Smart Healthcare System for Cardiovascular Diseases Diagnosis

Ahmed Tolba, Ahmed Elmasry, Yara Mohamed ,and Amany Salah
Zagazig, Sharqia, 20912018100221@fci.zu.edu.eg, ahmedalnasr22@gmail.com,

ym4299170@gmail.com,amanyazam08@gmail.com

Abstract **—The World Health Organization (WHO) reported that cardiovascular diseases account for 40% of total deaths in Egypt and 32% of all deaths worldwide . Cardiovascular disease is a major public health concern with significant social and economic implications in terms of healthcare needs, lost productivity, and premature death. According to the Egyptian Vision 2030, healthcare is one of the main pillars for achieving sustainable development. So we decided to develop an intelligent CAD system to improve the diagnosis of cardiovascular diseases in the Egyptian healthcare system. This project presents an innovative edge intelligence system for the automated diagnosis of cardiovascular disease. The system emphasis delivering real-time and low-cost processing of cardiovascular biomarkers to accelerate the initial diagnosis. Then, an elegant service is also provided to afford automated, interpretable, and accurate diagnosis from cardiac biomarkers of critical cases Taking into account the preservation of patient data and protection of the system from hacking. Our system is designed to satisfy the pillars of healthcare sustainability according to the Egyptian Vision 2030. The simulation results obtained from two public benchmarks validate the effectiveness of these approaches in terms of accuracy (99%, 97%) and f1-scores (99%, 98%).**

Keywords—Deep learning (DL) , Federated Learning (FL) , Explainable Ai , Internet of Health things(IOHT) , Cyber-Attacks

I. Introduction

Cardiovascular disease (CVD) is a class of diseases that involve the heart or blood vessels.CVD includes coronary artery diseases (CAD) such as angina and myocardial infarction (commonly known as a heart attack).Other CVDs include stroke, heart failure, hypertensive heart disease, rheumatic heart disease, cardiomyopathy, abnormal heart rhythms, congenital heart disease, valvular heart disease, carditis, aortic aneurysms, peripheral artery disease, thromboembolic disease, and venous thrombosis.

Of many solutions to current problems using Artificial intelligence (Ai). In recent years, AI, especially deep learning, DL has shown promising performance for diagnosing and securing effective cardiac biomarkers in the healthcare system; For example, a convolutional neural network (CNN) is known for efficient feature extraction, long-term memory (LSTM) and gated recurring unit (GRU) for temporal dependency modeling. this demonstrated an exciting potential to detect abnormalities in the heart from particular biomarkers. For example, in clinical medicine, the **electrocardiogram** (ECG) is a critical tool for diagnosing a wide range of arrhythmias. Every year, more than 300 million ECGs are collected around the world, making it a vital tool in the everyday practice of clinical medicine. Besides, **Photoplethysmography** (PPG) is an emerging technology that enables non-invasive heart rhythm measurement through optical sensing. A PPG sensor detects blood volume changes in the microvascular bed of tissue using a low-intensity light. Nowadays, many existing wearable devices on the market have built-in PPG sensors. Moreover, physicians have performed a fundus examination in patients with hypertension to determine the presence and severity of retinal vascular damage as a means to estimate cardiovascular disease risk in addition **Echocardiogram**, or "echo," is a type of ultrasound scan used to check the heart and nearby blood. The existing DL solutions for the diagnosis of cardiovascular diseases from such biomarker data suffer from the black-box nature that limits its applicability for high-risk decisions in real-world healthcare. This in turn makes the doctors unable to trust such a solution. Internet of Health Things (IoHT) provides a great shift toward automated collection and diagnosis of patients' data making the computer-aided diagnosis (CAD) smarter and more efficient.The current AI-driven IoHT systems rely on cloud computing for storing and processing patients' data which implied high communication latency and raised many security and privacy concerns. Another important issue in the current solutions is that they assume a diagnosis of cardiovascular disease from single modality data, which unqualifies them for real-world applications.

