Machine Learning and End to End Deep Learning for Detection of Chronic Heart Failure from Heart Sounds

Mettu Shruthi
Department of ECE
Siddhartha Institute of Technology & Sciences,
Korremula Road, Narapally, Ghatkesar, Medchal-Malkajgiri, Telangana, India
shruthimudhiraj9@gmail.com

Addagatla Prashanth
Department of ECE
Institute of Aeronautical Engineering,
Dundigal, Hyderabad - 500 043, Telangana, India
aprashanth4060@gmail.com

*Srinivas Bachu
Department of ECE
Siddhartha Institute of Technology & Sciences,
Korremula Road, Narapally, Ghatkesar, Medchal-Malkajgiri, Telangana, India
bachusrinivas@gmail.com

2024 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST) | 979-8-3503-5137-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICRTCST61793.2024.10578348

***Abstract*— Heart failure that does not resolve over time is known as chronic heart failure (CHF). This debilitating ailment occurs when the heart is unable to pump blood to the body's organs and tissues at the pressures necessary for metabolism. The occurrence of CHF is increasing at a rate of 2% each year, reaching epidemic proportions in the population. One to two percent of the general population and ten percent of those 65 and up in the industrialised world are affected by CHF. At present, around 2% of the yearly healthcare expenditure goes towards the cost of diagnosing and treating CHF. The United States spent about $35 billion on CHF treatment in 2018, and that number is projected to grow in the coming decade. The 5-year survival rate of this population is just around 50%, which is a very bad prognosis for congestive heart failure (CHF), even if medical and device-based therapy techniques have improved significantly over the recent decades. In the normal progression of congestive heart failure (CHF), the patient will go through phases where they are feeling fine and showing no indications of fluid overload, and phases where symptoms and signs of systemic fluid overload are more noticeable, such as orthopnea, peripheral edoema, liver congestion, and pulmonary edoema. To successfully achieve a negative fluid balance and return to the compensatory state, patients typically need to be admitted to the hospital during the latter episodes and treated with intravenous drugs (diuretics, inotropes). If a treating physician could quickly identify when a patient's HF is getting worse, they may make adjustments to the patient's outpatient medical therapy and keep them out of the hospital.**

At present, a skilled doctor can tell whether heart failure is getting worse just by looking at the patient and looking for certain changes in their blood that are called biomarkers for heart failure. It is unfortunate that when a patient's CHF symptoms worsen, it usually indicates that they are already in the midst of a fully established episode of CHF and will likely need to be admitted to the hospital. Also, phonocardiography can detect distinctive alterations in heart sounds that occur in some patients when their heart failure worsens. Hence, this project employs a state-of-the-art end-to-end average aggregate recording model that incorporates features extracted from both machine learning and deep learning to detect chronic heart failure from phonocardiography (PCG) data. Results from both the proposed ChronicNet model and standalone ML/DL models were also evaluated.

Keywords—Chronic Heart Failure, Heart Sounds, Machine Learning.

I. Introduction

The clinical condition known as heart failure (HF) can have many causes [1]. This condition develops when the heart's pumping capacity is inadequate to supply the body's metabolic needs. The lifestyles of people with HF are drastically altered as a result of the changes they must endure. Heart failure is a top cause of hospitalisation for people aged 65 and up [3] and one of the top causes of death [2]. From 2012 to 2030, the number of individuals living with HF is projected to rise to almost 8 million, a 46% increase [4]. Patients who experience heart failure can be classified into three categories based on the left ventricular ejection fraction (EF): those with decreasing EF (HFrEF; EF < 40%), those with mid-range EF (HFmrEF; EF 40-49%), or those with preserved EF (HFpEF; EF > 50%) [5].

