International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Heart Disease Binary and Multiclass Classification Using Deep Learning Hybridized with Ensemble Learner

Pardeep Kumar[1]* **Ankit kumar[1]**

[1]*Computer Science and Applications, Baba Mastnath University Asthal Bohar, Rohtak, India*
* Corresponding author's Email: pardeephigheredu@gmail.com

Abstract: With growth in world's population over the last several decades, the provision of medical care has emerged as the component of human life that is most essential. A considerable number of people are losing their lives much too early due to heart disease on an annual basis. It is crucial to detect the illness at its earliest possible stage in order to lessen the chance of dying from heart disease. This should be done as soon as possible. The enormous amounts of information generated for diagnostic purposes have made it possible to construct complex learning-based models for the early, automated diagnosis of cardiac issues. This has been made possible as a result of the availability of this information. Because of this, there have been substantial gains made in the accuracy of diagnostics. As a consequence of this, there have been significant breakthroughs made in medical technology. The traditional approaches to machine learning are unable to generalize their conclusions to new datasets since these datasets were not part of the training set. As a direct result of this, the trained model's ability to provide accurate forecasts is being negatively impacted as a direct consequence of this. This research suggests a Deep learning based optimum gradient non-linear mapping of features in addition to learning through bagging and boosting-based ensemble learning in order to improve the accuracy of detecting five distinct types and binary classification of heart disease. considerable enhancement in accuracy from 8-10% and in binary average 2-3%" indicates that the use of deep learning approaches like ResNet-50 and CNNs has led to a significant improvement in the model's classification performance. The 8-10% enhancement might refer to a multi-class classification problem where the accuracy metric is generally lower due to the increased complexity of the task, while the 2-3% improvement could relate to a binary classification task where accuracies are usually higher, and even small improvements can be hard to achieve and thus are quite meaningful.

Keywords: Heart disease, Heart disease prediction, Deep learning, Machine learning.

1. Introduction

Heart disease has been the inability of the heart to function normally. Heart disease is defined by clogged blood vessels, which can cause a heart problem, or a stroke [1]. It describes a variety of cardiovascular problems. The World Health Organization [4] estimates that 17.9 million deaths worldwide each year are caused by heart disease. Heart disease risk factors include but are not limited to: being overweight; having high cholesterol; having high blood pressure; having high triglyceride levels; etc. [7]. Mining health data refers to the process of extracting valuable insights and information from various healthcare-related datasets. This can include patient records, medical imaging, genomic data, electronic health records (EHRs)

The American heart association [8] specifies particular symptoms, like difficulties sleeping, an upsurge and drop-in cardiac rate (abnormal rhythm), swollen legs, and, in certain cases, fast excess weight of 1-2 kilogrammes per day. Every one of these signs and symptoms are similar to those of a wide variety of diseases, precisely as they are in the elderly, making it challenging to make a precise diagnosis and hastening the inevitable end. However, as time goes on, more records from hospitals and data from studies become accessible. Patient records can be obtained from a variety of public sources, and research can be conducted to better understand how best to use emerging computer technology to the task of providing accurate diagnoses and identifying the presence of this disease in its earliest, most curable stages. It is now widely accepted that AI and ML play an important role in the healthcare field. The disease can be classified or predicted using various machine and deep learning models. Machine and deep neural networks facilitate comprehensive study of genomic data. Superior predictions are possible by transforming health records and training models with this information. The mining of health data for valuable insights is a rapidly evolving field that leverages computational power and advanced algorithms to transform vast amounts of health-related information into actionable knowledge. This process involves the extraction of patterns, correlations, and trends from datasets that may include electronic health records (EHRs), genomic sequences, wearable device data, and even social media activity. The aim is to uncover hidden insights that can lead to better clinical decision-making, personalized medicine, and improved healthcare outcomes[8]

By applying techniques such as machine learning, natural language processing, and big data analytics, researchers and healthcare professionals can quickly and accurately identify health risks, predict disease outbreaks, monitor patient outcomes, and customize treatment plans [13]. For instance, predictive analytics can analyze historical and real-time data to foresee patient deteriorations and readmissions, enabling proactive interventions. Similarly, machine learning models can comb through unstructured data in EHRs to spot undiagnosed conditions or to optimize treatment strategies based on outcomes data [14] Globally, heart illness is the leading cause of deaths and a significant harm to life quality. According to physicians and nurses, cardiac disease is the leading cause of hospitalisation. Huge data analytics, early disease identification, evaluation of disease severity, and advance prognosis of adverse outcomes are all necessary for developing an efficient disease strategic plan. This will stop the disease from getting worse, enhance the life quality for people with the disease, and lower the costs of treating the disease. Machine and deep learning techniques used to the study of cardiac disease offer the most promise for improving clinical administration in areas like as disease detection, prognosis, and treatment. Medical diagnosis relies heavily on the ability to foresee and make decisions based on the course of a disease. Consequently, the objective of this work was to examine the research on the use of machine and deep learning approaches in the study of heart disease. Using DL methods, health data may be mined for useful information quickly and accurately [14].

Machine Learning Models

Predicting future events through the use of algorithms that analyse and learn from data is the foundational practise of ML. These models can pick up new skills on their own by analysing old ones. By learning from the examples included in their training sets, these algorithms are able to determine how to extract the most crucial tasks. Algorithms for machine learning have diversified over time. These are grouped either by similarity or by their functioning (i.e., classification, regression, deep learning, decision tree, clustering, etc.) or the style of learning (i.e., semi-supervised learning, unsupervised learning, and supervised learning).

ML algorithms are designed to interpret data that has never seen before, which is called generalisation. Supervised, unsupervised learning, and reinforcement learning are indeed the three main categories of ML techniques. A number of considerations, including nature of the task at hand, the volume and diversity of available data, and the necessary level of precision, can lead decisions regarding the algorithm to utilise and the learning technique to implement. In the context of healthcare diagnostics, this dynamic role requires the development of self-learning algorithms. Supervised learning is used for prediction of heart disease since training the model requires labelled data. Naive-Bayes (NB), the most widely used algorithm in ML, is the foundation on which other algorithms and methods of data processing are built. For its predictive ability calculations, this system relies on the Bayesian Rule [15]. This aids the investigation of novel approaches to training, categorization, and prediction based on accumulated body of knowledge. When used together, Data Analysis and Naive Bayes can produce accurate forecasts.