Therefore, we decided to develop a computer-aided diagnosis (CAD) system to help doctors reach a better, faster and more accurate diagnosis of cardiovascular diseases with automated predictions and explainable results for the patient's heart condition, which will lead to improved cardiovascular disease diagnosis in Egypt. health care system. In this system, advanced, low-cost devices that capture different biomarkers of heart patients, then provide real-time diagnosis and prediction of the heart based on different biomarkers such as : Electrocardiogram & Echocardiogram **Electrocardiogram**

(ECG): This is a common non-invasive diagnostic signal that doctors use widely to diagnose cardiovascular disease. **Echocardiogram**, or "echo," is a type of ultrasound scan used to check the heart and nearby blood.

The CAD system can classify heart rhythms into **5** different ECG categories with explainable results using Explainable Ai. (**Normal Beat**,

Supraventricular premature beat, Premature ventricular contraction, Fusion of ventricular, Unclassifiable beat)

It's also able to diagnose patients' echo data like:

1).Classify echo.

2).Segment the left ventricle. (Cardiac Echo Image Segmentation technique that is performed routinely for disease diagnosis and treatment planning. One of the key benefits of applying segmentation is that it allows for a more precise analysis of anatomical data by isolating only necessary areas.)

3).Predict heart failure based on ejection fraction (Ejection fraction (EF) refers to how well your left ventricle (or right ventricle) pumps blood with each heartbeat. Most times, EF refers to the amount of blood being pumped out of the left ventricle each time it contracts. The left ventricle is the heart's main pumping chamber. EF is expressed as a percentage. An EF that is below normal can be a sign of heart failure . EF helps your doctor know how severe your condition)

A. Main challenges

- Efficiency: The diagnostic decisions should be highly-accurate Because accordingly, effective treatment is given to the patients .
- Distributed Nature: healthcare system is not centrally deployed on single device of organization .There are many patients in hospitals, so it is necessary to provide more than one device in each hospital to treat patients effectively and quickly.
- Privacy: medical ethics pertains to the usage of personal information of patients. Based on the health organization's protocols, patient data must be kept private, and only the hospital and patient can look at it.
- Informed consent: medical decisions should be easy to interpret for patients. It is necessary to help the doctor to understand the diagnosis resulting from the system to ensure the condition of the diagnosis, and in addition, the patient can participate in understanding the diagnosis in an easy way.
- Security: healthcare system is always vulnerable to cyber-attacks.

A. Primary Contribution

This study's contributions to solving the above mentioned challenges:

1) DL models (CNN , LSTM , GRU , DNN) with highly accurate results for detecting cardiac biomarkers & cyber-attacks.

2) Using interpretable DL framework for explanation results such as gradcam and shap methods.

3) Using Edge cloud computing such as jetson nano , raspberry pi arduino. Then, apply privacy preserved FL to enable collaborative and homogeneity of data.

B. Article Organization

The rest of this article are organized as follows. In Section II, related research studies. In Section III, an architectural design of our system. In Section IV opservation, In section V the proposed DL model. In Section VI, Datasets a. In Section VII, results, comparisons, analyses, and discussions. Finally, Section VIII concludes this article

II. LITERATURE REVIEW

Recently, increased research attention has been given to the intelligent discovery of heart disease. in
In this context,
1 - The automatic diagnosis of ECG is carried out using the convolutional neural network, and based on the diagnosis by the traditional method using the CNN model, the average classification accuracy was 98.33%. Sensitivity (SNS) and specificity (SPC), respectively, were 98.33% and 98.35%, respectively.

2 - Classification Model for Prediction of Heart Disease
The overall comparison is analyzed among the classifier to detect which tends to be more effective and efficient for dataset. The decision tree shows the accuracy of 75.10% in, random tree represents the accuracy of 69.90% , whereas random forest has accuracy of 75.60%

3 - Designing lightweight deep learning models for echocardiography view classification . They used VGG-16, DenseNet, and Resnet, are distilled to train a set of lightweight models The best accuracy of 89.0% is achieved by an ensemble of the three very deep models while we show an ensemble of lightweight models has a comparable accuracy of 88.1%

1) They didn't tune the hyperparameter enough to get the best results to be more accurate and efficient.

