SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I6P102

Volume 13 Issue 6, 12-27, June 2026
© 2026 Seventh Sense Research Group®

Original Article

Hybrid CNN-LSTM Model for Enhanced Cardiac Disease Risk Prediction

Pachipala Yellamma[1], Thadiboina Sai Teja[2]

[1]*Department of CSE, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, India.*

[1]*Corresponding Author : pachipala.yamuna@gmail.com*

Received: 05 March 2026 Revised: 04 April 2026 Accepted: 03 May 2026 Published: 27 June 2026

Abstract *- Cardiac Disease continues to be a foremost universal health issue every year, highlighting the need for early and accurate predictions, which are essential for preventing critical outcomes. Traditional clinical prediction models focus largely on structured clinical features such as Age, cholesterol, and blood pressure. This provides valuable insights for identifying patients at risk. However, structured characteristics largely fail to express the Unpredictable characteristics of their data. Existing approaches, such as predictive analytics models and other deep-learning architectures (like CNN +LSTM), have more limited capabilities because they either learn temporal patterns (LSTM) but lack strong feature extraction or extraction of spatial features, or do not model sequential relationships (CNN). A combined CNN+LSTM model for cardiac risk prediction on a clinical dataset is presented in this work. A well-established clinical dataset, such as the publicly available Clinical dataset, with 14 clinically validated features derived from real records, will be used as the model training point. Preprocessing will be performed on the dataset for the purpose of improving performance. Within this proposed architecture, the CNN layers will extract high-level feature patterns from clinical inputs, and the LSTM layers will learn structural dependencies that provide a relatively complete indicator of patient health to allow better decision-making and easier patient stratification. The experimental results indicated that the hybrid CNN+LSTM outperformed in prediction with 98.05% accuracy, 96.29% precision, and 95.56% recall, thereby supporting the conclusion that the integration of CNN and LSTM improves feature learning and better sequence learning improves Cardiac disease prediction.*

Keywords *- Clinical Dataset, Convolutional Neural Networks (CNN), Feature Extraction, Heart Disease Prediction, Long Short-Term Memory.*

1. Introduction

Heart disease is a major cause of global deaths, claiming millions of lives every year. The increased burden of cardiovascular disease has created increased demand for reliable and accurate prediction methods, which can allow timely diagnosis and treatment plans. Clinical decision making is based on structured clinical variables such as Age, cholesterol level, resting blood pressure, type of chest pain, and resting ECG results. Clinical variables are clinically meaningful, but predicting heart disease remains challenging because clinical data is highly nonlinear, heterogeneous, and includes hidden interactions not captured by standard modelling. Although structured clinical datasets are readily available, current prediction models, whether classical machine learning models or isolated deep learning models, struggle to capture the deep feature relationships and sequential dependences inherent in medical datasets [1]. Classical learning models assume linear feature behaviour, so they have difficulty in identifying complex clinical patterns. In terms of representation, deep neural network models such as CNN and LSTM are superior to classical machine learning models, but they suffer many of the same limitations when considered in isolation. CNNs can extract high-level spatial features but are unable to pick up the temporal or sequential dependencies among clinical attributes. LSTMs can capture sequential dependencies, but do not demonstrate an adequate feature extraction capacity for numerical clinical variables [2]. These limitations lead to limited accuracy, poor generalization, and unreliable risk prediction regarding morbidity or mortality.

Regardless, an effective heart disease prediction in the proposed model that can integrate both feature extraction and sequential learning is essential to improving accuracy in cardiac risk prediction. To meet this challenge, the current work represents a combined CNN+LSTM architecture designed for cardiac risk prediction utilizing a clinical dataset. The combined architecture leverages the merits of CNN and LSTM. CNN extracts complex nonlinear clinical patterns; LSTM learns sequential dependencies among medical attributes. The capabilities of this hybrid architecture provide a deeper and richer representation of patient health statuses.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Cardiovascular disorder is one of the leading causes of death across the globe and is therefore one of the major public health issues in both urban and rural areas. People who live in urban areas are more affected by cardiovascular disorders due to the effects of their lifestyles, which include poor diet, high levels of stress, and a lack of physical activity. Conversely, people living in rural areas encounter obstacles when accessing healthcare facilities and experience a lack of systems in place to detect health problems early; as a result, many instances of cardiovascular disease go undiagnosed until they are in the late stages. Therefore, early detection of cardiovascular disease is important for decreasing mortality rates and lowering the probability of someone experiencing serious health complications.

The main aim of the proposed work is to develop an efficient and accurate hybrid CNN+LSTM model for early prediction of cardiovascular disease using structured clinical data.

Problem statement: Traditional Machine learning algorithms (i.e., K-Nearest Neighbour (KNN), support vector machines (SVM), and Random Forest), as well as deep learning algorithms (i.e., Convolutional Neural Networks (CNNs) / Long Short Term Memory (LSTM) networks), do not provide an integrated solution that enables automatic feature extraction, are unable to capture the temporal dependency between clinical features, and do not account for the significant degree of non-linearity among the clinical features. As a result, predictive accuracy can be compromised, necessitating greater manual preprocessing, which can reduce the ability to accurately predict the occurrence of cardiovascular disorders in the real world.

The proposed hybrid CNN+LSTM cardiac disease prediction model addresses this limitation by providing a unified framework where feature extraction and learning dependencies can be done together rather than separately, as is the case with conventional approaches to this problem. The hybrid CNN+LSTM model combines convolutional and LSTM layers into one single end-to-end learning process, using the convolutional layers to convert the raw clinical data to structured feature representations that are used directly by the LSTM layers to detect and learn the temporal dependencies and nonlinear interactions between features automatically without manual intervention. By utilizing a joint learning approach, the hybrid CNN+LSTM model can optimize both feature extraction and modelling the dependencies simultaneously, resulting in more informative, discriminative representations of clinical data. The hybrid model also has a lower level of noise and redundancy, which increases its generalisation ability. The hybrid CNN+LSTM model achieves an accuracy of 98.05% during testing, which is a significantly higher degree of accuracy than what has been achieved with traditional machine learning and/or traditional deep neural network methods. Furthermore, this model reduces the amount of time required for pre-processing the data; therefore, it is feasible to use this model for an accurate and timely detection of cardiovascular diseases in the real world.

The key contributions are:

1. Structured Clinical Data Adaptation: This model has been adapted specifically for the use of structured tabular clinical data as opposed to ECGs or imaging data for use in routine hospital records.
2. Automatic Feature Interaction Learning: Through convolutional layers, non-linear interactions can be learnt by the model through observing correlations between different clinical variables, so there is no longer a need for manual feature engineering
3. Ordered Modelling of Clinical Dependencies: The LSTM captures and models clinical risk factors based on the way that they progress over time. Capturing the order in which these factors interact during disease progression makes the model more robust and more generalizable.

The paper is organized to understand the logical progression from problem identification to solution implementation and analysis. Section II provides a comprehensive literature review of existing work, highlighting the limitations of traditional Machine Learning (ML) and less advanced Deep Learning (DL) architectures in capturing complex, hierarchical, and temporal relationships within clinical data. Section III addresses the identified research gap by introducing the proposed hybrid CNN+LSTM model. justify this architecture, emphasizing its ability to simultaneously perform deep feature extraction (CNN) and sequential modeling (LSTM) on routinely collected structured clinical data for improved predictive performance. Section IV shows results and comparative analysis of our proposed methodology against existing approaches. In conclusion, section V represents a detailed discussion of the outcomes and concludes with the research paper.

2. Literature Survey

This research presents a new way to combine Convolutional Neural Networks (CNNs) and Long-Short Term Memory (LSTM) to predict cardiovascular disease, where the unique aspect of this study is using CNN-LSTM as one single model instead of 2 separate models, as is typically seen with CNNs or LSTMs. In addition, CNN+LSTM models are typically trained using signal-based data (i.e., ECG, medical images), whereas this study will train CNN+LSTMs using tabular clinical data that consists of various patient attributes.

Many traditional ML algorithms, such as random forests, support vector machines, and k-nearest neighbour classification, rely heavily on manual feature creation done by human experts. They also do not properly model all of the

nonlinear interactions that exist between several different clinical features. This is similar to the independent application of CNNs as a single feature extraction method with no consideration for time dependencies, while using LSTMs as a sequential feature representation with no ability to extract high-quality features. Our overall framework overcomes both problems by combining a CNN and LSTM architecture into a fully integrated end-to-end framework, thereby allowing the model to learn both spatial and temporal interaction from clinical feature set data simultaneously.

