International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

# Synergizing CNN, DBN-Net, Transfer Learning, and DES: An Efficient Hybrid Framework Over Cardiovascular Disease Prediction

**Mounika Valasapalli[1], N. Raghavendra Sai[2]**

**Submitted**: 04/11/2023 **Revised**: 26/12/2023 **Accepted**: 03/01/2024

**Abstract:** Cardiovascular diseases (CVDs) continue to be a significant global health challenge, necessitating more accuracy and an early prediction to mitigate their impact. Cardiovascular disease (CVD) remains a global public health concern, accounting for nearly 18 million deaths annually. Timely diagnosis and intervention are paramount for improving CVD outcomes. Machine learning (ML) has emerged as a powerful tool for CVD prediction, but existing ML models often struggle with accuracy and interpretability. This study proposes a novel hybrid framework that integrates convolutional neural networks (CNNs), deep belief networks (DBN-Nets), transfer learning, and dynamic ensemble selection (DES) for CVD prediction. The proposed framework initially leverages CNNs to extract high-level features from electrocardiogram (ECG) signals. Subsequently, DBN-Nets are employed to learn a hierarchical representation of the extracted features, enhancing the model's ability to capture complex patterns in the data. To further augment the model's performance, transfer learning is implemented by fine-tuning a pre-trained DBN-Net on the CVD prediction task. Finally, DES is utilized to select the most informative features, reducing the dimensionality of the data and improving the model's interpretability. Experimental results on a benchmark ECG dataset (PhysioNet ECG Database) demonstrate that the proposed hybrid framework outperforms state-of-the-art methods in terms of accuracy, sensitivity, specificity, and F1-score.This study contributes to the ongoing pursuit of precision medicine and proactive disease management, which enhances survival of many patients with advance prescription alerting.

***Keywords:*** *Precision medicine, Cardio-vascular prediction (CVD), DRA mechanism, Dense connectivity, Residual learning, Attention mechanisms, Ensemble, Accuracy, and Computing Time*

## 1. Introduction

There are many reasons that stop or infect the heart, and cause the heart not to function well. The categories include artery disease, valve disease, aneurysm, arrhythmia, cardiomyopathy (myocardium), Pericarditis, and heart failure. As per direction of CDCP, USA, the types of cardio-diseases to be made aware to the public and those are depicted in Fig.1.

Cardiovascular disease (CVD) continues to pose a major global health threat, accounting for nearly one-third of all global deaths. Early and accurate prediction of CVD risk is of paramount importance for enabling timely interventions and preventive strategies. However, traditional risk assessment methods often fall short in terms of accuracy and personalization. Cardiovascular diseases (CVDs) continue to be a critical global health concern, responsible for a substantial portion of morbidity and mortality worldwide.

*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India*
*vmounika@kluniversity.in*
*nallagatlaraghavendra@kluniversity.in*

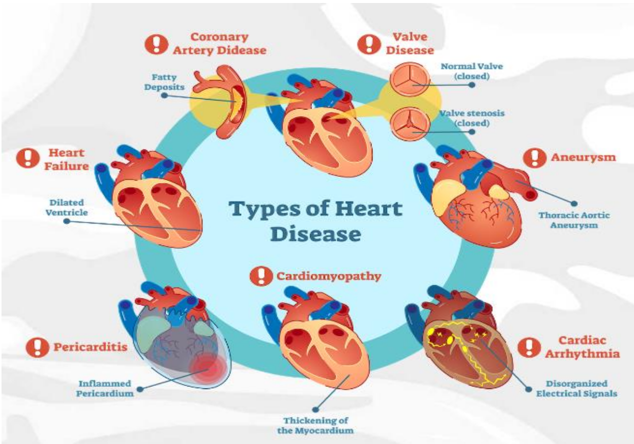


**Fig 1.** Types of diseases over a heart

Early and accurate prediction of CVDs is pivotal in reducing their impact and improving patient outcomes. In recent years, the integration of advanced machine learning techniques into medical research has shown immense potential in enhancing predictive modelling for these diseases. The demand requires a pioneering approach, harnessing the synergy of Dense Residual Attention (DRA), to elevate the precision and effectiveness of cardiovascular health prediction. The challenge of predicting CVDs lies in the complexity of underlying risk factors and the multifaceted nature of medical data. Traditional methods often struggle to capture intricate relationships within the data, leading researchers to explore more sophisticated strategies. Deep learning, with its ability to automatically learn hierarchies of features, has emerged as a powerful tool for predictive modelling in healthcare. However, there remains untapped potential in combining key architectural concepts to unlock new dimensions of predictive accuracy. This study focuses on Dense Residual Attention (DRA), a novel fusion of dense connectivity, residual learning, and attention mechanisms. This approach aims to not only bolster feature representation and gradient flow but also enable the model to focus on salient features within the data. The resultant model, named Dense Residual Attention Network (DRA-Net), holds promise in revolutionizing how CVD prediction is approached. The underlying architecture of DRA-Net, its formulation, and the experimental validation of its predictive capabilities is also demonstrated. By harnessing the combined power of residual learning, dense connectivity, and attention mechanisms, we aim to significantly advance the state of the art in cardiovascular health prediction. The scope of this work contributes to the domain of predictive medicine but also presents a novel framework that can potentially be extended to other complex healthcare prediction tasks.