2) They ignored data heterogeneity during training which is common in cardiac biomarkers.

3) Ignore the lack of interpretable results and focus only on the diagnosis and presentation of the results, but on what basis the diagnosis was made or the attack was determined.

4) They ignore failure to secure the system from hacking. Medical devices are known to be more vulnerable to cyber attacks.

III. SYSTEM DESIGN

The proposed research project aims to take the advantage of recent advances in DL and IoHT to develop a smart CAD for cardiovascular diseases in IoHT. In the following, we explain the details of the research plan of this project through different stages:

Step 1: this stage will start deep investigation and analysis of electrocardiography and echocardiogram datasets and its associated features, then, we provide an analytical review and comparatively analyze the previously designed DL models for efficient diagnosis. A common limitation of the current literature is the low generalization ability and black-box nature. To this end, a novel interpretable DL framework that combines the transformer modules to empower the learning capability from health records.

Step 2 :This stage will provide a solution for the black-box nature of deep learning approaches that prevent the doctors to depend on them for making a decision. In particular, the proposed two explanation methods (visually and textually) for a different levels of stakeholders. To ensure perfect interpretations, a human-in-the-loop strategy will be followed to involve medical experts to assess and evaluate the quality of interpretations according to the medical properties of the cardiovascular data.

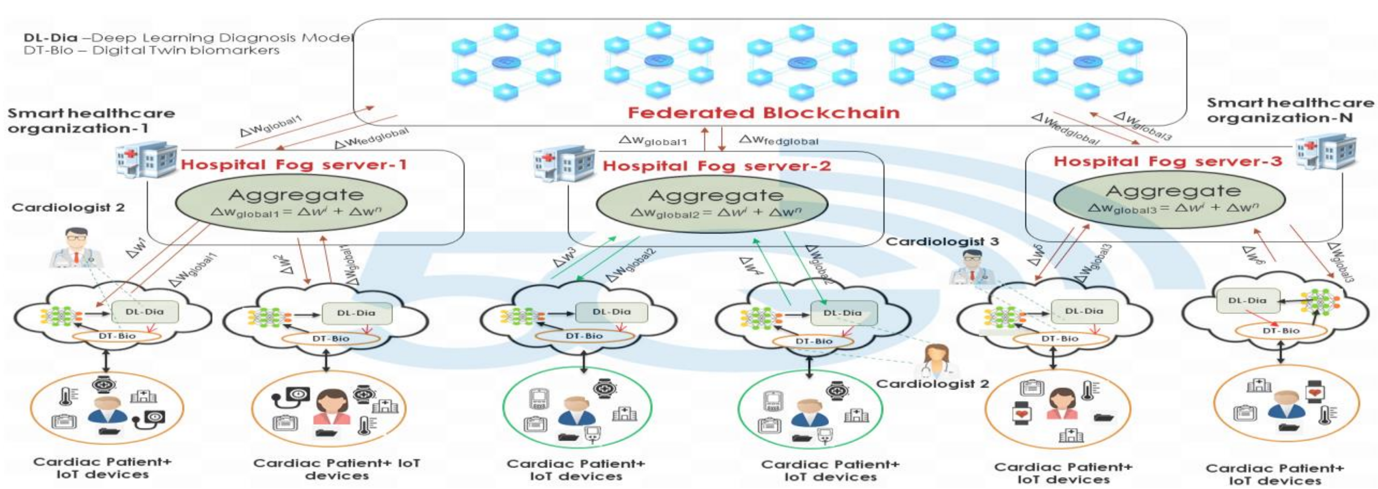
In particular, the proposed two explanation methods (visually and textually) for
a different levels of stakeholders. To ensure perfect interpretations, a human-in-the-loop strategy will be followed to involve medical experts to assess and evaluate the quality of interpretations according to the medical properties of the cardiovascular data.

