposed models were evaluated on two publicly available HD datasets, including the Cleveland HD dataset, and a large public HD dataset (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA). Additionally, the proposed system was measured through comparison with conventional systems concerning sensitivity, Matthews Correlation Coefficient (MCC), F1-measure, accuracy, precision, AUC, and specificity. The promising accuracy achieved through the proposed system is 98.86%. The results demonstrated that this approach proved more accurate in its predictions than previous research. These outcomes suggest that the proposed HDNN system has great potential to be embedded into healthcare systems to develop advanced and reliable HD prediction models that can significantly contribute to medical diagnosis and improve patient care.

INDEX TERMS Cardiovascular disease, heart disease prediction, Cleveland heart disease dataset, deep learning, hybrid deep neural networks, CNN-LSTM.

I. INTRODUCTION

Over the past decade, the prevalence of HD, also known as Cardiovascular Disease (CVD), has experienced a substantial surge, making it the primary cause of death in most countries globally. HD comprises a variety of factors that affect the structure or function of the heart, making it difficult for medical practitioners to make a prompt and accurate diagnosis. To help doctors identify HD/CVD faster and more effectively, it is vital to use digital technologies. As a result, the

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamed Elhoseny.

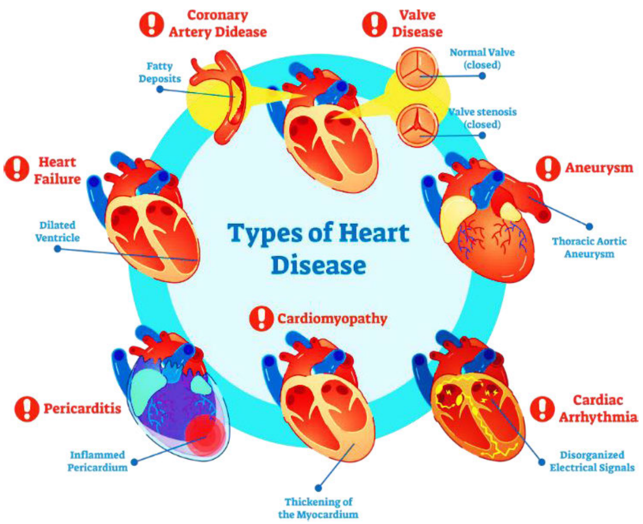

FIGURE 1. Types of HD (Figure source [2]).

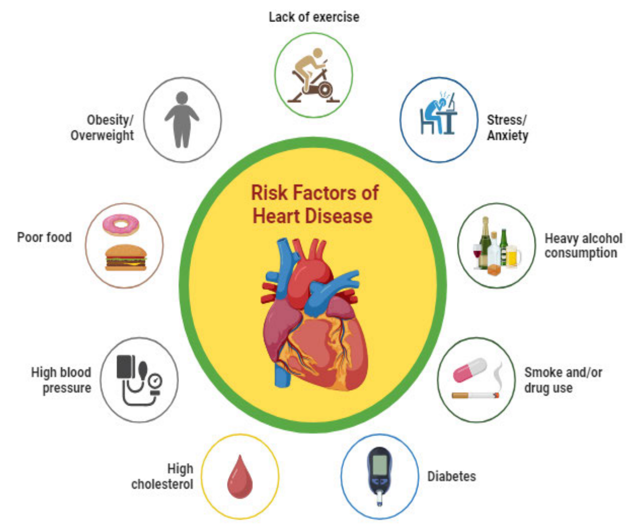

FIGURE 2. Cardiovascular risk factors: Generated by biorender.com.

incorporation of computerized technologies in HD diagnosis has become crucial to support doctors in making faster and more accurate assessments. These technological developments can support the fight against this widespread and potentially fatal disorder by raising diagnostic accuracy, enabling prompt therapies, and eventually improving patient outcomes [1].

As per the WHO, CVDs/HD is recognized as the leading cause of worldwide mortality, claiming approximately 17.9 million lives annually. This category encompasses various heart and blood vessel disorders, including coronary HD, cerebrovascular disease, rheumatic HD, and other related conditions. HD encompasses various types, including coronary artery HD, congestive heart failure, heart valve HD, cardiomyopathy, heart arrhythmia, and pericarditis as depicted in Figure 1 [2], and its origins often stem from multiple factors, including stress, the intake of poor foods, living a sedentary lifestyle, overweight/obesity, lack of exercise, high blood pressure, smoking, diabetes, and binge drinking alcohol, etc. [3], [4], and can be illustrated in Figure 2. Notably, heart attacks and strokes account for more than 80% of CVD-related fatalities, and a frightening one-third of these deaths occur in those under the age of 70. In order to lessen the impact of HD/CVDs on public health, there is an urgent need for effective HD prevention, early detection, and improved management. Therefore, early detection of heart failure is essential since it gives researchers the chance to test and create efficient pharmaceutical and lifestyle therapies. This is particularly significant due to the fact it may help prevent or delay the progression of HD, lowering the risk of mortality [5].

A subset of machine learning (ML) known as ''deep learning'' (DL) often considers multiple layers of information-processing stages in hierarchical structures. DL approaches have always been of interest because of their accuracy in problem-solving [6], [7]. A good example of the use of DL in intelligent healthcare systems is the processing of medical images [8], electronic health records [9], gene research [10], and disease detection in text data [11], which are all carried out with DL. For the prediction and diagnosis of diverse diseases, many expert systems and DL techniques have recently gained prominence in medical decision support systems [12], [13], [14], [15], [16].

Motivated by the recent advancements in DL for medical decision support systems [17], [18], this work introduces an enhanced and novel Hybrid Deep Neural Networks (HDNNs) system, utilizing a larger and a smaller dataset to design the system effectively. By integrating DL approaches into the existing ML methodologies, researchers can potentially achieve more accurate and reliable predictive methods. DL technologies can effectively process and analyze large volumes of medical data, leading to enhanced precision in detecting HD, predicting patient outcomes, and supporting clinical decision-making. The current state of research in the field of HD prediction reveals that none of the existing studies have applied DL approaches such as deep ANN, LSTM, CNN, and HDNN e.g., CNN with LSTM models for HD prediction. A robust HD prediction system using HDNN offers several advantages over traditional ML methods. However, it is essential to acknowledge that traditional ML methods can still be valuable for HD prediction, especially when data is limited, or interpretability is a crucial factor. In some cases, simpler models like logistic regression or decision trees may be preferred, especially when the focus is on interpretability. Ultimately, the choice between HDNNs and traditional ML models depends on the specific requirements of the HD/CVD prediction task, the availability of data, computational resources, interpretability needs, and the level of complexity in the relationships within the data.

Thus, developing a robust HD/CVD prediction system using HDNNs involves combining multiple neural network architectures to extract and learn relevant features from the input data. In this study, a combination of CNN and LSTM networks, along with additional Dense layers, to create the hybrid architecture will be leveraged. For instance, we use a combination of a 1D CNN, an LSTM layer, and additional Dense layers to create an HDNN for HD/CVD prediction. To operate large-scale medical datasets. As the amount of available data grows, the HDNN can be able to amend and operate well without compromising computational efficiency. For this, we compile and train the model using the two HD datasets, including the UCI Kaggle Cleveland HD dataset [19], and a comprehensive HD dataset such as (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA) dataset [20]. After training, the model's performance will be evaluated by applying various metrics, mainly precision, recall, accuracy, MCC, F1-measure, specificity, ROC curve, and AUC score.

The main contributions of the proposed HDNN are defined as follows:

1. To design a new robust HD prediction model that achieves high accuracy and diagnoses HD effectively. The HDNN is employed to apply its feature learning capabilities and non-linear technology to capture complex patterns and relationships in HD datasets, leading to enhanced prediction accuracy.
2. To outperform traditional ML methods, namely Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbor (KNN), and Random Forest (RF) in HD prediction tasks. By applying deep ANN, LSTM, CNN, and a hybrid CNN-LSTM, the model will validate superior predictive performance compared to conventional ML systems, particularly when dealing with complex medical records.
3. To exploit the inherent ability of HDNN (CNN-LSTM) to automatically learn relevant features from complex data. The algorithm can find significant and informative features that contribute to HD prediction without the requirement for manual feature engineering.
4. To create an HD prediction system that can be effectively integrated into clinical practice. The proposed HDNN system can provide accurate and reliable predictions that can assist healthcare specialists in making informed decisions, facilitating earlier diagnosis, and improving patient outcomes.
5. To evaluate and compare the performance of the proposed HDNN against existing state-of-the-art HD prediction approaches, including conventional ML methods.

The rest of the article is categorized in the following manners: Section II discusses a background study conducted on recent literature for HD/CVD prediction and classification. Section III discusses the research gap and Section IV demonstrates the proposed model architecture in detail. Section V presents experimental results conducted on the two benchmark datasets and their evaluations. Finally, Section VI provides conclusions.

II. RELATED WORK

Researchers have used various data mining approaches, ML methodologies, and neural networks to develop an HD prediction approach [13], [16]. For example, Tarawneh and Embarak [21] designed an HD prediction mechanism based on the hybrid approach utilizing data mining techniques, and Nalluri et al. [22] presented a chronic HD prediction employing data mining methods, including XGB and LR algorithms. Gokulnath and Shantharajah [23] designed an optimized feature selection model using a genetic algorithm and SVM to diagnose HD. Li et al. [24] developed ML based diagnosis system for HD in e-healthcare. They used ANN, SVM, DT, Logistic Regression (LR), K-Nearest Neighbours (KNN), and Naive Bayes (NB) for designing the diagnosing system. Compared to the other methodologies, the SVM method achieved an accuracy of 92.37%. Bharti et al. [25] built the HD prediction system by leveraging ML and DL methods. For method building, they used the UCI Cleveland HD dataset to compare the results and analysis. The DT method accomplished a specificity of 83.1%, and a sensitivity of 82.3%. Atallah et al. [26] conducted research work on an ensemble method using a majority voting ensemble model for predicting HD in medical decision systems. The model was trained on medical test data of healthy and ill patients recorded in a local clinic. Their ensemble voting achieved an accuracy of 90% using a hard voting ensemble technique.

These days, data analysis in the healthcare industry may save lives by enhancing medical diagnostics. Additionally, several data mining technologies are now accessible to scientists and can be utilized to carry out studies and experiments owing to the enormous advancements in artificial intelligence. For this, Tougui et al. [27] compared mining frameworks, including MATLAB, RapidMiner, Orange, Knime, Weka, and Scikit-Learn, applying ML methods, mainly LR, KNN, SVM, NB, ANN, and RF in order to classify HD. The models were trained on the UCI ML HD dataset consisting of 303 instances and 13 features. For the models' performance, they used three evaluation criteria such as accuracy, specificity, and sensitivity. Mahmud et al. [28] examined various ML models, mainly SVM, KNN, LR, RF, DT, and XGB. Researchers used a Kaggle dataset with 70,000 different data values for the analysis. The experimental findings illustrated that the RF framework had the best disease prediction accuracy, with a value of 84.03%. Lutimath et al. [29] leveraged SVM to predict the HD from the HD patients dataset designed by the UCI ML repository. Pawlovsky [30] introduced an ensemble method using KNN for diagnosing HD. The ensemble has been employed with two formations. Their ensemble method achieved an average accuracy of 85%, which was assessed with the HD Cleveland data. Kavitha et al. [31] designed an ML method for predicting HD by using the UCI HD Cleveland repository and regression and classification approaches. For model implementation, RF, DT, and a hybrid of RF and DT were employed. The hybrid approach obtained an accuracy of 88.7% by predicting

HD. Another related work by Shah et al. [32], in which the authors proposed an ML-based model for HD prediction. For model training, they used NB, DT, KNN, and RF models by using the UCI HD Cleveland dataset.