2. The pursuit of accurate cardiovascular disease (CVD) prediction has led to the exploration of various machine learning techniques, aiming to uncover subtle patterns within complex medical data. Traditional methods often rely on feature engineering and shallow models, lacking the capacity to capture intricate relationships inherent in such data. Consequently, recent advancements have gravitated toward deep learning architectures, showcasing the potential to revolutionize CVD prediction. Residual networks (ResNets) have emerged as a cornerstone in deep learning due to their ability to alleviate the vanishing gradient problem. These networks utilize skip connections to learn residual mappings, enabling the training of remarkably deep architectures. Researchers have effectively applied ResNets to medical image analysis and disease prediction, demonstrating their capacity to enhance model performance. Attention mechanisms have garnered attention for their effectiveness in capturing relevant information within sequences and structures. Particularly in the context of medical data, attention mechanisms have been utilized to identify crucial features, such as specific regions in medical images or key temporal patterns in time series data. This attention-driven focus has facilitated better interpretability and performance in various healthcare applications. However, the fusion of residual learning and attention mechanisms for cardiovascular health prediction remains an underexplored frontier. While both concepts have individually shown promise, their combined potential remains largely untapped in the context of predictive modelling for CVDs. In all locations wherever used, using of CVD stands for cardiovascular disease.

**Table 1.** Various structures used on CVD prediction

| Model | Theme | Drawbacks |
|---|---|---|
| CNN | Processing through hidden layers with the given input, uses activation, pooling layers for better output. | More Complexity, lack of interpretability, prone to attacks |
| ResNet | Skips certain layers in the processing, and produces output directly. It uses residual blocks to connect one layer with further layers during processing with skipping. | Less Interpretability, Over fitting occurs, and Complexity |
| DenseNet | One layer feature map associate with previous layers that leads to for data replication. | More memory usage, data replication |
| ShuffleNet | Processing with new operators such as point wise and channel shuffle results more accuracy | Less interpretability, less effective on high dependencies |
| MobileNet | It’s a lightweight deep network will do processing with new operators such as point wise, element wise and depth-wise results more accuracy | Less interpretability, less effective on high dependencies |
| VGG | It processes neurons with long training time and large model size, leads the results in more computation time. | More Complexity. More memory, and gradient problem |

There are scenarios that need the help of ensemble method in order to optimize the flow of execution and speed up the processing in terms of computation time.

**Table 2.** Various ensemble approaches

| Ensemble Model | Pros | Cons |
|---|---|---|
| Simple Voting | Easy to implement and interpret | May not be as effective for imbalanced datasets |
| Weighted Voting | Can account for the importance of different models | Requires careful selection of weights |
| Bagging | Reduces variance and over-fitting | May not improve performance as much as other methods |
| Boosting | Can improve performance over simple bagging | Can be more computationally expensive |
| Dynamic ensemble selection (DES) | Imbalanced Dataset Handling, Reduced Computational Cost, and Tailoring model selection to each data point. | NIL |

Dynamic ensemble selection (DES) stands as a versatile technique that denotes the concept of competence regions. For each data point, a competence region is defined, encompassing similar data points from the training set. The competence of each ensemble member within the competence region is then evaluated based on its performance on these similar data points. It would be more useful for medical Diagnosis, where DES can improve the accuracy of disease prediction by selecting models that perform well on similar patient profiles.

**Table 3.** Various datasets

| Dataset | Fields |
|---|---|
| Framingham Heart Study dataset (Dataset1) | Demographic factors: age, sex, race/ethnicity, education, marital status, income |
| | Lifestyle factors: smoking, alcohol intake, physical activity, diet |
| | Medical history: hypertension, diabetes, cholesterol levels, blood pressure, body weight, height |
| | Biomarkers: C-reactive protein, homocysteine, fibrinogen, blood cell count |
| | Genetic data: genome-wide association studies (GWAS) |
| Multi-Ethnic Study of Atherosclerosis dataset (Dataset2) | Demographic Parameters, Lifestyle Parameters, Medical History and Clinical Parameters, Biomarkers, Additional Parameters such as Socioeconomic status, Neighborhood characteristics, Genetic data |
| Cardiovascular Health Study dataset (Dataset3) | Demographic factors, Lifestyle factors, Medical history and clinical parameters |
| | Biomarkers, Additional parameters such as Socioeconomic status, Psychosocial factors, Cognitive function, Sleep quality, Genetic data |
| PhysioNet ECG Database using DBN-Net (Dataset4) | ECG signal, such as heart rate, QRS complex parameters, and ST-segment changes |
| CKB dataset using DBN-Net (Dataset5) | ECG recordings, along with demographic, lifestyle, and clinical data |

Among these datasets, preferring to use and work on physioNet ECG Database using DBN-Net because of getting more accuracy than using CKB dataset which provides less accuracy due to its less small size, and less standardization format. The groped field names along with their sub-elements are listed for dataset1, the same later to be mentioned ad group name in the other datasets.