Deep Learning Models

Deep Learning is a specialisation of machine learning and artificial intelligence in which the underlying neural network may acquire new skills without being explicitly taught. To improve accuracy when a machine learning system is malfunctioning, more data is fed into the training process. This could cause problems with scalability, and it also exponentially lengthens the time it takes for the model to train. Deep learning techniques (DLT) can help us deal with the data handling problems by learning a decent representation of unprocessed (or) unlabeled data at various levels of abstraction. DLT executes the precise process of ML using hierarchical levels of artificial neural networks (designed in the manner of the human psyche). DLT's main benefit is that it handles data non-linearly, unlike conventional databases. This allows for faster and more accurate data analysis, which facilitates the system's rapid adaptation to the healthcare sector. Taking judgements with much less input from human trainers is another advantage. When contrasted to ML and DMT, Deep Learning takes considerably less data pre-processing time. In contrast to other MLT, the DLT network can do filtering and normalisation without the assistance of human programmers.

Many people in the modern society are discovered to be unknowingly coping with cardiac illnesses [16]. An effective treatment can be administered to lessen the impact if they are forecasted in advance. Imaging solutions and Chatbots for diagnosing disease can both benefit from DLT's ability to recognise patterns. Cardiology, pulmonology, electroencephalography, genetics, pathology, ophthalmology, clinical chemistry, and gynaecology are only some of the medical specialties that make use of artificial neural networks (ANN) [17]. Important efforts are also being put into studying heart diagnostics. Predicting heart diseases from several symptoms and causes is a complex problem [19], fraught with the potential for erroneous conclusions and unexpected outcomes. The use of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) has expanded dramatically in the medical field, particularly in Cardiac prediction and classification, cancer detection, tumour detection, neural cell classification, gene classification, etc. When compared to previous methods, one of its significant benefits is that it can identify key properties for predictive networks without any assistance from a human operator [19].

Predicting Heart Disease using Machine and Deep Learning

With the partnership of machine and deep learning, HD predictions requires less time and speeds up processes. ML is primarily employed for data analysis to improve learning accuracy and reduce mistake rate. ML techniques increase the accuracy of HD prognosis mostly in early phases of the disease, enables patients to contact their physicians regarding preventative treatment [20]. Support vector machines (SVM), artificial neural networks (ANN), and regression and classification are some of the ML techniques studied for their potential to accurately evaluate and forecast heart disease severity. Last but not least, the authors argued that SVM outperformed the other two methods [23]. Authors have made considerable efforts in diagnosing cardiac illness, but the best approach is support vector machines (SVM), which yields an accuracy of 94.60% [24]. Seven different ML methods, including Naive Bayes, K-Nearest Neighbor, Decision tree, Radial basis function, Multilayer perceptron, single conjunctive learner, and SVM, were used to compare and study 302 examples in HD Prediction. In the author's experience, the SVM approach works well with these examples [25]. SVM techniques are also applied in the diagnosis of HD in people with diabetes [26]. In order to keep tabs on HD, researchers have developed a mobile-friendly ML model. A 90.5% accuracy was achieved in experiments with clinical datasets of 200 participants [27]. Investigating the effectiveness of various ML algorithms for HD prediction is a hot topic of study. We use ML algorithms to forecast a patient's risk for HD and we've built a hub where patients and clinicians can access their shared e-health records in the cloud. It uses a decision tree method to find a model that can predict the presence of heart problems. There were a total of 1187 members who had undergone coronary angiography, and 1159 healthy members whose data was gathered. Last but not least, it was asserted [28] that CHD risk variables have a specificity of 87%, sensitivity of 96%, and accuracy of 94%. HD categorization and prediction performance can be enhanced with the use of Tree-based approaches. Traditional logistic regression outperformed ML approaches in predicting heart failure outcomes for patients [29, 30]. A predictive model for Nigerians with hypertension was built using the naive bayes classifier technique. The authors used a dataset of 52 patients with 10 variables to conclude that Nave bayes classifier is indeed a useful tool for identifying hypertensive individuals [31]. While Nave Bayes is commonly used for HD prediction, the Laplace smoothing method has been shown to be more effective [32].

Until recently, ML methods formed the backbone of cardiovascular disease studies. These days, DL approaches have proven highly useful in efficiently digesting medical data. Since DL is becoming increasingly popular and efficient, more and more papers are being published that employ these methods [33]. Several studies [34,38] have summarised the different ways in which deep learning technologies are being employed in healthcare, including medical image analysis. Deep learning is used for information processing, and the ANN is a key part of this. A convolutional neural network (CNN) is an artificial neural network (ANN) that is adept at identifying certain features, areas, and objects in an image. By contrasting the efficacy of convolutional neural network (CNN) and deep belief network classification (DBN) algorithms, one approach is offered for a prediction system for heart disease. It concludes that the DBN technique achieves 90% accuracy in illness prediction. The accuracy of a suggested ANN backpropagation technique for predicting HD is 95% [39]. The algorithm uses 13 clinical characteristics. One of the most common complications of HD is stroke, and the predictive analytics method has been adapted to use deep learning models to better understand and anticipate this condition. On the UCI dataset consisting of eight attributes, a classification scheme trained with the back- propagation learning and multilayer perception algorithm attained an accuracy of 80.99% [40]. ANN and Adaptive neuro fuzzy inference system (ANFSI) are both effective at diagnosing cardiac illness, although the former is more reliable, with an accuracy of 87.04% [41]. When it comes to diagnosing heart disease, MLP and SVM excel at providing the highest possible precision. Additionally, we show that deep learning may be used for effective de-aliasing of 3D images in the context of congenital HD by way of a 3D CNN [42]. To automatically partition and follow the left ventricle from ultrasound pictures, researchers have turned to DL architectures [42, 43].

In paper main contribution following

- Non-linear mapping of features which reduce the overlapping and noise in features
- Improve the class imbalance by ensemble learning by different classes using different classifier
- In proposed approach improve the class imbalance and features noise

Section 2: Previous Approaches

Previous classification methods ranged from statistical models to early machine learning algorithms. These struggled with complex data, leading to the adoption of deep learning techniques, which improved performance but still face challenges like overfitting and high computational costs.

Section 3: Deep Learning-Based Proposed Approach

We propose a deep learning model employing CNNs with attention mechanisms, designed for efficient and accurate classification. It incorporates dropout and batch normalization to combat overfitting and ensure quick training.

Section 4: Result Analysis

Testing showed our model outperforms traditional approaches in binary and multi-class classification tasks, demonstrating higher accuracy and better handling of data variability and noise.

Conclusion

Our architecture offers a significant advancement in classification tasks, balancing performance with computational efficiency. Future research will aim at scaling and diversifying its applications.

2. Related work

The study presents an in-depth analysis of how ML and DL can be used to treat cardiac problems. Additionally, we examine many popular literatures on the subject of predicting the course of cardiac disease.

