AI Tech International Journal
Vol. 1, No. 1, July-December, 2023
Journal homepage: https://techaijournal.com

Detection of Left Ventricular Hypertrophy with Electrocardiographic Signals

Doddala Radhika[1], Dr. Rajesh Koolwal[2], Dr. Sikhakolli Gopi Krishna[3]

[1]Research Scholar, Department of Computer Science & Engineering, JS University, Shikohabad, Uttar Pradesh
[2]Associate Professor, Supervisor, Department of Computer Science & Engineering, JS University, Shikohabad, Uttar Pradesh
[3]Professor, Co-Supervisor & Professor, Sri Mittapalli College of Engineering, Guntur, Andhra Pradesh

Article Info

Article history:

Received July 09, 2023

Revised September 14, 2023

Accepted September 19, 2023

Published September 30, 2023

ABSTRACT

Left ventricular hypertrophy (LVH) is associated with an increased risk of cardiovascular disease and is a sign of subclinical organ damage. Electrocardiograms (ECGs) are cheap, non-invasive, and easy to repeat, according to medical experts. They are often the initial line of defense in the fight against heart disease. Nowadays, there are a number of criteria used to evaluate LVH by ECG. The RS peak voltage combination of the multi-lead ECG usually needs to be higher than one or more diagnostic thresholds in order to meet these requirements. Segmenting the ECG beats of each case (LVH or non-LVH) was done using 24 features that were automatically retrieved from the 12-lead ECG's R-peak and S-valley amplitudes. After that, a dataset with these attributes was used to train a back propagation neural network (BPN). In this way, we were able to create an algorithm that could identify LVH from ECG data. Until recently, echocardiogram (ECHO) was the go-to test for detecting LVH. In a group of people from Taiwan, there were 173 cases of LVH. We were able to detect 1466 ECG cycles of LVH patients after beat segmentation. This is because, due to changes in heart rate and other factors, each ECG sequence typically had 8 to 13 cycles (heartbeats). The findings showed that our BPN model for LVH detection has an accuracy of 0.961, a precision of 0.958, a sensitivity of 0.966, and a specificity of 0.956. Our BPN model outperforms seven other methods that use ECG criteria and other AI models that are based on ECG that have been reported for LVH identification in the past.

INTRODUCTION

The risk of cardiovascular illness and death is associated with left ventricular hypertrophy (LVH), an early stage of structu ral heart disease1. As a result of LVH, hypertension patients experience subclinical end organ damage and a worse prognosis for cardiovascular diseases 1-4. Several methods are now available for determining LVH, including electrocardiography (ECG), echocardiography (ECHO), and cardiac magnetic resonance imaging (cMRI).5. Detecting LVH with an electrocardiogram is a quick and repeatable process. To diagnose LVH by electrocardiogram (ECG), there are a number of criteria6. Two of these, the Cornell and Sokolow-Lyon criteria, look at

the total voltage amplitudes of the S and R waves, which indicate left ventricular depolarization. Transthoracic echocardiography (TTE) can directly evaluate the posterior wall thickness and left ventricular septum (7, 8), but the accuracy of the measurement relies on the experience of the clinicians doing the procedure (5). Because it is fast, simple, and reproducible, electrocardiograms (ECGs) are a good tool for screening for left ventricular hypertrophy (LVH) in the general public, especially in healthy demographics. The overall sensitivity and precision continue to be poor, despite recent reports of new ECG-based criteria or studies that use pre-existing criteria for LVH identification (9, 10).11–19. These criteria mainly rely on the integration of multi-lead peak and valley amplitudes of R and S waves, which are associated with ventricular depolarization, to provide a diagnosis of LVH. More recently, some have suggested AI or ML algorithms to identify LVH from ECG data alone or in conjunction with other variables (such as demographics, anthropometrics, etc.)20-27; however, overall detection results are still lacking.

The goal of this work is to develop a system that can identify LVH utilizing our proposed ECG signal processing technique and a machine learning model. This will improve the capabilities of ECG-based LVH identification.

MATERIALS AND METHODS

Data Research population. A cohort of patients attending the Cardiology department at Tri-service General Hospital Songshan Branch in Taipei, Taiwan, had their electrocardiogram (ECG) and transthoracic echocardiogram (TTE) data collected retrospectively from January 1, 2016, to December 31, 2017. We included those who seemed healthy on the outside but had abnormal electrocardiogram findings from a health evaluation and TTE9,10. Echocardiography was performed after an electrocardiogram (ECG) in all cases included in this study, with a one- to three-month interval between the two procedures. The cohort consisted of 856 men and 96 females; 779 of the patients did not have LVH (82% of the total) and 173 of the patients had LVH (18%). The data used in the study came from the Tri-service General Hospital Songshan Branch, a military facility in Taipei, Taiwan. Since men make up the vast bulk of Taiwan's military, the hospital sees a disproportionate number of military personnel. This explains why there is a gender gap in this study. All procedures were executed in accordance with the relevant laws and regulations. Following the American Heart Association's criteria for ECG standardization and interpretation, the ECG was measured using the Philips Page Writer TC30 machine. The TTE measurement was performed using Philips IE33 echocardiographic equipment in compliance with current guidelines28,29. The research was greenlit by the IRB at Tri-Service General Hospital in Taipei, Taiwan (No. TSGH 2-106-05-148). The current study design carries an exceedingly minimal risk, hence the board decided that informed consent was not necessary. Our study has already been filed with ClinicalTrials.gov under the number NCT0347395130. The data was also deidentified and treated in accordance with the privacy standards set out by HIPAA.