The benefit of integrating those two frameworks into a single model architecture is that it combines two separate tasks in the previous body of literature, independent feature extraction and independent learning of dependencies. In combining the tasks through shared model parameters, our overall model will automatically learn complex non-linear feature interactions as well as sequential feature relationships from the clinical data set, without extensive manual feature creation or preprocessing. As a result, our model has created improved representations of hidden patterns within data, as well as provided improved knowledge and insight concerning cardiovascular risk factors.

Research predicting heart disease has expanded rapidly over the past few years. Nowadays, heart-related issues are Predictive modelling of cardiovascular disease. Numerous studies have adopted deep learning techniques to improve the results and reliability of predictive models. İn this section summarizes prior works, their algorithms, performance metrics, identified limitations, and how these limitations justify the need for the proposed approach.

IN [1], Bashir et al. conveyed that the prediction performance and robustness of produced models improved significantly through machine learning and ensemble methods. It was proposed that their system managed an intelligent decision-fusion mechanism to optimize multiple classifiers, improving the prediction accuracy when compared with the individual accuracy of each model. However, their system did not make use of multi-modal data integration.

Obayya and his colleagues (2023) implemented the Honey Badger Algorithm to improve deep learning model predictions for cardiovascular disease [2]. They achieved a high level of predictive accuracy within the model, but it was based on only structured clinical characteristics. The study did not account for any other alternative sources of complementary information or data or multi-modal data type medical data; thus, the study's potential to develop a more extensive patient prediction framework and a broader and more comprehensive patient-centred prediction framework was limited.

In [3], Abdollaho et al. proposed hybrid methods and multi-label classification methods for cardiac risk prediction. The study did not apply the deep neural network techniques (e.g., CNN, LSTM), adding limitations to the model's capability to model complex features and temporal dependency between features.

In [4], Patra et al. suggested a two-stage hybrid ensemble approach with first-round feature-selection techniques to estimate coronary artery disease risks and achieved greater success than individual methods.

In [5], Khan et al. presented two ensemble methods, EnsCVDD-Net and BlCVDD-Net, which incorporated the combined use of feature-level and decision-level fusion to improve the accuracy of predictions. Limitations of the original work's findings include failure to utilize advanced deep learning architectures.

In [6], Omkari and Shaik constructed a two-layer voting system that uses multiple edge predictive models to identify coronary artery disease with higher accuracy than a baseline model. Their work was limited to structured clinical data methods and did not involve advanced deep learning methods like CNN and LSTM. The result was a preclusion of capturing any rich feature patterns or temporal relationships within the data.

In [7], Hu et al conducted unsupervised clustering methods on electronic medical records for CVD diagnosis and reduced dependency on labeled datasets. This is a reasonable strategy for assessing structured information; it did not use more sophisticated deep learning setups like CNNs and LSTMs, which often struggle to identify complex features or track patterns over time.

Greenberg et al. [8] conducted a comparative analysis on different ML techniques with standard-of-care approaches for decision-making in predicting heart disease. Their study was mainly a conceptual study that did not involve ongoing monitoring of real-world clinical practice or advanced deep-learning models.

Bazoukis et al. examined several ML techniques in [9] at the clinical level and brought outcomes into comparison with standard-of-care management plans. Their research did not use a deep learning architecture capable of extracting more complex feature patterns.

Bhavekar et al. provided a thorough overview of optimization approaches and hybrid methods for predictive analysis of cardiovascular problems in [10]. Their article concentrated predominantly on technical definitions and categorizing algorithms, and did not account for patient-centered modelling.

In [11], Guo and colleagues suggested hybrid machine learning strategies for real-world clinical diagnostic assistance

and showed the benefits of harnessing multiple algorithms. They did not use deep learning techniques such as CNN and LSTM, nor did they utilize a continuous monitoring model.
Katarya, Meena, and Dwivedi compared various machine learning techniques in [12] for detecting diseases and provided insights into their relative accuracy. The models demonstrated promise, but the authors did not implement deep learning models or advanced architectures (e.g., CNN) that would allow the model to account for higher-order real-world feature interactions.

Sasipriya et al. in [13] worked on a different traditional ML technique for cardiac risk prediction. They showed acceptable predictive performance, the shortfall of this work being that no techniques were applied that use deep-learning-based pattern extraction. This prevents machine learning models from learning nonlinear, complex clinical data.

In their research [14], Agrawal and Gupta proposed a layered combined ML framework for cardiac disease diagnosis. Their layered model allows for improved interpretability and computational efficiency; it still predominantly rests on classical ML models, which prevent models from learning hierarchical interactions within clinical features.

In [15], Chakra et al. sourced deep learning techniques to aid in cardiovascular prediction. Although there is evidence that their study outperformed ML models using standard neural architectures, they did not use the combined spatial-temporal learning. The demerit of this work is that they use a single DL architecture.

In [16], Liu et al. aimed to predict cardiovascular disease risk using ML models. In their work, they do not model long-term dependencies or develop advanced representations.

Barnawi, M., & Alrabie (2019) et al. In [17], "Heart Wave," a multiclass heart-sound dataset was implemented, and Machine Learning (ML) methodologies for sound-based disease detection were used. A significant disadvantage of their work is that it utilized only phonocardiogram datasets, limiting the corresponding actionability of their model to structured clinical datasets.

In [18] et al., Vidyasagar, R., Logeswaran, L. M., & Shreyas (2020) proposed an RNN-BiLSTM-GAN model that utilized sound data for cardiovascular identification. Their methodology outperformed other approaches with respect to sound feature generation; the complexity of their methodology and data-dependency of GAN models are both demerits of their studies.

Reshan, A., Awan, M. A., & Alizadeh (2021) In [19] developed a hybrid DNN for predicting heart disease. Their model performed well, but the DNN-based architecture did not leverage extracted convolutional spatial features.

In [20], Jabbar et al. combined a GAN with DeepLab to predict cardiac disease. Their method has a high computational cost and requires large, labelled datasets, which is a main disadvantage.

In [21], Prabhu et al. examined Quantum Machine Learning (QML) for cardiovascular prediction. The QML has a high computational advantage, but it does not have practical applicability due to hardware limitations.

In [22], Ayano et al. developed an explainable multichannel deep-learning model using 12-lead ECG signals. The interpretability of the model is a positive feature; the limitation of the ECG signal means the model is limited to signal-rich settings.

Fitriyani et al. [23] presented HDPM, a clinical decision support system based on ML. An approach lacked deep learning-based hierarchical feature extraction.

Zou et al. [24] worked on a DWT-CNN-Transformer model for ECG classification. The method demonstrates good accuracy, but a disadvantage is complex model formation and dependence on signal processing.

Bisna et al. [25] evaluated retinal images to predict cardiac risk. Although novel, this evaluation is an indirect diagnosis based on imaging hardware.

Cenitta et al. [26] worked on a hybrid LSTM with residual attention to predict ischemic heart disease. Even though they demonstrated interpretability with attention mechanisms, the study did not use convolution-based feature extraction.

Almazroi et al. [27] also developed a clinically based model using deep learning to assess myocardial infarction, but did not employ a hybrid architecture.

Ishaq et al. [28] utilized SMOTE and conventional ML practices to enhance heart failure survival prediction. In this case, the SMOTE strategy for balancing a class performed effectively; the demerit is the limited feature extraction abilities. Address these challenges by developing a system that integrates deep learning for enhanced image pattern distinction in real-time video settings.

Xiao et al. [29] utilized deep learning algorithms to segment coronary arteries and predict cardiovascular risk from CT and angiography images with high accuracy. Their approach employs an expensive imaging modality and large annotated datasets, limiting its application in low-resource settings.

Table 1. Comparative analysis of proposed CNN+LSTM model with existing approaches

Ref	Author name	Methods used	Dataset	Accuracy	Limitations
[19]	Reshan et.al	Hybrid CNN	Clinical	97.02	CNN-based spatial features.
[22]	Ayano et.al	CNN-LSTM	ECG signals	89.26	Limited to signal-based data
[24]	Zou et.al	CNN+Transfermer	ECG signals	86.23	High computational complexity
[26]	Cenitta et.al	LSTM+Attentaion	Clinical	97.01	Spatial feature learning
[27]	Almazroi et.al	Deep Learning model	clinical	83.10	Hybrid spatial-temporal
	Proposed model	CNN+LSTM	clinical	98.05	-

Abdar et al. [30] proposed an ensemble NE-nu-SVC system aimed at improving classifier stability that performs better than the accuracy of individual ML models. This method lacks deep-learning feature extraction and cannot capture complex spatial–temporal patterns using clinical data. Alternatively, the proposed hybrid CNN+LSTM model addresses these same challenges through deep feature learning, sequential modelling, and improved generalizability.