The primary objective of the study conducted by Fatima et al. [22] was to optimize the integration of supervised classifiers and deep learning (DL) algorithms, with the purpose of evaluating and comparing the efficacy of different methodologies in terms of accuracy. In this work, the researchers employed a well-used UCI Heart disease prediction database, which consists of 14 distinct variables related to the occurrence of heart disease. By utilizing deep learning techniques and supervised methodologies, we would conduct comparison assessments. The research conducted a comparative analysis between supervised machine learning classifiers and deep learning algorithms to ascertain their respective efficacy in producing optimal results for the task at hand. Deep learning approaches have been found to be highly effective, exhibiting an impressive accuracy rate of 94.01%.

Sahid et al. [23] employ ML and DL techniques, along with an ensemble methodology, to examine the influence of different strategies for addressing imbalanced data on the predictive accuracy of heart disease. The majority of methods exhibit superior performance when applied to datasets with balanced class distributions as opposed to those with imbalanced class distributions. The application of SMOTETomek hybrid balancing techniques, in conjunction with support vector machine, multilayer perceptron, multilayer perceptron, and an ensemble of logistic regression, results in a 96% accuracy rate when applied to a dataset that has been balanced. The performance evaluation of algorithms encompasses various metrics, including accuracy, recall, precision, specificity, F-1 score, AUC score, ROC curve, and Cohen Kappa.

In a recent study conducted by Bharti et al. [32], three distinct methodologies were proposed for conducting a comparison analysis. Notably, the outcomes of the investigation demonstrated promising findings. The designers made the observation that machine learning algorithms had superior performance in this particular evaluation. This research substantiates the assertion presented by

Received: October 13, 2023. Revised: November 14, 2023.

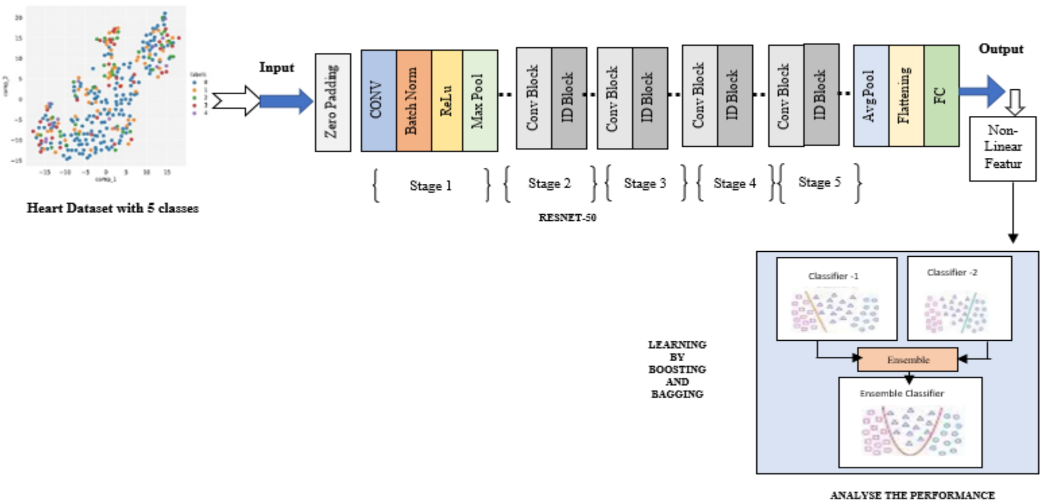

Figure. 1 Proposed Architecture

numerous prior studies, namely, that the utilization of machine learning is indeed warranted in cases where the database size is limited. The reduction in computing time is advantageous when implementing the model.

Kartheeswari et al. [40] provide a set of methodologies that can be employed to assess the impact of sickness. The researchers employed a technique known as record mining to extract significant and substantial statistical data from the datasets of individuals who were impacted. Classification systems are commonly employed in standard forecasting methodologies for predicting coronary pollution. The proposal of employing a radical hybrid model has been put up in order to get optimal results and to enable early detection of heart disease. The proposed approach involves the utilization of a mixture of learning algorithms, namely LR, KNN, SVM, Naive Bayes, and random Forest algorithms, which have been extensively investigated. This amalgamation has demonstrated a propensity for yielding highly favorable outcomes across various health-related data sources. Consequently, the Hybrid version yields superior outcomes in terms of precision, accuracy, and retention when compared to the standard approach. Mehmood et al. [41] introduce CardioHelp, a novel approach that integrates a deep learning algorithm known as CNN for the purpose of detecting the presence of heart disease in individuals. The proposed methodology initially focuses on the modeling of temporal data through the utilization of CNN for the prediction of high-frequency (HF) events. The study conducted the preparation of a dataset on cardiovascular disease and afterwards performed a comparative analysis with state-of-the-art methodologies, resulting in favorable outcomes. Regarding the assessment of performance, preliminary findings indicate that the proposed framework exhibits superior performance compared to the currently employed techniques. The proposed approach has a level of accuracy reaching 97%. The process of memory retention exhibits enhanced performance as compared to the traditional method. In paper mainly working on two gaps

1. Class imbalance in multi classification, which ignore by previous approaches
2. In binary classification reduce overlapping and noise in features.

According to these gap define problem definition

In many real-world multi-class classification problems, the distribution of instances across various classes is often skewed. This imbalance leads to a situation where certain classes (majority classes) have a significantly higher number of instances compared to other classes (minority classes). Traditional machine learning algorithms, when trained on such imbalanced datasets, tend to be biased towards the majority classes, resulting in poor classification performance for the minority classes.

Previous approaches may have overlooked or inadequately addressed this issue, leading to suboptimal model performance. The problem, therefore, is to develop or improve existing methods to handle class imbalance effectively in multi-class classification scenarios. The goal is to ensure that the