what LVH stands for. The standard left ventricular hemodynamic (LVH) can be determined by echocardiography by using the Devereux formula to the left ventricle mass index (LVMI). On top of that, LVH meets LVMI requirements of over 95 g/m2 for women and over 115 g/m2 for men19,28. We used the definition in our earlier study as well.system architecture (number 31). The LVH detection system architecture is illustrated in Fig. 1. The proposed signal processing method for diagnosing LVH using 12-lead ECG data involved the use of three separate machine learning models, which were then compared extensively to seven different kinds of criteria. From 12-lead ECG raw data, 24 features were automatically selected using the signal processing approaches indicated by the machine learning models employed in this study. These features comprise the amplitudes of the R peak and S valley. Our goal in doing this is to improve the overall performance of ECG-based LVH detection predictions.

Guidelines for LVH specifications. In order to assess LVH, several ECG-based criteria have been published. In order to compare the effectiveness of seven different kinds of ECG-based criteria with machine learning models for LVH identification, they are listed in Table 1. analyzing electrocardiogram signals. Rhododendrons and Sphalerons are present. The initial step in developing a machine-learning model for the LVH detection system was to ensure that each R peak and S valley in an electrocardiogram sequence could be automatically retrieved. In wave detection approaches for ECG signals, the R peak is often located first, and then the peaks and valleys of succeeding waves are located by comparing their relative positions and attributes with those of other waves37–41. As an example, the initial point of lowest voltage generally follows a R peak and is located in a S valley. This study used electrocardiogram (ECG) equipment from Philips to measure the signal. The 500 Hz sample frequency and 11 s sampling period allowed for the acquisition of 5500 sampling points of each lead in every occurrence. With the help of MATLAB (Mathwork Inc.), the ECG signals were read and processed.

A common representation for the raw data of an ECG series is X = [x1, x2, x3, …, x5500]. If we take X as an example, we can see that the least distance between two consecutive R peaks is miR, the minimum peak amplitude of all R waves is mpa, the lowest point of the ith S wave is Si, and the peak of the ith R wave is Ri. Our algorithm for finding R peaks and S troughs is summarized here..

Setmpa=0.025mV
SetmiR=300(Thisisthenumberofpointsandstandsfor0.6s)
SearchforRpeaksfromthestartingpoint$x_1$ofECGdatasequencebythefollowingconditions:
x_j>mpa
(x_{j-1}<x_j)and(x_{j+1}<x_j)
If point x_j meets conditions (a) and (b), then it is set as R_i
Search for the next peak R_{i+1} from potential points with index ≥ (j+miR) and meet two criteria (a) and (b) described in step 3
Repeat step 4 until the data index (j+miR) > the end index (=5500 here)
Search for the minimum amplitude within 0.12s forward from each R_i and set the point as S_i
Search for the minimum amplitude m backward within 0.25 s from each R_i. If m<S_i, then update S_i as m, and search for the maximum amplitude within 0.12 forward from the updated S_i and update that point as R_i
Repeat step 7 for each R peak preliminarily detected in step 5