Step 3: This stage of the project will emphasize designing a federated approach for preserving the privacy of interpretable DL during the training from cardiac healthcare records. On the other hand, the encryption approaches will present key-encryption and decryption methods to avoid leaking the private patient's information. The design of this approach will lessen the possible trade-off between privacy and performance data. Federated Learning (FL) aims to collaboratively train a ML model while keeping the data decentralized.

Where:
• initialize model
• each party makes an update using its local dataset
• parties share local updates for aggregation
• server aggregates update and sends back to parties
• parties update their copy of the model and iterate
In FL, data is naturally distributed and generated locally Data is not independent and identically distributed and it is imbalanced.

Step 4 :
Protect healthcare system from cyber-attacks

Fig 1. Federated learning: Applying Federated learning with aggregation function Average



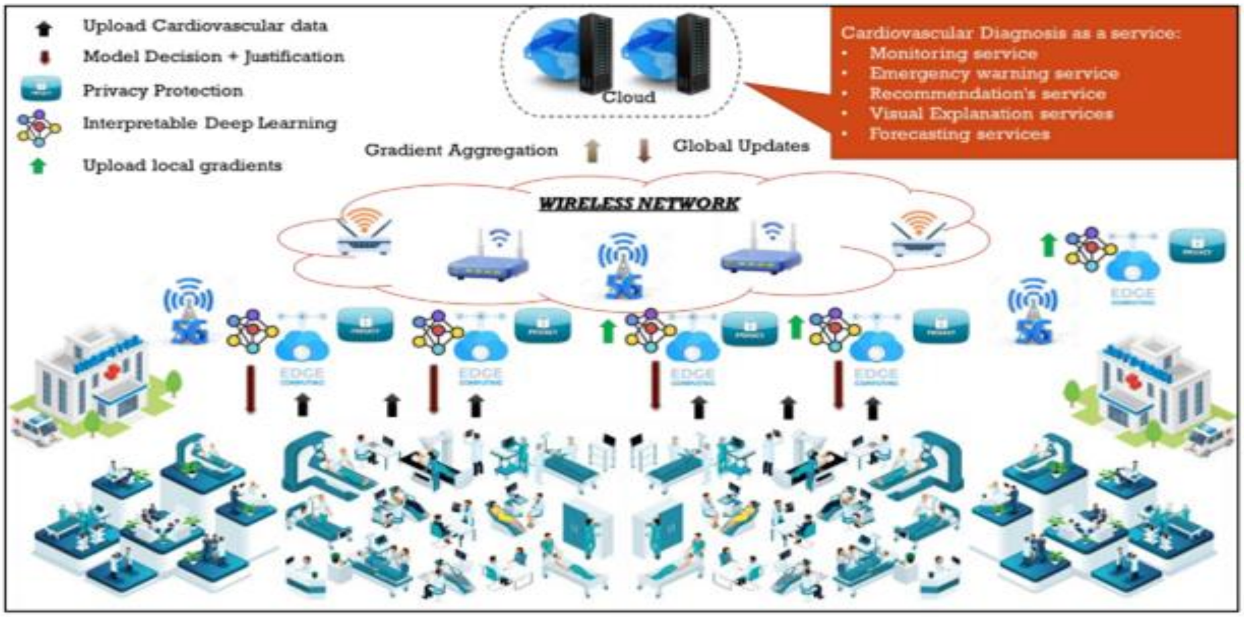


Fig 2.Proposed System

Hospitals connect to server IP and the model begins training without access to any of the data being shared

The server (Ministry of Health) will be able to initiate model training using hospital data without needing access to the data, minimizing risk of patient privacy breach

The server device will have a user-friendly interface allowing selection of needed option for the process (number of client, training details, etc.)