Patients with HFpEF or those in the early phases of the disease may make HF diagnosis more difficult. It may be more challenging for elderly persons, those with chronic lung illness, and overweight people to recognise certain signs and markers [3]. The most valuable diagnostic tools for those suspected of having HF are standard laboratory testing, electrocardiograms (ECGs), natriuretic peptides, and echocardiography. In their guidelines for the treatment of heart failure, the European Society of Cardiology (ESC) stressed the importance of closely monitoring cardiovascular and co-morbid conditions, as well as biomarkers and biomarkers involving the renin-angiotensin-aldosterone system, as well as remote monitoring (through an implanted device when indicated), structured telephone support, multidisciplinary care, and therapies utilising these three systems. Some such approaches include the use of diuretics, beta blockers, sodium-glucose co-transporter 2 inhibitors (for HFrEF), antagonists of mineralocorticoid receptors, and beta blockers. Although it is recommended to monitor vital signs such as weight, blood pressure, pulse, food, and symptoms daily, most people with HF who undertake therapy usually have follow-up appointments in a doctor's office 2-12 times a year.

Many potential avenues exist for ML to improve healthcare efficiency. Medical professionals may be able to benefit from prognostic models when deciding how to treat their patients. Diagnostic models also have other applications, such as screening, risk stratification, and test prescription. This lessens financial strain, conserves materials, and makes doctors' jobs easier. Due to the huge monetary burdens connected with HF management and the increasing frequency of the disease, the diagnosis and treatment of the condition continue to be top objectives.

Figure 1. The research's digital stethoscope.

A high-quality digital stethoscope with Bluetooth connectivity is used to record the sounds, as illustrated in Figure 1. We have utilised 152 cardiac sounds collected from 122 individuals, including 23 with a history of medically-confirmed heart problems. Filtering, segmenting, feature extraction, and ML classifier stacking are the four steps that make up the process. The overall accuracy achieved by the applied methodology was 96%, with a negative class F-score of 0.97 for healthy classes and a positive class F-score of 0.87 for unhealthy classes.

We target the detection of the CHF condition (compensated vs. decompensated) using the analysis of heart sound recordings, in addition to differentiating between healthy individuals and CHF patients. Our work expands upon previous research that showed promise on a small dataset for differentiating between healthy persons and patients experiencing a decompensated CHF episode using a combination of expert characteristics and machine learning classifiers.

II. Literature Review

The proper function of the heart can be compromised by a variety of medical illnesses collectively known as cardiac diseases. Many various kinds of heart problems exist, such as heart failure (HF), coronary artery disease (CAD), irregular heart rhythms, and countless more. Myocardial infarction (MI) and chest pain (angina) are the most prevalent symptoms of cardiovascular disease (CVD), however there are many more. This condition is characterised by constricted or blocked blood channels. Heart problems manifest as angina, chest pain, dyspnea, irregular heartbeats, and chest discomfort [6]. The heart chambers are affected by HF, a chronic condition. Intracardiac pressure rises when a person's heart is unwell and cannot pump blood normally enough to the rest of the body. In response to the heart's inability to provide enough blood to all organs and tissues, the kidneys cause fluid to accumulate, which causes localised edoema and, ultimately, airway blockage. Worldwide, 26 million adults are dealing with congestive heart failure, making it one of the most pressing healthcare issues of our day. Heart disease is responsible for the deaths of about 17.9 million people annually, or 31% of all fatalities globally [7].

Factors that cannot be changed, such as gender, family history, and advanced age, raise the likelihood of heart failure; factors that can be changed, such as cholesterol levels, smoking, blood pressure, and obesity, decrease this likelihood [8]. In order to gain a better grasp of HF, it is helpful to first familiarise oneself with the most frequent forms of the condition. In this context, Figure 1 shows the four cardiac chambers that work together to pump blood.

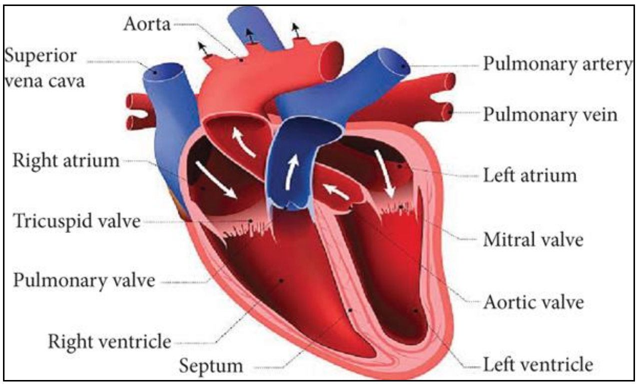

Figure 2. Anatomy of the heart.