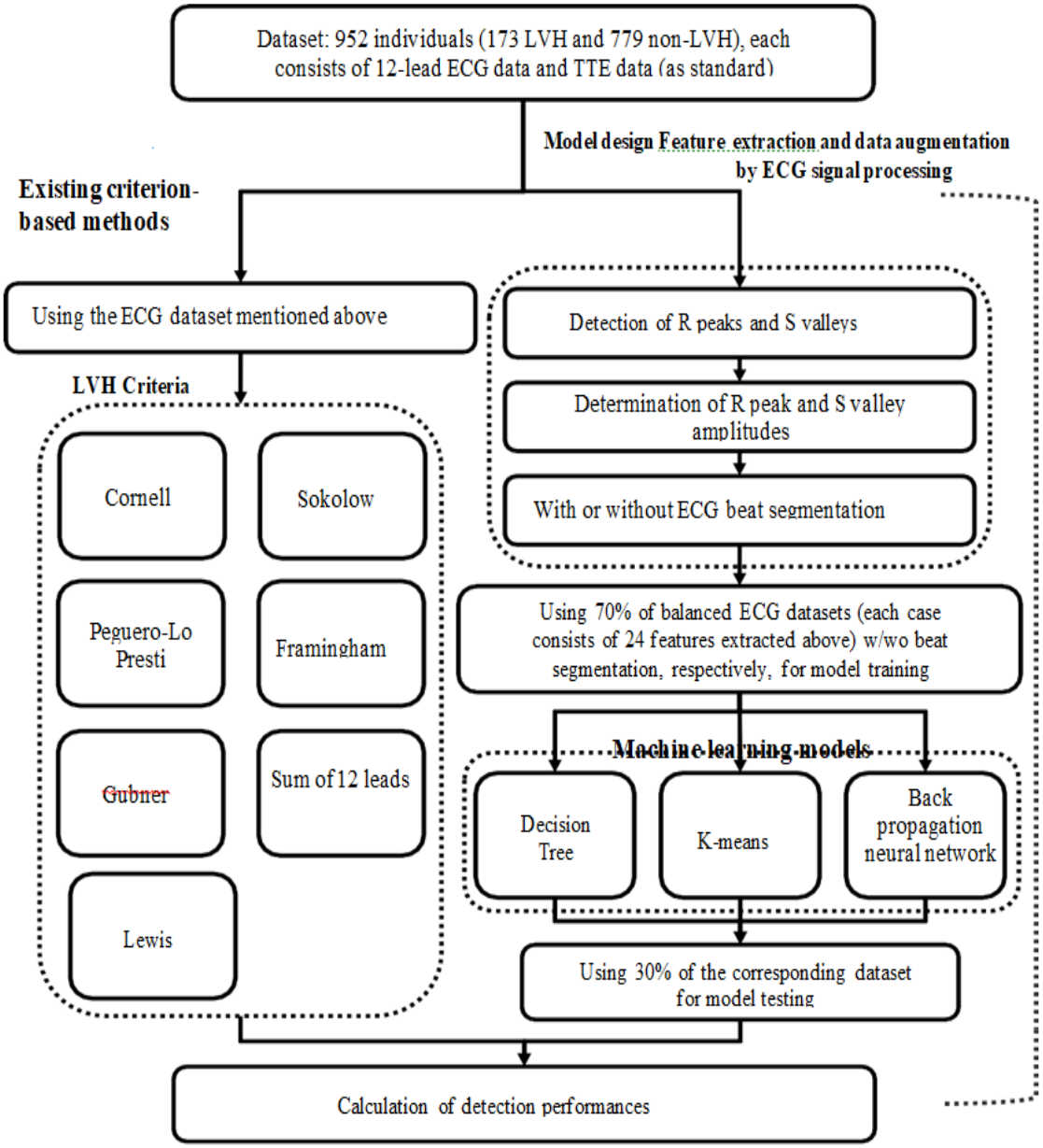

Figure1. The system architecture of LVH detection.

Table1. Sevenmethods, ECG-based criteria, for assessing LVH.

Criterion name	ECG criteria for assessing LVH
Cornell32	SV3+RaVL>2.8mV(man) SV3+RaVL>2.0mV(woman)
Sokolow7	SV1+max(RV5orRV6)≥3.5mV
Peguero11	SD(deepest S wave in any lead)+SV4≥2.8mV(man)SD (deepest S wave in any lead) +SV4≥2.3 mV (woman)
Framingham33	RaVL>1.1mV,RV4orRV5orRV6>2.5mV,SV1orSV2orSV3>2.5mV, max(SV1orSV2)+max(RV5orRV6)>3.5 mV, or RI+SIII>2.5 mV
Gubner34	RI+SIII≥2.2mV
$Sumof12_{leads}$35	Sum of max(R,S)amplitudeineachofthe12lead≥17.9mV
Lewis36	(RI+SIII)−(RIII+SI)>1.7mV

The preceding steps are understood in this way. First, the minimum peak amplitude (mpa) is adjusted to limit the minimum amplitude of a R wave peak that can be recorded. This allows the ECG signal to filter out peaks with too tiny amplitudes. An research was conducted to determine mpa = 0.025 mV. Secondly, there is a limit to how fast a human heart can beat; normally, it is between 60 and 100 ppm, or 1 to 0.6 s each beat. The minimum interval between two adjacent R peaks, or miR, is utilized to restrict the amount of time that two adjacent R peaks can be recognized consecutively and to prevent the subsequent R peak from being too close to its neighbor. The R peak is also often a local maximum; this study employed the criteria developed by mpa and miR to identify R peaks (steps 3-5).

Also, when each R peak is detected (step 6), you may utilize the fact that the S valley usually appears 0.12 seconds later as a feature to identify it. On rare occasions, nevertheless, the greater peak amplitude of the T wave in an electrocardiogram (ECG) can cause some T peaks to be incorrectly recognized as R peaks. To avoid this situation, the algorithm will search for an S-valley with a smaller amplitude than this one within 0.25 seconds of each R peak that has been found in the preliminary phase. If that is the case, the point will replace the previously defined S valley. Once the S valley is established, the R peak will be found at the point with the largest amplitude in the 0.12 s window before this valley. After this point is located, it will replace the previously recognized R peak (step 7). After the initial analysis (step 8), this procedure will be carried out again for each R peak that was detected. Afterwards, the algorithm will be finished such that it can automatically recognize S valleys and R peaks.

Finding the amplitudes of the R peak and the S valley is done. Following the detection of R peaks and S valleys, an average of 8-13 ECG cycles per ECG signal sequence were recovered in this experiment. This was because there was a chance that artifacts or tiny amplitude cycles could arise in some leads when ECG monitoring, and heart rate changes also played a role. Due to their potential for being too loud or lacking information, the first and last cycles were deleted initially. Next, we determined the R and S amplitudes for every lead by averaging the amplitudes of the remaining ECG cycles' R peaks and S valleys. This process has the potential to make the automatically computed R and S amplitudes more comparable to the actual manual measurements taken by the doctor.

