Automatic Right Ventricular Hypertrophic Detection Integrating Electrocardiography-based QRS Biomarkers with Machine Learning

Marion Taconné[1], Valentina D A Corino[1], Luca Mainardi[1]

[1] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy

Abstract

Despite its widespread use, the electrocardiogram (ECG) exhibits limited sensitivity in detecting right ventricular hypertrophy (RVH), primarily due to the overshadowing effects of the left ventricular activation. This study addresses this diagnostic challenge by extracting morphological QRS biomarkers from 12-lead recordings to classify RVH patients versus healthy controls. Leveraging a publicly available database comprising 9,001 patients (101 RVH, 8,900 control), we extracted features including width, amplitudes, slopes between fiducial points, and Hermite transform coefficients. Utilizing logistic regression, random forest, and support vector machine algorithms following sequential feature selection, our classifiers achieved a minimum accuracy of 88% on an independent validation dataset of 1,456 individuals (69 RVH, 1,387 control). Notably, logistic regression and random forest demonstrated valuable sensitivity, reaching 85% and 87%, respectively. The three or four selected features align with clinical recommendations, underscoring their potential utility in enhancing RVH detection via ECG biomarkers, driven by machine learning algorithms.

1. Introduction

The electrocardiogram (ECG) is often the initial diagnostic tool employed in patient screening, given its widespread availability, affordability, and repeatability. However, its utility in detecting right ventricular hypertrophy (RVH) is limited due to several factors. Primarily, the sensitivity of the ECG for detecting RVH is relatively low [1]. This limitation stems from the anatomical positioning of the right ventricle and the predominance of the left ventricular activation vector, which tends to overshadow RV forces on the ECG [2]. Consequently, the current recommendations lack a standardized set of criteria for the definitive diagnosis of RVH using ECG alone.

Furthermore, the diagnostic reliability of ECG is hindered by the absence of consensus regarding the most effective criteria for identifying RVH. Despite its widespread use, ECG's efficacy in RVH diagnosis is notably inferior to imaging modalities such as echocardiography, magnetic resonance imaging, or more recently, the three-dimensional echocardiography. These techniques offer superior accuracy and are less prone to the confounding effects of various comorbid conditions, including congenital heart disease, valvular heart disease, and chronic pulmonary disease. However, all imaging modalities are more expensive and time-consuming than an ECG recording.

Given these challenges, exploring alternative approaches for RVH detection, particularly leveraging ECG biomarkers, is necessary. Besides classical biomarkers such as width, amplitudes, and slopes between the fiducial points, features based on Hermite transform have shown interest [3]. Therefore, the goal of our work is to analyze QRS biomarkers extracted from 12-lead recordings to perform classification, highlighting the most discriminative features in differentiating RVH patients from a healthy population.

2. Methods

2.1. Database

PTB-XL ECG database from Physikalisch-Technische Bundesanstalt, Brunswick, Germany, were used in this study [4] for model building. This publicly available database contains clinical 12-lead ECGs from 18,885 patients of 10s in length, sampled at 500Hz. From this database, we extracted ECGs labeled as 'NORM' and 'RVH'. We exclude RVH patients with other hypertrophic labels and check if patient labeled as normal does not present hypertrophic label either. Due to bad ECG quality (less than 3 beats not correlate with the median beat: ¡ 0.8), some patients were also excluded. After these steps, 101 RVH and, 8900 control patients were included in the training and test phase of the study.

As validation database, the Georgia 12-lead ECG Challenge Database from the Emory University, Atlanta, Georgia, USA was used. Similarly, it is composed of 10s length ECGs, sampled at 500Hz. The same selection approach was applied with control patients labeled as 'sinus rhythm'.

To make the validation more realistic, patients with RVH label which also have LVH hypertrophic or atrial enlargement label were included. Finally, 69 RVH patients and 1387 controls were included in the validation database.

Both databases are uploaded on the Physionet challenge database repository [5].

2.2. QRS Biomarkers Extraction

In order to perform classification, several representative features were extracted from the 12-lead ECG signal.

2.2.1. Morphological Biomarkers

Several morphological biomarkers were computed directly from the 12-lead ECG signals (Fig.1), including:

- Median fiducial point amplitude: P, Q, R, S, J, T.
- Ratio of median fiducial amplitudes: R/P and R/T.
- Median of the interval length: PR, PS, PT, QT, QRS, RS, ToT, TTe.
- Slope ascending and descending of the T wave.
- negative percentage of QRS.