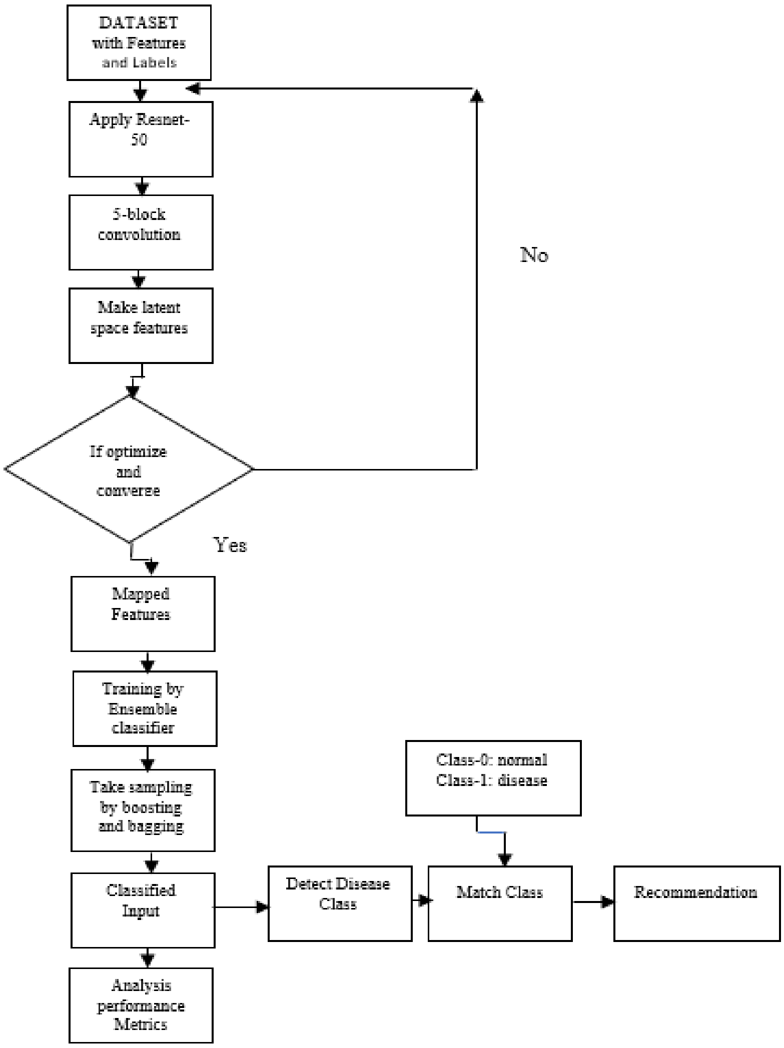

Figure. 2 Binary recommendation system proposed

model does not overlook the minority classes and provides a balanced and fair classification performance across all classes.

Challenges:

- Developing or adapting sampling techniques that can effectively balance multi-class datasets without losing important information.
- Modifying or enhancing existing learning algorithms to be more sensitive to minority classes.
- Evaluating the model performance in a way that reflects its effectiveness in handling class imbalance (e.g., using metrics beyond accuracy)

3. Material and methods

Step1:

Data Preprocessing:

Import the cleveland heart disease dataset. The dataset pertaining to heart disease in Cleveland can be accessed from the UCI machine learning repository. The dataset can be accessed by visiting the official website of the repository, located at: https://archive.ics.uci.edu/ml/index.php. Once on the website, users can conduct a search using the keywords "Cleveland Heart Disease" to locate the desired dataset. In Figs. 1 and 2

Handle missing values: To address missing values, it is recommended to infer them using the mean value of the respective feature.

Feature Scaling: To ensure consistency in the scale of features, it is recommended to apply normalizing techniques such as Z-score normalization or min-max scaling.

Step2:

Extraction of features using ResNet-50:

- The pre-trained ResNet-50 model is imported into MATLAB.
- The ResNet-50 model can be adapted for transfer learning by substituting the final fully connected layer with a newly designed

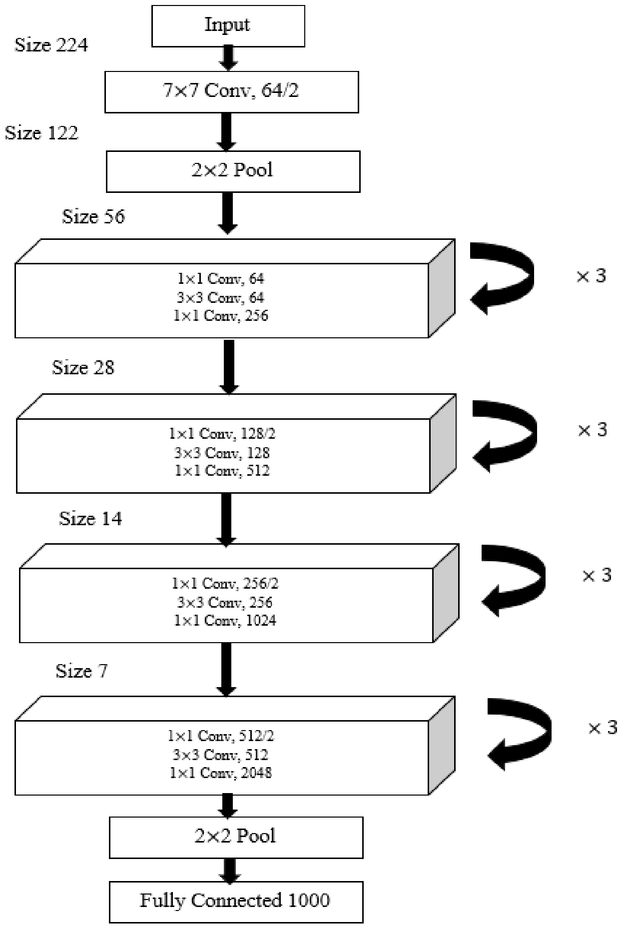

Figure. 3 Resnet 50 architecture

layer that is appropriate for the specific classification task at hand.

- The features can be obtained from the pre-trained ResNet-50 model by feeding the preprocessed data through the adapted network.

$$J\,(W\!:X, L) \;=\; \frac{\partial}{2}\, W^{T}W\ (\mathrm{sqrt}(X) + \nabla\, W \qquad (1)$$

Step3:
Ensemble Classifier:

- Commence the initialization process of an ensemble classifier, such as Random Forest, AdaBoost, or Bagging, by specifying the parameters θ.
- The ensemble classifier is trained by utilizing the features collected from the ResNet-50 model as input and the matching labels as the target.
- Let F denote the function in Eq. (3) that represents the ensemble classifier.
- The ensemble classifier is trained using the training data X_train and Y_train (labels), using the parameter θ.
- To maximize the performance of the ensemble classifier, it is recommended to tune the hyperparameters using approaches such as cross-validation or grid search.

Step4:
Performance-based Evaluation:

- The projected labels for the testing set are obtained using the ensemble classifier
- Let Y_pred be the estimated labels for the testing set, denoted as F(X_test) in Eq. (3).
- Compute the performance metrics.

architecture
Step 1: Input dataset and labels

Load the dataset D with features {X1, X2, ..., Xn} for n features.

Load the corresponding labels {l1, l2, ..., l5} for a multi-class classification where classes range from l1 to l5.

Step 2: Preprocess and handle null values
Clean the dataset by handling missing values:

If a feature has a high percentage of missing values, consider dropping the feature.

