






From signal averaging to single-beat analysis: a statistical learning approach for ventricular late potential detection

 Andrej Novak^{1,2,3,4,*},
 Ivan Zeljković^{1,5},
 Ante Lisičić¹,
 Ana Jordan¹,
 Nikola Pavlović¹,
 Šime Manola¹

¹Dubrava University Hospital, Zagreb, Croatia

²University of Zagreb, Faculty of Science, Department of Physics, Zagreb, Croatia

³University of Vienna, Department of Mathematics, Vienna, Austria

⁴Luxembourg School of Business, Luxembourg, Luxembourg

⁵Catholic University of Croatia, Zagreb, Zagreb, Croatia

KEYWORDS: ventricular late potentials, time-series classification.

CITATION: *Cardiol Croat.* 2024;19(11-12):404. | <https://doi.org/10.15836/ccar2024.404>

***ADDRESS FOR CORRESPONDENCE:** Andrej Novak, Klinička bolnica Dubrava, Avenija Gojka Šuška 6, HR-10000 Zagreb, Croatia. / Phone: +385-91-2998-616 / E-mail: andrej.novak@phy.hr

ORCID: Andrej Novak, <https://orcid.org/0000-0002-7828-4870> • Ivan Zeljković, <https://orcid.org/0000-0002-4550-4056> • Ante Lisičić, <https://orcid.org/0000-0002-4365-9652> • Ana Jordan, <https://orcid.org/0000-0001-5610-6259> • Nikola Pavlović, <https://orcid.org/0000-0001-9187-7681> • Šime Manola, <https://orcid.org/0000-0001-6444-2674>

Introduction: Ventricular late potentials (VLPs) are small, high-frequency signals found at the end of the QRS complex, often associated with areas of fibrosis in the heart. Detecting VLPs is crucial for identifying patients at risk of ventricular arrhythmias and sudden cardiac death. Traditional methods, such as signal-averaged electrocardiograms (SAECGs), rely on the aggregation of multiple beats, but these techniques have limitations in sensitivity and specificity.¹⁻³ *Aim:* To develop and evaluate statistical learning models capable of detecting VLPs from a single cardiac beat, potentially reducing the need for time-consuming signal-averaging methods and improving the accuracy of real-time VLP identification.

Methods: We employed a range of statistical learning models—Decision Tree, Random Forest, Time-Series Forest, and MiniRocket—to classify ECG beats as containing VLPs or not. The dataset consisted of 4500 beats across three leads (II, V2, V6), with equal representation of beats containing synthetic VLPs. Feature extraction included both time-domain and frequency-domain features, with special emphasis on the Fast Fourier Transform (FFT) and feature engineering to enhance model performance. The interpretability of the models was assessed using SHapley Additive exPlanations (SHAP) to analyze feature importance across all models.

Results: The Random Forest model achieved the highest accuracy, outperforming other models across all leads, with particularly strong performance in lead V2, where it reached an accuracy of 97.9%. The Decision Tree and TimeSeries Forest models also demonstrated reasonable performance, with the Decision Tree achieving an accuracy of 84.7% on lead V2 and the TimeSeries Forest showing a lower but consistent performance, with an accuracy around 50% across all leads. MiniRocket, while fast, showed the least consistent results, especially in capturing the subtle features associated with VLPs. Feature importance analysis revealed that frequency-domain features, particularly those derived from the FFT and its first derivative, were the most influential in detecting VLPs across models. SHAP further confirmed that these features had the greatest impact on model predictions, particularly in distinguishing beats containing VLPs from those without. Models trained on individual leads consistently outperformed those trained on all leads combined, with lead-specific characteristics playing a significant role in improving classification accuracy.

Conclusion: Machine learning models, particularly Random Forest, show strong potential in detecting VLPs from single cardiac beats. Training models on individual leads yields better classification performance than using combined leads, confirming that lead-specific characteristics are critical for accurate VLP detection. The findings suggest that incorporating frequency-domain features, especially those derived from FFT, is essential for enhancing model performance. This study demonstrates the feasibility of moving from traditional signal averaging to real-time, single-beat VLP detection using advanced statistical learning approaches, offering a more efficient and accurate method for risk stratification in patients prone to ventricular arrhythmias.

LITERATURE

- Gatzoulis KA, Arsenos P, Trachanas K, Dilaveris P, Antoniou C, Tsiachris D, et al Signal-averaged electrocardiography: Past, present, and future. *J Arrhythm.* 2018 May 28;34(3):222-229. <https://doi.org/10.1002/joa3.12062>
- Zandi AS, Moradi MH. Quantitative evaluation of a wavelet-based method in ventricular late potential detection. *Pattern Recognition.* 2006 Jul 1;39(7):1369-79. <https://doi.org/10.1016/j.patcog.2006.01.012>
- Pavlov M, Barić D, Novak A, Manola Š, Jurin I. From statistical inference to machine learning: A paradigm shift in contemporary cardiovascular pharmacotherapy. *Br J Clin Pharmacol.* 2024 Mar;90(3):691-699. <https://doi.org/10.1111/bcp.15927>

RECEIVED:
October 13, 2024

ACCEPTED:
October 31, 2024