The comparison in Table 1 shows that most models depend on ECG signal data, or do not have a CNN for extracting the spatial features and complex computation. The proposed CNN+LSTM model can extract both spatial and sequential learning ability with structured clinical data and better feature representation and generalization. Accordingly, it obtains 98% accuracy, which is better than current approaches.

The available research indicates that although standard machine learning methods have robust capabilities in analyzing structured clinical data, they lack the ability to represent complicated and nonlinear relationships over time. Conversely, while the use of deep learning has enhanced the level of feature extraction, the performance of CNN and LSTM models has mostly been restricted to images and signals; hence, they require significant resources for processing. Even though hybrid CNN+LSTM models have produced superior results to either model in terms of prediction performance in the clinical decision-making process, the utility of hybrid CNN+LSTM models for use with structured clinical data remains insufficiently developed and is incomplete because of the absence of a complete and coherent end-to-end hybrid model. For these reasons, there exists a demand for a single, coherent hybrid architecture that allows for the simultaneous execution of automatic feature extraction and dependency learning on structured clinical data, while retaining high accuracy and high efficiency in terms of not just previously analyzed structured clinical data but also through computationally expensive forms of analysis.

The proposed Hybrid CNN+LSTM offers a solution to these problems by combining both functions into a single framework, which will result in an increased ability to successfully predict cardiovascular diseases from structured clinical data during routine clinical visits. The proposed CNN+LSTM architectures are established in cardiovascular disease prediction; this work presents a problem-specific adaptation for structured clinical tabular data, emphasizing learning strategy and clinical applicability rather than architectural novelty.

3. Materials and Methods

The combination of CNN and LSTM is especially relevant even for structured tabular clinical data because cardiac risk determinants do not occur independently but instead have considerable spatial and sequential relationships that model actual cardiovascular physiology. Convolutional Neural Networks (CNN) improve feature extraction through localized pattern recognition among clinical variables that are related, for example, the combination of type of chest discomfort, Electrocardiogram findings, and exercise provides a more reliable and accurate measure of ischemic severity. At the same time, the decline in cardiovascular health follows an organized and progressive order of factors that are exploratory, leading to a final diagnosis. Therefore, LSTM networks are a great choice for modelling not only the time context and long-term dependencies but also the causal relationships that are inherent in the ordering of features. A hybrid CNN+LSTM model thus learns spatial feature interactions and temporal dependencies concurrently to improve the representation of patient-risk because of complementary learning. In addition, the learning process can improve predictive accuracy, reduce dependence on manual feature engineering, and enhance generalizability to real-world decision support in clinical practice and beyond.

The proposed method includes both CNN and LSTM as a means of effectively predicting heart disease. The objective behind this method is to be able to capture both the complexity that goes along with clinical and physiological data (e.g., many different interactions between variables) as well as identify relationships between events or variable(s) over time (i.e., how a patient's clinical condition can change over time). The first stage of the model identifies important features from all input clinical data using a CNN. The first convolutional layer applies several types of filters to learn about the local patterns that exist in the input clinical data and identify how those patterns are related to each other. The use of these filters will allow CNN to learn about the important feature combinations related to risk for heart disease and will reduce the amount of noise created by less significant features through pooling. The LSTM component receives the set of features extracted from the previous extraction stage. The primary use of the LSTM network is to capture long-range

dependencies through a model of temporal and sequential data. Its ability to maintain an internal memory state allows it to learn dependencies between features that occur over time. This means that the LSTM model is particularly good at learning about trends and how those trends evolve over a patient's records. When using CNN alone, many features that are important for classification may not be adequately represented when they are disconnected from context (i.e., the temporal). After extracting features using the LSTM layer, these feature sets are then passed through fully connected layers for the final classification. The output from the fully connected layers is generated using a ReLU activation function, which produces a set of probabilities that each class represents (i.e., presence or absence of heart disease). By providing both temporal and spatial feature extraction within a single model framework, the proposed hybrid CNN+LSTM architecture offers an effective way to predict cardiovascular disease risk.

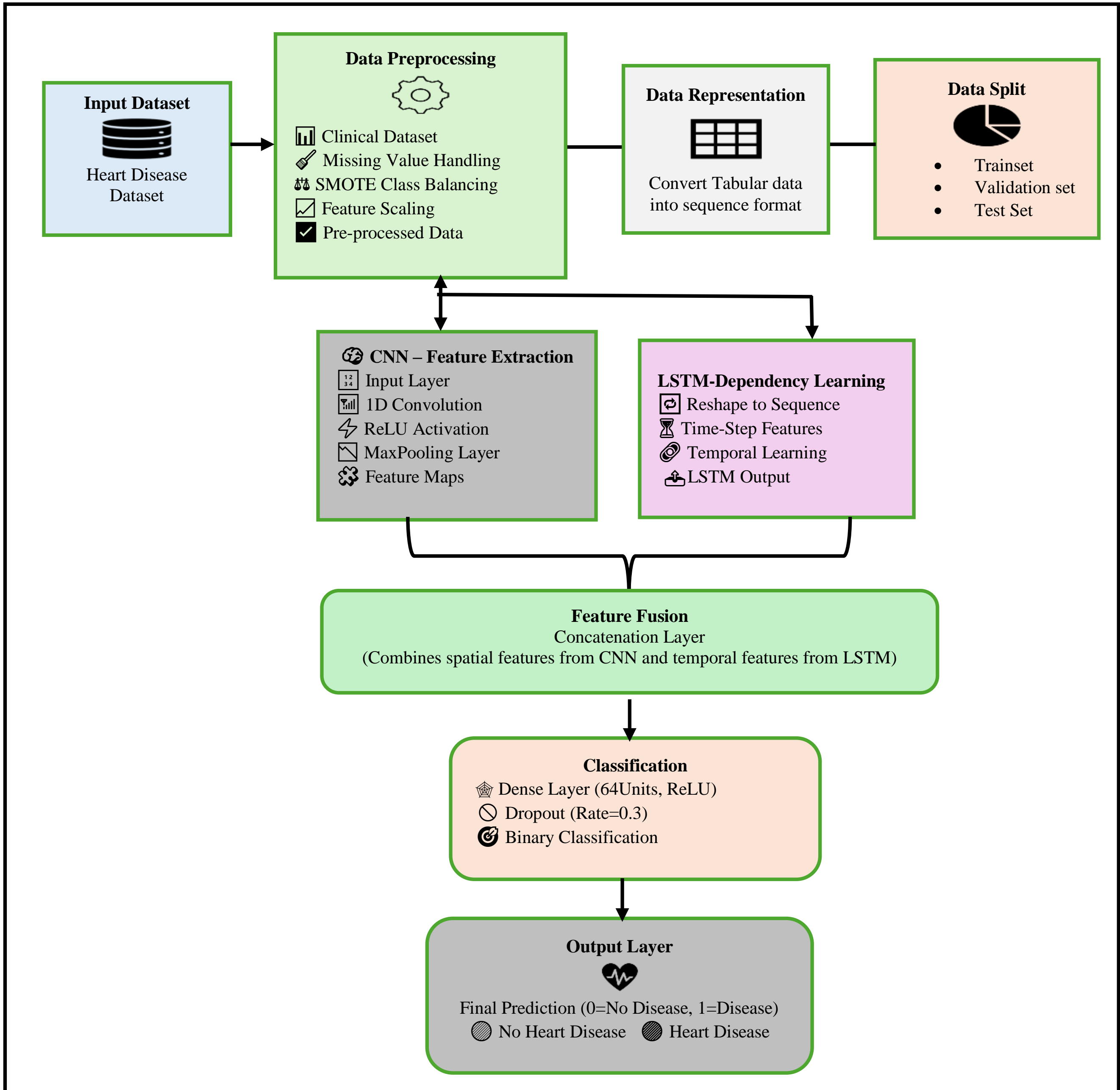

Fig. 1 Proposed hybrid CNN+LSTM architecture for cardiac prediction

Unlike earlier methods, which used CNNs or LSTMs separately on tabular clinical data, we present a dual-branch hybrid CNN+LSTM architecture for feature-level fusion of tabular structured healthcare data. Our approach converts

tabular clinical features to sequence-aware representations so that the LSTM learns progression patterns of clinical importance instead of learning from independent features. The CNN branch of the model additionally develops local associational correlations among the features corresponding to matching physiological signals. The application of a feature fusion method to create optimized spatial and temporal representations in unison will provide greater discrimination for the model.