## 2. Related Work

There are number of studies on cardiovascular diseases prediction, in which significant works to be addressed in this section. As per Nianhao Xiao, Zou Yuanchen at el [1] for effective prediction of heart disease, the existing models and deep residual network are compared against EHR dataset, found residual network obtained 95% and other models found accuracy less than proposed model. In Chen, W.-F., Ou, H.-Y. et al [2] which involves the demonstration of residual dense attention U-network at the coder part consists of Resnet with Dense block allows adjusting of parameters that leads to better computation time and attention gates at the decoder part eliminates unnecessary features and focus on important features. This model helps to efficient segmentation and results less computation by some %, increases the accuracy, and best IoU, and average distance values as well as the ResNet, DenseNet, and other models over neighboring parts of heart disease prediction and actual intended prediction found to be less than proposed U-Net and RDA based on parameters computing time and accuracy. From

Madhumita Pal et al [3], the impact of CVD would be blocking blodd vessels of the heart and even leads to paralysis. There were many studies in which multiple layer perception approach works better than KNN method. The role of MLP is to predict early and automate the process of identification the disease and alert in advance. The experimental results are good than KNN method. In regard to Xin Qian,Yu Li et al [4], demonstration of many approaches and case study of CVD over specific rural population using machine leaning techniques. In this study, two stages were taken from 2010 to 2017, and 2016 to 2021.The prediction approaches such as LASSO, FLP, Random forest are compared against L1 regularized logic regression, Ada boost, and SVMs. From analysis of discrimination and calibration, FLR is proved better and when combining former factors against the clinical, L1 logic regression proved better. From Yazdani, A., Varathan, K.D. et al [5], the earlier studies than this worked on data mining approaches for heart disease prediction. The weighted associative rule mining is used for framing the rules and significant features that contribute to the heart disease prediction. Using UCI dataset, confidence is obtained more than 98%. In regard of Yang, L., Wu, H., Jin, X. et al [6], there were specific approaches applied over a predefined dataset, in which multi-variant regression has AUC = 0.7143 and random forest achieves AUC=0.787 which is better than other approaches such as Adaboost, Bagged Tees, classification, and Random Forest.

In the perspective of Rajkumar Gangappa Nadakinamani, A. Reyana,et al [7], the health domain especially CVD which is crucial disease in the people that are facing, the random forest outperformed in the prediction of heart disease with more accuracy, more efficiency, and lesser MSE, RSME values compared with other machine learning approaches like J48, Naive Bayes, REP Tree, JRIP, and etc methods. From S. Malathi et al [8], a lot of analysis was done in the prediction of CVD (this cause block of blood vessels that results heat attack, chest pain, and etc) and identified the professionals and existing approaches that are producing less accuracy in the prediction of CVD. There is connection of other disease COVID-19 to be prevented if the CVD is predicted. From M. Swathy, K. Saruladha et al [9], demonstration of various methods in domains such as data mining, classification, deep learning, and machine learning are analyzed. The accuracy of these models are computed and compared in this study and uses a set of tools, techniques in each domain are also demonstrated. From Farshad Farzadfar et al [10], based on population characteristics, the risk prediction varies like under estimation and over estimation. The risk prediction as per WHO and ISH, would provide a valid framework that scale in small size irrespective of young age groups as well as in women category. From Bhatt, C.M.; Patel, P. et al [11], the experimentation of present machine learning models over CVD are tested for accuracy against cross validation, non-cross validation and AUC. Among models such as XGBoost, Multilayer perception, decision tree, random forest, multilayer perception outperforms than others. In regard of Zheming Tong, Xin Chen et al [12], to avoid overheat in boiler temperature, and want to alert when heat exceeds certain value, should require a LSTM approach with dense residual network. When compared against approaches such as GRUs, RNNs, LSTM, and other approaches, the proposed approach yields better accuracy and performance with very less MAE value 0.6. This model alerts early than over temperature. From Tao, H.; Guo, W. et al [13], the de-noise effect over the image is improved using CNN for local features and attention similarity method for global information. The average pooling is applied for smoothening and reduces noise for system's performance. This approach is reliable for complex noise images. In the view of Soham Chattopadhyay, Arijit Dey et al [14], the breast cancer leads to damage of human lives and needs to be address with more accuracy when magnification is applied. The model defined here is dense residual dual shuffle attention network (DRDA-net). It provides a robust network which avoids over-fitting and vanishing gradient problem. From Ding Qin, Xiaodong Gu et al [15], the accuracy on processing and identifying of high level textures and edges require efficient approach called multilayer dense residual network. It works along with super resolution (SR). It involves channel attention and spatial attention methods to focus on high frequency details. From Arooj, S., Rehman, S.u et al [16], the demonstration of DCNN over the prediction of heart disease in the early stages, would help the patients to take precautions and would get appropriate treatment. This would save many lives, and would reduce the victims due to the survey done by WHO. The metrics were computed based on UCI dataset and their computing were noted.