These 19 biomarkers were automatically computed for all the 12 leads after a step of fiducial points localization. Both the fiducial points localization and the computation of the average QRS for each lead were done thanks to the open-source ECGDeli package [6] executed on MATLAB R2023b.

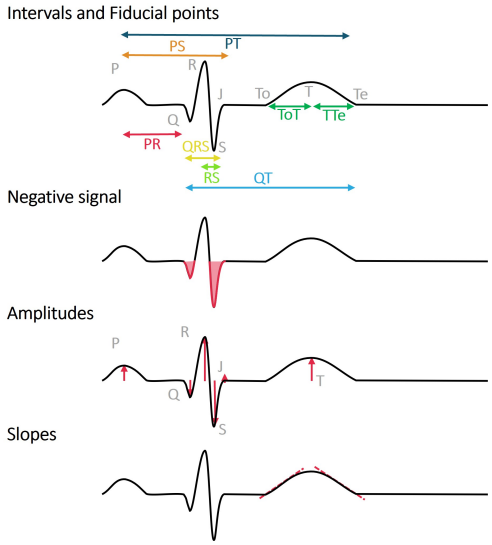

Figure 1: Morphological features extraction: Interval width, negative % of the signal, amplitude of the fiducial points and slopes of the waves.

2.2.2. Hermite transform

After creating an average QRS complex for each lead, we computed its Hermite transform [3]. These mathematical functions are able to provide a compact description of the QRS morphology. Indeed, thanks to their shape, they are able to recover the majority of the QRS waveforms. Few Hermite functions are needed to describe the QRS. Figure 2 represents the first four Hermite functions (Φ_i) used in this study. Each QRS complex is approximated by a linear combination q of these functions:

$$q(t) = \sum_{i=0}^{N=3} a_i \Phi_i(t) \tag{1}$$

where a_i are the coefficient of the linear combination. These Hermite coefficients were extracted for each lead and each patient of the database.

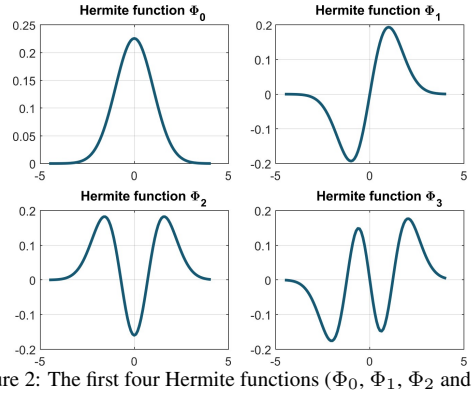

Figure 2: The first four Hermite functions (Φ_0, Φ_1, Φ_2 and Φ_3).

In addition, the root mean square error (RMSE) between the approximated QRS ($q(t)$) and the original QRS signal ($s(t)$) was computed on the 12 leads.

$$RMSE(s,q) = \sqrt{\frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (s(t_i) - q(t_i))^2} \tag{2}$$

With the 4 first Hermite coefficients and the RMSE computed on the 12 leads, (19+5)x12 leads = 288 biomarkers were in total extracted.

2.3. Machine Learning Algorithms

RVH classification was performed using three supervised machine learning algorithms: logistic regression (LogReg), random forest (RF) [7] and support vector machine (SVM) [8]. Figure 3 sums up the different steps applied for the three classifiers. The PTB-XL was separated into training and testing sets encompassing the 70% and 30% of cases respectively. Due to the imbalance of the database, the majority label of the training set was randomly downsampled to be equal to the minority label.

2.3.1. Feature Selection Process

The sequential floating forward selection (SFFS) process was applied on the training set thanks to the Python

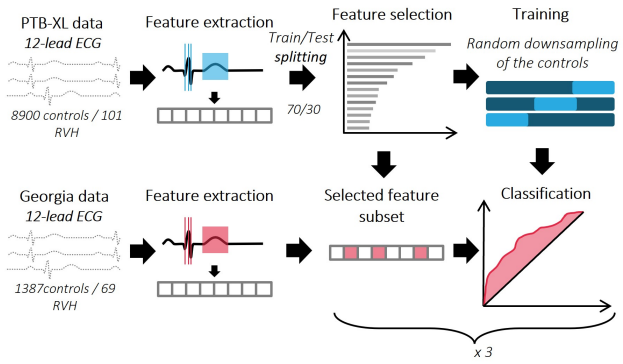

Figure 3: Methodological steps separated in training-test (top) and validation phase (bottom).

library Scikit-Learn [9, 10]. At each stage, the best feature was added based on the cross-validation score of the classifier: the area under the curve (AUC). The Monte Carlo cross-validation (CV) was done with 5 folds on the training set and the maximum of features was set to 30. The optimal combination was then selected as the best CV AUC score after the cycle of forward and backward selection repeated until the maximum of features was reached.