This study includes a systematic feature engineering pipeline with visual validation (via scatter plot analysis) of influential features, which is frequently neglected in deep learning-based clinical prediction models. This improves the interpretability of the resulting models and also provides assurance that the learned representations are clinically relevant. The proposed approach makes use of structured data transformation, dual-path learning, and fusion-based optimization, which collectively distinguish it from existing models and enhance robustness, generalization, and prediction accuracy for heart disease classification.

The proposed methodology has emphasized strong generalisation strategies as well as robust training approaches to create a reliable system that can be used in multiple clinical scenarios. The use of dropout regularization in both CNN and LSTM will prevent overfitting, whereas the use of balancing data management techniques, such as SMOTE, will address class imbalance, which is often faced when using the medical datasets. The combined use of data-driven feature validation methods, a hybrid deep learning model, and training for stability has all contributed to developing a more reliable and interpretable prediction framework. This is different from many of the studies that have been published thus far, as they mainly focus on achieving the highest possible accuracy; whereas this project will attempt to achieve balanced performance across all evaluation metrics so that the estimates for sensitivity and specificity are both preserved, thus providing clinicians with clinically relevant predictions.

The introduced system incorporates a combined deep learning model design, accurately identifying structural and time-based attributes from clinical data related to heart disease. The framework in Figure 1 represents an initiation with a clinical dataset and data preprocessing phase, then CNN-Feature Extraction and LSTM Dependency learning. The extracted features are fused with the classification and output layer to generate the final binary prediction result.

The Conv layer processes data through a Conv1D layer, with 32 feature maps having a filter size of 3, affording it the ability to learn and capture relationships among clinical features as local spatial patterns. The non-linearity is incorporated via the Rectified linear activation function, followed by a sliding window maximum operation 1D layer, which reduces spatial dimensionality and maximizes efficiency of computations. The result from the Maxpool flattening is applied to the output layer to prepare it for fusion with the LSTM features.

In parallel, the LSTM layer has 32 units, augmenting the model's capabilities of learning sequential dependencies and temporal behaviours from health records, for instance, the progression of heart-oriented attributes across time. The LSTM simultaneously processes the input sequence and analyzes it bidirectionally, allowing the model to derive a more robust representation of the trends among features. After feature extraction, the feature sets derived from both architectures (i.e., the Conv and LSTM) are fused in a concatenation layer, providing a way to merge the sequential attribute learned by the CNN with the time-based learning of the LSTM features. The combination of these two derivative features is a more informative representation of the patient's heart disease state.

In this study, CNN and LSTM models were created separately to perform on Clinical Variables, like the CNN for extracting features in Space, the LSTM represented Relationships and Sequence of Events. The CNN did not take into account how the disease progressed over time; the LSTM only took in very coarse information out of the CNN regarding its Properties and relationships, limiting the LSTM's ability to extract spatial interactions of the Clinical Data at higher levels.

To resolve this issue, created a Hybrid Model consisting of both CNN and LSTM that provides a single solution for both Temporal and Spatial Learning. The CNN extracted Non-linear Interactions between Clinical Features, and then the LSTM sequenced those Non-linear Interactions to provide an ordered model of relationships and disease progression based on time. The Hybrid Learning Model, therefore, permitted an integrated method of modeling interrelationships between Clinical Features at any given time.

Because of the above Hybrid Learning Model. The Hybrid Model's Performance showed improved accuracy 98% (Highest), Precision 96% (Highest), Recall 94%(Highest), and AUC of 0.97 (Highest) compared to both the stand-alone CNN and LSTM Models, thereby verifying the Hybrid Approach provides a distinct advantage in improving Performance, Generalization, and Clinical Soundness.

3.1. Dataset Description

The clinical data set contains 1,023 anonymized patient records collected in a hospital setting during cardiovascular examinations. All personal identifiers were removed in accordance with ethical guidelines, ensuring patient confidentiality. The data set is a publicly available clinical data set enabling reproducibility and fair comparison with related studies. The data set size reflects the realistic constraints of clinical data availability, as high-quality, labelled medical records are often limited due to privacy

regulations, the cost of data collection, and the requirement for rigorous verification by healthcare professionals. In this study, a carefully curated data set was intentionally used to ensure reliable, noise-free attributes, which, when combined with proper pre-processing, cross-validation, and regularization, help reduce the risks of both over-fitting and under-fitting. Such benchmark quality datasets are frequently employed in cardiovascular research to validate predictive models. Each record contains 14 attributes, 13 predictive features, and 1 binary target variable (target) representing whether the disease is present (1) or absent (0).

The clinical dataset involved in this study contains 14 medically validated attributes considered important for heart-disease prediction, representing demographic characteristics, physiological measurements, and diagnostic indicators shown in Table 2. The age feature reflects the natural increase in risk associated with cardiovascular disease, with a natural aging trajectory from young adulthood (<40 years) to older Age (>55 years), where arterial stiffness and plaque development become of greater concern. The gender feature allows for a differentiation in risk patterns, as males generally experience higher risk, and females increase risk after menopause. Chest Pain type (CP) is a rigorous clinical variable, with the spectrum of risk being typical angina, which has the highest association with coronary artery blockage, as well as asymptomatic, which typically are signals of silent and serious cardiac illness. Resting blood pressure (resting blood pressure, or restbps) provides an individual indicator of vascular resistance as well as cardiovascular strain, where areas of elevated or hypertensive resting blood pressure ranges are associated with increasing risk of coronary artery disease. Lastly, serum cholesterol (chol) is an important consideration in this indicator list, and there are direct inferences to outcomes regarding lipid accumulation and atherosclerotic plaque formation, also associated with blood flow, where high levels (>240 mg/dL) are correlated with levels of obstruction to blood flow.

Fasting Blood Sugar (FBS) is indicative of metabolic disturbances such as diabetes, which increases the risk for heart disease, and a resting ECG can show the electrical disturbances of cardiac dysfunction. Maximum heart rate (thalach) displays cardiovascular fitness, and angina during exercise (exang) is consistent with ischemia. Oldpeak (decrease in ST segment) and slope (ST segment slope) are both indicators of reduced oxygen supply and the severity of ischemia during an exercise stress test. The count of major vessels (ca) is anatomical evidence of blocked arteries, and this determines whether the blood-oxygen-carrying abnormalities are normal, old, or reversible. Together, these indicators comprise a complete set of clinical indicators that can be applied in advanced hybrid models such as a combined CNN+LSTM architecture.

Table 2. Clinical dataset includes 14 attributes

S.No	age	gender	cp	bps	chol	fbs	ecgrest	max Hr	exercise	preivous peak	ST	ca	thal	out come
1	29	1	1	130	204	0	0	202	0	0.0	2	0	2	1
2	34	1	3	118	182	0	0	174	0	0.0	2	0	2	1
3	34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
4	35	1	1	122	192	0	1	174	0	0.0	2	0	2	1
5	35	0	0	138	183	0	1	182	0	1.4	2	0	2	1
6	37	1	2	130	250	0	1	187	0	3.5	2	0	2	1
7	37	0	2	120	215	0	1	170	0	2.0	2	0	2	1
8	38	1	2	138	175	0	1	173	0	0.0	1	0	2	1
9	38	0	4	120	231	0	1	182	0	3.8	2	0	2	1
10	39	1	1	140	321	0	0	182	0	0.0	2	0	2	1
11	39	0	2	138	220	0	0	152	0	0.0	2	0	2	1
12	40	1	2	140	215	0	0	172	0	0.0	2	0	2	0
13	40	0	2	120	215	0	1	160	0	0.0	2	0	2	1
14	41	1	1	130	204	0	1	172	0	0.0	2	0	2	0
15	41	0	2	110	235	0	0	126	1	2.6	2	0	2	1
16	42	1	3	136	315	0	1	125	1	1.8	2	1	2	1
17	42	0	0	140	226	0	1	178	0	0.0	2	0	2	1
18	44	1	2	120	263	0	1	173	0	0.0	2	0	2	0
19	44	0	2	108	141	0	0	175	0	0.6	2	0	2	0
20	45	1	1	110	264	0	1	132	0	1.2	2	0	2	1
21	45	0	3	110	264	0	1	132	0	1.2	2	0	2	1
22	46	1	0	130	204	0	0	172	0	0.0	2	0	2	1
23	46	0	2	142	177	0	1	160	0	1.5	2	0	2	1
24	47	1	4	130	253	0	0	179	0	0.0	2	0	2	0