Dissection of electrocardiogram impulses. There was not enough electrocardiogram data from 173 LVH cases in the initial dataset to train a machine learning model. This study employed the Pan-Tompkins methodology for beat segmentation to enhance the amount of ECG data, which was proposed in 43, to improve the detection performance. The technique for beat segmentation uses the identified R peaks as landmarks. We also removed the first and last cycles from each ECG series to keep the training and testing errors of the machine-learning model for LVH diagnosis (discussed in the next section) to a minimum. Following ECG beat segmentation, our investigation yielded an average of 8.47 beats.

LVH detection utilizing machine learning algorithms. Three machine-learning algorithms—a decision tree, k-means, and a back propagation neural network—were built using Python for the purpose of LVH identification in this study. Decision trees (DTs) are a kind of supervised machine learning that produce clear results quickly and are used in data science44 for tasks like categorization and regression. When compared to other machine-learning techniques, the decision tree's decision-making process is incredibly clear and easy to understand at every step.

To prevent the decision tree from rapidly growing during classification and causing overfitting, its depth is limited to 1-25. From all the trees, the one with the highest accuracy is selected as the ideal model. It is common practice to use k-fold cross

validation while training a machine learning model to verify the model's predictive performance45,46. In order to evaluate the correctness of each decision tree model at a specific depth, this work utilized tenfold cross validation. The entropy, shown in Eq. (1), is used to assess the performance of each decision tree classification node. In a binary classification problem, p is the number of positive cases and q is the number of negative cases in a node. In a perfect world, where every node has exactly one type of case, the entropy would be 0. It shows an ideal categorization. If, on the other hand, a node is associated with positive cases and negative ones, the entropy will be 1. This category represents the absolute worst case scenario.

$$\text{Entropy} = p \cdot \log 2p - q \cdot \log 2q \quad (1)$$

Use K-means. The widely-used k-means approach for clustering47 automatically assigns each data point to the appropriate category by continually calculating the distance between their neighbors. When it comes to clustering large amounts of data, the k-means method excels thanks to its simple concept and high computing efficiency. The algorithm has a fixed k-value. The beginning group centers of the k-means algorithm often significantly impact the categorization results. The k-means++ algorithm48, a variant of the original k-means method with an emphasis on increasing the distance between the original group centers, was employed in this study.

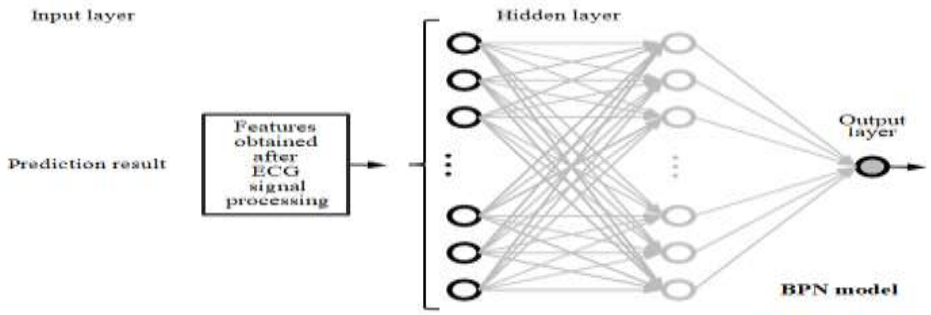

Figure2 The structure of back propagation neural network (BPN)

Artificial neural network trained with back propagation? Back propagation neural networks (BPNs) are one method developed to match the architecture of real neural networks. With the error back propagation algorithm50 as its training method, BPN is a multilayer feed forward neural network (Fig. 2). As a cornerstone for numerous AI applications, BPN continues to be one of the most popular machine learning algorithms in use today. Some examples of these include the following: the detection of COVID-19 and ischemic stroke; the diagnosis of hypertrophic cardio myopathy and hypertensive heart disease; the classification of arrhythmia disease (using ECG signals); the prediction of medical costs and the 30-day readmission risk for all causes using BPN-associated models; and many more health and medical events.

A neural network can also be seen as a nonlinear statistical method with a strong ability to learn and create connections between input and output nodes customized to a specific application. The fundamental architecture of the BPN consists of three layers: the input layer, the hidden layer, and the output layer. Using three processes (described in the "ECG signal processing" section and depicted in Fig. 1), 24 features were automatically obtained from 12-lead ECG data and employed in our BPN model's input layer. The MATLAB software was utilized for this operation. The amplitudes of the R peaks and S valleys of 12-lead ECG data were these 24 properties that were input characteristics of three LVH-detection models developed for this study. Obtaining optimal prediction performances requires a neural network with a certain number of hidden layers and nodes, but there is no solid design guideline for this task at the moment. After studying their detection capabilities and doing multiple experiments, we chose to adopt a single hidden layer with 26 nodes for the hidden layer structure of the BPN model. The presence or absence of LVH is also displayed by the output layer.

The data mean and standard deviation for each feature were set to 0 and 1, respectively, prior to model training, using the z-score normalization56. Using epoch = 400 and tenfold cross validation, the BPN model was trained.