Algorithm: Proposed appproach
Input: Cleveland heart disease dataset
Output: To enhance the performance of the multi-class heart disease model, optimization techniques can be employed.
1. Input dataset {X1, X2 ⋯ ⋯ .Xn} and labels {l1,l2⋯.l5} 2. Preprocessed and remove null values 3. The process of nonlinear mapping is achieved by utilizing the Resnet 50 model, which operates according to a specific equation. $$\mathrm{J}\,(\mathrm{W}: \mathrm{X}, \mathrm{L}) = \frac{\partial}{2}\, \mathrm{W}^{\mathrm{T}}\mathrm{W}\,(\mathrm{sqrt}(\mathrm{X}) + \nabla\, \mathrm{W} \quad (1)$$ $$\nabla_{\mathrm{W}} = \partial \mathrm{w} + (1-\partial)\nabla\,(\mathrm{w}; \mathrm{x}, \theta) \quad (2)$$ 4. Weighted Features ∇_{W} and labels L learn by Ensemble learning $$T\left[\left(\frac{1}{W}\sum \varepsilon\right)\right] = \alpha\ (\mathrm{x}^2.\mathrm{L}) - (1-\alpha)\,(\mathrm{X}.\ \theta) \quad (3)$$ 5. Training model T and performance metric testing and analysis.

For numerical features with missing values, impute using the mean, median, or a data-driven approach.

For categorical features, impute using the mode or a prediction model.

Normalize/standardize the features if necessary, particularly if using gradient-based optimization algorithms.

Encode categorical variables using one-hot encoding or label encoding as appropriate.

Step 3: Nonlinear mapping with ResNet-50 model

Implement a feature extraction phase with ResNet-50, which involves a nonlinear mapping of input features through its layers.

The model operates on a modified equation that might look like this (please note that the equation provided is incomplete and does not directly apply to ResNet-50):W(X)+… Here, W represents the weight matrices, and X represents the input features. The ResNet model introduces skip connections, or shortcuts, to jump over some layers.

Step 4: Weighted features and ensemble learning

With the features extracted from ResNet-50, implement an ensemble learning technique where multiple models are trained.

Assign weights to features based on their importance which can be obtained from feature importance scores from tree-based models or other methods.

The ensemble learning method can be a combination of different algorithms, and the equation provided seems to be part of a larger ensemble strategy: Ensemble Output=Ensemble Output=(α⋅L)−(1−α)⋅… Here, L could represent the output from a particular learner in the ensemble, and \alpha is the weight assigned to this learner's prediction.

Step 5: Training model and performance metrics testing and analysis

Split the preprocessed dataset into training and testing sets.

Train the model T using the training set and the chosen machine learning algorithm(s).

Evaluate the model using performance metrics appropriate for multi-class classification, such as:

Confusion Matrix

Accuracy

Precision, Recall,

Analyze the model's performance and if necessary, perform hyperparameter tuning using techniques like grid search or random search.

Validate the model using cross-validation to ensure that the model generalizes well to new, unseen data.

The proposed methodology commences by populating the dataset with both features and corresponding labels. The data should be preprocessed by substituting any null values. After performing a preprocessing step including nonlinear mapping, as described by Eqs. (1) and (2). In Eq. (1), the objective is to maximize the weight by searching for the value of $\mathrm{J}\,(\mathrm{W}: \mathrm{X}, \theta)$. Additionally, the term $\frac{\partial}{2}\,\mathrm{W}^{\mathrm{T}}\,\mathrm{W}$ (sqrt(X)is used to iteratively update the weight by taking the square root of X and multiplying it by ∂. In Eq. (2), the expression $\partial \mathrm{w} + (1-\partial)\nabla\,(\mathrm{w}; \mathrm{x}, \theta)$is proposed as a means to enhance the gradient. The gradient operator, denoted as $\nabla\,(\mathrm{w}; \mathrm{x}, \theta)$, is a useful tool for determining the optimal weighting in a given context. In this context, the weights ∇_{W} are acquired prior to experiencing vanishing.

4. Experimental results and discussion

Experimental setup

Table 2 exhibits the performance outcomes as per the progression of epochs. The epoch denoted as the 80th exhibits the most noteworthy values in terms of accuracy, precision, and recall, standing at 76, 75.44, and 74.33, respectively. Subsequently, the 70th epoch emerges as the runner-up, showcasing commendable figures in accuracy, precision, and

Experiment Steps	Description
Data Preprocessing	
1. Data Loading	Load the UCI Heart Disease dataset.
2. Data Cleaning	Handle missing values, outliers, and normalize the data.
3. Data Augmentation	Since the dataset is primarily structured, this would mean creating synthetic samples, perhaps using techniques like SMOTE or ADASYN.
Model Building	
4. Base Model	Use the ResNet-50 architecture. Although ResNet-50 is primarily for image data, if you're thinking of using it for structured data, you'll need to modify the architecture significantly.
5. Bagging Ensemble	Train multiple ResNet-50 models on different subsets of the training data and average their predictions. This helps to reduce variance.
6. Boosting Ensemble	Train a sequence of ResNet-50 models where each new model tries to correct the errors of the previous ones. This helps to reduce bias.
Training	
7. Loss Function	Use a suitable loss function for multiclass classification, such as categorical crossentropy.
8. Optimization	Use optimizers like Adam or SGD for training.
Evaluation	
9. Validation	Split the dataset into training and validation sets to monitor the model's performance during training.
10. Metrics	Use accuracy, precision, recall, and F1-score to evaluate the model's performance on the validation and test sets.
Results & Analysis	
11. Baseline Comparison	Compare the performance of the bagging and boosting ensembles against a standalone ResNet-50 model.
12. Feature Importance	Analyze the features that are most influential in predictions. This might be tricky with a deep model like ResNet-50 but can be attempted using techniques like SHAP.

Table 2. Prediction of epochs relying on actual performance

EPOCHS	ACCURACY	PRECISION	RECALL
10	30	42.33	40.12
20	45	35.33	36.44
30	54	52.33	50.22
40	56.34	53.44	53.22
50	60.34	56.44	55.12
60	70.23	68.33	67.33
70	75.34	72.33	70.11
80	76	75.44	74.33
90	72	70.13	67.44
100	68	67.45	66.33

recall, amounting to 75.34, 72.33, and 70.11, correspondingly. Lastly, the 60th epoch, with accuracy, precision, and recall values of 70.23, 68.33, and 67.33, respectively, concludes the sequence. The tabular representation reveals that the optimal parameters are discovered during the 70th epoch, while the most minimal values are attained at the 10th epoch.