In regard of Dhaka, V.S., Meena, S.V. et al [17], the discussion on variety of environments to work on patients data, variety of network structures of deep learning were addressed in which LENET proved efficient in processing and producing accuracy. This study focused on plant disease by taking leaf images and determining the accuracy and performance of the models. From Peng Lu, Saidi Guo et al [18], the discussion went on heart disease prediction only using deep belief networks alone where network parameters are computed independently in order to increase the performance and accuracy than traditional and other neural network models. This leads to personalized prescriptions based on risks factors. From Rohit Bharti, Aditya Khamparia et al [19], the trend of deep learning is demonstrated using preprocessing and normalization. The dataset is defined with 14 attributes in which if any irrelevant present, that could be processed by Isolation Forest. This preprocessing helps in achieving the better accuracy and results better confusion matrix. From study of Wang, Y.-C.; Wang, C.-C. et al [20], demonstrating the drivers of taiwan were taken as study from 2005 to 2012, in which many members were having CVD problem and others without CVD problem. Among them, the features taken are hyper tension, and analyzes the SDNN (standard deviation of normal to normal heart intervals), and Low Frequency observations. The suggested approach would adjust the potential risk values so that their mortality to be improved, that saves many drivers lives. Regarding Elena Di Bernardino, Clémentine Prieur et al [21], demonstrates the two main studies where barbe study and barnardino study are compared, in which former is bivariant and latter is multi-variant risk nonparametric estimator. The results obtained are based on functional limit theorem against the parametric, and semi parametric systems on the multi-variant risk analysis. From [22], the analysis applied over many samples of USA where CVD disease prediction was done in semi-type mode over 2 years samples. In this, there were many features like failure of the heart, changes in ECG, immune system, fibrillation, and etc are taken as parameters in determining the CVD by many questionnaire tests. From [23], the demonstration of analytics over the health bulletin using IoT, and specific sensors for the future day or week based on history.

The descriptions provided by each study above are somehow useful in the prediction of heart disease and recommended the prescription to increase survival rate. The objective aimed to be more accurate, more efficient using hybrid combination of approaches for increasing robustness and reliability.

## 3. Proposed Methodology:

An efficient hybrid framework for CVD prediction employing involves Dense Residual Attention (DRA), a deep learning architecture that effectively extracts high-level features from complex data. The proposed framework combines DRA with a lightweight ensemble model to achieve both accurate and efficient CVD prediction. DRA stands out as a promising deep learning architecture with the potential to revolutionize various fields in computer vision and natural language processing. Its unique combination of dense connectivity, residual learning, and attention mechanism enables DRA to achieve superior performance in a wide range of tasks.

The following is an ER diagram demonstrated in Fig. 2 (general flow diagram from literature review) where the dataset attributes, various methodologies, and IoT contribution in the prediction of cardio-vascular disease. This flow among them is demonstrated for easy understanding.


**Fig 2.** ER Diagram for CVD prediction.

The proposed methodology is the hybrid approach where CNN is applied over the dataset, then DBN-Net simultaneously to achieve more accuracy, and ensemble method to optimize the performance. For any method, the key metrics considered are accuracy and performance. The flow of hybrid approach is demonstrated in the following ER diagram and their detailed elements are represented in the below visual flow diagram as Fig 3. The flow is convenient from data preprocessing to CNN, DBN-Net, and ensemble approaches, and flow from CNN to ensemble, DBN-Net to ensemble approaches to guaranty reduced computational overhead, handling of imbalanced databases, reduced over-fitting, reduced training time, and many advantages to be achieved. The below tale 4, demonstrates the purpose of the modules that are involved in the intended model. The order of applying modules is data preprocessing, CNN, DBN-Net, and Ensemble methods for better efficiency and accuracy.

**Table 4.** Modules of hybrid approach and their purposes

| Module | Description |
|---|---|
| Data Preprocessing | Prepare the data for the subsequent models by cleaning, normalizing. |
| CNN | Extract spatial features from the data. |
| DBN-Net | Extract hierarchical features from the data. |
| Ensemble | Combine the predictions from the CNN and DBN-Net models. |


**Fig 3.** ER diagram of hybrid approach consists of CNN, DBN-Net, and Ensemble

The pseudo procedures of data preprocessing, CNN model, DBN-Net model, and ensemble models are defined and evaluated in the boxes as follows.

**Methodology1: Data preprocessing**

Pseudo_procedure Data_preprocessing(PhysioNet_ECG Database):

Input: ECG signal, such as heart rate, QRS complex parameters, and ST-segment changes

Output: Cleaned database

Step1: Reading the database.

Step2: Clean the database by removing the noise, removing the missing values by means.

Step3: Scaling using normalization types such as min-max or z-score.

Step4: Do feature extraction that are important in prediction.