Moreover, Almazroi et al. [33] designed a framework using the DL-based method ANN to predict HD. For model evaluation, they used four HD datasets, including Cleveland, Hungarian, Switzerland, and Long Beach. For model comparison, they used ML techniques and a DL-based ANN model. The experimental results showed that in all datasets DL achieved the highest accuracy of 83%. This shows a system for supporting medical decisions is performing at an appropriate level. This performance is a result of the DL model's hidden layers, where the error rate is minimized. The suggested DL framework's greater accuracy suggests that it may forecast HD effectively. Mohan et al. [34] developed an ML-based model by finding significant features for predicting HD. The prediction system was established with an ML approach such as hybrid RF with a linear approach and achieved an accuracy of 88.7%. Similarly, Chowdhury et al. [35] discussed an ML-based model for predicting HD by using an HD dataset consistent with 564 instances and 18 features. They collected the dataset by personally visiting hospitals and healthcare facilities in the Sylhet region of Bangladesh. Using the benchmark dataset, they trained the model by applying ML approaches including, DT, KNN, LR, NB, and SVM. After evaluation and comparison among the models, the SVM accomplished better with an accuracy score of 91%. Rani et al. [36] built a hybrid framework that can detect HD using the clinical parameters of the heart patient. In their work, a hybridized feature selection method incorporating a genetic approach and recursive feature elimination has been employed. For the development of the model, SVM, NB, LR, RF, and Adaptive boosting techniques were applied. For model training, the Cleveland HD dataset was used. Among the ML techniques, the model achieved the best results with the RF with an accuracy of 86.6%.

Moreover, Goyal [37] proposed HD prediction methods Lion Optimization-Based Feature Selection (LOFS)-ANN, LOFS-SVM, and LOFS-DT utilizing the LOFS technique and ML algorithms. The datasets utilized were from the UCI repository. However, Tasnim and Habiba [38] leveraged ML techniques, including NB, SVM, KNN, DT, LR, ANN, and RF to estimate the likelihood of coronary HD. Scientists have been working tirelessly to develop an innovative healthcare system. A computerized system that can identify HD risk could be considered a major accomplishment. They used the dataset from the UCI ML repository to evaluate their suggested strategy. The feature selection technique enhances the performance of traditional ML methods. Among the ML models, RF model with PCA achieved the best accuracy of 92.85% for classifying HD. Kadhim and Radhi [39] presented an ML-based model using RF, SVM, KNN, and DT algorithms. Compared to the ML models the RF performed better than other models with an accuracy score of 94.958%. Hamdaoui et al. [40] established a clinical decision support system for predicting HD to help clinicians with diagnosis and make better decisions. For model building, they employed ML algorithms including, NB, KNN, SVM, RF, and DT for predicting HD. They conducted several experiments to predict HD using the UCI HD dataset, and the result shows that NB outperforms employing both cross-validation and train-test split methods with an accuracy of 82.17%, and 84.28%, respectively. Amin et al. [41] built an HD prediction mechanism based on ML methods that can predict HD effectively. Prediction models were developed using seven classification methods, including LR, KNN, NB, DT, SVM, NN, and Vote (a hybrid method with NB and LR). After evaluation, their results show that the HD prediction voting model attained an accuracy of 87.4%. Bizimana et al. [42] designed an ML-based HD prediction approach that makes use of a variety of data scaling techniques, split ratios, ideal parameters, and ML technologies using UCI HD data. Similarly, Saboor et al. [43] designed ML based HD prediction system. To validate and compare ML models, they employed Classification And Regression Trees (CART), Adaptive Boost Classifier (ABC), LR, ETC, MNB, SVM, RF, LDA, and Extreme Gradient Boosting (XGB).

Numerous ML techniques have been used by scholars in the field of HD diagnosis and prediction. These conventional ML techniques have demonstrated notable efficacy in producing predictive strategies for HD. However, these prediction models' accuracy can be further increased with the introduction of DL and HDNN technology.

III. RESEARCH GAP

Early diagnosis plays a crucial role in achieving the goal of predicting HD promptly. In the pursuit of diagnosing and predicting HD/CVD, existing studies have employed ML techniques. In the current era, HD remains a leading cause of mortality globally, posing significant challenges for CVD prediction in clinical data analysis. The continuous growth in the size and complexity of medical datasets in the healthcare industry imposes the use of automated systems based on DL technologies to assist medical authorities/specialists in making precise and efficient decisions.

However, one of the major challenges faced by ML approaches is the decline in accuracy when dealing with large datasets. To address this issue, this research aims to introduce innovative feature optimization and classification methodologies for HD prediction by using small datasets with 303 instances and large datasets with 1190 instances, enabling medical practitioners to achieve early and accurate disease diagnosis. This research follows a similar path but introduces an enhanced and innovative HDNN approach, utilizing a larger and a small dataset to train the model effectively. The current state of research in the field of HD prediction reveals that none of the current research studies have applied DL approaches such as deep ANN, LSTM, CNN, and HDNN e.g., CNN with LSTM models for HD prediction. However, traditional ML and data mining approaches

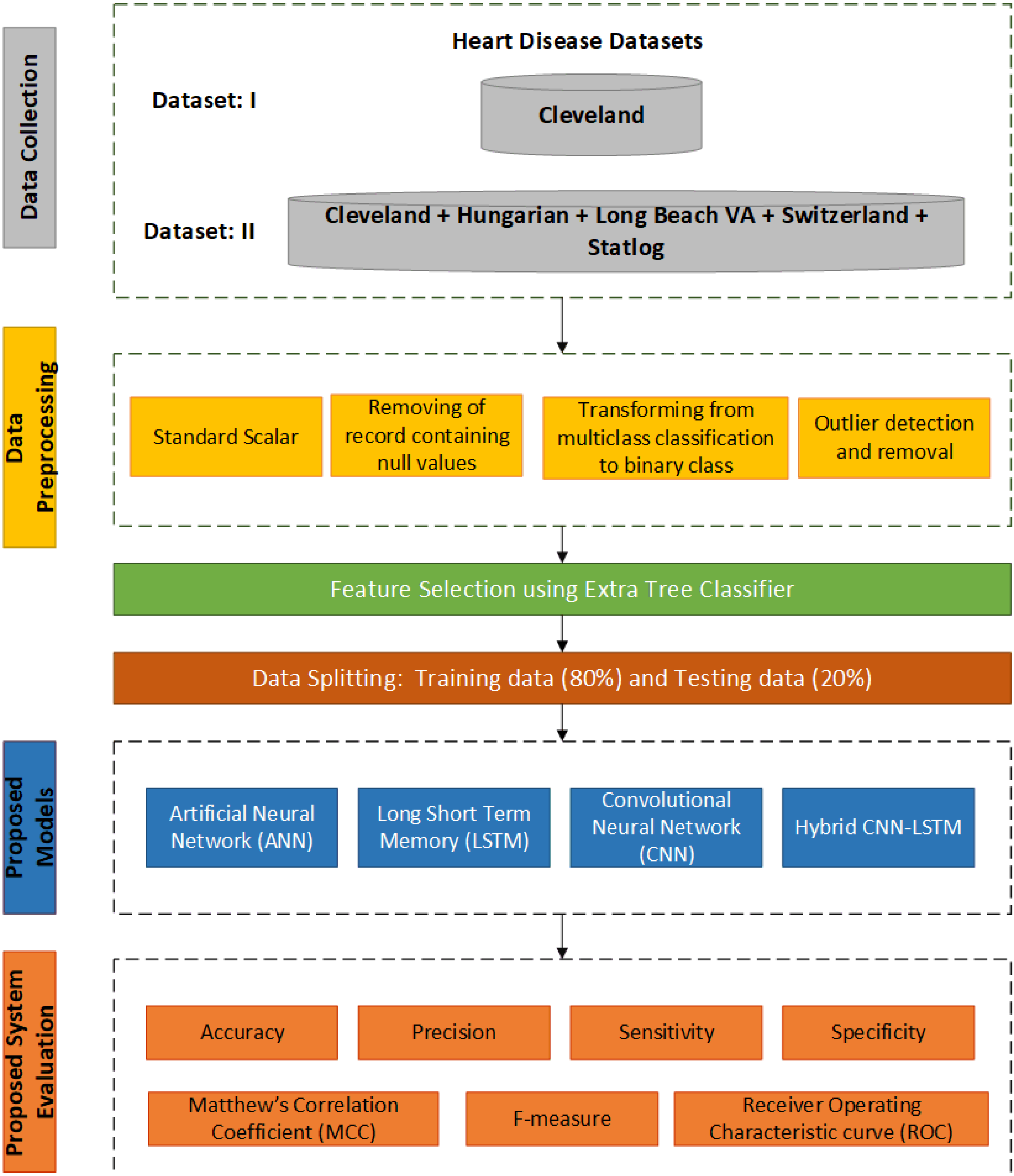

FIGURE 3. Overall flow diagram of the proposed system.

have shown promising results in predicting and classifying HD into HD positive and HD negative. Motivated by the promising development in DL methods, the main objective of this study is to design an HDNN system capable of predicting HD effectively. To achieve this goal, the proposed approach employs deep ANN, LSTM, CNN, and hybrid CNN with LSTM with multiple layers to achieve high analytical and prediction accuracy. To refine HD data and ensure its quality, the proposed HDNN system also considers data imputation strategies.

To evaluate the performance of the HDNN system, two benchmark datasets on HD are applied. It is expected that the HDNN method will greatly increase the accuracy of HD diagnosis when applied to these datasets. Two distinct comparisons are presented in the study to demonstrate the efficacy of the proposed HDNN. One comparison is carried out with the individual method, where the hybrid CNN-LSTM model is tested against deep ANN, LSTM, and CNN methods. The other comparison compares the performance of the proposed system with conventional ML techniques using deep neural networks. By applying DL approaches, the goal of this work is to improve HD diagnosis capabilities and forecasting accuracy. It is expected that the application of data imputation and DL methods will enhance our understanding of HD detection and make early and precise diagnosis easier. The objective is to investigate efficient methods to promote clinical diagnosis and early treatment, and to construct an HDNN-based HD prediction system associated with the satisfaction of health specialists. The improved performance of the proposed DL methodology over current techniques will be illustrated by a comparison with traditional and HDNNs. Furthermore, this research has the potential to contribute to the field of HD prediction by exploring the capabilities of HDNN techniques and improving the accuracy and effectiveness of HD diagnosis hopefully ultimately advancing clinical diagnosis and early treatment.

IV. PROPOSED HEART DISEASE PREDICTION METHODOLOGY

This section demonstrates HD prediction using proposed models such as deep ANN, LSTM, CNN, and a combination of CNN-LSTM architectures. Figure 3 demonstrates the working flow of the proposed HDNN system. The proposed system is implemented using the Anaconda framework (Jupyter Notebook) [44], Python programming language, and fundamental ML/DL libraries, including Scikit-learn [45], NumPy [46], and TensorFlow [47], etc. Furthermore, a brief demonstration of the suggested HD prediction system is provided in the following subsections.