IV. OBSERVATIONS

After comparing multiple research papers (ARXIV, Multimodal Fusion, Ensemble, etc.) we found that results ranged from 93 to 97.5 . We decided to improve on the codes and data by using futuristic techniques including:

- Resampling Data such as SMOTE and ADASYN.
- Variational Mode Decomposition.
- Hyperparameter Optimization Methods (Optuna, GridSearch)
- Repeating Ensemble Methods for higher accuracy.
- K-Fold & Stratified Kfold

We also noticed that it would be beneficial to include echocardiogram reading classification to this project as it would improve diagnosis accuracy

VI. DATASETS

Different from current homogeneous datasets, which do not reflect the real-world behaviors of IoT data, the proposed model is evaluated using the MIT-BIH ECG , Stanford-echonet and NSL_KDD datasets

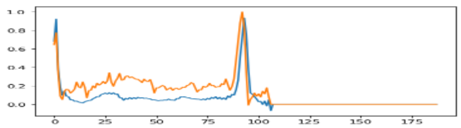
1 – ECG MIT-BIH Time series DataSets

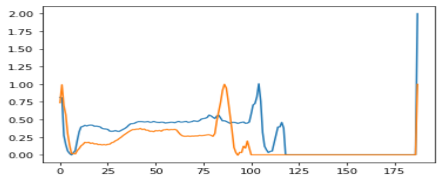
With the MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%) at Boston's Hospital; the remaining 25 recordings were selected from the same set to include less common but clinically significant arrhythmias that would not be well-represented in a small random sample.

The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10-mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database.

The dataset could easily achieve this project's main aim, which is classifying and monitoring ECG Patterns using Deep Neural Networks models. The results of running the models on this dataset can be found below:

Sample from mit-bih visualization

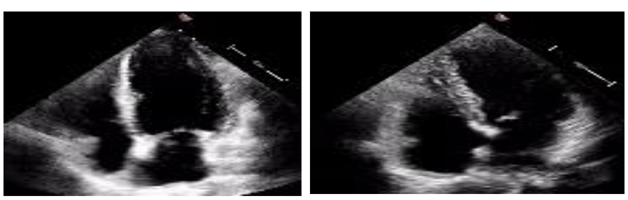




2 – Echo Stanford Dataset

Echocardiogram Videos: A standard full resting echocardiogram study consists of a series of videos and images visualizing the heart from different angles, positions, and image acquisition techniques. The dataset contains 10,030 apical-4-chamber echocardiography videos from individuals who underwent imaging between 2016 and 2018 as part of routine clinical care at Stanford University Hospital. Each video was cropped and masked to remove text and information outside of the scanning sector. The resulting images were then down sampled by cubic interpolation into standardized 112x112 pixel videos.
This data was used to classify Echo patients with determination of the locations of the disease by applying the segmentation. We were able to predict the ejection fraction, which we predict increases or decreases in the occurrence of a heart attack. The results were satisfactory for the application

Sample from Stanford Dataset echonet

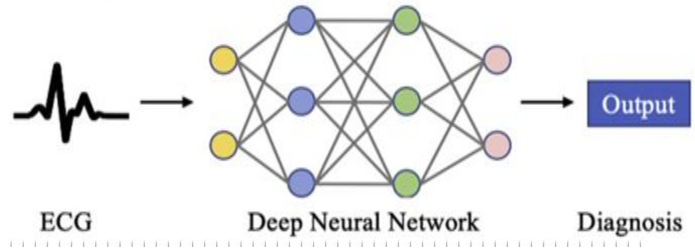


3 – NSL-KDD Datasets
NSL-KDD is a data set suggested to solve detect attacks
Sample of Data

dst_host_same_src_port_rate	dst_host_srv_diff_host_rate	dst_host_serror_rate	dst_host_srv_serror_rate	dst_host_rerror_rate	dst_host_srv_rerror_rate	labe
0.17	0.00	0.00	0.00	0.05	0.00	norn
0.88	0.00	0.00	0.00	0.00	0.00	norn
0.00	0.00	1.00	1.00	0.00	0.00	neptu
0.03	0.04	0.03	0.01	0.00	0.01	norn
0.00	0.00	0.00	0.00	0.00	0.00	norn

