Rule-Based ECG Quality Classification Using Spectral and Morphological Features

First A. Author, Second B. Author, and Third C. Author *School of Biomedical Engineering*
Universitat Politècnica de Catalunya
Barcelona, Spain
first.author@upc.edu, second.author@upc.edu, third.author@upc.edu

Abstract—Objective: This study presents an automated electrocardiogram (ECG) quality classification system designed for mobile health applications, where signal quality varies significantly due to electrode placement, motion artifacts, and environmental interference.

Methods: We developed a rule-based classifier using 16 interpretable quality rules derived from signal processing theory. The system employs the Pan-Tompkins algorithm for QRS detection, extracts features across seven categories (SNR, baseline, morphological, spectral, statistical, time-frequency, and lead correlation), and implements spectral parameters from the psdpar.m methodology. Classification decisions combine critical rule requirements with a 70% majority voting scheme. The classifier was evaluated on the PhysioNet/Computing in Cardiology Challenge 2011 dataset comprising 998 twelve-lead ECG recordings.

Results: Using an 80/20 stratified held-out validation approach, the classifier achieved 81.0% accuracy, 89.3% precision, 85.8% recall, 64.4% specificity, and 87.5% F1 score on the test set (200 records). Key discriminating features included signal-to-noise ratio (10.96 dB for acceptable vs. 0.68 dB for unacceptable ECGs), cardiac power ratio (0.94 vs. 0.63), and spectral centroid (10.66 Hz vs. 6.74 Hz).

Conclusion: The proposed rule-based approach provides interpretable ECG quality assessment with performance suitable for real-time mobile applications.

Significance: Unlike black-box machine learning methods, this system explains rejection reasons to clinicians, enabling targeted signal acquisition improvements in mHealth ECG devices.

Index Terms—ECG quality assessment, signal processing, Pan-Tompkins algorithm, spectral analysis, mobile health, rule-based classification

I. INTRODUCTION

THE widespread adoption of mobile electrocardiogram (ECG) devices has enabled continuous cardiac monitoring outside clinical settings. However, ECG signals acquired using portable devices frequently suffer from quality degradation due to poor electrode contact, patient movement, and electromagnetic interference [1]. Poor quality ECGs can lead to misdiagnosis or missed detection of cardiac abnormalities, presenting a significant challenge for mobile health (mHealth) applications.

The PhysioNet/Computing in Cardiology Challenge 2011 addressed this problem by tasking participants with developing automated methods to classify ECG recordings as clinically acceptable or unacceptable for interpretation [2]. The challenge dataset comprises 998 twelve-lead ECG recordings, each 10 seconds in duration, sampled at 500 Hz. Expert

TABLE I
DATASET CHARACTERISTICS

Property	Value
Total records	998
Acceptable	773 (77.5%)
Unacceptable	225 (22.5%)
Duration	10 seconds
Sampling rate	500 Hz
Leads	12-lead ECG

cardiologists labeled 77.5% of recordings as acceptable and 22.5% as unacceptable based on signal quality assessment.

Previous approaches to ECG quality classification have employed various methodologies. Xia et al. achieved 93.2% accuracy using entropy-based features with multi-stage decision rules [3]. Li and Clifford obtained 92.6% accuracy using higher-order statistical moments with support vector machines [4]. Hayn et al. developed a real-time system achieving 87.3% accuracy using SNR-based heuristics [5]. While machine learning approaches achieve higher accuracy, they often function as black boxes, providing no explanation for rejection decisions.

This paper presents a rule-based ECG quality classifier that prioritizes interpretability alongside performance. Our contributions include: (1) implementation of the Pan-Tompkins algorithm for robust QRS detection, (2) extraction of spectral parameters based on the psdpar.m methodology from biomedical signal analysis, (3) a 16-rule classification framework with physiologically motivated thresholds, and (4) validation using held-out test data to demonstrate generalization capability.

II. METHODS

A. Dataset

The PhysioNet/Computing in Cardiology Challenge 2011 dataset (set-a) contains 998 standard twelve-lead ECG recordings [2]. Each recording is 10 seconds in duration with a sampling frequency of 500 Hz (5000 samples per recording). The twelve leads comprise limb leads (I, II, III), augmented leads (aVR, aVL, aVF), and precordial leads (V1–V6).

For validation, we employed an 80/20 stratified held-out split, ensuring proportional class representation in both sets. The training set (798 records) informed threshold selection based on domain knowledge, while the test set (200 records) provided unbiased performance evaluation.

B. Preprocessing

ECG signals undergo preprocessing to remove noise while preserving diagnostically relevant information.

Bandpass Filtering: A fourth-order Butterworth bandpass filter (0.5–100 Hz) removes baseline wander below 0.5 Hz and high-frequency noise above 100 Hz, consistent with American Heart Association recommendations [6]. Zero-phase filtering (forward-backward application) prevents phase distortion:

$$y[n] = \text{filtfilt}(b, a, x[n]) \tag{1}$$

Notch Filtering: A second-order IIR notch filter at 50 Hz ($Q = 30$) eliminates powerline interference:

$$H(z) = \frac{1 - 2\cos(\omega_0)z^{-1} + z^{-2}}{1 - 2r\cos(\omega_0)z^{-1} + r^2 z^{-2}} \tag{2}$$

where $\omega_0 = 2\pi \times 50/f_s$ and r determines the notch bandwidth.