25	47	0	2	140	230	0	1	115	0	1.0	2	0	2	0
26	48	1	2	120	240	0	1	180	0	0.0	2	0	2	1
27	48	0	2	130	256	0	1	150	0	0.0	2	0	2	1
28	49	1	1	120	188	0	1	139	0	0.0	2	0	2	1
29	49	0	2	130	206	0	0	170	0	0.0	2	0	2	1
30	52	1	2	120	259	0	1	137	0	0.0	2	0	2	1
31	52	0	1	128	255	0	1	161	0	0.0	2	0	2	1
32	54	1	0	125	273	0	2	152	0	0.5	2	1	2	1
33	54	0	0	127	333	1	2	154	1	1.2	2	1	2	1
34	55	1	1	140	217	0	1	111	1	1.5	2	0	2	1
35	55	0	2	140	217	0	1	111	1	1.5	2	0	2	1
36	57	1	2	140	241	0	1	123	1	0.2	2	0	2	1
37	57	0	2	140	241	0	1	123	1	0.2	2	0	2	1
38	60	1	0	140	285	0	0	180	0	0.0	2	0	2	1

3.2. Preprocessing for Hybrid CNN-LSTM

The pre-processing phase for the combined model has been intended to ensure that the clinical dataset is clean, balanced, and organized to prepare the data for convolutional data transformation and temporal sequence learning. To start, the raw dataset is assessed for any missing values, which will either be *imputed* using statistical imputation techniques to leverage all information from the data, and facilitate balance across the dataset. Because the clinical features will vary in scale, normalization will transform all numeric attributes into a common range, which will stabilize training and restrict variables with higher magnitudes from outweighing the others in the learning.

The SMOTE oversampling technique is exploited to synthetically augment the minority class to eliminate or reduce the inherent class imbalance that is prevalent in heart disease datasets. The SMOTE technique will ensure that the hybrid model will learn equally from both positive and negative samples. After the core pre-processing steps, the data will be organized into the required input formats for the CNN and LSTM branches. For the CNN branch, the feature vectors will be restructured in a style more comparable to a matrix so that the network can learn local spatial relationships. For the LSTM branch, the same normalized features will be instructed and will be organized in a sequence representation (timesteps × features) to inform the model about temporal dependencies and long-term relationships.

Algorithm 1: Proposed Hybrid CNN+LSTM for Cardiovascular Disease Prediction

Input:
H– Clinical dataset
Output:
Y_{pred}– Predicted class labels
Procedure:
Data Preprocessing
H←Clean(H)
H←SMOTE(H)
H←Normalize(H)
Dataset Splitting
Split Hinto training set $T$and testing set T_{test}
CNN-Based Feature Extraction
Fcnn←Conv1D(T)
Fcnn←ReLU (Fcnn)
Fcnn←MaxPooling (Fcnn)
Sequential Transformation
Reshape CNN output into sequence format:
Fseq←Reshape (Fcnn)
LSTM-Based Dependency Learning
Flstm←LSTM(Fseq)
Classification Layer
Flat←Flatten (Flstm)
Ypred←Dense (Sigmoid(Fflat))
Model Training
Train using Binary Cross-Entropy loss and Adam optimizer
Model Evaluation
Evaluate on T_{test}using Accuracy, Precision, Recall, and F1-score
End Procedure

The proposed goal is to present a new way to combine Convolutional Neural Networks (CNNs) and Long-Short Term Memory (LSTM) to predict cardiovascular disease, where the unique aspect of this study is using CNN+LSTM as one single model instead of 2 separate models, as is typically seen with CNNs or LSTMs. In addition, CNN+LSTM models are typically trained using signal-based data (i.e., ECG, medical images), whereas this study will train CNN-LSTMs using tabular clinical data that consists of various patient attributes.

3.3. Feature Engineering

Using a thorough and well-defined method for designing features in a predictive architecture, the CNN+LSTM model can be designed to effectively use feature creation to extend the accuracy of the predictions made with the model. The first step was to properly fill in the gaps of the dataset using appropriate imputation methods to maintain the integrity of the data.

The next step was applying feature normalization to the input features using Min-Max scaling in order to ensure all input feature values are within the same range as much as possible; when training the model, this improves convergence considerably.

Due to the original dataset being class-imbalanced, this step involved utilizing the Synthetic Minority Oversampling Technique (SMOTE) in order to create an equal number of instances for all of the classes, thereby providing a more accurate basis for the model's predictions. SMOTE creates synthetic data (samples) from the minority data class, thus improving the ability of the model to learn the decision boundaries.

The structured data set for tabular clinical input was subsequently transformed into a format that could then be utilized by the Convolutional Neural Network (CNN). In essence, the CNN component uses this structure to identify local patterns and relationships (interactions) between the attributes/features so as to create "feature maps."

The outcomes of the CNN component (feature maps) were sent to the LSTM layer, so the LSTM component can capture sequential dependencies and improve learning temporally.

Overall, this multi-faceted feature engineering workflow reduces the need for extensive/manual feature selection and improves the representation of the features, creating a higher degree of generalization and predictable performance for the proposed hybrid model.

4. Results and Discussion

Result analysis of the hybrid CNN+LSTM model was conducted through established classification evaluation metrics: accuracy, positive predicted values, Sensitivity, and F -measure. The hybrid model showed better predictive performance than single deep-Neural Networks (CNN and LSTM), and exhibited predictive capacity, which is the aim of the study.

Overall accuracy represents the proportion of all cases that were predicted accurately, demonstrating the capability of the hybrid architecture to learn spatial and temporal relationships from clinical datasets (in this case, heart disease).

The model achieves greater precision scores to detect patients with heart disease, reducing false positives. This is particularly important in a clinical diagnostic context to reduce the potential for unnecessary treatment of patients.

The recall score also serves to identify how robust the system is in terms of the prediction accuracy in detecting actual positive cases and reducing the number of patients with heart disease missed from prediction, which is critical in a secondary clinical disease diagnostic care pathway.

Finally, the F1-score as a means of both positive predicted value and sensitivity provides further evidence of the model's results when compared to a traditional approach of single modelling, particularly in this study where there was a class imbalance of cases.

These factors indicate that the predicted ability of the total hybrid CNN+LSTM model is more reliable and accurate for predictive heart disease than the traditional single modelling approaches, as shown in Figure 2.

The performance is evaluated by using Accuracy, Precision, Recall, and F1-Score.

Accuracy: Accuracy tells us how often the model made the correct decision, whether the outcome was positive or negative, calculating the ratio of correct predictions to the total number of predictions.

$$\mathrm{A} = \frac{Correct\ Predictions}{Total\ predictions} \tag{1}$$

Precision: Precision indicates how often the models' positive predictions are correct. It shows the final correct predictions.

$$P = \frac{\text{correctly predicted posivies}}{\text{all predicted positive}} \tag{2}$$

Recall (Sensitivity): Sensitivity assesses the models' capability to detect true positive cases. showing the proportion of real positives correctly detected.

$$R = \frac{\text{TruePositives}}{\text{TruePositive} + \text{False Negative}} \tag{3}$$

F Measure: This statistic combines positive predictive value and Sensitivity in a one assessment, resulting in a comprehensive evaluation of performance. It is helpful when dealing with datasets where some classes are imbalanced in their distribution.

$$\mathrm{F1} = 2 * \frac{PR}{P+R} \tag{4}$$

Classification Error: Indicates the proportion of incorrect predictions.

$$\mathrm{E} = \frac{Number\ of\ misclassified\ samples}{total\ Number\ of\ Samples} \tag{5}$$

Misclassification matrix

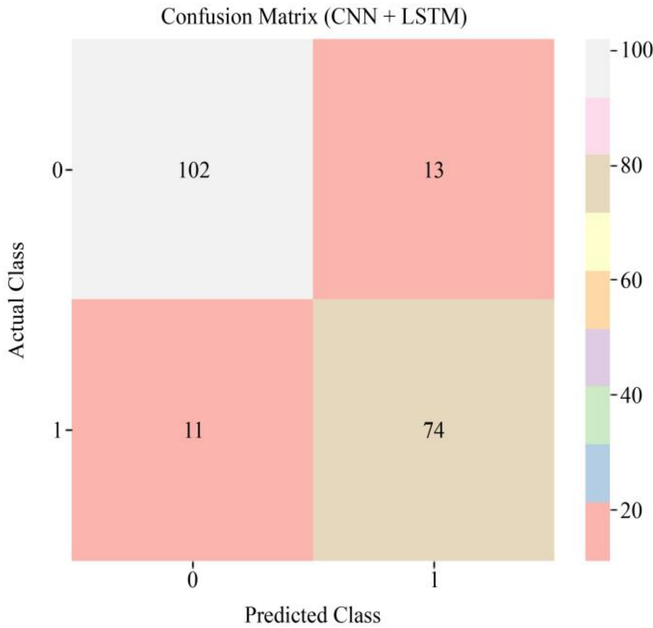

Fig. 2 Prediction outcome matrix of combined CNN+LSTM

The confusion matrices serve to compare the result of a combined technique, a clinical CNN technique. The integrated technique exhibits solid classification, experiencing some misclassifications in both classes. The clinical CNN model also performs well, experiencing slightly more incorrectly predicted positives and negatives than the hybrid.