The healthcare industry has recently produced massive amounts of patient data. Nevertheless, this data is not being utilised effectively by researchers and practitioners in order to diagnose the disease. Quality of service (QoS), which guarantees accurate and prompt disease diagnosis that leads to competent treatment of patients, is a big problem in the healthcare industry. It is unacceptable to have impaired diagnosis because it leads to harmful effects [9].

Major Types of Heart Diseases

Myocardial ischemia describes this condition [10]. A heart attack, which occurs when one or more arteries are partially or completely blocked, is an unavoidable consequence of this condition. Each of the four chambers of the human heart—the right and left atriums—serves a different purpose: taking blood from the outside world and pumping it throughout the body. The right ventricle collects blood that has lost some of its oxygen content and pumps it to the lungs. Following its exit from the lungs, the left ventricle carries oxygenated blood to all parts of the body [11–13]. The left ventricular chamber is the most common site of cardiac injury that leads to heart failure. To help detect CAD, echocardiography analyses or tracks the heart for the onset of wall motion abnormalities and coronary artery disease (CAD) progression [14]. Wall motion scoring and LV measurement can detect CAD. Hence, it is crucial to monitor the LV to prevent long-term damage that can impact its size, shape, and function. Echocardiography is a special kind of ultrasound imaging that records the heart's anatomy, function, and motion in three dimensions. In order to diagnose cardiac disease, echocardiography evaluates the heart's morphology and function[15].

One to two percent of healthcare expenditure goes towards treating heart failure, making it a big and increasing medical and economic concern globally. Heart failure has been more common in recent decades, and the ageing population in developed countries is likely to contribute to this trend. To better understand the reasons and difficulties that lead to the increasing expenses in heart failure, cost-of-illness studies might be very helpful in this context [16].

III. Proposed Methodology

Preprocessing, feature selection, and classification make up the three fundamental stages of the suggested methodology. Features with 50% or more missing values are removed from the preprocessing pipeline. Additionally, characteristics that have an uneven distribution of values are eliminated, and errors and outliers (such as 4.5 being recorded as 4.5) are identified and fixed for each feature. On top of that, the undersampling method is commonly used to cope with the class imbalance problem in datasets. A random subsample with a specified dispersion across class frequencies is generated by this technique. The most common and rarest class can only have a maximum "spread" of this kind. To maintain a balanced representation of all classes, this study employs a random undersample of the dominating class. Table 1 displays the 19 features used for HF diagnosis after missing value features and discrete features with an uneven distribution were removed.

Table 1. Features for HF diagnosis.

Category	Description
General demographic data	Age and gender
Classical cardiovascular risk factors	Hypertension
Personal history of cardiovascular disease	MI, CAD, and any arrhythmia (Arr) or paroxysmal atrial fibrillation (Afib) combined as Arr-Afib
Physical examination	BMI, SBP, DBP, and HR
Laboratory findings	BNP
Echocardiographic parameters	LVIDd, LVMI, LAVI, EA, E deceleration time, Ee, EF, and peak E-value

The second step involves applying feature selection to each and every characteristic in every category. Only the features that make it to the subset that is retained are then used for categorization. Feature selection methods are utilised in this study to evaluate the predictive power of feature subsets and the extent to which they are redundant. The goal is to favour feature sets that have low intercorrelation and high class correlation.

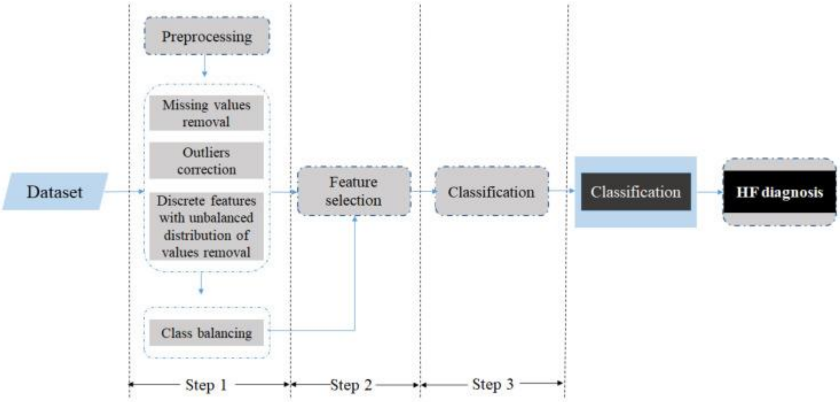

Figure 3. Methodology.