2.3.2. Model training

After the step of feature selection, the cross-validation method was redone 100 times on the PTB-XL database. At each iteration, a new downsampling of the majority label in the training set was done. Since the result appear satisfying in the PTB-XL test database, a final model was trained (after downsampling of the majority label) for each of the three classifiers on the complete database and validated on the Georgia database.

3. Results

Hermite Approximation: As all the other biomarkers, Hermite approximation of the QRS of each lead was computed for each patient. All the approximations do not fit perfectly the original ECG signal (a limited number of 4 Hermite functions was used) for all the patients, but this information is included in the RMSE score. Figure 4 presents two Hermite approximation for two QRSs, with different RMSE.

We can notice that Hermite functions better fit on signal without any fragmentation on the QRS. The simpler the QRS morphology is, the lower is the fitting error.

Feature Selection Process: Table 1 shows the features selected after the SFFS. Only few features were selected and for the three algorithm, the amplitude of the R peak in lead V1 appears as the main one. In second position, there is the RMSE of the Hermite approximation on lead II for the logistic regression and on lead aVR for the RF and

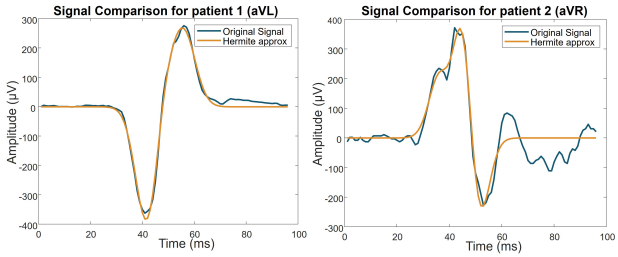

Figure 4: Two example of Hermite approximation for two RVH patients on different lead.

SVM. Then the T wave is represented with the descending and ascending slope respectively for the logistic regression and the RF and by its amplitude for the SVM. SVM have another feature selected at the third place: the amplitude of the S point on lead I. This reduced number of features selected by SFFS bodes well for the generalization ability of the future classifiers.

N	LogReg	RF	SVM
1	R amplitude (V1)	R amplitude (V1)	R amplitude (V1)
2	Herm RMSE (II)	Herm RMSE (aVR)	Herm RMSE (aVR)
3	T slope des (II)	T slope asc (I)	S amplitude (I)
4	-	-	T amplitude (V2)

Table 1: Features selected by the SFFS for the 3 classifiers: Logistic Regression (LogReg), Random Forest (RF) and Support Vector Machine (SVM).

Validation: After being trained on the PTB-XL database, the model was applied on the Georgia database. The validation step was repeated 100 times to verify that the random downsampling has no major effect on the training. Figure 5 provides the obtained mean ROC curves for the three type of algorithm and their standard deviation. The RF classifier presents the higher standard deviation through the repetition of training that differ due to the downsampling.

However, it is a classifier with very good performance with an accuracy of 88.0±3.6% with a sensitivity in predicting RVH of 87.0±3.1% and a specificity of 88.1±3.8%. The logistic regression classifier and the SVM have respectively an accuracy of 95.9±0.7% and 97.0±0.0%, sensitivity of 84.5±1.3% and 58.0±0.0% and specificity of 96.5±0.7% and 99.0±0.0%

4. Discussion

As previously mentioned, the reduced number of features selected for the classification produce particularly good results and underline the models' generalization capacity and the absence of overfitting.