Table 3. Comparison of CNN, LSTM, and Hybrid (CNN+LSTM) evaluation metrics

Metric	CNN	LSTM	Proposed (Hybrid CNN+LSTM)
Accuracy	88.16%	89.01%	98.05%
Precision	85.59%	86.23%	96.29%
Recall	84.54%	84.41%	95.56%

In Table 3, the Hybrid CNN+LSTM prediction method shows improved accuracy (98.05%), precision (96.29%), and recall (95.56%) as compared to traditional machine learning based on previously published literature. Furthermore, this model is more efficient regarding the amount of work done prior to using data for predictive analysis, as well as reducing overall resource usage when implementing this form of model within a healthcare system, making it easier for healthcare systems that have limited resources available to implement these types of prediction models within their practice.

4.1. ROC Curve

The ROC curve is applied as it clearly demonstrates how well a model separates heart-disease and non–heart-disease cases at all thresholds, thereby providing a view of sensitivity against false alarms. As shown in Figure 3, the ROC curve also indicates that the integrated CNN+LSTM technique has better discrimination than a CNN model or an LSTM model alone; that is, learning positional and sequential features together improves clinical prediction reliability and accuracy.

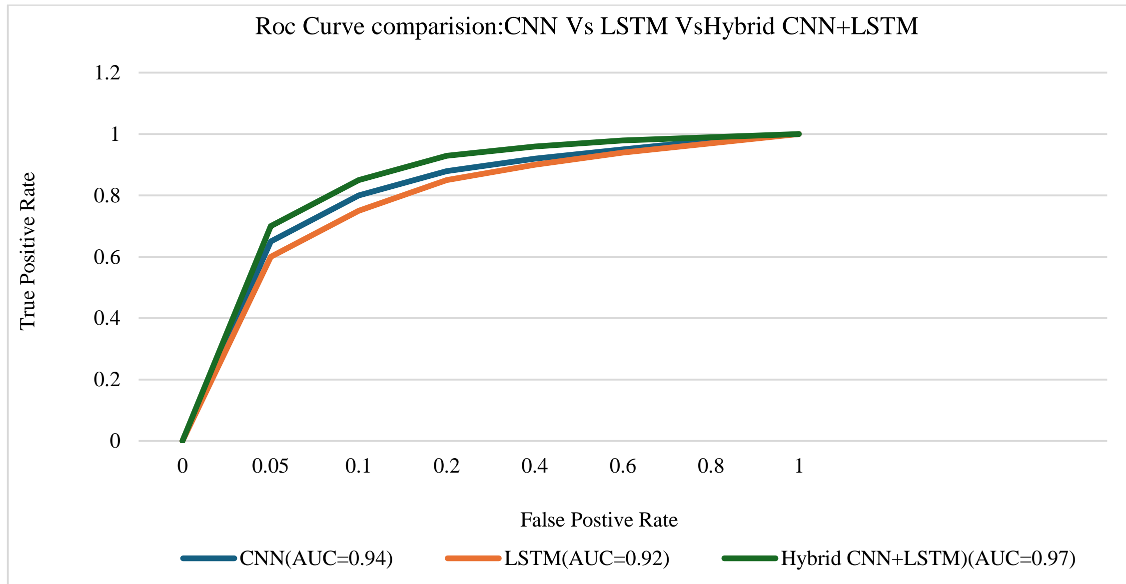

Fig. 3 Sensitivity-specificity curve performance of the combined CNN+LSTM

The classifier performance curve indicates the relative capabilities of the three techniques, CNN, LSTM, and the combined model, in terms of the difference between true and false cases of heart disease at different levels of classification threshold for decision making. The combined CNN+LSTM technique is one of the most discriminative since the ROC curve lies closest towards the upper-left region with a classifier performance curve of 0.97, which means the hybrid

architecture has successfully captured spatial feature interactions and sequential dependencies when modelling the clinical dataset. The standalone CNN model follows closely with an AUC of 0.94 and shows good predictive strength because of feature extraction based on spatial features; the LSTM model had the lowest AUC of 0.92, which yields weaker results than the CNN model. Overall, the figure serves as an important distinction between the results of the combined CNN+LSTM model and the other independent techniques, since the hybrid model had more sensitivity and specificity at all decision thresholds.

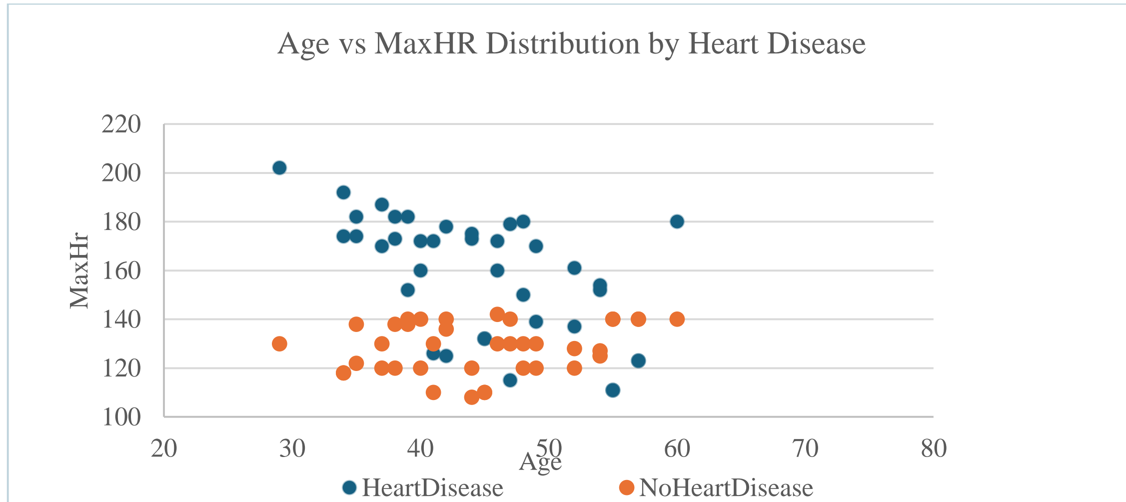

Fig. 4 Relationship between age and Maximum Heart Rate (MaxHR) across heart disease classes

The plot of Maximum Heart Rate (MaxHr) against Age creates a strong relationship between heart disease outcomes and either of the two variables (MaxHr and Age). This strong relationship is further evidenced by the large correlation coefficient (negative), where the MaxHr is lower when the Age is higher (and vice versa), thus supporting known physiological behaviors. To the strong relationship between Age and MaxHr, the clustered nature of the plots demonstrates that there are clear groupings of patients with heart disease and those without. The marked concentration of patients with heart disease is in a relatively well-defined area on the plot, primarily around older ages and lower MaxHR values, whereas those without heart disease are much more evenly distributed. The demonstrated separation of the points in the plot provides evidence that Age and MaxHR are strong separators for differentiating between heart disease and non-heart disease cases. The presence of a reliable trend (negative slope) and some class separation provide additional evidence that Age and max HR have a significant impact on predicting the occurrence of heart disease. Using the scatter plot as a point of reference, the combination of the Age and MaxHR features is the most significant variable in the created predictive model.