At last, the reduced feature subset is subjected to various classifiers during the classification step. The classifiers are subsequently evaluated using this methodology. You can see the suggested approach in Figure 1. Accuracy, responsiveness, and affinity are the metrics used to describe the results.

The following feature combination (Figure 4) was subjected to the aforementioned technique.

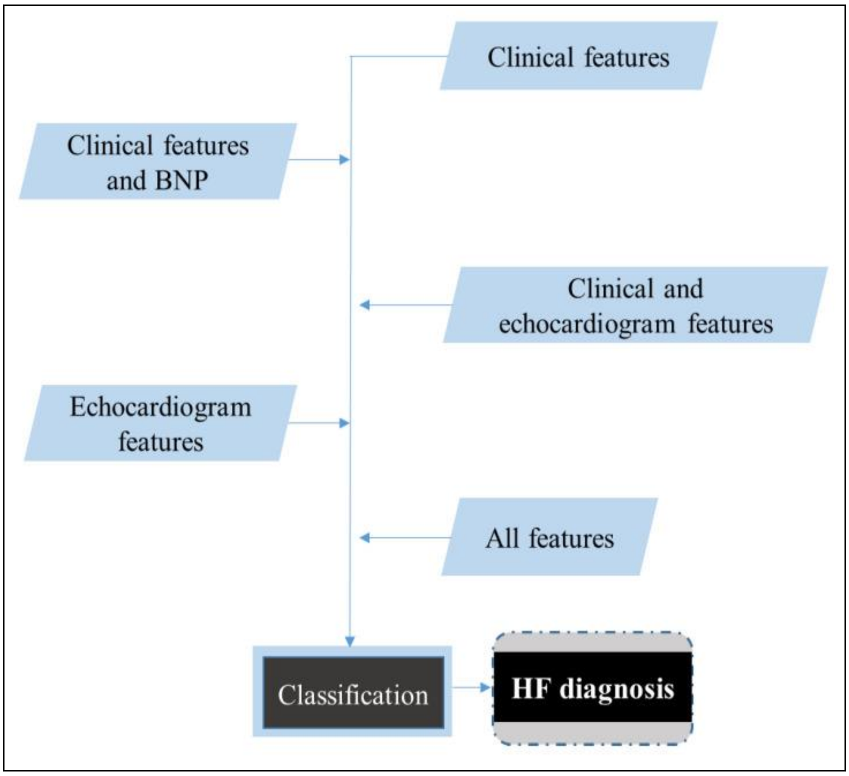

Figure 4. Classification for various sets of features.

IV. Results And Study

The HF diagnosis did not include people with acute HF or NYHA grade III-IV because it is not difficult to diagnose HF in patients with severe symptoms. The HF diagnosis dataset comprised 422 people; 162 had chronic HF and 260 did not. Feature selection's results for the HF diagnosis issue are shown in Table 2.

Table 2. HF diagnosis classification results.

Features Type	**Classifier**	**Accuracy %**	**Sensitivity %**	**Specificity %**
Clinical features	LMT	84.12	82.10	85.38
Clinical features and BNP	LMT	88.15	85.80	89.62
Clinical and echocardiographic features	ROT	90.76	93.21	89.23
Echocardiographic features	ROT	87.91	90.74	86.15
All features	ROT	91.23	93.83	89.62