F-statistic for determining if a feature is statistically significant. Several features extracted from 12-lead ECG data can be incorporated into the LVH detection model design. It is also possible to rank important characteristics and establish the feature importance between LVH and non-LVH groups using the F-statistic values of all features. When building the BPN model for LVH detection, this study used the F-statistic, which is defined as Eq. (2), to determine the relevance of each feature.:

$$F = \frac{\sum_{g=1}^{G} n_g(\bar{f}_g - \bar{f})^2/(G-1)}{\sum_{g=1}^{G}\sum_{h=1}^{n_g}(f_{hg} - \bar{f}_g)^2/(n-G)}$$

with n samples (each with p features) and ng representing the number of samples in the g-th group, f g denoting the mean of the g-th group, f denoting the mean of all samples, and f h denoting the feature value of the h-th sample in the g-th group. The variances within and between groups are contained in the numerator and denominator, respectively, of Eq. (2). When the F-statistic for a feature is high, it means that the LVH and non-LVH groups are significantly different.

Findings from the study

Validation of detection R-peaks and S-valleys. Table 2 shows the calculated mean-square errors (MSEs) and Pearson correlation coefficients (PCCs) between the observed and measured amplitudes of R peaks and S valleys in each 12-lead ECG signal sequence from 30 randomly selected individuals. This is to verify the effect of the algorithm on R peak and S valley detection (as discussed in section "ECG signal processing"). Low mean squared errors and high correlation coefficients are examples of how effectively the algorithm can identify R peaks and S troughs.

ECG performance criteria for LVH assessment. As mentioned in the "LVH criteria" section, this study employed seven different types of ECG criteria for LVH assessment. Table 3 displays the findings of these criteria. It is clear that the diagnostic accuracy of the criteria was inadequate, ranging from 0.50 to 0.62, for LVH. Despite having poor sensitivity and accuracy, the Lewis criterion had excellent specificity and precision. The Cornell criteria performance showed a similar pattern.

Effectiveness of ML models for LVH identification. This research used two distinct datasets to train machine learning models for LVH identification. Separate datasets contained 12-lead ECG data from 173 LVH and 173 non-LVH cases without ECG beat segmentation (section "ECG beat segmentation"). In contrast, 1466 LVH and 1466 non-LVH cases had ECG beat segmentation applied to them.

Table 2. MSEs and correlation coefficients between detected and actual amplitudes of R peaks and S valleys in each sequence of 12-lead ECG signals of 30 individuals

Lead	MSE ($10^{-8}V^2$)	Correlation coefficient
SII	0.00112	0.991
RaVR	0.00163	0.904
SIII	0.00172	0.977
SaVF	0.00213	0.989
RV3	0.00260	0.997
RI	0.00267	0.988
SI	0.00308	0.965
RaVL	0.00390	0.966
RV6	0.00409	0.995
RV2	0.00515	0.995
RII	0.00520	0.994
SaVL	0.00537	0.986
RV5	0.00548	0.997
RV4	0.00553	0.995
SV5	0.00584	0.975
SV6	0.00598	0.962
RaVF	0.00608	0.992
RV1	0.00673	0.989
RIII	0.00929	0.988
SV4	0.01158	0.947
SV3	0.02237	0.965
SV2	0.03751	0.975
SaVR	0.04108	0.945
SV1	0.05410	0.957

Table 3. Performances of 7 methods with different ECG criteria for assessing LVH. The bold and underlined values indicate the maximum and minimum values

Criterion name	Accuracy	Precision	Sensitivity	Specificity
Cornell	0.54	0.88	0.09	0.99
Sokolow	0.57	0.69	0.26	0.88
Peguero	0.57	<u>0.60</u>	0.44	0.71
Framingham	**0.62**	0.62	**0.66**	<u>0.59</u>
Gubner	<u>0.50</u>	0.67	<u>0.01</u>	0.99
Sumof12leads	**0.62**	0.65	0.53	0.71
Lewis	0.52	**1.00**	0.03	**1.00**

Between 173 individuals who had LVH and 173 individuals who did not. Once, there was a 7:3 ratio of training data to testing data. Three models' LVH detection capabilities are shown in Table 4.After utilizing ECG beat segmentation, all three models' related detection performances—specificity, accuracy, precision, and sensitivity—improved, with the exception of the k-means model. With accuracy(0.961), precision(0.958), sensitivity(0.966), and specificity (0.956), the BPN model outperforms the other two models that use ECG beat segmentation for detection.

Figure 3 shows the growth scenario of our decision tree model with ECG beat segmentation. This shows how similar the model is to the present ECG criteria for LVH diagnosis and how easy it is to understand for LVH identification. The seven nodes of the top three layers of the decision tree model utilize four-lead signals, which are likewise utilized by six of the seven ECG criteria listed in Table 1. The four-lead signals being referred to are RV4, RV5, RIII, and RI. RV4 and RV5 were also used in the Framingham criteria, the Sokolow-Lyon and Framingham criteria, the Lewis criterion, and the Gubner and Lewis criteria, respectively. Peguero and the total of the 12 leads, two of the seven criteria for measuring LVH listed in Table 1, naturally take into consideration all of the 12-lead ECG data.