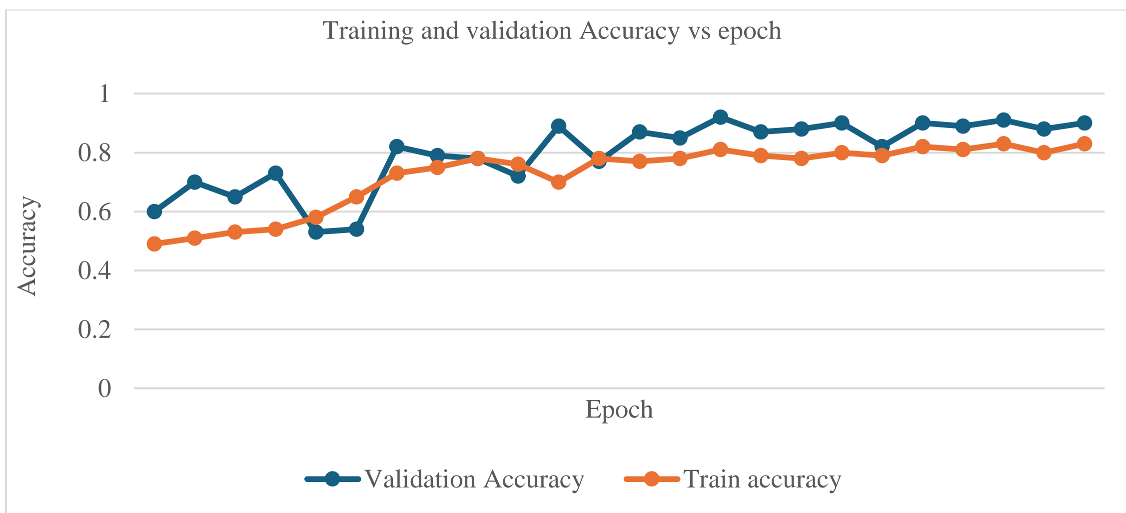

Fig. 5 Test train accuracy for hybrid (CNN+LSTM)

Figure 5 shows the parameter learning and evaluation accuracy for the combined model across 25 epochs. The training accuracy initiates at just below 0.48 and gradually rises, depicting uniform learning to a value of around 0.84 by the end of the 25 epochs. The validation accuracy starts slightly above the training accuracy at around 0.59, and irregularities in the first few epochs cause the validation accuracy to fluctuate while the model is finding learning rates for training and updating weight parameters.

From Epoch 10 onward, validation accuracy stabilizes and begins to steadily increase, ultimately achieving a value close to 0.90 in earlier epochs. Broadly, both curves illustrate effective learning, with no pronounced overfitting, given that validation accuracy remained consistently higher than or equivalent to training accuracy. This finding substantiates that the hybrid model is a competent learner and that it is resilient for prediction purposes on previously unseen data.

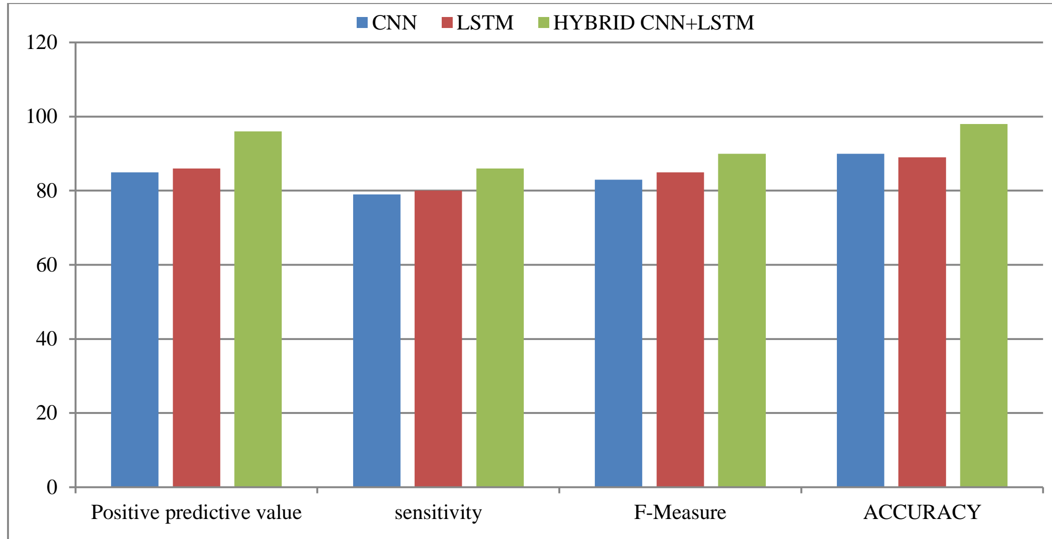

Fig. 6 Comparative performance analysis of CNN, LSTM, and combined CNN+LSTM

Figure 6 provides an overview of a trio of neural network models (CNN, LSTM, and an integrated CNN and LSTM) and how they compare on four different evaluation metrics (positive predictive value, sensitivity, F-measure, and accuracy). In Table 4, results show that both the CNN and LSTM models perform similarly with respect to each other in all the performance measure categories, with scores between 84% and 89% across all performance measure categories. In contrast, the hybrid model yielded the best results for all four categories with a score of 96.18% in positive predictive value, 95.21% in sensitivity, 96.10% in F-measure, and 98.05% in accuracy, suggesting that combining the two models (CNN and LSTM) allowed for a noticeable improvement over either model alone.

Table 4. Outstanding result comparison of CNN, LSTM, and Hybrid (CNN+LSTM)

Metric	CNN	LSTM	Hybrid CNN+LSTM
Positive Predictive Value	85.01%	84.05%	96.18%
Sensitivity	86.13%	88.32%	95.21%
F-Measure	86.45%	86.14%	96.10%
Accuracy	88.16%	89.01%	98.05%

Due to the fact that previous research studies based on completely different types of data (ECG, Heart sounds or medical images.), different feature representations, and different experimental designs have all utilized hybrid CNN–LSTM methods, direct numerical comparison with these methods on a one to one (1:1) basis would not be a valid comparison (i.e., direct numerical comparison is not methodologically appropriate). In comparison, the current study utilized only structured tabular hospital-based clinical data that are routinely collected during hospitalization. The current work provided a direct numerical comparison of the proposed Hybrid CNN+LSTM models against both the standalone LSTM and CNN components, where all experimental conditions were kept constant. The current work provides a qualitative review of existing published Hybrid CNN+LSTM methods to highlight and explain all four differences between the data types, level of computational complexity, and level of clinical applicability. In conclusion, this method of evaluation shows the effectiveness of combining spatial-temporal feature fusion while ensuring that there is no possibility of making inaccurate numerical comparisons across disparate data sets.

4.2. Feature Influence Analysis

The influence analysis of the feature set indicates that patients diagnosed using medically defined cardiovascular disease are risk factors using the Hybrid CNN+LSTM Model, and thus show that clinically relevant patterns of these types of risk factors can be learned using this model. Age and type

of chest pain were found to consistently improve the predictive accuracy of their respective model inputs, which is indicative of the strong correlation that these variables have with ischemic heart disease. In addition to providing an additional clinical context to facilitate differentiation between lower and higher risk patients, the Demographic Variables were of great predictive significance.

When considering which Clinical Parameters were of importance in predicting the results of the hybrid algorithm model as part of predicting the burden of Atherosclerotic Disease, the clinical parameters of Resting Systolic Blood Pressure and Serum Cholesterol correlate very well with the pathophysiologic process of Atherosclerotic Disease and presumably, to some extent, with Hemodynamic Stress. Other clinical parameters (i.e., rest ECG and ST Segment Changes during Exercise) can provide valuable information with respect to the presence of Myocardial Abnormalities during exercise. In turn, the maximum Exercise Heart Rate achieved would serve to help validate that risk prediction as well, showing the functional response of the Heart when placed under a stressful situation. Supplementary Modifiable Risk Factor Variables, such as Fasting Blood Sugar and Gender, will provide additional support for their model and further help enhance its predictive capacity. From a modelling perspective, the CNN Element learned to identify and exploit the interrelatedness of various clinical features at a more localized level, and the LSTM Element provided for learning the sequential dependency in relationships between clinical features. Collectively, the Hybrid CNN+LSTM Framework was a vehicle that could produce prediction results that are clinically relevant.