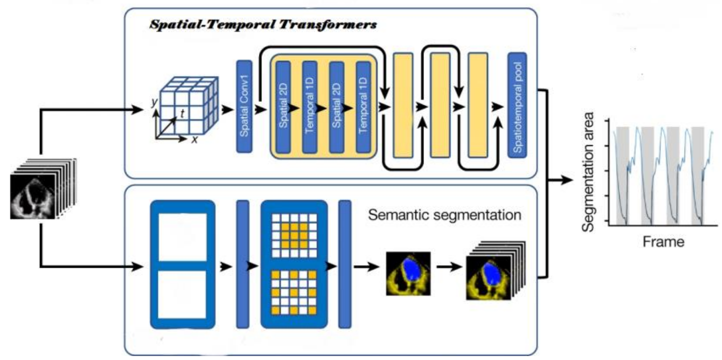
V. PROPOSED DL MODEL

A. Biomarker-based Diagnosis ECG



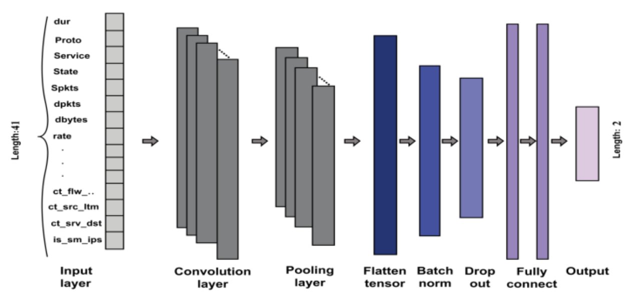
B. Echocardiogram-based Diagnosis

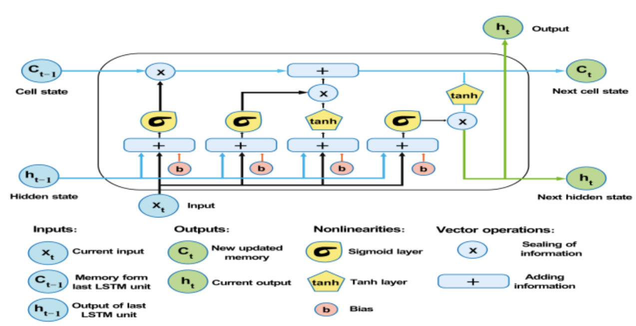
.



A. Securing Healthcare System

B. Based on Model cnn & Lstm





VII. Results

1 - Comparative Analysis

MIT-BIH

Method	Accuracy	F1-score
Proposed	**0.99**	**0.99**
RNN	**0.99**	**0.99**
DNN	**0.97**	**0.97**
LSTM	**0.97**	**0.98**

Echonet

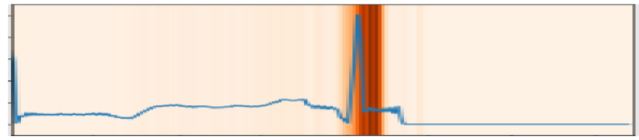
Method	AUC	Dice
U-Net++	0.95	0.88
EchoNet	0.97	0.92
U-Net	0.96	0.90
Proposed	0.99	0.93

NSL_KDD

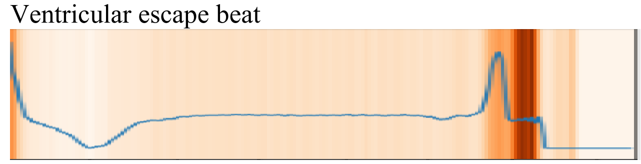
Method	Accuracy	Loss
Proposed	**98.7**	**0.09**