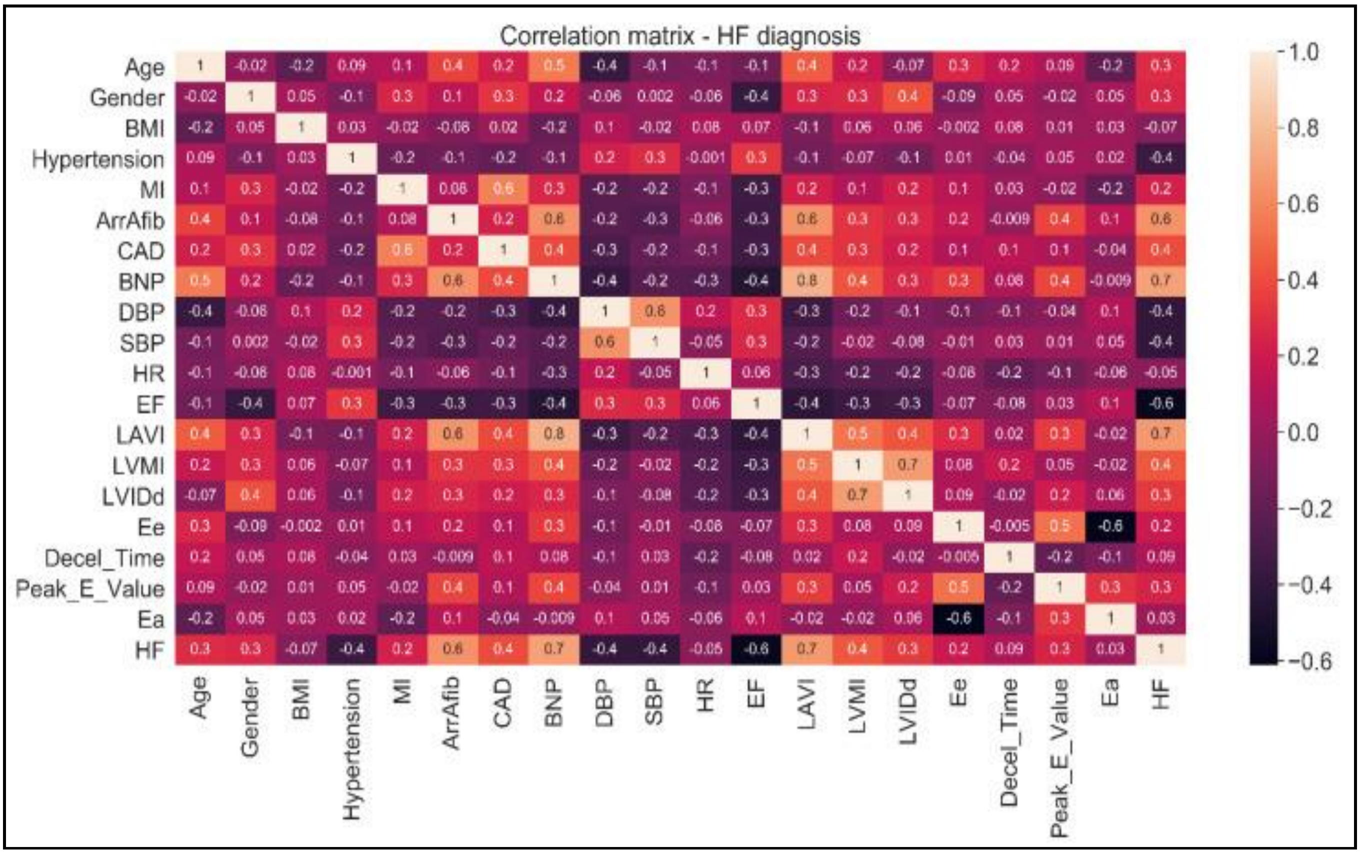

Figure 5. Correlation matrix.