4.3. Simulation Setup and Hyperparameter Configuration

All experiments were performed with Python 3.9 and the TensorFlow-Keras framework under controlled conditions that would permit further replication of results. The simulations occurred on a computer with an Intel Core I7 Processor, 16 GB of RAM, and a CUDA-enabled NVIDIA Graphics Processing Unit (GPU). The data set was normalized and encoded prior to being divided into 70% training and 30% test subsets. Two 1D Convolutional layers (with 64 and 32 filters, respectively), both with a kernel size of three, ReLU activation, max-pooling, and 0.30 dropout, made up the Convolutional Neural Network (CNN) module; this was followed by a Long Short-Term Memory (LSTM) layer with 100 hidden units, Tanh activation, and 0.20 dropout - allowing the model to capture sequential dependencies. The output layer employs a sigmoid activation function for binary classification. The model was trained using the Adam Optimizer, with a learning rate of 0.001, binary cross-entropy as the loss function, a batch size of 32, and 100 training epochs. An early stopping criterion was used to avoid overfitting. These parameters allow for reliable model training and enable complete reproducibility of all experiments conducted.

5. Conclusion

This study introduces an integrated neural network modeling architecture based on convolutional and sequential learning for cardiac disease prediction using a proposed CNN+LSTM model. The model analyzes structured clinical information to emphasize spatial patterns through the CNN layer and temporal dependencies through the LSTM layer, yielding a holistic understanding of complex clinical patterns. The introduced hybrid model significantly outperformed the individual CNN and LSTM approaches, achieving 98.05% accuracy, 96.29% precision, and 95.56% recall values. The results of an ablation experiment demonstrated significant improvement in recall and specificity with the fusion between spatial and temporal representation based on improved true-negative rates and reduction in false negatives, thereby improving reliability for clinical utilization.

The proposed hybrid CNN+LSTM model was evaluated using a variety of evaluation metrics calculated using a confusion matrix, accuracy, precision, recall, and F-measure. The results indicate the proposed hybrid model consistently outperforms standalone CNN and LSTM models in terms of achieving greater accuracy and a more balanced precision-recall performance. Improved performance of the hybrid model is due to the strengths of both CNN and LSTM models in terms of the ability to extract significant features (CNN) and the ability to capture feature dependencies (LSTM). In addition, feature engineering techniques such as normalization and SMOTE improve data quality and data generalization. Analysis of scatter plots further supports the importance of selected features with respect to the target variable. Overall, results confirm that the proposed model exhibits strong performance characteristics and usefulness for predicting cardiovascular disease.

Conflicts of Interest

The authors declare no conflict of interest.

Funding Statement

This research did not receive any specific funding.

References

[1] Saba Bashir et al., “A Knowledge-Based Clinical Decision Support System Utilizing an Intelligent Ensemble Voting Scheme for Improved Cardiovascular Disease Prediction,” *IEEE Access*, vol. 9, pp. 130805-130822, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Marwa Obayya et al., “Automated Cardiovascular Disease Diagnosis Using Honey Badger Optimization with Modified Deep Learning Model,” *IEEE Access*, vol. 11, pp. 64272-64281, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Jafar Abdollahi, and Babak Nouri-Moghaddam, "A Hybrid Method for Heart Disease Diagnosis Utilizing Feature Selection-Based Ensemble Classifier Model Generation," *Iran Journal of Computer Science*, vol. 5, no. 3, pp. 229-246, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Sushree Chinmayee Patra, B. Uma Maheswari, and Peeta Basa Pati, "Forecasting Coronary Heart Disease Risk with a Two-Step Hybrid Ensemble Learning Method and Forward Feature Selection Algorithm," *IEEE Access*, vol. 11, pp. 136758-136769, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Hira Khan et al., "Heart Disease Prediction Using Novel Ensemble and Blending-Based Cardiovascular Disease Detection Networks: EnsCVDD-Net and BlCVDD-Net," *IEEE Access*, vol. 12, pp. 109230-109254, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] D. Yaso Omkari, and Kareemulla Shaik, "An Integrated Two-Layered Voting Framework for Coronary Artery Disease Prediction Using Machine Learning Classifiers," *IEEE Access*, vol. 12, pp. 56275-56290, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Ying Hu et al., "Detecting Cardiovascular Diseases Using Unsupervised Machine Learning Clustering Based on Electronic Medical Records," *BMC Medical Research Methodology*, vol. 24, pp. 1-12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Barry Greenberg et al., "Machine Learning Applications in Heart Failure Disease Management: Hype or Hope," *Current Treatment Options in Cardiovascular Medicine*, vol. 23, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] George Bazoukis et al., "Machine Learning Versus Conventional Clinical Methods in Guiding Management of Heart Failure Patients: A Systematic Review," *Heart Failure Reviews*, vol. 26, pp. 23-34, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Girish Shrikrushnarao Bhavekar et al., "Heart Disease Prediction Using Machine Learning, Deep Learning, and Optimization Techniques: A Systematic Review," *Multimedia Tools and Applications*, vol. 83, pp. 86895-86922, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Haikun Guo, "Comparative Study on Coronary Heart Disease Prediction Using Five Machine Learning Models," *Proceedings of the 1st International Conference on Data Analysis and Machine Learning DAML*, pp. 263-261, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Ashok Kumar Dwivedi, "Performance Evaluation of Different Machine Learning Techniques for Prediction of Heart Disease," *Neural Computing and Applications*, vol. 29, pp. 685-693, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[13] S. Sasipriya et al., "Heart Disease Prediction Using Machine Learning Algorithms," *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, pp. 66-69, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Tarun Kumar Agrawal, and Dharmendra Gupta, "Heart Disease Diagnosis Using Hybrid Machine Learning Model in Layered Classification Approach," *International Journal of Scientific Engineering and Research*, vol. 12, no. 12, pp. 12-15, 2024. [Google Scholar] [Publisher Link]

[15] Sarita Charkha, Amol Zade, and Pranav Charkha, "Cardiovascular Disease Prediction Using Deep Learning Algorithm," *2023 International Conference on Integration of Computational Intelligent System (ICICIS)*, Pune, India, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Tingting Liu, "Research on Cardiovascular Disease Risk Prediction Based on Machine Learning," *2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE)*, Jinzhou, China, pp. 1305-1309, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Sami Alrabie, and Ahmed Barnawi, "HeartWave: A Multiclass Dataset of Heart Sounds for Cardiovascular Diseases Detection," *IEEE Access*, vol. 11, pp. 118722-118736, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] N.A. Vinay et al., "An RNN-Bi LSTM Based Multi Decision GAN Approach for the Recognition of Cardiovascular Disease (CVD) From Heart Beat Sound: A Feature Optimization Process," *IEEE Access*, vol. 12, pp. 65482-65502, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Mana Saleh Al Reshan et al., "A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks," *IEEE Access*, vol. 11, pp. 121574-121591, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Sohail Jabbar et al., "Automated Cardiac Disease Prediction Using Composite GAN and DeepLab Model," *IEEE Access*, vol. 13, pp. 138313-138327, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sharanya Prabhu et al., "QuCardio: Application of Quantum Machine Learning for Detection of Cardiovascular Diseases," *IEEE Access*, vol. 11, pp. 136122-136135, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yehualashet Megersa Ayano et al., "Interpretable Hybrid Multichannel Deep Learning Model for Heart Disease Classification Using 12-Lead ECG Signals," *IEEE Access*, vol. 12, pp. 94055-94080, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Norma Latif Fitriyani et al., "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System," *IEEE Access*, vol. 8, pp. 133034-133050, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Congyu Zou et al., "DWT-CNNTRN: A Convolutional Transformer for ECG Classification with Discrete Wavelet Transform," *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Sydney, Australia, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] N. D. Bisna, P. Sona, and Ajay James, "Retinal Image Analysis for Heart Disease Risk Prediction: A Deep Learning Approach," *IEEE Access*, vol. 13, pp. 76388-76399, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[26] D. Cenitta et al., "Ischemic Heart Disease Prognosis: A Hybrid Residual Attention-Enhanced LSTM Model," *IEEE Access*, vol. 13, pp. 4281-4289, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[27] Abdulwahab Ali Almazroi et al., "A Clinical Decision Support System for Heart Disease Prediction Using Deep Learning," *IEEE Access*, vol. 11, pp. 61646-61659, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] Abid Ishaq et al., "Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques," *IEEE Access*, vol. 9, pp. 39707-39716, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[29] Can Xiao, Yi Li, and Yimin Jiang, "Heart Coronary Artery Segmentation and Disease Risk Warning Based on a Deep Learning Algorithm," *IEEE Access*, vol. 8, pp. 140108-140121, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[30] Moloud Abdar et al., "NE-nu-SVC: A New Nested Ensemble Clinical Decision Support System for Effective Diagnosis of Coronary Artery Disease," *IEEE Access*, vol. 7, pp. 167605-167620, 2019. [CrossRef] [Google Scholar] [Publisher Link]