2 - Interpretability Analysis

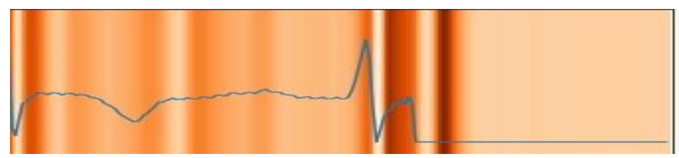
Normal Beat



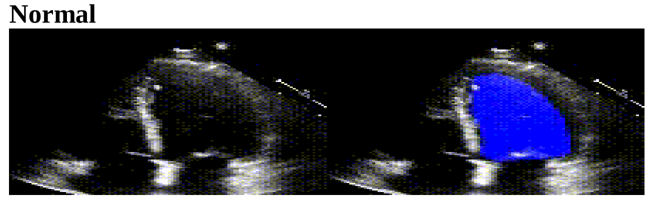
Ventricular escape beat



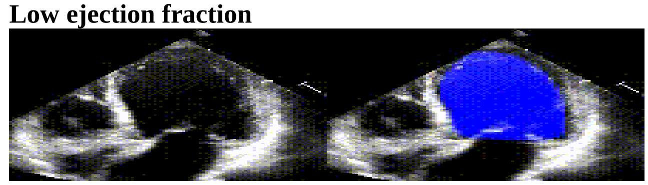
Premature ventricular



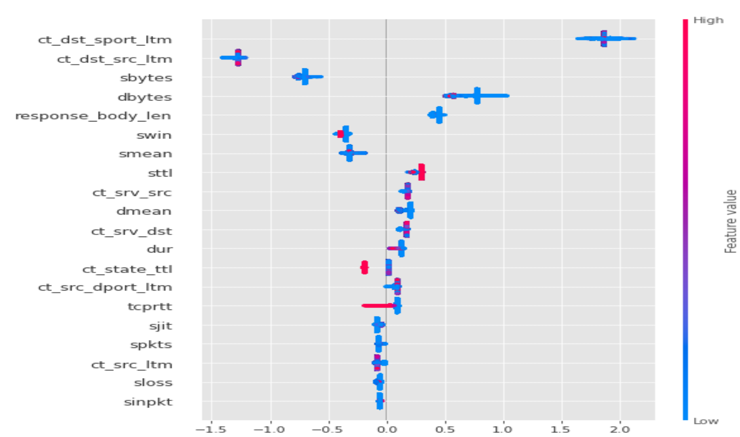
Normal



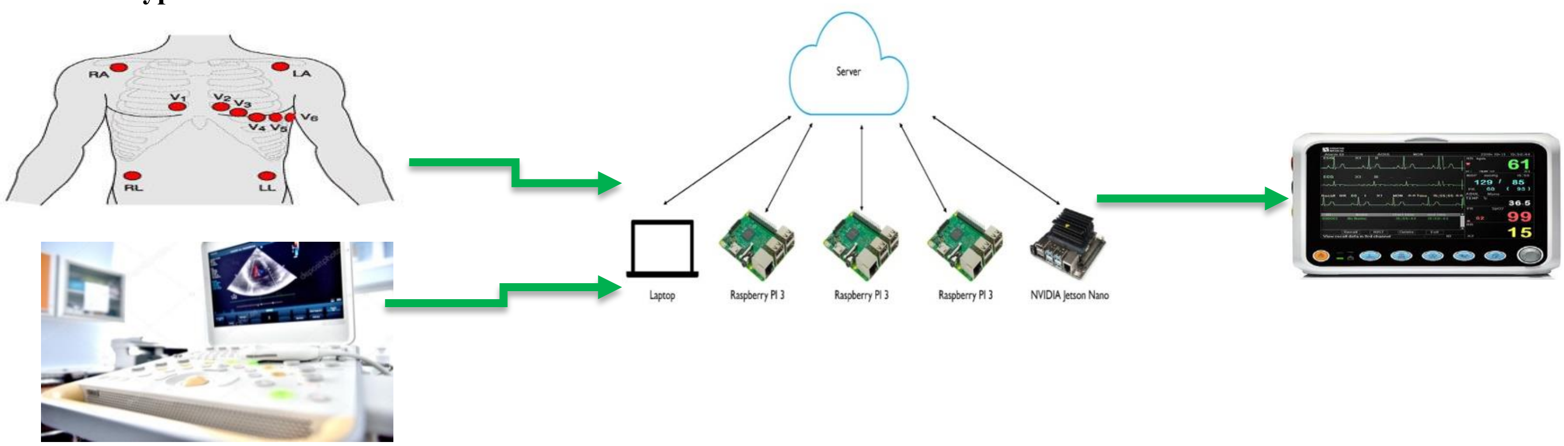
Low ejection fraction



Detect Important features from Attack



Prototype



VIII. Conclusion and Future Work

In conclusion, this article presented a novel federated DL approach for efficient cardiovascular disease and cyber-attack detection. Collaborative training in the edge-cloud environment revealed its efficiency while preserving the privacy of patients. This will encourage hospitals and doctors to use our system as it will keep patients data confidential, and The application of Explainable AI will provide doctors justification for the classification of cardiac patient.

The proposed system will enable ministries of health to collect health records of cardiac patients, to create a public benchmark database. This would be helpful in two dimensions.

- First, they will be able to analyze, manage, treat the citizens with such diseases in an efficient reliable way.
- second, the system will help researchers to study and monitor rare cases with higher reliability.

References

1. Automatic ECG Diagnosis Using Convolutional Neural Networkby Roberta AvanzatoORCID and Francesco Beritelli ORCID Department of Electrical, Electronic and Computer Engineering, University of Catania, 95125 Catania, Italy