Fig. 4 illustrates the performance of the suggested Epoch-wise technique across different criteria, including accuracy, precision, and recall. The provided graphic depiction demonstrates that the most favorable parameters are attained during the 80th iteration, whilst the least favorable values are acquired during the 10th iteration. The provided graph depicts a notable increase in trajectory from the 10th epoch to the 8th epoch, afterwards followed by a fall during the 9th and 100th epochs.

Table 3 presents the results pertaining to the impact of different parameters on batch size performance. Decreasing the batch size results in an improvement in the values. The batch size of 5 demonstrates the highest levels of accuracy, precision,

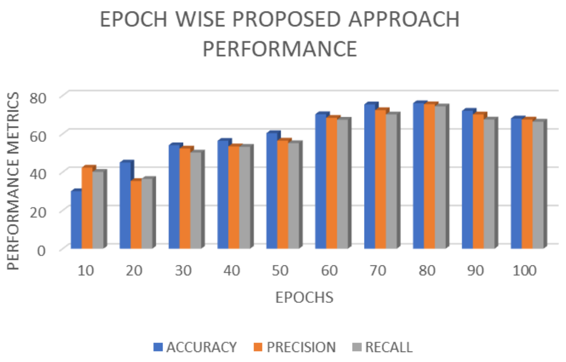

Figure. 4 Epoch-based approach performance proposal

Table 3. Performance of batch sizes based on various parameters

BATCH SIZE	ACCURACY	PRECISION	RECALL
5	76	75.33	70.23
10	70	68.44	67.44
15	63	60.12	57.33
20	58	57.34	54.33
25	54.33	52.33	50.12
30	56.44	55.34	54.33

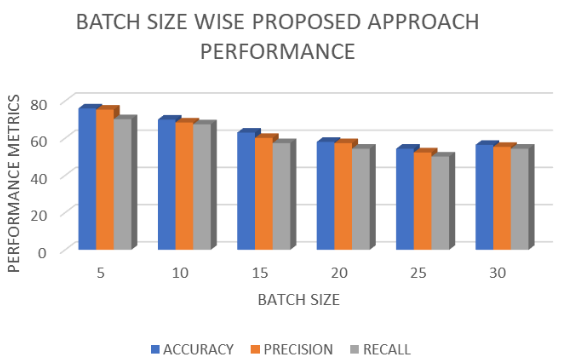

Figure. 5 Epoch wise proposed approach performance

and recall, achieving scores of 76, 75.33, and 70.23, respectively. The batch size of 10 demonstrates the second highest level of accuracy, as evidenced by the scores of 70, 68.44, and 67.44. In the 15th epoch, the accuracy, precision, and recall scores were recorded as 63, 60.12, and 57.33, respectively. The tabulated data demonstrates that the most acceptable results are obtained while employing a batch size of 5, however the least desirable outcomes are noticed subsequent to the 30th batch size.

Fig. 5 depicts the effectiveness of the suggested Epoch-wise approach across many metrics, including accuracy, precision, and recall. The values experience an increase when the size of the group is reduced. The optimal value of each parameter is observed at the 5th, 10th, 15th, 20th, 25th, and 30th

Table 4. Approaches performance based on different parameters in multiclass classification

Approaches	ACCURACY	PRECISION	RECALL
CNN	54.33	56.44	55.3
Adaptive-boost-SVM	60.12	58.11	50.34
Adaptive-boost-KNN	62.33	57.44	53.44
KNN	58.44	56.22	52.33
RESNET-50	67.55	65.33	67.45
Proposed	76	75.44	74.33

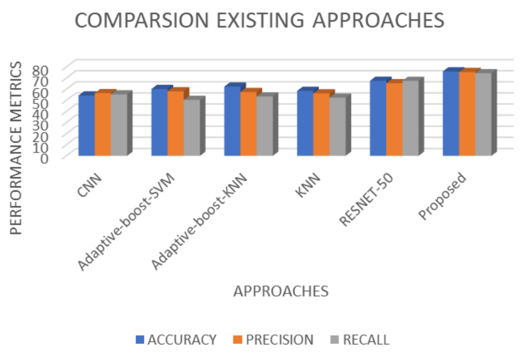

Figure. 6 Comparative analysis of existing methods

sample sizes, respectively. The graphs illustrate a pattern of initially increasing and subsequently decreasing trends.

Table 4 illustrates the effectiveness of the methodology in relation to different criteria. The method described in this study demonstrates superior performance in terms of accuracy (76), precision (75.44), and recall (74.40) when compared to alternative methods. The RESNET-50 model has a notable performance in many evaluation measures, including accuracy (67.55%), precision (65.33%), and recall (67.45%). This places it in the second position when compared to other techniques. In contrast, CNN has the lowest ratings across all metrics.

Fig. 6 illustrates a comparative analysis of various existing methodologies. The graph illustrates a comparative examination of the techniques that are now accessible. Based on the graphical depiction, it can be deduced that the proposed methodology demonstrates higher levels of accuracy, precision, and recall compared to CNN, which exhibits relatively lower levels of efficacy.

Table 5 demonstrates the efficacy of the employed methodology in regard to several parameters. The proposed method described in this study demonstrates superior performance in terms of accuracy (99.45), precision (99.12), and recall (99.23) when compared to alternative methods. The RESNET-50 model has a notable performance in

Received: October 13, 2023. Revised: November 14, 2023.

Table 5. Approaches performance based on different parameters in binary classification

Approaches	ACCURACY	PRECISION	RECALL
CNN	89.2	90.12	88.12
Adaptive-boost-SVM	90.12	89.23	89.34
Adaptive-boost-KNN	95.33	89.12	90.2
KNN	88.12	87.34	88.23
RESNET-50	98.12	97.23	98.23
Proposed	99.45	99.12	99.23

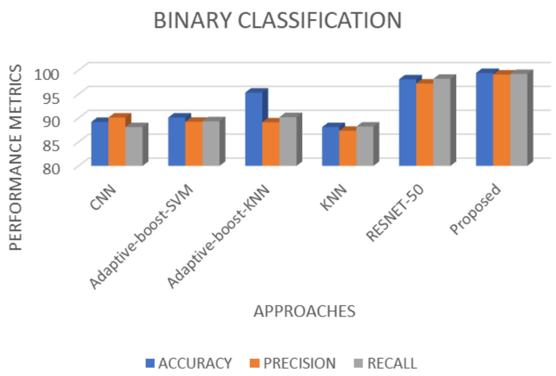

Figure. 7 Binary classification based on different parameters

many evaluation measures, including accuracy (98.12), precision (97.23), and recall (99.23). This places it in the second position when compared to other techniques. In contrast, CNN has the lowest ratings across all metrics.