Baseline Wander Removal: A two-stage median filter with windows of 200 ms and 600 ms estimates and subtracts the baseline drift component.

C. QRS Detection

We implemented the Pan-Tompkins algorithm [7] for QRS complex detection, consisting of five processing stages:

1) **Bandpass Filter (5–15 Hz):** Isolates the QRS frequency band while attenuating P and T waves.
2) **Derivative Filter:** Highlights the steep slopes characteristic of QRS complexes:

$$y[n] = \frac{1}{8}(-x[n-2] - 2x[n-1] + 2x[n+1] + x[n+2]) \tag{3}$$

3) **Squaring:** Makes all values positive and nonlinearly amplifies large derivatives:

$$y[n] = x[n]^2 \tag{4}$$

4) **Moving Window Integration:** Smooths the signal using a 150 ms window:

$$y[n] = \frac{1}{N} \sum_{i=0}^{N-1} x[n-i] \tag{5}$$

5) **Adaptive Thresholding:** Dual thresholds adapt to signal and noise levels with a 200 ms refractory period to prevent double detection.

Fig. 1 illustrates the Pan-Tompkins processing stages for acceptable and unacceptable ECG examples.

D. Feature Extraction

Features were extracted across seven categories based on approaches from Challenge 2011 winners [3]–[5]:

1) Signal-to-Noise Ratio (SNR):

$$\text{SNR} = 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \text{ dB} \tag{6}$$

where P_{signal} is the power in QRS regions and P_{noise} is the power in non-QRS regions.

2) Baseline Features: Baseline power ratio, standard deviation, and range quantify low-frequency drift.

TABLE II
SPECTRAL PARAMETERS FROM PSD

Parameter	Description
f_{peak}	Peak frequency: $\arg\max(P_{xx})$
f_{mean}	Spectral centroid: $\sum f \cdot P_{xx} / \sum P_{xx}$
f_{median}	Median frequency at 50% cumulative power
f_{std}	Spectral spread: $\sqrt{\sum(f-\mu)^2 \cdot P_{xx} / \sum P_{xx}}$
f_{iqr}	Interquartile range: $f_{q75} - f_{q25}$
h_{Shannon}	Normalized entropy: $-\sum p\ln(p)/\ln(N)$
Kurtosis	Spectral peakiness: $\mu_4/\sigma^4 - 3$
Flatness	Wiener entropy: $\sqrt[N]{\prod P_{xx}}/(\sum P_{xx}/N)$

3) Morphological Features: Beat count, heart rate, RR interval statistics, and RR coefficient of variation (CV) assess rhythm characteristics:

$$\text{RR}_{CV} = \frac{\sigma_{RR}}{\mu_{RR}} \tag{7}$$

4) Spectral Features: Cardiac power ratio (power in 0.5–40 Hz relative to total), high-frequency noise ratio (power above 40 Hz), and spectral entropy characterize frequency content.

5) Statistical Features: Skewness, kurtosis, and zero-crossing rate capture signal distribution properties.

6) Time-Frequency Features: Using short-time Fourier transform (STFT), we compute temporal variance, stationarity index, and artifact frame ratio to assess signal consistency over time.

7) Lead Correlation: Mean inter-lead correlation quantifies consistency across the 12 leads, as all leads measure the same cardiac source:

$$r_{ij} = \frac{\text{Cov}(L_i, L_j)}{\sigma_{L_i}\sigma_{L_j}} \tag{8}$$

E. Spectral Parameters (psdpar.m)

Following the methodology from biomedical signal analysis [8], we compute 12 spectral parameters from the power spectral density (PSD) estimated via periodogram (Table II).

Fig. 2 shows PSD comparison between acceptable and unacceptable ECGs. Fig. 3 presents box plots of spectral parameters by class.

F. Classification Rules

The classifier applies 16 rules based on physiological constraints and signal quality theory (Table III).

Decision Logic: An ECG is classified as acceptable if: (1) All critical rules pass (SNR, beat count, cardiac power), AND (2) At least 70% of all rules pass.

This two-tier approach ensures that fundamental quality requirements are met while allowing tolerance for marginal failures in non-critical rules.