A. HEART DISEASE DATASET

The HD datasets used in the suggested HDNN prediction system are taken from the widely used and freely accessible ML UCI data collection, which has been endorsed by numerous academics. In this work, two benchmark HD datasets are utilized, including the Cleveland HD dataset, and a comprehensive HD dataset such as (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA). These datasets contain particular characteristics that will be utilized to decide whether or not to diagnose HD in the patient. The output feature's label attribute has categorical values in the form of HD positive and HD negative values. HD negative values were changed to 0, and HD positive ones were changed to 1.

The detail of the two datasets is given as follows:

1) DATASET-I: CLEVELAND DATASET

This dataset on HD, which includes 2 classes, 14 attributes, and 303 occurrences, was made available by the Medical Centre and the Cleveland Clinic Foundation and may be obtained in the UCI repository [19]. A multivariate dataset with 76 variables, the Cleveland HD dataset was taken into consideration for this investigation. The Cleveland HD in particular has 76 properties and 302 occurrences. However, only 14 out of 76 traits are used in all reported research. These 14 characteristics are regarded as the Cleveland database's subset. The 13 attributes are all input attributes, and they stand alone. The final column contains an output feature, which is essentially a label attribute, and it is based on the input characteristics. For example, if it is 0, the patient is detected with HD negative and if it is 1 it means that the patient is diagnosed with HD positive.

2) DATASET-II: (SWITZERLAND + CLEVELAND + STATLOG + HUNGARIAN + LONG BEACH VA)

This HD dataset was designed by fusing five well-known HD datasets that are freely accessible on IEEE Data Port [20]. It may also be found in the Kaggle repository [48]. In this dataset, five HD datasets are integrated over 11 features which yield it the largest HD dataset accessible so far for research uses. The dataset is a combination of (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA) dataset with important features such as maximum heart rate achieved, serum cholesterol, chest pain type, fasting blood sugar, and so on. This dataset on HD comprises 11 features, 1190 instances, and 2 classes. It is a collaborative effort involving contributions from the Cleveland Clinical Foundation, the Medical Center, the Long Beach V Clinical Foundation, the Hungarian Institute of Cardiology, Switzerland, and the Stalog (Heart) Foundation. Table 1 outlines the total number of instances for each dataset.

TABLE 1. Total number of instances in HD datasets.

HD dataset types	Number of instances
Hungarian	294
Switzerland	123
Stalog (Heart)	270
Long Beach VA	200
Cleveland	303
Total	1190

B. PREPROCESSING

Data preprocessing was done following data collection. The Cleveland dataset had 6 records with missing values. The dataset was reduced from 303 to 297 records by eliminating all the records with missing values. Next, the multiclass values of the predicted attribute for the presence of HD in the dataset were converted to binary values (0 for absence;

1 for the presence of HD). To execute the data preparation operation, all diagnosis values between 2 and 4 were transformed to 1. Thus, the diagnosis values in the resulting dataset are merely 0 and 1, with 0 denoting the absence of HD and 1 denoting its existence.

Since most medical data is discontinuous, data standardization is crucial to converge the data's features. Data must also be normalized or standardized before the implementation of DL approaches. One common procedure for data standardization is z-score normalization, which uses the attribute's mean (μ_i) and standard deviation (σ_i) to normalize the attribute's data. Data standardization is a method of transforming various types of data into a format that is normalized and uniform. When μ_i and σ_i are taken into account as the mean and standard deviation of the *ith* attribute of a dataset, the z-score z_{ij} for the *jth* instance, *ij* is determined in Equation 1.

$$z_{ij} = \frac{x_{ij} - \mu_i}{\sigma_i} \tag{1}$$

C. FEATURE SELECTION

The method of feature selection involves picking features from a wide range of available qualities or features in order to reduce computational latency and complexity while boosting accuracy. Using the Model Characteristics property, the importance of each feature in HD datasets can be determined. Every function of the outcomes is given a score by feature value, demonstrating its relevance and impact on the performance variable. A higher score denotes greater importance or suitability of the feature. To extract the most important features from the dataset, in this study, the Extra Tree Classifier (ETC) is exercised using the Gini relevance technique. This model is particularly useful for feature significance analysis as it is equipped with a built-in class that enables feature importance assessment, especially in Tree-Based models. By using the ETC, we can identify the key features that significantly contribute to the system's predictive performance and gain valuable insights for our analysis. The significance of various features may have changing values as a result of the randomness of feature samples as can be investigated in Table 2.

D. DATA SPLITTING

Data splitting is a technique that involves dividing a dataset into smaller subsets [8] and the normalized preprocessed HD data is partitioned into two chunks (train, and test set). The HD dataset is partitioned into an 80:20 holdout validation technique, with 80% of the statistics being employed to train the proposed HDNN system and 20% being preserved for model evaluation (testing). Where, Train set is a statistic of data samples employed for learning, to train or fit the parameters of the model (e.g., a real dataset is applied to train the model). A Test set is a statistic of data samples that are applied only to evaluate the efficiency of the HDNN system.

TABLE 2. Feature importance in HD datasets.

Feature	Ranking
ST slope	0.16502219208594232
chest pain type	0.12464289848775935
max heart rate	0.10104757501424239
cholesterol	0.0917755431380437
exercise angina	0.13491029293081083
oldpeak	0.09079997817697384
age	0.07816976039185529
resting bp	0.07691931331560171
sex	0.06263122053450898
resting ECG	0.03856771210569106
fasting blood sugar	0.03551351381857052
ST slope	0.16502219208594232

E. PREDICTION MODELS

After determining the features, the models were built using the four DL prediction and categorization techniques, including ANN, LSTM, CNN, and Hybrid CNN-LSTM. The proposed models are described in the following sections.

1) ARTIFICIAL NEURAL NETWORK

An ANN [49] is a computational system designed with a large number of simple but well-connected processing units as shown in Figure 4. Variations in external inputs are used to process these units. The ANN architecture proposed in this study was developed by applying multiple weighted hidden layers directed by a feed-forward network with a back-propagation algorithm. This approach is commonly used in many applications [50], including image processing, speech recognition, robotic control, sentiment analysis, forecasting, and power system safety and control management are all areas where ANNs are best suited. ANN can be linked to the human brain. The human brain is made up of several neurons that function quickly. Each input (bit or signal) of

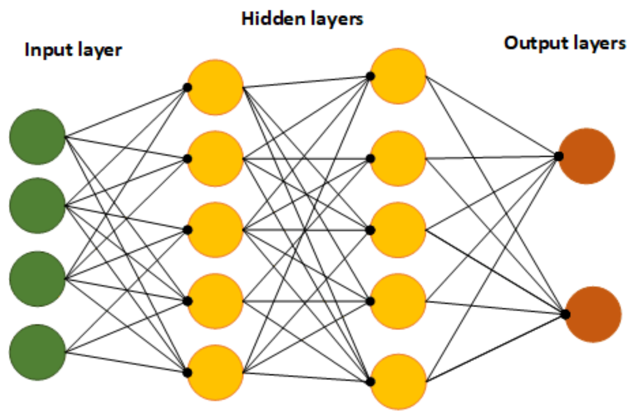

FIGURE 4. Structure of ANN.

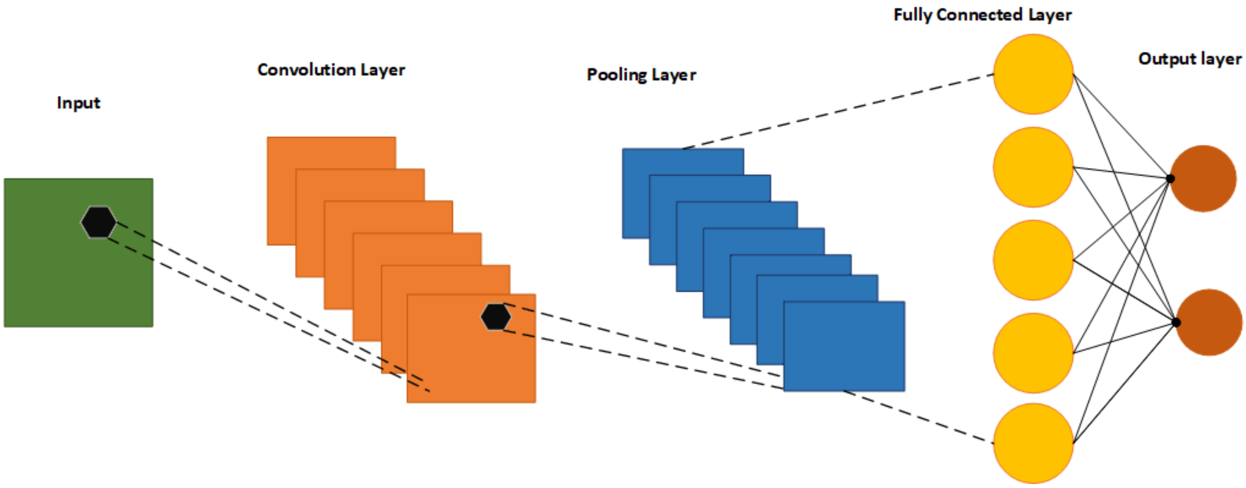

FIGURE 5. Structure of CNN.

data passes via a neuron, where it is interpreted, computed, and processed before being passed on to the next neuron cell. While each neuron or node's overall processing speed is sluggish, the network as a whole is extremely quick and effective. ANNs designed in this study along with the input layer, hidden layers, and neurons are listed in table. The first hidden layer has the same number of nodes as the input layer, as recommended by the best practices. Five neurons make up the second hidden layer, and two neurons make up the output layer for the prediction of HD cases. The activation function for the hidden layers is a ReLU algorithm, while the output layer uses a softmax activation function (e.g., binary classification).

2) CONVOLUTIONAL NEURAL NETWORK

A CNN is a DL technology capable of learning directly from data. CNNs prove to be exceptionally valuable in recognizing patterns within images, thereby identifying various objects [8], [51]. Moreover, these networks can also be extremely advantageous when it comes to categorizing non-image data such as audio, time series, and signal data. The primary concept behind CNNs is to extract local attributes from higher-level data streams and pass them to lower levels to detect more complex features. A typical CNN design involves convolution, pooling, and fully connected layers, as shown in Figure 5.

In the convolution layer, a collection of kernels is utilized to generate feature mappings in the form of a tensor. These kernels use "stride(s)" to convolve the information, resulting in output volumes of integer sizes. However, due to striding, the sizes of the convolutional layer's input volumes decrease. To preserve zero padding is required to fill the input space with zeros when the input volume has low-level features.

The Rectified Linear Unit (ReLU) function introduces nonlinearity into the feature maps. It calculates the activation by setting the threshold input to zero. The pooling layer is used to downsample a specified input dimension to reduce the number of variables. One common technique in pooling is max pooling, which retains the greatest value within a given input area. Finally, the Fully Connected layer functions as a classifier, enabling data classification based on the information obtained from the convolution and max pooling layers.