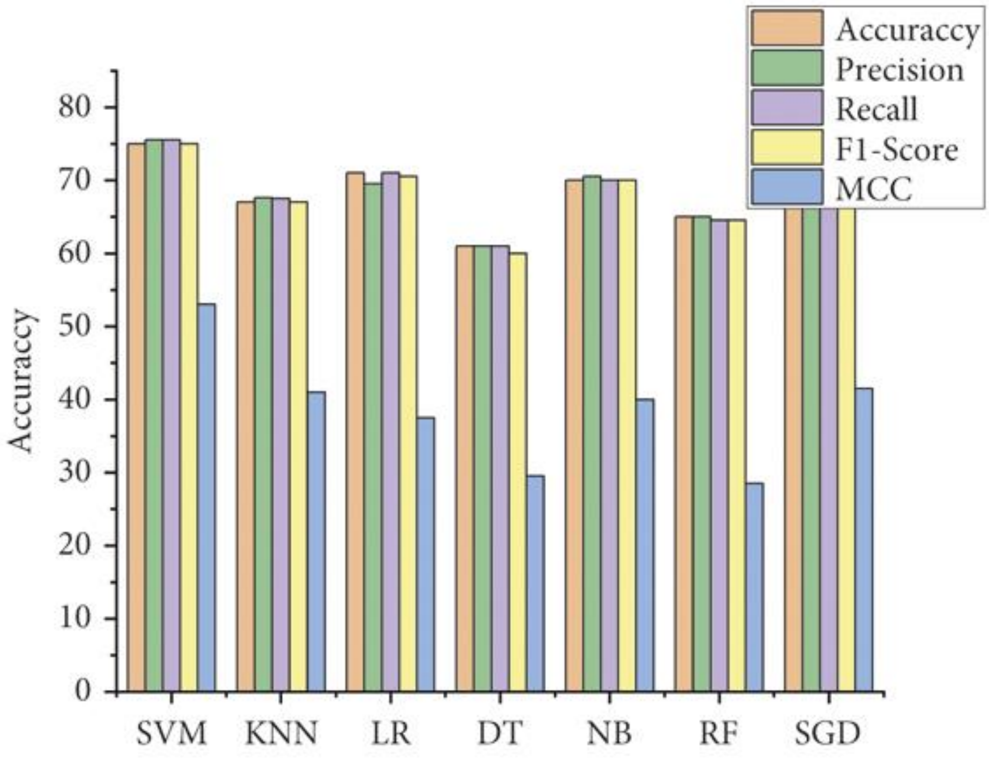

Figure 6. Result of the classifier with full feature

Table 2 Classifier performance before feature selection.

Classifier		Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	MCC (%)	Time complexity (sec)
SVM	1	75	80	73	76	51	10.4
	0	75	71	78	74	55	
	Overall	75	75.5	75.5	75	53	
KNN	1	67	64	69	66	35	16.7
	0	67	71	66	68	47	
	Overall	67	67.6	67.5	67	41	
LR	1	71	76	69	72	42	12.2
	0	71	63	73	69	33	
	Overall	71	69.5	71	70.5	37.5	
DT	1	61	56	62	58	22	19.9
	0	61	66	60	62	37	
	Overall	61	61	61	60	29.5	
NB	1	70	75	68	71	41	24.7
	0	70	66	72	69	39	
	Overall	70	70.5	70	70	40	
RF	1	65	68	64	66	30	17.1
	0	65	62	65	63	27	
	Overall	65	65	64.5	64.5	28.5	
SGD	1	69	76	66	71	39	14.4
	0	69	62	72	66	44	
	Overall	69	69	69	68.5	41.5	

Figure 5 shows the correlations between all features and between each feature and the class (HF result). Table 2 and Figure 6 demonstrate that the SVM performs admirably, with accuracy of 75%, precision of 75.5%, recall of 75.5%, F1-score of 75%, MMC of 53%, and time complexity of 10.4 seconds. For the KNN classifier, we tried a range of K values; the optimal result across all rounds was 67% accuracy, 67.6% precision, 67.5% recall, 67% F1-score, 41% MCC, and 16.7 seconds of time complexity. A time complexity of 12.2 seconds was attained by the LR classifier, which had the following metrics: 71% recall, 69.5% precision, 70.5% F1-score, and 37.5% MCC. A 61% F1-score, a 61% recall, a 61% precision, a 29.5% MCC, and a 19.9-second time complexity were all attained by the DT classifier. An accuracy of 70%, precision of 70.5%, recall of 70%, F1-score of 70%, MCC of 40%, and time complexity of 24.7 seconds were all attained by the NB classifier. With a time complexity of 17.1 seconds and a recall of 64.5%, F1-score of 64.5%, MCC of 28.5%, and precision of 65%, the RF classifier accomplished a lot. The SGD classifier has a temporal complexity of 14.4 seconds, an F1-score of 68.5%, a precision of 69%, a recall of 69%, and an MCC of 41.5%.