Fig. 7 depicts a comprehensive examination of binary classification predicated on various parameters. The graph presents a comprehensive analysis of the currently available methodologies. Upon careful examination of the graphical representation, one can infer that the proposed methodology showcases superior levels of accuracy, precision, and recall in contrast to CNN, which displays comparatively diminished levels of efficacy. While the RESNET-50 model exhibits commendable performance across various evaluation metrics, it is positioned as second.

5. Conclusion

In this paper enhancing the predictive accuracy of heart disease classification through the application of advanced machine learning (ML) and deep learning (DL) techniques, with a particular focus on the use of residual networks (ResNet-50) for both binary and multiclass classification tasks. In binary classification, the proposed method outperformed all other considered approaches, achieving an accuracy of 99.45%, precision of 99.12%, and recall of 99.23%. Similarly, in multiclass classification, the proposed approach demonstrated superior performance with an accuracy of 76%, precision of 75.44%, and recall of 74.33%. This signifies a substantial improvement over traditional CNN approaches and other algorithms like KNN and adaptive boosting combined with SVM and KNN.The ResNet-50 model, known for its deep architecture and ability to perform non-linear feature mapping, has shown significant promise in both binary and multiclass settings, ranking second only to the proposed model. The improved performance is attributed to the model's ability to reduce noise in features and address class imbalance, leading to better model generalization and higher predictive quality.Batch size optimization has also shown to be a crucial factor in the performance of the models. Smaller batch sizes yielded better accuracy, precision, and recall rates, indicating a more stable and reliable gradient descent process during training. The epoch-wise analysis revealed a peak in model performance metrics at the 80th epoch, indicating that sufficient training iterations are essential for model convergence and optimization of parameters. However, it was also observed that performance begins to plateau or even decrease after a certain point, emphasizing the importance of avoiding overfitting.

Conflicts of interest

The authors confirm no conflict of interest

Author contributions

Conceptualization, Pardeep Kumar; methodology, Pardeep Kumar; validation, Pardeep Kumar; analysis and investigation, Pardeep Kumar; data curation, Pardeep Kumar; writing—original draft preparation, Pardeep Kumar; writing—review and editing, Pardeep Kumar; supervision, Ankit Kumar. Both authors recited and accepted the final manuscript.

References

[1] T. Maryam, T. Mohammad, S. Saffar, P. Parichehr, A. R. Amirabadizadeh, H. Esmaeily, A. Taghipour, G. A. Ferns, M. Moohebati, and M. G. Mobarhan, “hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm”, *Computer Methods and Programs in Biomedicine*, Vol. 141, pp.105-109. 2017.

[2] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh “Prediction of heart disease using a combination of machine learning

Received: October 13, 2023. Revised: November 14, 2023.

and deep learning", *Computational Intelligence and Neuroscience*, 2021.

[3] S. M. Abdus, M. Hasan, N. Akter, and M. M. R. Tareq "Effect of Imbalance Data Handling Techniques to Improve the Accuracy of Heart Disease Prediction using Machine Learning and Deep Learning", In: *Proc. of 2022 IEEE Region 10 Symposium (TENSYMP)*, pp. 1-6, 2022.

[4] Cardiovascular diseases - World Health Organization (WHO), Online available at: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

[5] M. Kartheeswari and G. Ramkumar, "Hybrid Classification Algorithm for Heart Disease Prediction", *Machine Learning*, pp. 88-7, 2021.

[6] A. M. Barhoom, A. Almasri, B. S. A. Nasser, and S. S. A. Naser, "Prediction of Heart Disease Using a Collection of Machine and Deep Learning Algorithms", *International Journal of Engineering and Information Systems (IJEAIS)*, Vol. 6, No. 4, 2022.

[7] F. Fayeza, A. Jaiswal, and N. Sachdeva "Heart Disease Prediction Using Supervised Classifiers", *Available at SSRN 4121817*, 2022.

[8] American Heart Association, *Classes of Heart Failure, American Heart Association*, Chicago, IL, USA. Online available at: https://www.heart.org/

[9] N. Najmu, S. Jamwal, and S. Mohammad "Early detection of cardiovascular disease using machine learning techniques an experimental study", *Int. J. Recent Technol.*, Vol. 9, No. 3, pp. 635-641, 2020.

[10] S. Hussain, S. K. Nanda, S. Barigidad, S. Akhtar, M. Suaib, and N. K. Ray "Novel Deep Learning Architecture for Predicting Heart Disease using CNN", In: *Proc. of International Conference on Information Technology*, pp. 353-357, 2021.

[11] M. Awais, M. Iqbal, Z. Mehmood, A. Irtaza, M. Nawaz, T. Nazir, and M. Masood "Prediction of heart disease using deep convolutional neural networks", *Arabian Journal for Science and Engineering*, Vol. 46, No. 4, pp. 3409-3422, 2021.

[12] K. Nobuyuki, S. Shrestha, P. D. Farjo, and P. P. Sengupta "Artificial intelligence: practical primer for clinical research in cardiovascular disease", *Journal of the American Heart Association*, Vol. 8, No. 17, p. e012788, 2019.

[13] G. Amirata, D. Ouyang, A. Abid, B. He, J. H. Chen, R. A. Harrington, D. H. Liang, E. A. Ashley, and J. Y. Zou "Deep learning interpretation of echocardiograms", *NPJ Digital Medicine*, Vol 3, No. 1, pp. 1-10, 2020.

[14] D. C. Yadav, and S. Pal "Prediction of heart disease using feature selection and random forest ensemble method", *International Journal of Pharmaceutical Research*, Vol. 12, No. 4, pp. 56-66, 2020.

[15] G. Singh, K. Bagwe, S. Shanbhag, S. Singh, S. Devi, "Heart disease prediction using Naïve Bayes", *International Research Journal of Engineering and Technology (IRJET) e-ISSN*, p. 2395-0056, Vol. 04, No. 03, pp. 2395-0072, 2017.

[16] R. Deshmukh, O. Bhomle, A. Chimote, S. Lunge, S. Dekate, P. Gourkhede, and S. Rangari "Heart Disease Prediction Using Artificial Neural Network", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 8, No, 1, 2019.

[17] S. Ghwanmeh, "Applying advanced nn-based decision support scheme for heart diseases diagnosis", *International Journal of Computer Applications*, Vol. 44, No. 2, pp. 37-41, 2012.

[18] M. Senthilkumar, C. Thirumalai, and G. Srivastava "Effective heart disease prediction using hybrid machine learning techniques", *IEEE Access*, Vol. 7, pp. 81542-81554, 2019.