3) LONG SHORT-TERM MEMORY

LSTM belongs to the domain of DL. It falls under the category of recurrent neural networks and is renowned for its ability to grasp long-term dependencies, particularly in tasks involving sequence prediction [11] and [50]. As it progresses, it takes input and transmits it to others. The LSTM's cells perform a variety of tasks. LSTM has a memory state that can remember information and learn long-term dependencies for long periods. As a result, LSTMs have proven to be a valuable tool in various domains, including speech recognition, time series forecasting, and natural language processing.

LSTM has a significant advantage over RNN as it incorporates a cell state to store long-term information. This allows data from previous time steps to be retained and connected to data in the current time step within the LSTM network. To achieve this, LSTM employs three gates: the input gate, the forget gate, and the output gate. The present input is denoted as $\mathrm{i_t}$, while $\mathrm{C_t}$ and $\mathrm{C_{t-1}}$ represent the current and previous cell states, respectively. Similarly, $\mathrm{H_t}$ and $\mathrm{H_{t-1}}$ represent the past and present outputs.

Figure 6 illustrates the internal architecture of an LSTM, showcasing how the gates and cell state facilitate memory retention and information flow throughout the network. The LSTM layer receives the feedback obtained from the dropout layer. The calculation is trumped up of four parts: an input gate (i_t), a forget gate (f_t), an output gate (o_t), and a new memory container (c_t). To integrate the performance of forward and backward, the element-wise computation is calculated based on Equations (2 – 5):

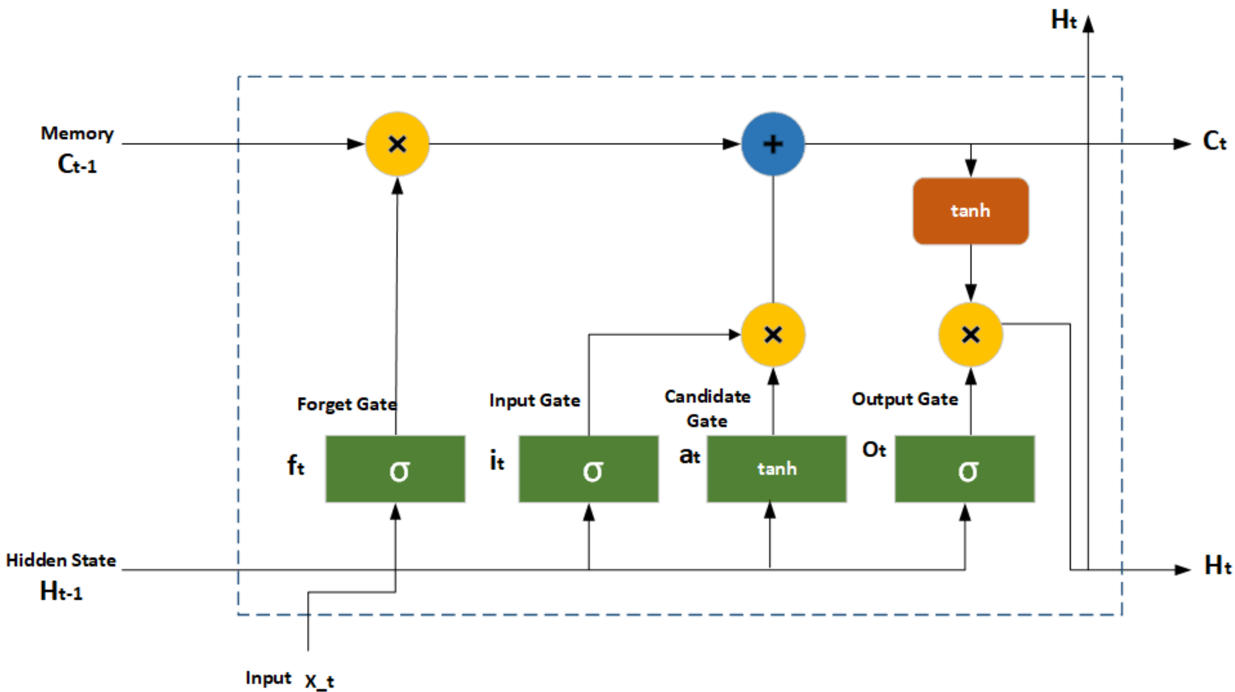

FIGURE 6. Structure of LSTM.

An LSTM receipts a current input (X_t), a previous state (h_{t-1}), does some calculation (Equations. 2–5), and afterward combines data in the arrangement of a hidden state ($\vec{h}$), as shown below:

$$f_t = \sigma \left(U_f h_{t-1}, W_f X_t + b_f\right) \tag{2}$$
$$i_t = \sigma \left(U_i h_{t-1}, W_i X_t + b_i\right) \tag{3}$$
$$a_t = tanh \left(U_c h_{t-1}, W_c X_t + b_c\right) = \tan\left(\hat{a}_t\right) \tag{4}$$
$$o_t = \sigma \left(U_o h_{t-1}, W_o X_t + b_o\right) \tag{5}$$

4) HYBRID CNN-LSTM NETWORK

This study proposed a novel strategy for predicting HD instances using multivariate HD datasets, including Cleveland and a comprehensive (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA). The model involved combining CNN-LSTM networks, where the CNN was responsible for capturing intricate audio information, while the LSTM served as the prediction model. To design the CNN-LSTM mechanism with multivariate HD datasets, features of HD-related symptoms were extracted from raw the datasets and used for training.

The implied hybrid model for HD predicted, as shown in Figure 7, comprised 30 layers: 18 convolutional layers, 12 pooling layers, 1 fully connected layer, 1 LSTM layer, and an output layer using the softmax approach. Each convolutional block included a pooling layer, two to three 2D CNNs, and one convolutional block, followed by a dropout layer with a 20% dropout rate. The features, initiated by the ReLU technique, were extracted using a convolutional layer with a 3 × 3 kernel size. The dimensions of the input multivariate HD features were reduced through a max-pooling layer with 2 × 2 kernels. The resulting feature map was then passed to the LSTM layer in the final phase to extract patient data. The output shape was molded to comprehend the convolution section (none, 8, 8, 512). The LSTM layer's input size was lowered by operating a reshaping method (16, 512). The architecture analyzed the temporal features before passing the multivariate HD features via a fully connected layer to classify each instance into two categories such as HD positive, or HD negative (no HD).

V. RESULTS AND DISCUSSION

This section summarizes the experimental findings for all four models' performance in terms of prediction, optimization, and computational cost for two HD datasets used in medical diagnosis. The efficiency of the mentioned HDNN system is assessed through the metrics for accuracy, precision, sensitivity, MCC, specificity, f-measure, and ROC/AUC. These metrics are determined by equations 5-11 [52]:

$$Accuracy = \frac{TP + TN}{FN + FP + TP + TNF} \tag{6}$$

$$Recall = \frac{TP}{FN + TP} \tag{7}$$

$$Preccision = \frac{TP}{FP + TP} \tag{8}$$

$F - measure$

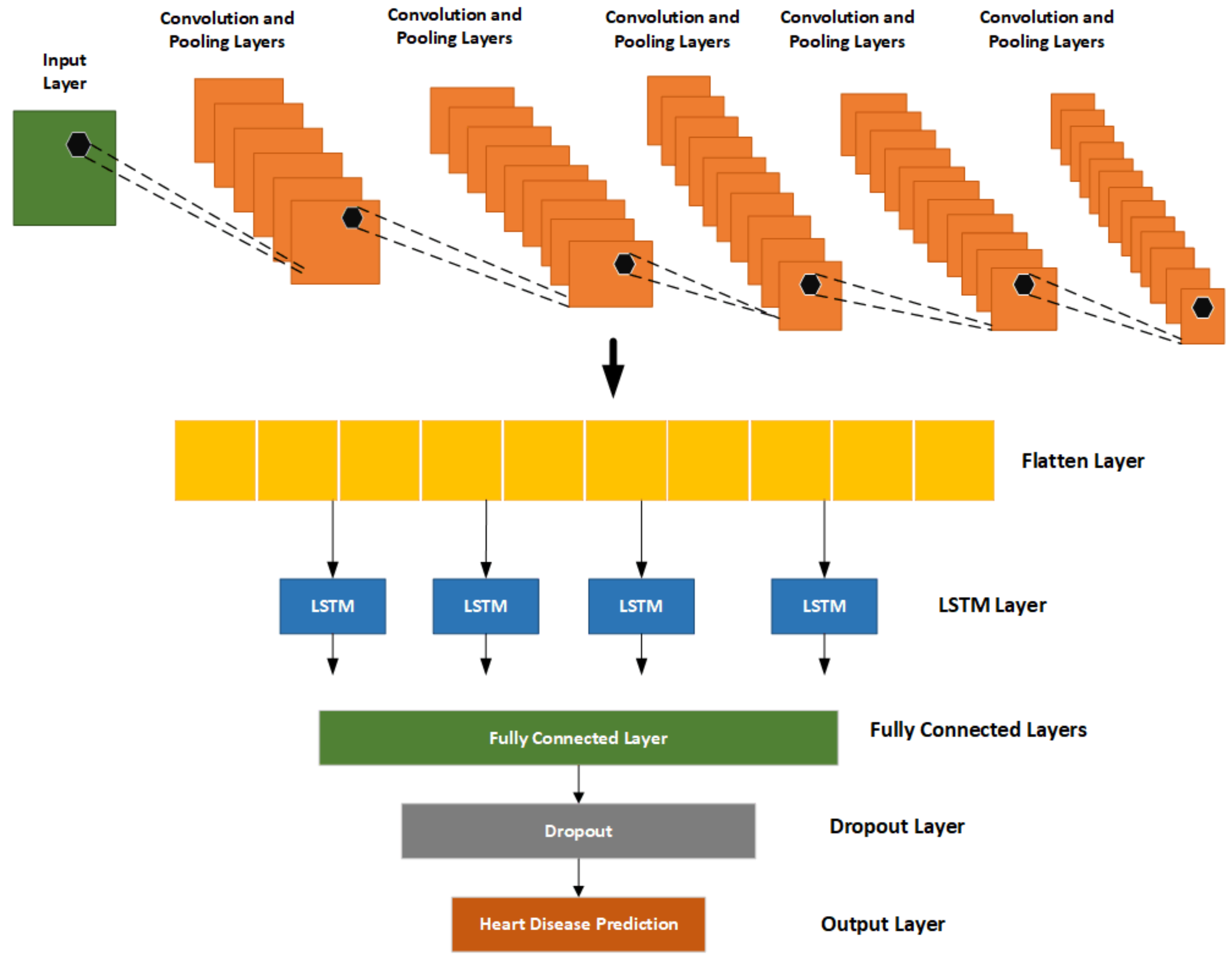

FIGURE 7. Hybrid structure of CNN-LSTM.

$$= 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{9}$$

Specificity

$$= \frac{\text{TN}}{\text{FN} + \text{TN}} \times 10 \tag{10}$$

MCC

$$= \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP) \times (TP+FN) \times (TN+FP) \times (TN+FN)}} \tag{11}$$

The experimental analysis of the results obtained with different combinations of preprocessing techniques and the performance of four ML methods (SVM, KNN, DT, and RF) and four DL models (ANN, LSTM, CNN, and CNN-LSTM) are presented in Table 3 and Table 4.