The importance of characteristics. All twenty-four characteristics, namely the R peak and S valley amplitudes, retrieved from the ECG signal processing method's output (section "ECG signal processing") were utilized to construct machine learning models for the present study. Eq. (2) states the significance level for each feature, which was evaluated using the F-statistic57,58. Table 5 displays the ordered results of the computed F-statistics for each feature. (Since it examines all 24 features obtained from 12-lead ECG data, we disallow the sum-of-12-leads technique here.) Some of the criterion approaches in Table 1 use five out of the top six features (not including RaVF) with the highest F-statistic values.
A BPN model's training and testing phases involved sequentially removing features from consideration according to their F-statistic values. We determined the relationship between the number of adopted attributes and the accuracy of the updated BPN model, as demonstrated in Figure 4. After narrowing the features down to seven, the accuracy decreased by less than 10%. The results demonstrated that the top seven characteristics listed in Table 5 were most strongly linked to LVH.

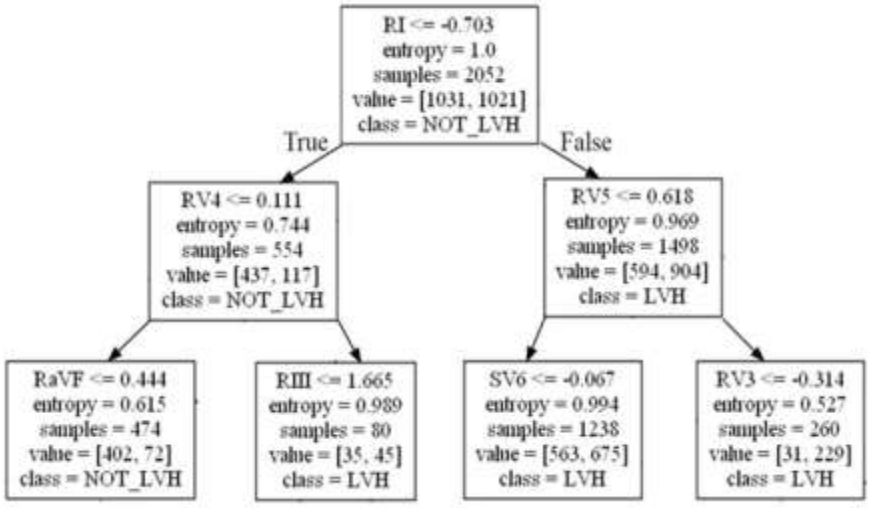

Figure 3: The first three layers of the decision tree after ECG beat segmentation.

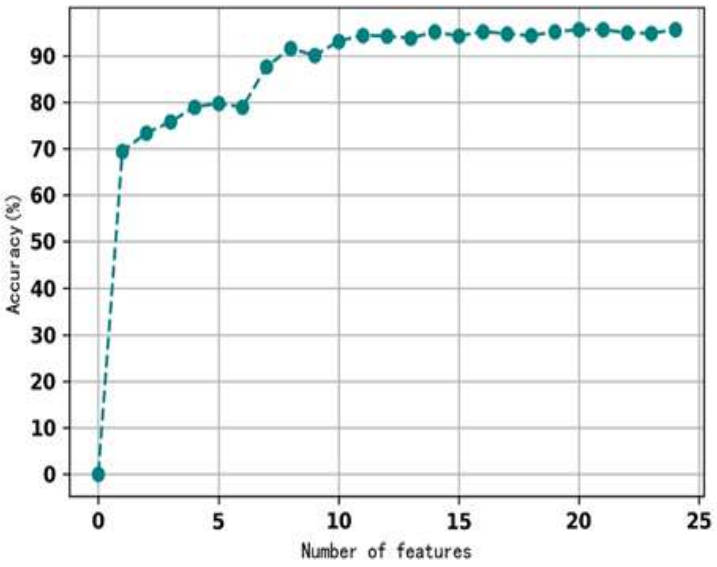

Figure 4: Accuracy variation of the number of features

DISCUSSIONS

Proficiency in detecting in this study. This section details the methods and results of the machine learning models used to assess left ventricular hypertrophy, with an emphasis on the criteria-based approaches. Table 1 shows that there are seven different methods, with 2, 3, 12, 9, 2, 12, and 2 leading to each of them. Only two or three leads of 12-lead ECG signals are considered by four of the seven criterion approaches based on ECG when LVH is being evaluated. The other three criteria mentioned earlier—Peguero, Framingham, and Sum of 12 leads—consider all or nine leads when assessing LVH. Essentially, the evaluation criterion is thresholded discrepancies. In contrast to machine learning models like BPN employed in this study, more comprehensive recognition for evaluating LVH can be integrated into the model during training using complete 12-lead ECG signals. In addition, we utilized the ECG signal processing method (described in section "ECG signal processing") to segment the ECG beats and automatically determine the R-peak and S-valley amplitudes before building machine learning models for the present study.