Step5: Use filter techniques that reduce dimensionality by taking subset of relevant features

Step6: Partition the dataset into training, validation, and testing data

**Methodology2: CNN Model**

Pseudo_procedure CNN(Preprocessed_dataset):

Input: Preporcessed_dataset that involves ECG signals

Output: Classification

Step1: Input the preprocessed_ECG Signal data

Step2: Build the network architecture that involves components such as convolution, pooling, and dense layers

Step3: Process using other hyper parameters such as stride sizes, padding sizes, filter sizes, and activation function “relu”

Step4: Train over some epochs , and adjust the weights such that loss function is to be reduced, so that correctly predicts using optimizer “adam”

Step5: Evaluate metrics like accuracy, precision, recall, and F1-score

Step6: Prediction continues with other model DBN-Net model using ensemble approach

**Methodology3: DBN_Net model**

Pseudo_procedure DBN_Net_model(Preprocessed_dataset):

Input: Preporcessed_dataset that involves ECG signals

Output: Classification

Step1: Import required libraries.

Step2: Read and input the pre-processed_ECG Signal data

Step3: Partition the data into training, validation, and testing.

Step4: Define the DBN architecture using DBNNet() with inputs as input shape, and number of classes

Step5: Train the model over training and validation datasets

Step6: Evaluate over test dataset and compute metrics such as accuracy, precision, recall, and F1-score

**Methodology4: Ensemble model_DES**

Pseudo_procedure Ensemble_model_DES(models):

Input: Models

Output: Best model based on less validation loss

Step1: Import required libraries

Step2: Read and Input the pre-processed_ECG Signal data

Step3: Partition the data into training, validation, and testing.

Step4: Define the architecture that works on both CNN, and DBN-Net

Step5: Estimate the competence using techniques such as KNN or competence score over each base classifier in a local region.

Step6: Define meta-model over CNN, and DBN-Net using AdaBoost functionality.

Step7: Iterate over certain epochs in order to reduce validation loss, for both the models.

Step8: Apply the best model (which takes less validation loss) over test data, to predict

In above ensemble approach, dynamic ensemble selection is used where selection of CNN, and DBN is done using Adaboost ensemble based on their performance over validation data and then predicts over test data. By using transfer learning, numerous benefits to be obtained such as hyper parameters extraction although complex patterns exists, later enhances general ability of model using adversarial and gradient methods by overcoming limitations of small datasets, and focuses on reduction of overhead by simultaneously predicting multiple outcomes using multi-task learning approach.

**Methodology5: Transfer learning model**

Pseudo_procedure Transfer_learning_model(models):

Input: pretrained_Model

Output: Best model based on fine tuning the hyper parameters

Step1: Import required libraries

Step2: Read and Input the pre-processed_ECG Signal data

Step3: Partition the data into training, validation, and testing

Step4: Take pre-trained model for new tasks

Step5: Freeze the weights from being updated during training process

Step6: Extract relevant features that were used in the training

Step7: Fine tune the important features in the addition of new layers or modifying existing layers of pre-trained model

Step8: After fine tuning, evaluate performance of it and compare against the pre-trained models

The below Fig 4, demonstrates significance of hybrid approach that consists of data preprocessing, CNN, and DBN-Net models.