Table 3 reports the scores of key performance metrics concerning accuracy, specificity, precision, MCC, sensitivity, f-measure, and AUC, for different ML and DL methods using the Cleveland dataset. The last row of Table 3 highlights the best-performing model. However, determining the best model solely based on the highest value of one performance metric may not be sufficient. Therefore, other metrics like precision, sensitivity, MCC, specificity, f-measure, ROC, and AUC are also calculated to assess the efficiency of the HDNN. Furthermore, Table 4 presents the results of the proposed models when using a larger public HD dataset, which comprises data from (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA). Notably, the DL models, particularly the HDNN CNN-LSTM, demonstrated superior performance compared to other models in this comprehensive dataset. Following the results in Tables 3 and 4, it can be observed that the DL models, especially the HDNN CNN-LSTM, surpassed the other approaches in terms of various performance metrics. This underscores the efficiency and effectiveness of the suggested HDNN for HD diagnosis and prediction, making it a promising approach in the field of medical research and decision-making support.

Figure 8 illustrates the graphical comparison of the ML algorithms. The comparison indicates that the RF model surpassed other models like SVM, KNN, and DT in terms of precision, sensitivity, MCC, specificity, and f-measure when using the Cleveland dataset. Figure 9 indicates the

TABLE 3. Evaluation of the proposed DL models using a comprehensive HD dataset (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA).

Model	Accuracy	Precision	Sensitivity	MCC	Specificity	F-measure	AUC
SVM	89.07%	0.8947	0.9083	0.7790	0.8691	0.9015	0.8791
KNN	89.82%	0.8888	0.91603	0.7789	0.8598	0.9022	0.8885
DT	88.80%	0.8515	0.8787	0.7758	0.8700	0.8969	0.8603
RF	92.17%	0.9381	0.9191	0.8423	0.9250	0.9285	0.9120
ANN	93.21%	0.9408	0.9408	0.8629	0.9220	0.9408	0.9514
LSTM	94.65%	0.9465	0.9191	0.9003	0.9441	0.9533	0.9578
CNN	95.02%	0.9521	0.9546	0.9495	0.9571	0.9563	0.9686
CNN-LSTM	97.75%	0.9857	0.9887	0.9660	0.9787	0.9718	0.9885

TABLE 4. Evaluation of the proposed DL models using a comprehensive HD dataset (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA).

Model	Accuracy	Precision	Sensitivity	MCC	Specificity	F-measure	AUC
SVM	89.07%	0.8671	0.8224	0.7811	0.8224	0.9051	0.8844
KNN	88.65%	0.8714	0.9312	0.7712	0.8815	0.9003	0.8815
DT	86.24%	0.8950	0.8579	0.7223	0.8682	0.8761	0.8331
RF	89.93%	0.9112	0.9112	0.7949	0.8837	0.9112	0.8974
ANN	94.53%	0.9402	0.9618	0.8896	0.9252	0.9509	0.9685
LSTM	96.64%	0.9793	0.9595	0.9325	0.975	0.9693	0.9865
CNN	96.86%	0.9722	0.9722	0.9296	0.9574	0.9722	0.9878
CNN-LSTM	98.86%	0.9913	0.9874	0.9705	0.9942	0.9983	0.9978

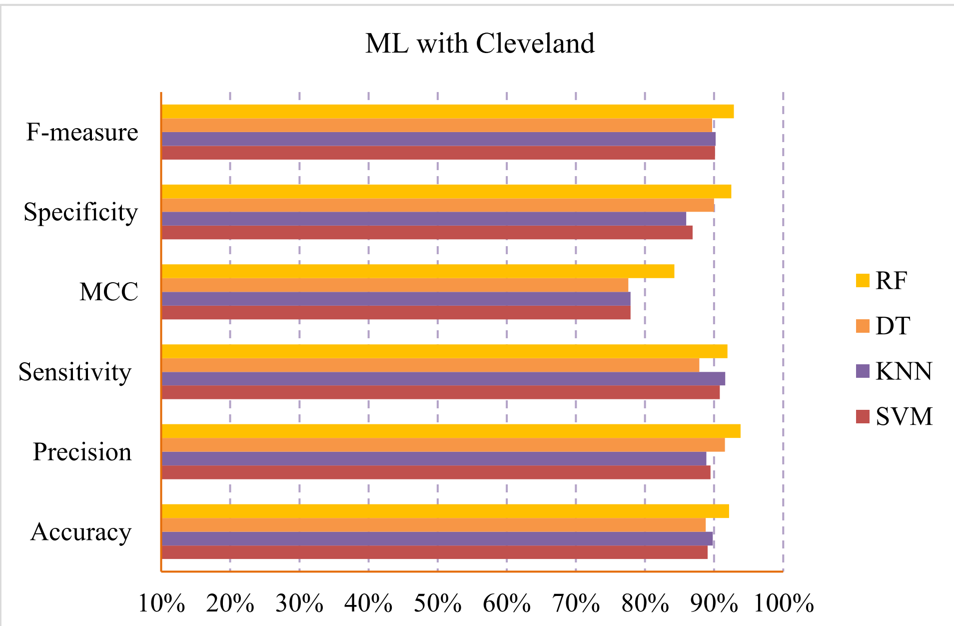

FIGURE 8. Graphical comparison of ML models using the Cleveland dataset.

graphical comparison of the proposed DL technologies using the Cleveland dataset. The comparison reveals that the CNN-LSTM model achieved superior performance in terms of precision, sensitivity, MCC, specificity, and f-measure when compared to other models like ANN, LSTM, and CNN.

For the more extensive dataset (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA), Figure 10 exhibits the graphical comparison of ML models. The RF model once again performed better compared to SVM, KNN, and DT models when considering precision, sensitivity, MCC, speci-

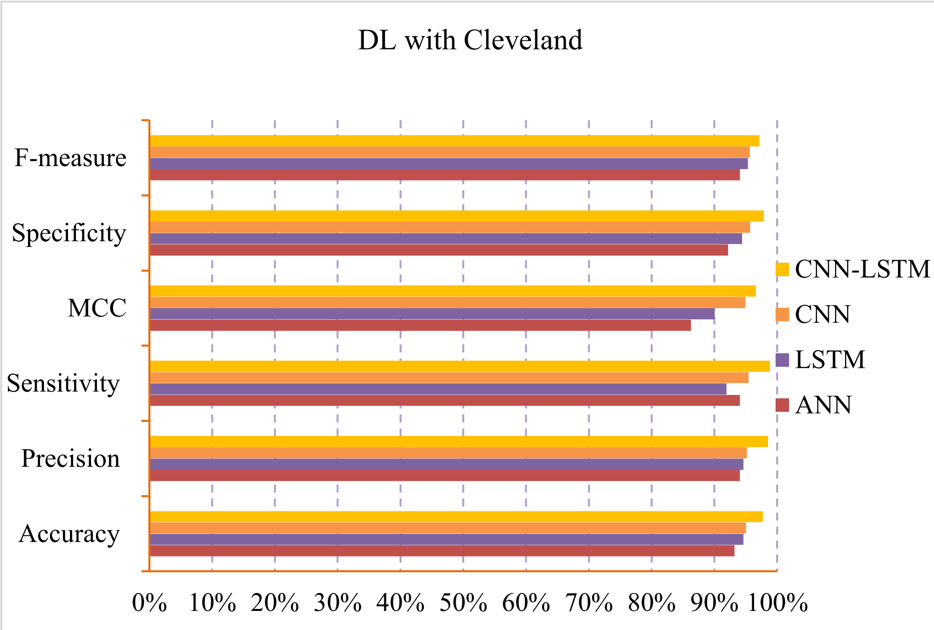

FIGURE 9. Graphical comparison of DL models using the Cleveland dataset.

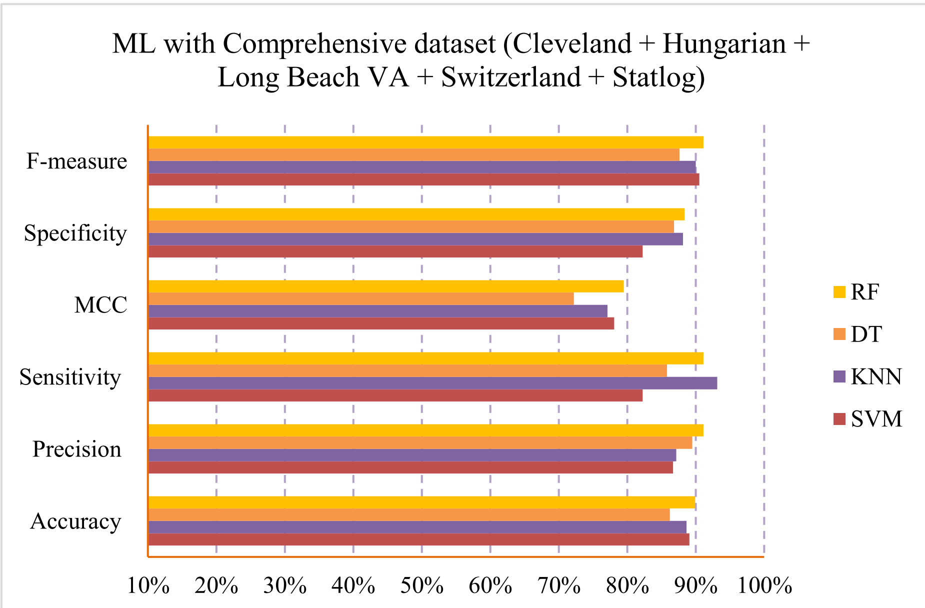

FIGURE 10. Graphical comparison of DL models using a comprehensive dataset. (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA.)

ficity, and f-measure. In Figure 11, the graphical comparison of the anticipated DL models is presented using the comprehensive dataset (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA). Similar to the Cleveland dataset, the CNN-LSTM model showed better performance in terms of precision, sensitivity, MCC, specificity, and f-measure when compared to other models like ANN, LSTM, and CNN.

The statistical analysis indicates that the results obtained from the proposed HDNNs are highly significant compared to other techniques. These findings highlight the effectiveness of the proposed ML and DL models in predicting HD, making them promising approaches for medical decision support systems.