The LVH is a subtle indicator of ventricular overload, which can be detected with electrocardiogram (ECG) methods such the voltage criteria11. Nevertheless, the amount of indirect manifestation in the R peak and S valley of an LVH subject's ECG cycles might differ slightly, and the specifics of the waveforms from different ECG cycles of the same subject can also differ11. Using the single-cycle ECG signal alone to determine if a person is LVH, as the seven methods in Table 1 often do, might not be comprehensive enough. The aforementioned description can be one of the typical reasons why the seven ECG criteria shown in Table 1 are not very good at detecting LVH. Regardless of the number of ECG cycles averaged for any criterion-based strategy, the average process may change the connection between the R peak and S valley of each cycle.

The flip side is that we tried to build an LVH identification model using precise ECG cycle signals and the hidden information in each LVH participant's ECG cycle. Section "ECG beat segmentation" explains how we were able to get an average of 8.47 times as many ECG cycles from each LVH subject following the procedure. With an average increase of 7.47 cycles, these extra ECG cycles may include more information specifics. So, after ECG beat segmentation, we can train the BPN model using the more comprehensive and detailed dataset that contains 8.47 times as many ECG cycles (1021 positive and 1031 negative cases). Accordingly, the BPN model may learn more about LVH and non-LVH from the larger dataset. Consequently, the generated BPN model demonstrated enhanced detection capabilities throughout the test phase. Table 4 shows that a BPN model with detection outcomes better than a BPN model without beat segmentation was constructed using the same dataset (118 positive and 124 negative cases), and this explains why beat segmentation is important. Furthermore, this clarifies why the seven previously employed methods (Table 3) using criteria established from fewer leads and a single cycle of ECG signals considerably underperformed when compared to the BPN model's detection results (the penultimate row in Table 4) using 12-lead ECG signals and ECG beat segmentation..

The article compares current methods for detecting LVH using electrocardiograms. Table 6 compiles the results of previous studies that have described innovative methods or machine learning models for LVH identification using ECG signals. Here is an overview of the seven most important studies' criterion approaches based on ECG signals (12,14-18,23). In order to detect LVH in Chinese individuals older than 65 years old, the Sokolow-Lyon, Cornell, and Cornell Product (CP) criteria were employed in Ref. 16. It would not make a difference if CP got the top result (AUC = 0.62). According to Ref. 17, LVH can be detected using a combination of six different ECG criteria methods. The detection performances for LVH in a hypertension group were slightly greater using this combo method compared to the general population. Over the Cornell voltage, Sokolow-Lyon voltage, and Romhilt-Estes criteria, the Peguero-Lo Presti criteria outperformed them for the detection of LVH in patients aged 70 years or older and 83.1% with hypertension (Ref. 18). They were somewhat better than our results in Table 3 (Peguero) in terms of accuracy, sensitivity, and specificity (0.665, 0.519, and 0.821, respectively), which met the Peguero-Lo Presti criterion.

On the other hand, summarizes the machine learning techniques for LVH detection using ECG data that have been reported in previous studies (20-22, 24-27). In Ref. 27, the random forest model was constructed using ECG data together with K-nearest neighbor and Z-score for LVH identification. In reference 20, the BART-LVH criterion was suggested for LVH identification. Age, sex, height, systolic and diastolic blood pressures, and 21 electrocardiogram variables were among the 26 attributes used to construct it using Bayesian additive regression trees. In reference 21, we saw a decision tree model for LVH identification that used six features from ECG signals and logistic regression to reduce dimensionality.

In reference 24, a model for detecting LVH was trained using 87 ECG features and then constructed as a six-layer deep neural network (DNN). Referring back to Ref. 25, we see an ENN model that incorporates CNN and DNN for LVH forecasting. The ECG characteristics and demographic information (height, weight, age, and sex) were sent to the DNN, while the 8-s 12-lead ECG signals were received by the CNN (two-dimensional data, 4000 × 12). After that, the results of the CNN and DNN were combined to create the LVH detection output.

A convolutional neural network (CNN) was trained using the full 12-lead electrocardiogram (ECG) waveform, which lasts for 10 seconds, in order to quantify the mass of the left ventricle (LV) (Ref. 26). The LV mass as measured by cardiac MRI served as the gold standard. Thanks to their CNN model, LVH can be predicted indirectly. In 2021, Ref.22 released a number of LVH detection models that were generated using four different machine learning algorithms. Among the models tested, GLMNet (penalized logistic regression with the ElasticNet penalty) achieved the highest area under the curve (AUC) of 0.873, indicating the best detection performance. Ten more features retrieved from ECG signals and twenty-four features of R peaks and S valleys derived from 12-lead ECG signals were utilized to construct the GLMNet model.

Earlier studies indicated a number of other LVH detection strategies, and Table 6 summarizes their detection capabilities. Results for detection ranged from 0.29 to 0.96 percent, from 0.34-0.446 percent, and from 0.661 to 0.851 percent, according to the aforementioned machine learning models proposed in previous studies. Previous research found substantial variation in the sensitivity, specificity, and accuracy of these models. With scores of 0.961 for accuracy, 0.958 for precision, 0.966 for sensitivity, and 0.956 for specificity, our BPN model produced reliable predictions. Table 6 lists seven models that have been recorded in other studies (20-22, 24-27) and compares them to our BPN model when it comes to identifying LVH using ECG data.