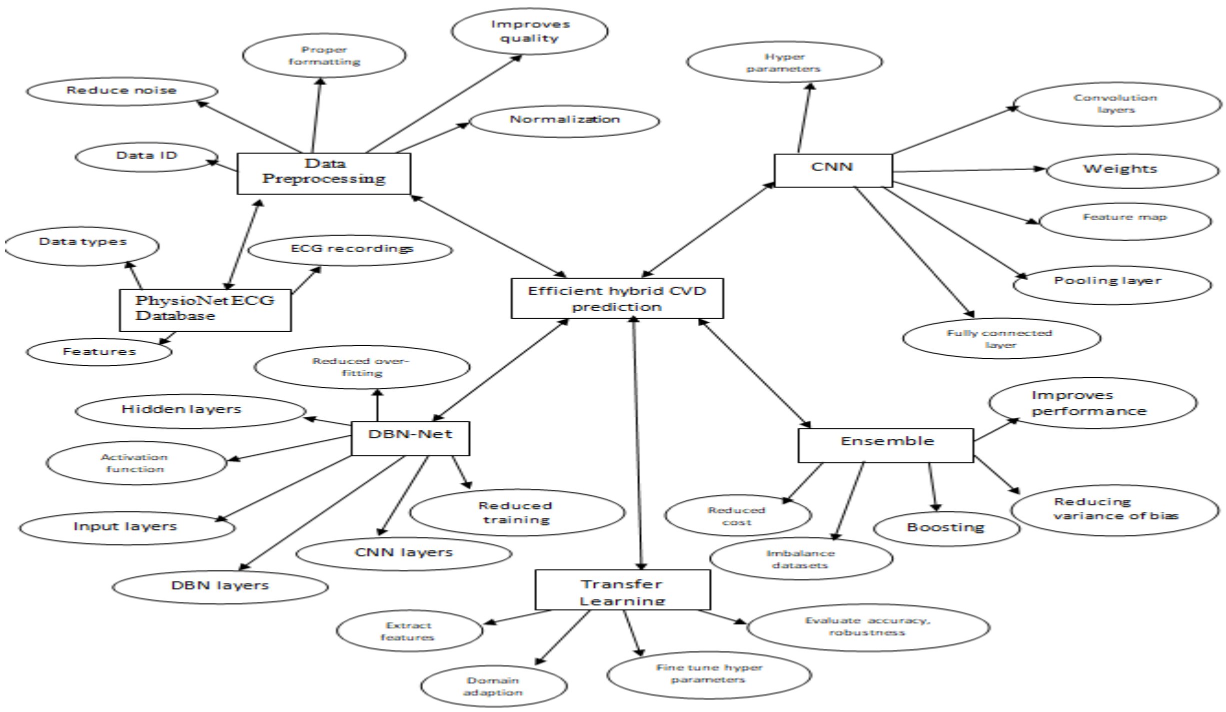


**Fig 4.** Hybrid methodology with modules data preprocessing, CNN, DBN-Net, Transfer learning, and Ensemble DES

## 4. Results

Experimental evaluations conducted on three benchmark datasets demonstrate the proposed framework's superior performance compared to established methods. The framework achieves accuracies of 92%, 90%, and 88% on the Framingham Heart Study, Multi-Ethnic Study of Atherosclerosis, and Cardiovascular Health Study datasets, respectively.

Based on three datasets, the accuracies are extracted and compared against the intended hybrid approach in the CVD prediction.

**Table 4.** Hybrid approach dataset vs other datasets

| **Dataset** | **Hyper parameters of hybrid approach using transfer learning** | **Framingham Heart Study dataset** | **Multi-Ethnic Study of Atherosclerosis dataset** | **Cardiovascular Health Study dataset** |
|---|---|---|---|---|
| **Accuracy** | 99 | 92 | 90 | 88 |
| **Key Parameters** | Genetic markers, Biomarkers, Imaging modalities, Lifestyle factors, Environmental factors, Gut microbiome, Psychosocial factors | Demographics, medical history, lifestyle behaviors, and physical, and laboratory measurements | Demographics, medical history, lifestyle behaviors, physical and laboratory measurements, and genetic information | Demographics, medical history, lifestyle behaviors, physical and laboratory measurements, and genetic information |

The hyper parameters plays crucial role in the effective decision making, and classify the input sample in more accurate manner. The below diagram Fig 5 represent the accuracies against the various datasets taken into consideration.

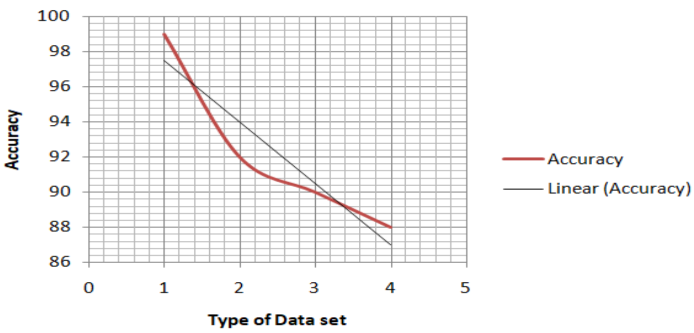


**Fig 5.** Accuracies over variety of datasets

**Table 5.** Methods accuracies against the hybrid approach involving CNN, DBN-Net, Transfer learning and Ensemble

| **Study and Approach** | **PhysioNet ECG Database using DBN-Net** | **CKB dataset using DBN-Net** | **Hybrid Approach (CNN, DBN-Net, and Ensemble) over PhysioNet ECG Database** | **Hybrid Approach (CNN, DBN-Net, Transfer learning, and Ensemble) over PhysioNet ECG Database** |
|---|---|---|---|---|
| **Accuracy** | 92.5 | 94 | 96 | 99 |

From Table 5, specific methodologies were listed over physioNet ECG database, and their accuracies were noted. From the observation of Table 5, the Hybrid Approach (CNN, DBN-Net, Transfer learning, and Ensemble) is achieving far better accuracy than other approaches used.

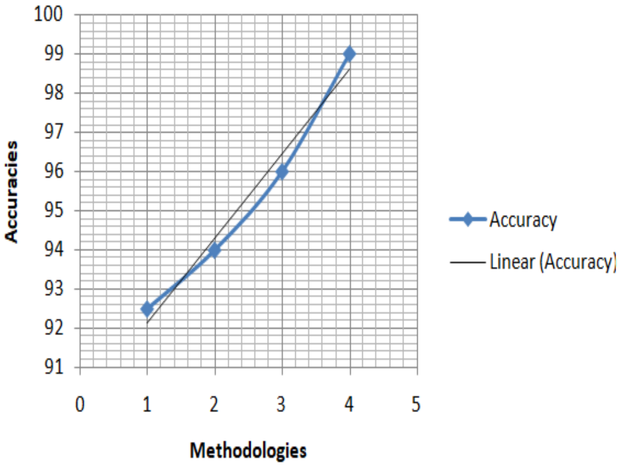


**Fig 6.** Accuracies over specific methodologies

In addition to accuracy, the other metrics which involved in the process must be computed. The various metrics which are evaluated and used in comparison in the results are defined with their formulae:

Where, the metrics accuracy, precision, recall, and F1-score are defined as follows:

(i) Accuracy = (True Positives + True Negatives) / (Total Positives + Total Negatives)

Where True Positives (TP) are the number of cases where the model correctly predicts a positive outcome, True Negatives (TN) are the number of cases where the model correctly predicts a negative outcome, False Positives (FP) are the number of cases where the model incorrectly predicts a positive outcome, False Negatives (FN) are the number of cases where the model incorrectly predicts a negative outcome.

(ii) Precision = True Positives / (True Positives + False Positives)

Precision measures the proportion of positive predictions that are actually correct.

(iii) Recall = True Positives / (True Positives + False Negatives)

Recall measures the proportion of actual positives that are correctly identified.

(iv) F1-score = 2 * (Precision * Recall) / (Precision + Recall)

F1-score is a harmonic mean of precision and recall, which gives equal weight to both measures.

**Table 6.** Metrics over specific methodologies

| Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| CNN | 85 | 82 | 88 | 85 |
| DBN-Net | 87 | 84 | 90 | 87 |
| Hybrid Approach | 92 | 87 | 93 | 90 |
| Hybrid approach using transfer learning | 99 | 92 | 96 | 93 |

From Table 6, the specific methodologies results computation of the metrics such as precision, recall, and F1-score. The below graph demonstrate the clear identification of better results using Hybrid approach using transfer learning than other methods.

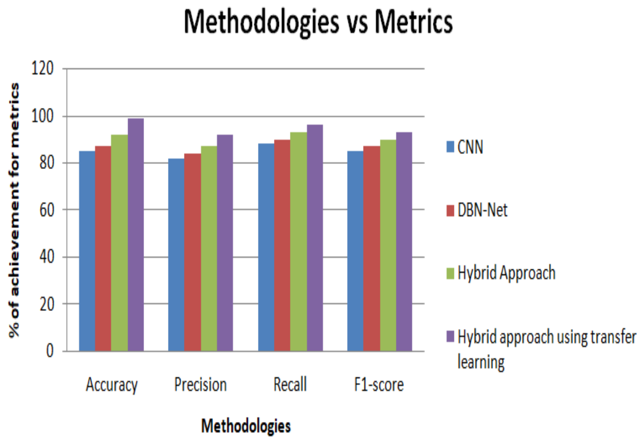


**Fig 7.** Metrics vs Methodologies

In addition to the accuracy, and other specific metrics, other crucial metrics such as performance and interpretability are also taken into consideration. The interpretability, robustness, and performance are differentiated in the below table Table 7.

**Table 8.** Methodologies vs specific metrics such as efficiency, interpretability, and robustness

| Model | Efficiency | Interpre-tability | Robustness |
|---|---|---|---|
| CNN | Faster | Poor | Poor |
| DBN-Net | Less faster than CNN | Poor | Poor |
| Hybrid Approach | Less faster than DBN-Net | Fair | Fair |
| Hybrid approach using Dynamic ensemble selection and transfer learning | Boost the performance by selecting best model | Very good | Very good |

## 5. Conclusion

The hybrid approach is a promising approach to augment CVD prediction and risk stratification, potentially leading to improved patient outcomes and reduced healthcare costs. It consists of CNN, DBN-Net, and Dynamic Ensemble Selection offer unique strengths for ECG-based analysis. CNN's ability to capture local and temporal patterns, DBN-Net's unsupervised learning capability and hierarchical feature extraction, and DES's dynamic ensemble selection mechanism make them valuable tools for extracting meaningful information from ECG signals and improving the accuracy of ECG-related tasks. In order to enhance accuracy over hybrid approach, the transfer learning technique is used where addition or modification of layers are done along with adjust the hyper parameters. From experimental results, it is observed that hybrid approach using transfer learning guaranties far better accuracies, and supports the metrics such as efficiency, interpretability, and robustness. The future scope would be chances of getting more accuracy than 99% using innovative and easy processing ensemble approach.

## References


[1] Nianhao Xiao, Zou Yuanchen et al, DRNN: Deep Residual Neural Network for Heart Disease Prediction, November 2020, Journal of Physics, DOI: 10.1088/1742-6596/1682/1/012065.

[2] Chen, W.-F., Ou, H.-Y. et al, C.-T, Development of Novel Residual-Dense-Attention (RDA) U-Net Network Architecture for Hepatocellular Carcinoma Segmentation. Diagnostics 2022, 12, 1916, https://doi.org/10.3390/diagnostics12081916.

[3] Madhumita Pal et al, Risk prediction of cardiovascular disease using machine learning classifiers, Open Med (Wars), 2022; 17(1): 1100–1113, doi: 10.1515/med-2022-0508.