V. Conclusion

This research presents a new approach to HF diagnosis that makes use of ML techniques. In this groundbreaking research, we mimicked clinical practice to determine how different feature types affected categorization accuracy. High levels of specificity (89.62%), accuracy (91.23%), and sensitivity (93.83%) were achieved for the HF diagnosis when all feature categories were used for categorization. By reducing the number of diagnostic tests needed, efficiency is enhanced by selecting a limited set of features through feature selection. In addition, our method still produces respectable results when limited to clinical criteria alone, and this is true even when we omit some of the other features.

References

[1] Braunwald E., Zipes D.P., Libby P., Bonow R.O. *Heart Disease: A Textbook of Cardiovascular Medicine.* 7th ed. Elsevier Science; Philadelphia, PA, USA: 2004.

[2] Sabate E. *Adherence to Long-Term Therapies, Evidence for Action.* World Health Organization; Geneva, Switzerland: 2003.

[3] Franzén K., Saveman B.-I., Blomqvist K. Predictors for Health Related Quality of Life in Persons 65 Years or Older with Chronic Heart Failure. *Eur. J. Cardiovasc. Nurs.* 2007;6:112–120. doi: 10.1016/j.ejcnurse.2006.06.001

[4] Mozaffarian D., Benjamin E.J., Go A.S., Arnett D.K., Blaha M.J., Cushman M., de Ferranti S., Després J.-P., Fullerton H.J., Howard V.J., et al. Heart Disease and Stroke Statistics—2015 Update. *Circulation.* 2015;131:e29–e322. doi: 10.1161/CIR.0000000000000152.

[5] McDonagh T.A., Metra M., Adamo M., Gardner R.S., Baumbach A., Böhm M., Burri H., Butler J., Čelutkienė J., Chioncel O., et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur. Heart J.* 2021;42:3599–3726. doi: 10.1093/eurheartj/ehab368. [PubMed] [CrossRef] [Google Scholar]

[6] Ali L., Niamat A., Khan J.A., Golilarz N.A., Xingzhong X., Noor A., Nour R., Bukhari S.A.C. An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart

Failure. *IEEE Access*. 2019;7:54007–54014. doi: 10.1109/ACCESS.2019.2909969.

[7] M. Cotter, “A systems view of cardiovascular disease,” 2019.

[8] E. Bathrellou, M. D. Kontogianni, E. Chrysanthopoulou et al., “Adherence to a dash-style diet and cardiovascular disease risk: the 10-year follow-up of the Attica study,” *Nutrition and Health*, vol. 25, no. 3, pp. 225–230, 2019.

[9] W. S. Colucci, “Patient education: Heart failure (beyond the basics),” 2021.

[10] L. Ali, S. U. Khan, M. Arshad, S. Ali, and M. Anwar, “A multi-model framework for evaluating type of speech samples having complementary information about parkinson’s disease,” in *2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pp. 1–5, Swat, Pakistan, 2019.

[11] L. Ali, C. Zhu, Z. Zhang, and Y. Liu, “Automated detection of parkinsons disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 7, pp. 1–10, 2019.

[12] L. Ali, C. Zhu, M. Zhou, and Y. Liu, “Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection,” *Expert Systems with Applications*, vol. 137, pp. 22–28, 2019.

[13] L. Ali, Z. He, W. Cao, H. T. Rauf, Y. Imrana, and M. B. Bin Heyat, “Mmdd-ensemble: A multimodal data–driven ensemble approach for parkinson’s disease detection,” *Frontiers in Neuroscience*, vol. 15, 2021.

[14] L. Ali, C. Zhu, N. A. Golilarz, A. Javeed, M. Zhou, and Y. Liu, “Reliable Parkinson’s disease detection by analyzing handwritten drawings: construction of an unbiased cascaded learning system based on feature selection and adaptive boosting Model,” *IEEE Access*, vol. 7, pp. 116480–116489, 2019.

[15] W. Akbar, W. P. Wu, S. Saleem et al., “Development of hepatitis disease detection system by exploiting sparsity in linear support vector machine to improve strength of adaboost ensemble model,” *Mobile Information Systems*, vol. 2020, Article ID 8870240, 9 pages, 2020.

[16] Wladimir Lesyuk, "Cost-of-illness studies in heart failure: a systematic review 2004–2016", Research article Open access Published: 02 May 2018