The results proved that the suggested HDNN system outperformed the state-of-the-art systems. When compared to

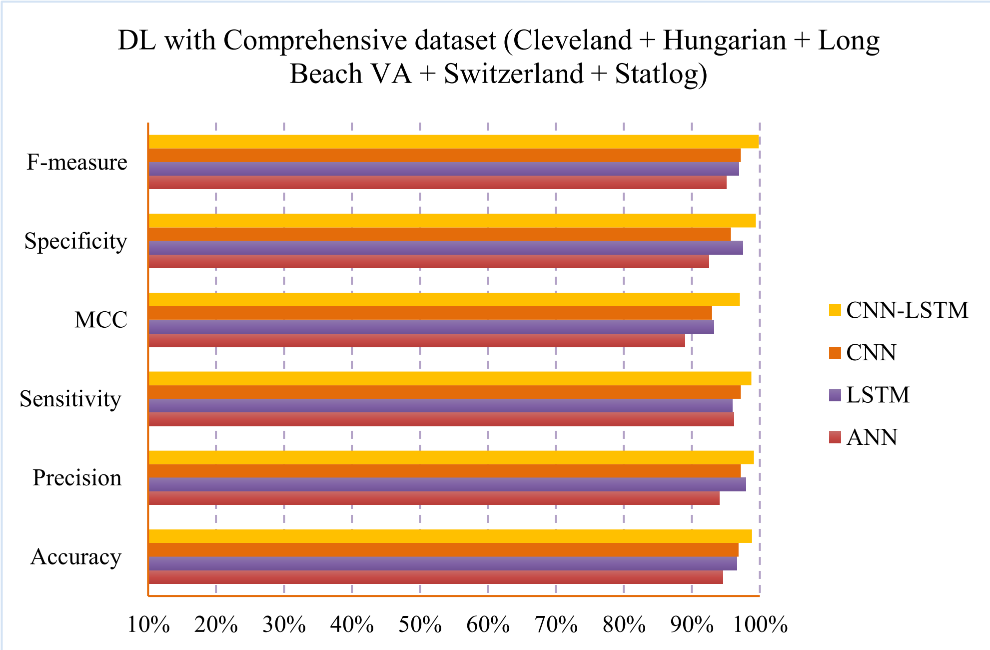

FIGURE 11. Graphical comparison of DL models using a comprehensive dataset. (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA.)

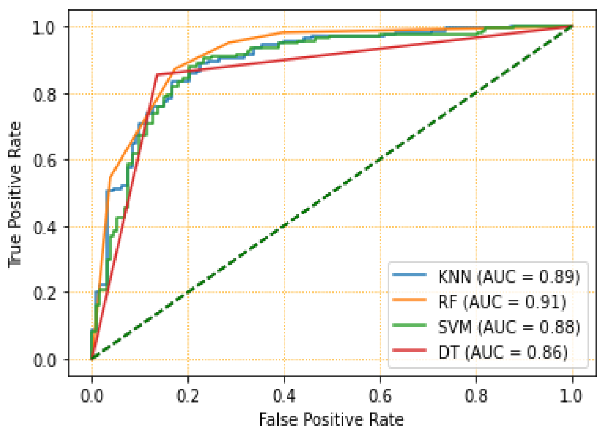

FIGURE 12. ROC/AUC score of ML models using Cleveland.

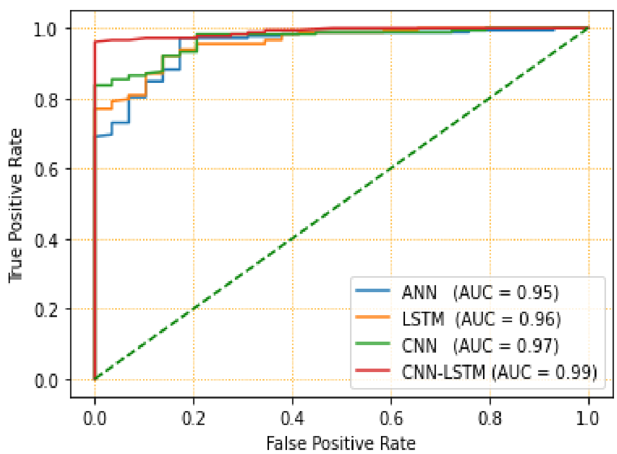

FIGURE 13. ROC/AUC score of DL models using Cleveland.

traditional ML approaches and existing state-of-the-art systems, the proposed CNN-LSTM model achieved significantly higher accuracy, precision, sensitivity, MCC, specificity, f-measure, and AUC values of 97.75%, 0.9857, 0.9887, 0.9660, 0.9787, 0.9718, and 0.9885, respectively, when using the Cleveland HD dataset.

Additionally, when the comprehensive HD dataset, including data from (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA), was used, the CNN-LSTM model continued to demonstrate superior performance. It achieved even higher accuracy, precision, sensitivity, MCC, specificity, f-measure, and AUC values of 98.86%, 0.9913, 0.9874, 0.9705, 0.9942, 0.9983, and 0.9978, respectively. This suggests that the CNN-LSTM model is highly effective in predicting HD and surpasses existing methods when using a more diverse and extensive dataset.

Figure 12 displays the AUC scores for ML models like SVM, KNN, DT, and RF, which are 0.8791, 0.8885, 0.8603, and 0.9120, respectively. These ML models also exhibit good performance, but the CNN-LSTM model still outperforms them with a higher AUC score. On the other hand, in Figure 13, the results demonstrate that the CNN-LSTM model outperforms other DL models, including ANN (95.00%), LSTM (96.00%), and CNN (97.00%), achieving an impressive AUC score of 99.00%. The CNN-LSTM strategy

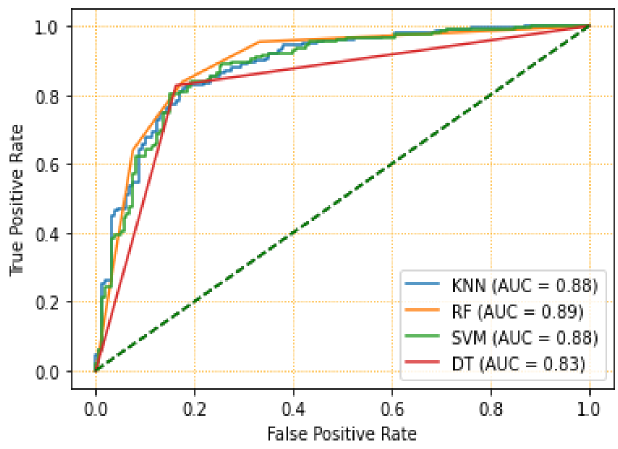

FIGURE 14. **ROC/AUC score of ML models using a comprehensive dataset. (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA.)**

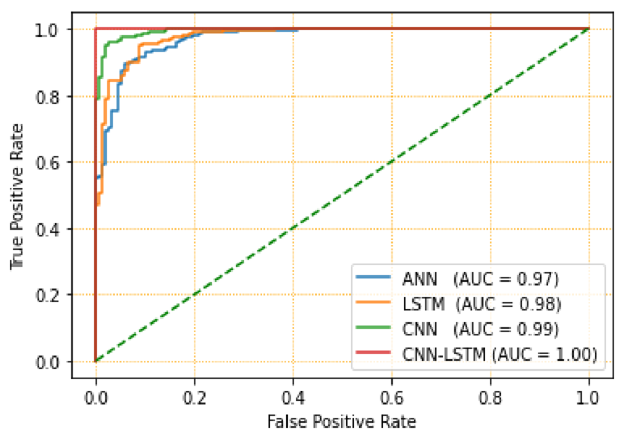

FIGURE 15. **ROC/AUC score of DL models using a comprehensive dataset. (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA.)**

shows its superiority in distinguishing between true positive and false positive instances in the dataset.

Furthermore, in Figure 14 and Figure 15, the ROC curves illustrate the performance of the proposed ML and DL models when using a comprehensive HD dataset that includes data from (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA). In this case, the hybrid CNN-LSTM model achieves the highest AUC score of 1.00%, surpassing the AUC scores of ANN, LSTM, and CNN, which are 97.00%, 98.00%, and 99.00%, respectively. Figure 14 also presents the ROC scores for ML models such as SVM, KNN, DT, and RF, which are 0.8844, 0.8815, 0.8331, and 0.8974, respectively. These findings highlight the exceptional accuracy of hybrid CNN-LSTM, ANN, CNN, and LSTM in both datasets, suggesting their potential as useful methods for HD diagnosis and prediction.

A comparison of the suggested ML and DL models for the two datasets is shown in Figure 16. The outcomes show that the ANN, CNN, LSTM, and hybrid CNN-LSTM performed better in both datasets, indicating that these models can be beneficial for diagnosing and predicting HD. For both datasets, the CNN-LSTM system performed exceptionally well, consistently achieving accuracy values exceeding 95%. This indicates the HDNN system's excellence in serving as a medical decision support system. The high accuracy can be attributed to the hidden layers of the HDNN system, which help in reducing the error rate and improving the system's overall predictive capabilities. These findings indicate that the proposed hybrid CNN-LSTM model exhibits superior performance in both datasets, making it a highly effective choice for the task at hand. The results from Figure 16 illustrate strong evidence of the capabilities of the proposed ML and DL methods in enhancing HD/CVD prediction accuracy. These results emphasize the significance of leveraging advanced HDNNs computational methods, such as CNN-LSTM, in medical research to improve patient care and support clinical decision-making in the field of HD/CVD.

Additionally, the performance of the proposed system was compared with existing studies. The proposed system's state-of-the-art comparison was conducted with other existing works, and the results are listed in Table 5. The outcomes demonstrated that the hybrid CNN-LSTM achieved a maximum accuracy of 98.86% and 97.75%. These simulated results proved that the CNN-LSTM system outperformed the existing techniques, showcasing superior performance in comparison. While most existing works are applied to a single HD dataset, the proposed system was evaluated on two different HD datasets, each having different ranges and types of values and features. The outcomes revealed that the suggested HDNN such as CNN-LSTM consistently demonstrated high performance across all the datasets. This suggests that the system is not biased towards any specific dataset or range of values and can be employed effectively for various datasets for HD prediction at an early stage.

To the best of our knowledge, the accuracy attained by the CNN-LSTM architecture is the highest among all the state-of-the-art systems reported in the literature for both datasets. This suggests that the anticipated CNN-LSTM mechanism surpasses the performance of other existing algorithms, making it a significant advancement in the field of HD prediction.

A. LIMITATIONS

The research work has some limitations, such as its poor compatibility with several feature selection techniques, its susceptibility to datasets that contain a large amount of missing data, and its lack of comprehensive testing on real-world datasets. Moreover, the lack of deep ensemble learning methods in the present model highlights an important path toward improving the research's predictive power.

B. FUTURE RESEARCH DIRECTIONS

There are various methods to improve this study and deal with its shortcomings. In the future, we want to further generalize the system to make it compatible with different feature selection algorithms and more resistant to other datasets with

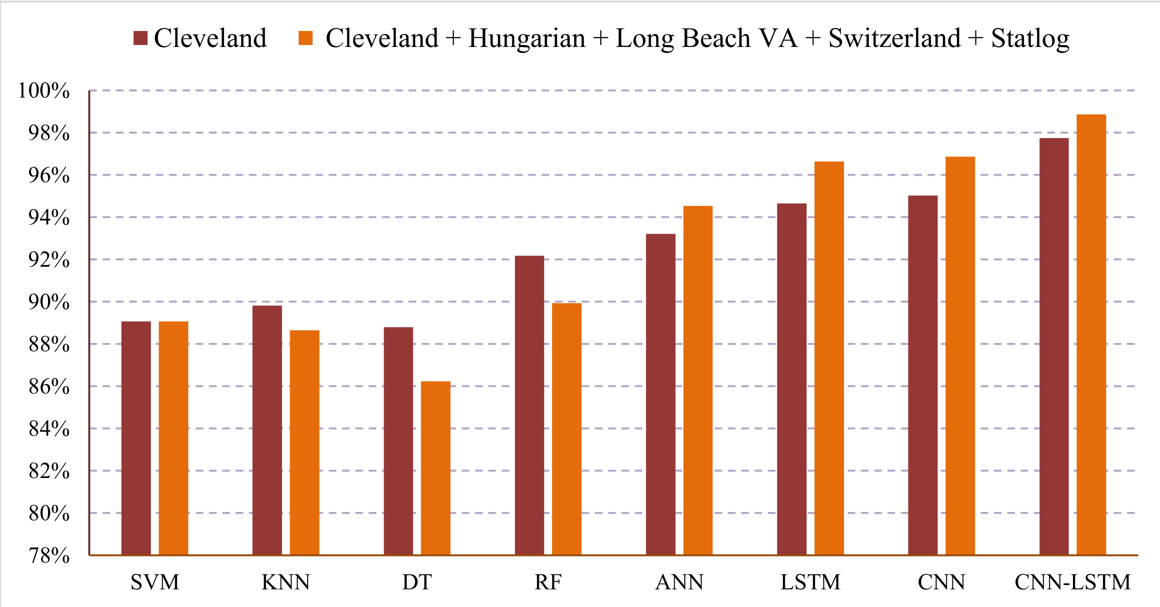

FIGURE 16. Comparison of the proposed ML and DL models for both datasets.