Also, they discovered that machine learning outperformed multiple conventional ECG criterion approaches for LVH detection (Refs. 12, 20, 22, 24-26). As shown in Table 6, our BPN model outperforms the criteria employed in previous studies in terms of accuracy, precision, sensitivity, and specificity. Our BPN model also had more consistent and improved accuracy, precision, sensitivity, and specificity. Possible explanations include the following: (1) utilizing a balanced dataset to design our BPN model, which may reduce the gap between the specificity and sensitivity of a model's predictive performances (46,59); (2) preprocessing a 12-lead ECG to reliably and automatically extract the amplitudes of R peaks and S valleys; and (3) training our BPN model with an abundance of ECG cycles (8.47 cycles per sequence on average), which enabled the model to learn and detect subtle differences among each subject's ECG cycles.

CONCLUSIONS

For the purpose of creating the machine-learning model, we used a straightforward and practical preprocessing technique for RS amplitude detection and ECG beat segmentation to extract 24 features from a 12-lead ECG. As a result, the developed BPN model may achieve exceptional LVH prediction performance and surpass previously published models and standards.

REFERENCES

[1]. Vakili, B. A., Okin, P. M. & Devereux, R. B. Prognostic implications of left ventricular hypertrophy. Am. Heart J. 141(3), 334–341 (2001).
[2]. Schmieder, R. E. End organ damage in hypertension. Dtsch. Arztebl. Int. 107(49), 866 (2010).
[3]. Roush, G. C. et al. Hydrochlorothiazide vs chlorthalidone, indapamide, and potassium-sparing/hydrochlorothiazide diuretics for reducing left ventricular hypertrophy: A systematic review and meta-analysis. J. Clin. Hypertens. 20(10), 1507–1515 (2018).
[4]. Hollenberg, N. K. Management of hypertension and cardiovascular risk. Am. J. Med. 90(2), S2–S6 (1991).
[5]. Armstrong, A. C. et al. LV mass assessed by echocardiography and CMR, cardiovascular outcomes, and medical practice. JACC Cardiovasc. Imaging 5(8), 837–848 (2012).
[6]. Casale, P. N. et al. Electrocardiographic detection of left ventricular hypertrophy: Development and prospective validation of improved criteria. J. Am. Coll. Cardiol. 6(3), 572–580 (1985).
[7]. Sokolow, M. & Lyon, T. P. The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads. Am. Heart J. 37(2), 161–186 (1949).
[8]. Molloy, T. J., Okin, P. M., Devereux, R. B. &Kligfield, P. Electrocardiographic detection of left ventricular hypertrophy by the simple QRS voltage-duration product. J. Am. Coll. Cardiol. 20(5), 1180–1186 (1992).
[9]. Liu, C.-W. et al. The dose–response effects of uric acid on the prevalence of metabolic syndrome and electrocardiographic left ventricular hypertrophy in healthy individuals. Nutr. Metab. Cardiovasc. Dis. 29(1), 30–38 (2019).
[10]. Liu, C.-W. et al. Hyperuricemia is associated with a higher prevalence of metabolic syndrome in military individuals. Mil. Med.183(11–12), e391–e395 (2018).
[11]. Peguero, J. G. et al. Electrocardiographic criteria for the diagnosis of left ventricular hypertrophy. J. Am. Coll. Cardiol. 69(13), 1694–1703. https://doi.org/10.1016/j.jacc.2017.01.037 (2017).
[12]. Chen, Y. et al. Performance of a novel ECG criterion for improving detection of left ventricular hypertrophy: A cross-sectional study in a general Chinese population. BMJ Open 11(9), e051172 (2021).
[13]. Maanja, M., Schlegel, T., Kozor, R., Bacharova, L., Wong, T. C., Schelbert, E. B. &Ugander, M. Improved evaluation of left ven- tricular hypertrophy using the spatial QRS-T angle by electrocardiography. medRxiv (2022).
[14]. Tavares, C. A. M. et al. Usefulness of ECG criteria to rule out left ventricular hypertrophy in older individuals with true left bundle branch block: An observational study. BMC Cardiovasc. Disord. 21(1), 1–9 (2021).
[15]. Lv, T. et al. The association between ECG criteria and Echo criteria for left ventricular hypertrophy in a general Chinese population.
[16]. Ann. Noninvasive Electrocardiol. 26(5), e12880 (2021).
[17]. Zhang, W. et al. Consistency of left ventricular hypertrophy diagnosed by electrocardiography and echocardiography: The Northern Shanghai Study. Clin. Interv. Aging 14, 549 (2019).
[18]. Kühl, J. T. et al. Left ventricular hypertrophy identified by cardiac computed tomography and ECG in hypertensive individuals: A population-based study. J. Hypertens. 37(4), 739–746 (2019).
[19]. Tavares, C. D. A. M. et al. Clinical applicability and diagnostic performance of electrocardiographic criteria for left ventricular hypertrophy diagnosis in older adults. Sci. Rep. 11(1), 1–10 (2021).
[20]. Keskin, K. et al. Assessment of a new electrocardiographic criterion for the diagnosis of left ventricle hypertrophy: A prospective validation study. North. Clin. İstanbul 7(3), 231 (2020).