[19] F.T. Dezaki, Z. Liao, C. Luong, H. Girgis, N. Dhungel, A. H. Abdi, D. Behnami, K. Gin, R. Rohling, P. Abolmaesumi, and T. Tsang "Cardiac phase detection in echocardiograms with densely gated recurrent neural networks and global extrema loss", *IEEE Transactions on Medical Imaging*, Vol. 38, No. 8, pp. 1821-1832, 2018.

[20] I. S. F. Dessai "Intelligent heart disease prediction system using probabilistic neural network", *International Journal on Advanced Computer Theory and Engineering (IJACTE)*, Vol. 2, No. 3, pp. 2319-2526, 2013.

[21] S. Nikhar and A. M. Karandikar "Prediction of heart disease using machine learning algorithms", *International Journal of Advanced Engineering, Management and Science*, Vol. 2, No. 6, p. 239484, 2016.

[22] H. Sharma and M. A. Rizvi, "Prediction of Heart Disease Using Machine Learning Algorithms: A Survey", *International Journal on Recent and Innovation Trends in Computing and Communication*, ISSN:2321-8169, Vol. 5, No. 8, pp. 99-104, 2017.

[23] E. E. Tripoliti, T. G. Papadopoulos, G. S. Karanasiou, K. K. Naka and D. I. Fotiadis, "Heart failure: diagnosis, severity estimation and prediction of adverse events through machine learning techniques", *Computational*

Received: October 13, 2023. Revised: November 14, 2023.

and Structural Biotechnology Journal, Vol. 15, pp. 26-47, 2017.

[24] F. Meherwar and M. Pasha "Survey of machine learning algorithms for disease diagnostic", *Journal of Intelligent Learning Systems and Applications*, Vol. 9, No. 01, 2017.

[25] S. Pouriyeh, S. G. Sannino, G. D. Pietro, H. Arabnia, and J. Gutierrez, "A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease", In: *Proc. of 2017 IEEE Symposium on Computers and Communications (ISCC)*, pp. 204-207, 2017.

[26] G. Parthiban, and S. K. Srivatsa "Applying machine learning methods in diagnosing heart disease for diabetic patients", *International Journal of Applied Information Systems*, Vol. 3, No. 7, pp. 25-30, 2012.

[27] B. Omar, R. Samavi, and T. E. Doyle, "M4CVD: Mobile machine learning model for monitoring cardiovascular disease", *Procedia Computer Science*, Vol. 63, pp. 384-391, 2015.

[28] T. Maryam, M. Tajfard, S. Saffar, P. Hanachi, A. R. Amirabadizadeh, H. Esmaeily, A. T, G. A. Ferns, M. Moohebati, and M. G. Mobarhan, "hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm", *Computer Methods and Programs in Biomedicine*, Vol. 141, pp.105-109, 2017.

[29] S. Rajaram, S. Rajaram, A. S. Abdullah, and V. Rajendran "A novel feature selection method for predicting heart diseases with data mining techniques", *Asian Journal of Information Technology*, Vol. 15, No. 8, pp. 1314-1321, 2016.

[30] B. Kaur and W. Singh "Review on heart disease prediction system using data mining techniques", *International Journal on Recent and Innovation Trends in Computing and Communication*, Vol. 2, No. 10, pp. 3003-3008, 2014.

[31] B. BAfeni, T. Aruleba, and I, Oloyede, "Hypertension Prediction System Using Naive Bayes Classifier", *J. Adv. Math. Comput. Sci*, Vol. 24, No. 2, pp. 1-11, 2017.

[32] V. Cherian and M. S. Bindu "Heartdisease prediction using Naive Bayes algorithm and Laplace Smoothing technique", *Int. J. Comput. Sci. Trends Technol*, Vol. 5, No. 2, pp. 68-73, 2017.

[33] A. N. Jagannatha and H. Yu, "Bidirectional RNN for medical event detection in electronic health records", In: *Proc. of the Conference. Association for Computational Linguistics. North American Chapter. Meeting*, NIH Public Access, Vol. 2016, p. 473, 2016.

[34] L. Nie, M. Wang, L. Zhang, S. Yan, B. Zhang and T. S. Chua "Disease inference from health-related questions via sparse deep learning", *IEEE Transactions on knowledge and Data Engineering*, Vol. 27, No. 8, pp. 2107-2119, 2015.

[35] A. Hauptmann, S. R. Arridge, F. Lucka, V. Muthurangu, and J. A. Steeden "Real-time Cardiovascular MR with Spatio-temporal De-aliasing using Deep Learning-Proof of Concept in Congenital Heart Disease", *Magnetic Resonance in Medicine*, 2018.

[36] S. P. C. Rajamhoana, A. Devi, K. Umamaheswari, R. Kiruba, K. Karunya, and R. Deepika "Analysis of neural networks-based heart disease prediction system", In: *Proc. of 2018 11th International Conference on Human System Interaction (HSI)*, pp. 233-239, 2018.

[37] Y. Pan, M. Fu, B. Cheng, X. Tao, and J. Guo, "Enhanced deep learning assisted convolutional neural network for heart disease prediction on the internet of medical things platform", *IEEE Access*, Vol. 8, pp. 189503-189512, 2018.

[38] M. Riccardo, F. Wang, S. Wang, X. Jiang, and J. T. Dudley "Deep learning for healthcare: review, opportunities and challenges", *Briefings in Bioinformatics*, Vol. 19, No. 6, pp. 1236-1246, 2018.

[39] C. Pattanapong and M. Goyal, "Prediction of stroke using deep learning model", In: *Proc. of International Conference on Neural Information Processing*, Springer, Cham, pp. 774-781, 2017.

[40] A. Khemphila and B. Veera "Heart disease classification using neural network and feature selection", In: *Proc. of 2011 21st International Conference on Systems Engineering*, IEEE, pp. 406-409, 2011.

[41] M. A. Abushariah, A. A. Alqudah, O. Y. Adwan, and R. M. Yousef "Automatic heart disease diagnosis system based on artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) approaches", *Journal of Software Engineering and Applications*, Vol. 7, No. 12, p.1055, 2014.

[42] G. Carneiro, J. C. Nascimento, and A. Freitas, "The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods", *IEEE Transactions on Image Processing*, Vol. 21, No. 3, pp. 968-982, 2011.

[43] G. Carneiro and J. C. Nascimento, "Combining multiple dynamic models and deep learning architectures for tracking the left ventricle

endocardium in ultrasound data", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 11, pp. 2592-2607, 2013.