TABLE 5. Comparison of the proposed DL models' performance against the existing work.

Work	Heart Disease Dataset	Model	Accuracy
[25]	Cleveland (14, 303)	LR, KNN, SVM, RF, DT, DL	83.3%, 84.8%, 83.2%, 80.3%, 82.3%, 94.2%
[26]	Heart Disease UCI dataset	Hard voting ensemble	90.00%
[27]	Cleveland (14, 303)	NB	84.51%
[30]	Cleveland (14, 303)	KNN	85.00%
[31]	Cleveland (14, 303)	RF+DT	88.70%
[33]	Hungarian (294, 12)	ANN	81.02%
[33]	Cleveland (14, 303)	ANN	82.49%
[33]	Switzerland (13, 123)	ANN	60.06%
[33]	Long Beach VA (13, 200)	ANN	60.00%
[34]	Heart Disease UCI dataset	RF with a linear model	88.70%
[37]	Heart Disease UCI dataset	LOFS-ANN	90.5%
[53]	Hungarian + Cleveland + Long Beach VA + Switzerland + Statlog (11, 1190)	Stacked ensemble classifier with Extra Trees Classifier, RF, XGB	92.34%
[54]	Heart Disease UCI dataset	RF	83%
Proposed	Cleveland (14, 303)	Deep ANN, LSTM, CNN, CNN+LSTM	93.21%, 94.65%, 95.02%, 97.75%
Proposed	Combined [(Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA) (11, 1190)]	Deep ANN, LSTM, CNN, CNN+LSTM	94.03%, 96.24%, 96.86%, 98.86%

significant amounts of missing data. By carrying out the same experiment on an extensive real-world dataset, this study can be developed further. To evaluate various combinations of DL approaches for HD/CVD prediction, additional investigation can be performed. New feature selection techniques can also be used to gain a wider perspective on the important features and boost prediction accuracy. Another approach is to employ deep ensemble learning techniques to strengthen and improve the model.

VI. CONCLUSION

The heart is crucial for all organisms. The prediction of HD requires greater accuracy, validity, and reliability because even a minor error can result in exhaustion or even death. HD-related deaths are common, and the number of these deaths is rising rapidly each year. Early diagnosis plays a crucial role in achieving the goal of predicting HD promptly. In the pursuit of diagnosing and predicting HD, existing studies have employed ML techniques. This research follows a similar path but introduces an enhanced and innovative HDNN system, utilizing a larger and a small dataset to train the model effectively. The research showcases the effectiveness of the extra tree classifier method, which aids in identifying relevant features for HD prediction. Furthermore, the proposed HDNN system employs deep ANN, LSTM, CNN, and hybrid CNN with LSTM with multiple layers to achieve high diagnosing and prediction accuracy. The approach also incorporates data imputation techniques to refine the data and ensure its quality. The proposed system was trained on two HD datasets such as Cleveland and a comprehensive HD (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA) dataset. Finally, the discussed system was measured through comparison with conventional systems concerning sensitivity, MCC, F1-score, accuracy, precision, and specificity. The outcomes showed a superior prediction of the proposed HDNN system compared to the state-of-the-art. Compared to the traditional ML and existing state-of-the-art, the proposed CNN-LSTM achieved the best accuracy, precision, sensitivity, MCC, specificity, f-measure, and AUC of 97.75%, 0.9857, 0.9887, 0.9660, 0.9787, 0.9718, and 0.9885, respectively, when using Cleveland HD dataset. Furthermore, using a comprehensive HD (Switzerland + Cleveland + Statlog + Hungarian + Long Beach VA), the CNN-LSTM surpass accuracy, precision, sensitivity, MCC, specificity, f-measure, and AUC of 98.86%, 0.9913, 0.9874, 0.9705, 0.9942, 0.9983, and 0.9978, respectively.

REFERENCES

[1] N. C. Long, P. Meesad, and H. Unger, ''A highly accurate firefly based algorithm for heart disease prediction,'' *Expert Syst. Appl.*, vol. 42, no. 21, pp. 8221–8231, Nov. 2015, doi: 10.1016/j.eswa.2015.06.024.

[2] *UDMI*. Accessed: Jul. 24, 2023. [Online]. Available: https://www.udmi.net/cardiovascular-disease-risk/

[3] P. K. Mehta, S. Gaignard, A. Schwartz, and J. E. Manson, ''Traditional and emerging sex-specific risk factors for cardiovascular disease in women,'' *Rev. Cardiovascular Med.*, vol. 23, no. 8, p. 288, 2022.

[4] J. M. Alves-silva, M. Zuzarte, and H. Gir, ''The role of essential oils and their main compounds in the management of cardiovascular disease risk factors,'' *Molecules*, vol. 26, no. 3506, pp. 1–26, 2021.

[5] C. Zhou, A. Hou, P. Dai, A. Li, Z. Zhang, Y. Mu, and L. Liu, ''Risk factor refinement and ensemble deep learning methods on prediction of heart failure using real healthcare records,'' *Inf. Sci.*, vol. 637, Aug. 2023, Art. no. 118932, doi: 10.1016/j.ins.2023.04.011.

[6] F. Ali, S. El-Sappagh, S. M. R. Islam, D. Kwak, A. Ali, M. Imran, and K.-S. Kwak, ''A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion,'' *Inf. Fusion*, vol. 63, pp. 208–222, Nov. 2020, doi: 10.1016/j.inffus.2020.06.008.

[7] S. P. Barfungpa, H. K. D. Sarma, and L. Samantaray, ''An intelligent heart disease prediction system using hybrid deep dense Aquila network,'' *Biomed. Signal Process. Control*, vol. 84, Jul. 2023, Art. no. 104742, doi: 10.1016/j.bspc.2023.104742.

[8] S. Amin, B. Alouffi, M. I. Uddin, and W. Alosaimi, ''Optimizing convolutional neural networks with transfer learning for making classification report in COVID-19 chest X-rays scans,'' *Sci. Program.*, vol. 2022, pp. 1–13, May 2022, doi: 10.1155/2022/5145614.

[9] S.-C. Huang, A. Pareek, S. Seyyedi, I. Banerjee, and M. P. Lungren, ''Fusion of medical imaging and electronic health records using deep learning: A systematic review and implementation guidelines,'' *NPJ Digit. Med.*, vol. 3, no. 1, p. 136, Oct. 2020.

[10] B. He, L. Bergenstråhle, L. Stenbeck, A. Abid, A. Andersson, Å. Borg, J. Maaskola, J. Lundeberg, and J. Zou, ''Integrating spatial gene expression and breast tumour morphology via deep learning,'' *Nature Biomed. Eng.*, vol. 4, no. 8, pp. 827–834, Jun. 2020.

[11] S. Amin, A. Alharbi, M. I. Uddin, and H. Alyami, ''Adapting recurrent neural networks for classifying public discourse on COVID-19 symptoms in Twitter content,'' *Soft Comput.*, vol. 26, no. 20, pp. 11077–11089, Oct. 2022, doi: 10.1007/s00500-022-07405-0.

[12] T. Si, J. Bagchi, and P. B. C. Miranda, ''Artificial neural network training using metaheuristics for medical data classification: An experimental study,'' *Expert Syst. Appl.*, vol. 193, May 2022, Art. no. 116423, doi: 10.1016/j.eswa.2021.116423.

[13] A. Najafi, A. Nemati, M. Ashrafzadeh, and S. H. Zolfani, ''Multiple-criteria decision making, feature selection, and deep learning: A golden triangle for heart disease identification,'' *Eng. Appl. Artif. Intell.*, vol. 125, Oct. 2023, Art. no. 106662, doi: 10.1016/j.engappai.2023.106662.

[14] S. P. Barfungpa, L. Samantaray, H. Kumar Deva Sarma, R. Panda, and A. Abraham, ''D-t-SNE: Predicting heart disease based on hyper parameter tuned MLP,'' *Biomed. Signal Process. Control*, vol. 86, Sep. 2023, Art. no. 105129, doi: 10.1016/j.bspc.2023.105129.

[15] N. V. MahaLakshmi and R. K. Rout, ''Effective heart disease prediction using improved particle swarm optimization algorithm and ensemble classification technique,'' *Soft Comput.*, vol. 27, no. 15, pp. 11027–11040, Aug. 2023, doi: 10.1007/s00500-023-08388-2.

[16] M. Swathy and K. Saruladha, ''A comparative study of classification and prediction of cardio-vascular diseases (CVD) using machine learning and deep learning techniques,'' *ICT Exp.*, vol. 8, no. 1, pp. 109–116, Mar. 2022, doi: 10.1016/j.icte.2021.08.021.

[17] B. R. Reddy and R. L. Kumar, ''Classification of health care products using hybrid CNN-LSTM model,'' *Soft Comput.*, vol. 27, no. 13, pp. 9199–9216, Jul. 2023, doi: 10.1007/s00500-023-08279-6.

[18] A. Rayan, S. Holyl Alruwaili, A. S. Alaerjan, S. Alanazi, A. I. Taloba, O. R. Shahin, and M. Salem, ''Utilizing CNN-LSTM techniques for the enhancement of medical systems,'' *Alexandria Eng. J.*, vol. 72, pp. 323–338, Jun. 2023, doi: 10.1016/j.aej.2023.04.009.

[19] UCI Machine Learning Repository. *Heart Disease Cleveland Daatset*. Accessed: Jul. 20, 2023. [Online]. Available: https://archive.ics.uci.edu/dataset/45/heart+disease

[20] M. Siddhartha. (2020). *Heart Disease Dataset (Comprehensive)*. IEEE Dataport. Accessed: Jul. 20, 2023. [Online]. Available: https://ieee-dataport.org/open-access/heart-disease-dataset-comprehensive

[21] M. Tarawneh and O. Embarak, ''Hybrid approach for heart disease prediction using data mining techniques,'' in *Proc. 7th Int. Conf. Emerg. Internet, Data Web Technol. (EIDWT)*. UAE: Springer, 2019, pp. 447-454.

[22] S. Nalluri, R. V. Saraswathi, S. Ramasubbareddy, K. Govinda, and E. Swetha, ''Chronic heart disease prediction using data mining techniques,'' in *Proc. 3rd Data Engineering and Communication Technology (ICDECT)*. Singapore: Springer, 2020, pp. 903–912.