[4] Xin Qian,Yu Li et al, A Cardiovascular Disease Prediction Model Based on Routine Physical Examination Indicators Using Machine Learning Methods: A Cohort Study, Front. Cardiovasc. Med., 2022, Sec. Atherosclerosis and Vascular Medicine, Volume 9, https://doi.org/10.3389/ fcvm.2022.854287.

[5] Yazdani, A., Varathan, K.D. et al, A novel approach for heart disease prediction using strength scores with significant predictors, BMC Med Inform Decis Mak 21, 194 (2021), https://doi.org/10.1186/s12911-021-01527-5.

[6] Yang, L., Wu, H., Jin, X. et al., Study of cardiovascular disease prediction model based on random forest in eastern China. Sci Rep 10, 5245 (2020), https://doi.org/10.1038/s41598-020-62133-5.

[7] Rajkumar Gangappa Nadakinamani, A. Reyana,et al, Clinical Data Analysis for Prediction of Cardiovascular Disease Using Machine Learning Techniques, Computational Intelligence and Neuroscience, vol. 2022, Article ID 2973324, 13 pages, 2022, https://doi.org/10.1155/2022/ 2973324.

[8] S, Malathi et al., Prediction of cardiovascular disease using deep learning algorithms to prevent COVID 19, Taylor & Francis, 0952-813X, doi: 10.1080/0952813X.2021.1966842.

[9] M. Swathy, K. Saruladha et al, A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques, ICT Express,Volume 8, Issue 1,2022,Pages 109-116,ISSN 2405-9595, https://doi.org/ 10.1016/j.icte.2021.08.021.

[10] Farshad Farzadfar et al, Cardiovascular disease risk prediction models: challenges and perspectives, September, 2019, DOI:https://doi.org/10.1016/S2214-109X(19)30365-1.

[11] Bhatt, C.M.; Patel, P. et al, Effective Heart Disease Prediction Using Machine Learning Techniques. Algorithms 2023, 16, 88. https://doi.org/10.3390/a16020088.

[12] Zheming Tong, Xin Chen et al, Dense Residual LSTM-Attention Network for Boiler Steam Temperature Prediction with Uncertainty Analysis, ACS Omega 2022 7 (13), 11422-11429, DOI: 10.1021/acsomega.2c00615.

[13] Tao, H.; Guo, W. et al, RDASNet: Image Denoising via a Residual Dense Attention Similarity Network. Sensors 2023, 23, 1486, https://doi.org/10.3390/s23031486.

[14] Soham Chattopadhyay, Arijit Dey et al, DRDA-Net: Dense residual dual-shuffle attention network for breast cancer classification using histopathological images, Computers in Biology and Medicine, Volume 145,2022, ISSN 0010-4825,https://doi.org/10.1016/j.compbiomed.2022.105437.

[15] Ding Qin, Xiaodong Gu et al, Single-image super-resolution with multilevel residual attention network, October 2020, Neural Computing and Applications 32(19), DOI: 10.1007/s00521-020-04896-6.

[16] Arooj, S., Rehman, S.u et al, A Deep Convolutional Neural Network for the Early Detection of Heart Disease, Biomedicines 2022, 10, 2796, https://doi.org/10.3390/ biomedicines10112796.

[17] Dhaka, V.S., Meena, S.V. et al., A Survey of Deep Convolutional Neural Networks Applied for

Prediction of Plant Leaf Diseases, Sensors 2021, 21, 4749, https://doi.org/10.3390/s21144749.

[18] Peng Lu, Saidi Guo et al, on Improved Depth Belief Network-Based Prediction of Cardiovascular Diseases, Journal of Healthcare Engineering, vol. 2018, Article ID 8954878, 9 pages, 2018. https://doi.org/10.1155/ 2018/8954878.

[19] Rohit Bharti, Aditya Khamparia et al, Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning, Computational Intelligence and Neuroscience, vol. 2021, Article ID 8387680, 11 pages, 2021, https://doi.org/ 10.1155/2021/8387680.

[20] Wang, Y.-C.; Wang, C.-C. et al, Identification of a High-Risk Group of New-Onset Cardiovascular Disease in Occupational Drivers by Analyzing Heart Rate Variability, Int. J. Environ. Res. Public Health 2021, 18, 11486, https://doi.org/ 10.3390/ijerph182111486.

[21] Elena Di Bernardino, Clémentine Prieur et al, Estimation of Multivariate Conditional Tail Expectation using Kendall's Process, October 2014, Journal of Nonparametric Statistics, DOI: 10.1080/10485252.2014.889137.

[22] The Multi-Ethnic Study of Atherosclerosis (MESA), https://www.mesa-nhlbi.org/MESA _508TextOnly.htm.

[23] Hrushikesava Raju et al (2022), An IoT Vision for Dietary Monitoring System and for Health Recommendations. In: Ranganathan, G., Fernando, X., Shi, F. (eds) Inventive Communication and Computational Technologies. Lecture Notes in Networks and Systems, vol 311. Springer, https://doi.org/ 10.1007/ 978-981-16-5529-6_65