Article Google Scholar

2. Classification Model for Prediction of Heart Disease
Ritu Chauhan, Rajesh Jangade & Ruchita Rekapally
Article Google Scholar

3.Designing lightweight deep learning models for echocardiography view classification . Hooman Vaseli, Zhibin Liao, Amir H. Abdi, Hany Girgis, Delaram Behnami, Christina Luong, Fatemeh Taheri Dezaki, Neeraj

4. Nourhan Bayasi; Temesghen Tekeste; Hani Saleh; Baker Mohammad; Ahsan Khandoker, 2016. ‘LowPower ECG-Based Processor for Predicting Ventricular Arrhythmia’, IEEE Transactions On Very Large Scale Integration (VLSI) Systems, 24(5): 1962-1974.
5. Tien-En Chen, Shih-I Yang, Li-Ting Ho, Kun-Hsi Tsai, Yu-Hsuan Chen, Yun-Fan Chang, Ying-Hui Lai, Syu-Siang Wang, Yu Tsao, and Chau-Chung Wu, 2016. ‘S1 and S2 Heart Sound Recognition using Deep Neural Networks’, IEEE Transactions on Biomedical Engineering.

6. Nima Karimian, Zimu Guo, and Mark Tehranipoor, Domenic Forte, 2016. ‘Highly Reliable Key Generation from Electrocardiogram (ECG)’, IEEE Transactions on Biomedical Engineering.

6. Farzad Khosrow-khavar, Kouhyar Tavakolian, Andrew Blaber and Carlo Menon, 2016. ‘Automatic and Robust Delineation of the Fiducial Points of the Seismocardiogram Signal for Non-invasive Estimation of Cardiac Time Intervals’, IEEE Transactions On Biomedical Engineering.
7. Weihua Pei, He Zhang, Yijun Wang, Xuhong Guo, Xiao Xing, Yong Huang, Yuxing Xie, Xiaowei Yang, and Hongda Chen, 2016. ‘Skin-potential variation insensitive dry electrodes for ECG recording’, IEEE Transactions On Biomedical Engineering.