[23] C. B. Gokulnath and S. P. Shantharajah, ''An optimized feature selection based on genetic approach and support vector machine for heart disease,'' *Cluster Comput.*, vol. 22, no. S6, pp. 14777–14787, Nov. 2019, doi: 10.1007/s10586-018-2416-4.

[24] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, ''Heart disease identification method using machine learning classification in e-healthcare,'' *IEEE Access*, vol. 8, pp. 107562–107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[25] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Comput. Intell. Neurosci.*, vol. 2021, Jan. 2021, Art. no. 8387680, doi: 10.1155/2021/8387680.

[26] R. Atallah and A. Al-Mousa, "Heart disease detection using machine learning majority voting ensemble method," in *Proc. 2nd Int. Conf. New Trends Comput. Sci. (ICTCS)*, Oct. 2019, pp. 1–6, doi: 10.1109/ICTCS.2019.8923053.

[27] I. Tougui, A. Jilbab, and J. El Mhamdi, "Heart disease classification using data mining tools and machine learning techniques," *Health Technol.*, vol. 10, no. 5, pp. 1137–1144, Sep. 2020, doi: 10.1007/s12553-020-00438-1.

[28] T. Mahmud, A. Barua, M. Begum, E. Chakma, S. Das, and N. Sharmen, "An improved framework for reliable cardiovascular disease prediction using hybrid ensemble learning," in *Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE)*, Feb. 2023, pp. 1–6, doi: 10.1109/ECCE57851.2023.10101564.

[29] N. M. Lutimath, C. Mouli, B. K. B. Gowda, and K. Sunitha, "Prediction of heart disease using hybrid machine learning technique," in *Paradigms of Smart and Intelligent Communication, 5G and Beyond*, A. Rai, D. K. Singh, A. Sehgal, and K. Cengiz, Eds. Singapore: Springer, 2023, pp. 277–293.

[30] A. P. Pawlovsky, "An ensemble based on distances for a kNN method for heart disease diagnosis," in *Proc. Int. Conf. Electron., Inf., Commun. (ICEIC)*, Jan. 2018, pp. 1–4, doi: 10.23919/ELINFOCOM.2018.8330570.

[31] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, and R. S. Suraj, "Heart disease prediction using hybrid machine learning model," in *Proc. 6th Int. Conf. Inventive Comput. Technol. (ICICT)*, Jan. 2021, pp. 1329–1333, doi: 10.1109/ICICT50816.2021.9358597.

[32] D. Shah, S. Patel, and S. K. Bharti, "Heart disease prediction using machine learning techniques," *Social Netw. Comput. Sci.*, vol. 1, no. 6, p. 345, Nov. 2020, doi: 10.1007/s42979-020-00365-y.

[33] A. A. Almazroi, E. A. Aldhahri, S. Bashir, and S. Ashfaq, "A clinical decision support system for heart disease prediction using deep learning," *IEEE Access*, vol. 11, pp. 61646–61659, 2023.

[34] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019.

[35] M. N. R. Chowdhury, E. Ahmed, M. A. D. Siddik, and A. U. Zaman, "Heart disease prognosis using machine learning classification techniques," in *Proc. 6th Int. Conf. Converg. Technol. (I2CT)*, Apr. 2021, pp. 1–6, doi: 10.1109/I2CT51068.2021.9418181.

[36] P. Rani, R. Kumar, N. M. O. S. Ahmed, and A. Jain, "A decision support system for heart disease prediction based upon machine learning," *J. Reliable Intell. Environments*, vol. 7, no. 3, pp. 263–275, Sep. 2021, doi: 10.1007/s40860-021-00133-6.

[37] S. Goyal, "Predicting the heart disease using machine learning techniques," in *Proc. ICT Anal. Appl. (ICT4SD)*. Singapore: Springer, 2023, pp. 191–199.

[38] F. Tasnim and S. U. Habiba, "A comparative study on heart disease prediction using data mining techniques and feature selection," in *Proc. 2nd Int. Conf. Robot., Electr. Signal Process. Techn. (ICREST)*, Jan. 2021, pp. 338–341, doi: 10.1109/ICREST51555.2021.9331158.

[39] M. A. Kadhim and A. M. Radhi, "Heart disease classification using optimized machine learning algorithms," *Iraqi J. Comput. Sci. Math.*, vol. 2023, pp. 31–42, Feb. 2023.

[40] H. E. Hamdaoui, S. Boujraf, N. E. H. Chaoui, and M. Maaroufi, "A clinical support system for prediction of heart disease using machine learning techniques," in *Proc. 5th Int. Conf. Adv. Technol. Signal Image Process. (ATSIP)*, Sep. 2020, pp. 1–5, doi: 10.1109/ATSIP49331.2020.9231760.

[41] M. S. Amin, Y. K. Chiam, and K. D. Varathan, "Identification of significant features and data mining techniques in predicting heart disease," *Telematics Informat.*, vol. 36, pp. 82–93, Mar. 2019, doi: 10.1016/j.tele.2018.11.007.

[42] P. C. Bizimana, Z. Zhang, M. Asim, and A. A. A. El-Latif, "An effective machine learning-based model for an early heart disease prediction," *Biomed Res. Int.*, vol. 2023, Jan. 2023, Art. no. 3531420, doi: 10.1155/2023/3531420.

[43] A. Saboor, M. Usman, S. Ali, A. Samad, M. F. Abrar, and N. Ullah, "A method for improving prediction of human heart disease using machine learning algorithms," *Mobile Inf. Syst.*, vol. 2022, Jan. 2022, Art. no. 1410169, doi: 10.1155/2022/1410169.

[44] *Anaconda*. Accessed: Aug. 4, 2023. [Online]. Available: https://anaconda.org/

[45] *Scikit-Learn*. Accessed: Aug. 4, 2023. [Online]. Available: https://scikit-learn.org/stable/

[46] *NumPy*. Accessed: Aug. 4, 2023. [Online]. Available: https://numpy.org/

[47] *TensorFlow*. Accessed: Aug. 4, 2023. [Online]. Available: https://www.tensorflow.org/tutorials/keras/keras_tuner

[48] M. Siddhartha. *Heart Disease Dataset (Comprehensive)*. Kaggle. Accessed: Aug. 4, 2023. [Online]. Available: https://www.kaggle.com/datasets/sid321axn/heart-statlog-cleveland-hungary-final

[49] S. Amin, M. I. Uddin, M. A. Zeb, A. A. Alarood, M. Mahmoud, and M. H. Alkinani, "Detecting information on the spread of dengue on Twitter using artificial neural networks," *Comput., Mater. Continua*, vol. 67, no. 1, pp. 1317–1332, 2021, doi: 10.32604/cmc.2021.014733.

[50] S. Amin, M. I. Uddin, S. Hassan, A. Khan, N. Nasser, A. Alharbi, and H. Alyami, "Recurrent neural networks with TF-IDF embedding technique for detection and classification in tweets of dengue disease," *IEEE Access*, vol. 8, pp. 131522–131533, 2020, doi: 10.1109/ACCESS.2020.3009058.

[51] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[52] P. K. Shrivastava, M. Sharma, P. Sharma, and A. Kumar, "HCBiLSTM: A hybrid model for predicting heart disease using CNN and BiLSTM algorithms," *Measurement, Sensors*, vol. 25, Feb. 2023, Art. no. 100657, doi: 10.1016/j.measen.2022.100657.

[53] A. Tiwari, A. Chugh, and A. Sharma, "Ensemble framework for cardiovascular disease prediction," *Comput. Biol. Med.*, vol. 146, Jul. 2022, Art. no. 105624, doi: 10.1016/j.compbiomed.2022.105624.

[54] V. Chang, V. R. Bhavani, A. Q. Xu, and M. Hossain, "An artificial intelligence model for heart disease detection using machine learning algorithms," *Healthcare Anal.*, vol. 2, Nov. 2022, Art. no. 100016, doi: 10.1016/j.health.2022.100016.

MANA SALEH AL RESHAN received the B.S. degree in information systems from King Khalid University, Abha, Saudi Arabia, in 2007, the M.S. degree (Hons.) in computer, information, and network security from DePaul University, Chicago, IL, USA, in 2011, and the Ph.D. degree in computer science from The Catholic University of America (CUA), Washington, DC, USA, in 2019. He was a Teaching Assistant with the College of Computer Science and Information Systems, Najran University, Saudi Arabia, from 2007 to 2009. He was a Lecturer with the College of Computer Science and Information Systems, Najran University, in 2012, where he is currently an Assistant Professor and the Head of the Network Engineering Department. His current research interests include computer network and security, system security, wireless and mobile security, body area networks, and cloud security.

SAMINA AMIN received the M.S. degree in computer science from the Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan, in 2020, where she is currently pursuing the Ph.D. degree. She has published several articles in reputed journals. Her research interests include machine learning, deep learning, data science, data mining, natural language processing, social media analysis, neural networks, recurrent neural networks, convolutional neural networks, image processing, sentiment analysis, health informatics, recommender systems, and reinforcement learning. She has received the Gold Medal from the Kohat University of Science and Technology. She serves as a reviewer for different journals.

MUHAMMAD ALI ZEB is currently pursuing the Ph.D. degree with the Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan. He is a Lecturer with the Institute of Computing, Kohat University of Science and Technology. He has published several papers in reputed journals and conferences. His research interests include machine learning, deep learning, social media analysis, natural language processing, sentiment analysis, data science, visualization, topic modeling, neural networks, image processing, and temporal graphs neural networks.

ADEL SULAIMAN received the master's degree in software engineering from the University of St. Thomas, USA, and the Ph.D. degree in computer science from Iowa State University, USA. He is currently an Assistant Professor with the College of Computer Science and Information Systems, Najran University. He is also the Vice Dean for quality and development in the deanship of e-learning and leads the Research Unit, College of Computer Science, Najran University. His research interests include human–computer interaction, large-scale simulations, mixed reality, scientific visualization, and data science.

HANI ALSHAHRANI received the bachelor's degree in computer science from King Khalid University, Abha, Saudi Arabia, the master's degree in computer science from California Lutheran University, Thousand Oaks, CA, USA, and the Ph.D. degree from Oakland University, Rochester, MI, USA. He is currently an Associate Professor of computer science and information systems with Najran University, Najran, Saudi Arabia. His current research interests include smartphones, the IoT, crowdsourcing security, and privacy.

ASADULLAH SHAIKH (Senior Member, IEEE) received the B.Sc. degree in software development from the University of Huddersfield, U.K., the M.Sc. degree in software engineering and management from the University of Gothenborg, Sweden, and the Ph.D. degree in software engineering from the University of Southern Denmark. He was a Researcher with UOC, Barcelona, Spain. He is currently a Professor, the Head of research and graduate studies, and the Coordinator of seminars and training with the College of Computer Science and Information Systems, Najran University, Najran, Saudi Arabia. He has vast experience in teaching and research. He has more than 170 publications in the area of software engineering in international journals and conferences. His current research interests include UML model verification, UML class diagrams verification with OCL constraints for complex models, formal verification, and feedback technique for unsatisfiable UML/OCL class diagrams. He is also an Editor of the *International Journal of Advanced Computer Systems and Software Engineering* (IJACSSE) and an International Advisory Board of several conferences and journals. Further details can be obtained using www.asadshaikh.com.