Moreover, the features automatically selected by the SFFS process well connect with the one proposed in clinical recommendations. In fact, in the literature, numerous criteria are derived from the R amplitude in V1 but also of S [2]. So it is not a surprise to find this criterion on

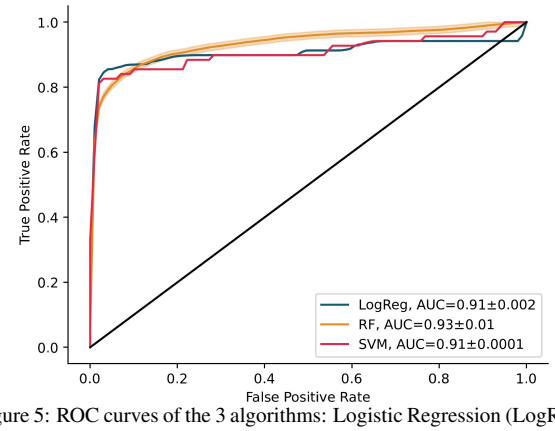

Figure 5: ROC curves of the 3 algorithms: Logistic Regression (LogReg), Random Forest (RF) and Support Vector Machine (SVM), on the validation database.

top of each SFFS. These tall R waves in right precordial leads suggest pressure overload and could be associated with right axis deviation, as well discussed in the literature. In third position for LogReg and RF and in fourth for SVM, we found the T waves. This matches with the frequently associated ST depression and T wave inversion clinically observed in these patients.

With a higher fitting error observed for patients with RVH, which ranks second in the feature selection list, we can hypothesize that this error might indicate the presence of fragmented QRS. The presence of fragmented QRS represents distortion of signal conduction and depolarization, which is related to myocardial scar or myocardial fibrosis.

Although the SVM algorithm provides the best accuracy score, its lack of sensibility made it unsuitable for further studies. However, the two other algorithms present very interesting specificity and sensitivity scores. The quality of the results adding to the reduced number of features used and extracted automatically has to be noticed and augurs well for future studies.

One of the main limitation is the lack of details on the labelization process of both of the database.

Future work should consider a multi-class classification, introducing left ventricular hypertrophy as well as valvular overload or enlargement. Moreover, integrating more complex patient diagnosis with multiple diagnosis labelization must be considered. More complex classification methods could also be investigated.

5. Conclusion

12-lead ECG from two major publicly available databases were analyzed to proposed a sensitive classifier for RVH detection. The algorithms focused on the ECG morphology based on classical extracted biomarker and Hermite function derived features. Classification results achieved a minimum accuracy of 88% with two algorithms that reached sensitivity of 85%. This finding highlights the benefice of machine learning to gather clinical findings in the ECG-based automatic detection of RVH.

Acknowledgment

This research study is part of the Project SMASH-HCM funded by the European Union under GA #101137115

References

[1] Nikus K, Pérez-Riera AR, Konttila K, Barbosa-Barros R. Electrocardiographic recognition of right ventricular hypertrophy. Journal of Electrocardiology 2018;51(1):46–49. ISSN 15328430.

[2] Hancock EW, Deal BJ, Mirvis DM, Okin P, Kligfield P, Gettes LS. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: Part V: Electrocardiogram changes associated with cardiac chamber hypertrophy: A scientific statement from the American Heart Association Electrocardiography. Circulation 2009;119(10). ISSN 00097322.

[3] Sörnmo L, Börjesson PO, Nygårds ME, Pahlm O. A Method for Evaluation of QRS Shape Features Using a Mathematical Model for the ECG. IEEE Transactions on Biomedical Engineering 1981;BME-28(10):713–717. ISSN 15582531.

[4] Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset. Scientific Data dec 2020; 7(1). ISSN 20524463.

[5] Reyna MA, Alday EA, Gu A, Liu C, Seyedi S, Rad AB, Elola A, Li Q, Sharma A, Clifford GD. Classification of 12-lead ECGs: The PhysioNet/Computing in Cardiology Challenge 2020. Computing in Cardiology 2020;2020-Septe. ISSN 2325887X.

[6] Pilia N, Nagel C, Lenis G, Becker S, Dössel O, Loewe A. ECGdeli - An open source ECG delineation toolbox for MATLAB. SoftwareX 2021;13(100639). ISSN 23527110.

[7] BREIMAN L. Random Forests LEO. Machine Learning 2001;45:5–32.

[8] Chang CC, Lin CJ. LIBSVM: A Library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2001;2(3):1–40. ISSN 21576904.

[9] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dibourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 2011; 12(9):2825–2830.

[10] Ferri FJ, Pudil P, Hatef M, Kittler J. Comparative Study of Techniques for Large-Scale Feature Selection. Machine intelligence and pattern recognition 1994;16:404–413.

Address for correspondence:

Marion Taconné (marionhelene.taconne@polimi.it)
Politecnico di Milano, Building 21, Via Camillo Golgi 39, 20133, Milan, Italy