TABLE III
CLASSIFICATION RULES

#	Feature	Threshold	Rationale
1	SNR	$\geq$ 5.0 dB	Minimum signal quality
2	Beat count	5–30 beats	Expected for 10s recording
3	Heart rate	30–200 bpm	Physiological limits
4	RR regularity	CV $\leq$ 0.5	Rhythm consistency
5	Baseline wander	ratio $\leq$ 0.3	Low-frequency drift
6	Cardiac power	$\geq$ 50%	Power in cardiac band
7	High-freq noise	$\leq$ 30%	Power above 40 Hz
8	Artifacts	$\leq$ 20% frames	STFT artifact detection
9	Stationarity	$\geq$ 0.3	Time-frequency stability
10	Lead correlation	$\geq$ 0.3	Multi-lead consistency
11	Spectral flatness	$\leq$ 0.4	Noise-like spectrum
12	Spectral spread	$\leq$ 25 Hz	Maximum f_{std}
13	Spectral kurtosis	$\geq$ 1.0	Spectral peakiness
14	f_{mean}	$\geq$ 8.5 Hz	Spectral centroid
15	f_{std}	$\geq$ 11.0 Hz	Minimum spread
16	f_{iqr}	$\geq$ 7.0 Hz	Frequency range

TABLE IV
CLASSIFICATION RESULTS ON HELD-OUT TEST SET

	Predicted		
Actual	Accept	Reject	Total
Acceptable	133 (TP)	22 (FN)	155
Unacceptable	16 (FP)	29 (TN)	45
Total	149	51	200

Metric	Value
Accuracy	81.0%
Precision	89.3%
Recall (Sensitivity)	85.8%
Specificity	64.4%
F1 Score	87.5%

III. RESULTS

A. Classification Performance

Table IV presents the confusion matrix and performance metrics on the held-out test set (n=200).

The classifier demonstrates high precision (89.3%), indicating that ECGs classified as acceptable are reliable for clinical interpretation. The lower specificity (64.4%) reflects a conservative approach that may accept some borderline-quality recordings.

B. Feature Comparison

Table V compares feature values between acceptable and unacceptable ECG classes.

The most discriminating features are SNR (10.96 vs. 0.68 dB), cardiac power ratio (0.94 vs. 0.63), and spectral centroid (10.66 vs. 6.74 Hz). Unacceptable ECGs exhibit higher variability in all features, reflecting diverse quality degradation mechanisms.

C. Rejection Analysis

The most frequent rejection reasons were: (1) Very low SNR (signal dominated by noise): 81 cases; (2) Low spectral kurtosis (flat spectrum): 74 cases; (3) Low cardiac frequency content: 73 cases; (4) Low spectral centroid (drift-dominated): 73 cases; and (5) Low inter-lead correlation: 62 cases.

This analysis demonstrates that the classifier can provide interpretable feedback to users about specific quality issues.

IV. DISCUSSION

A. Comparison with Challenge 2011 Results

Table VI compares our approach with Challenge 2011 winners.

Our classifier achieves lower accuracy than Challenge winners but offers several advantages: (1) complete interpretability of rejection decisions, (2) no training required (thresholds based on domain knowledge), (3) fast execution suitable for real-time mobile applications, and (4) modular design allowing threshold adjustment for specific use cases.

B. Limitations

The classifier exhibits lower specificity (64.4%) compared to sensitivity (85.8%), indicating a tendency to accept borderline-quality recordings. This behavior may be appropriate for screening applications where missing acceptable ECGs is more costly than accepting marginal recordings for human review.

Fixed thresholds may not generalize optimally across different recording devices or patient populations. Future work could incorporate adaptive threshold learning while maintaining interpretability.

C. Clinical Implications

The interpretable nature of rejection reasons enables targeted interventions: if an ECG is rejected due to baseline wander, the user can be instructed to remain still; if rejected due to low SNR, electrode contact should be checked. This feedback loop is not possible with black-box machine learning approaches.

TABLE V
FEATURE VALUES BY ECG QUALITY CLASS (MEAN ± STD)

Feature	Acceptable	Unacceptable
SNR (dB)	10.96 ± 4.12	0.68 ± 12.39
Beat count	12.47 ± 3.10	14.75 ± 6.19
Heart rate (bpm)	74.50 ± 17.43	84.94 ± 33.18
RR variability (CV)	0.08 ± 0.12	0.23 ± 0.22
Cardiac power ratio	0.94 ± 0.08	0.63 ± 0.42
f_{mean} (Hz)	10.66 ± 3.86	6.74 ± 6.77
f_{std} (Hz)	14.84 ± 5.88	10.29 ± 9.82
f_{iqr} (Hz)	9.13 ± 3.46	6.00 ± 7.43
Stationarity index	0.98 ± 0.03	0.89 ± 0.29

TABLE VI
COMPARISON WITH PHYSIONET CHALLENGE 2011 WINNERS

Method	Accuracy	Approach
Xia et al. [3]	93.2%	Entropy + rules
Li & Clifford [4]	92.6%	Moments + SVM
Hayn et al. [5]	87.3%	SNR heuristics
Moody [9]	89.6%	Rule-based
This work	**81.0%**	**16 rules + spectral**

V. CONCLUSION

We presented a rule-based ECG quality classifier achieving 81.0% accuracy on the PhysioNet Challenge 2011 dataset using 16 interpretable rules derived from signal processing theory. The system implements the Pan-Tompkins algorithm for QRS detection and extracts spectral parameters following the psdpar.m methodology.

Key findings include: (1) SNR is the most discriminating feature (10.96 dB for acceptable vs. 0.68 dB for unacceptable); (2) Spectral parameters (f_{mean}, f_{std}, f_{iqr}) provide complementary discrimination; and (3) Rule-based classification enables interpretable rejection explanations.

The approach is suitable for real-time mobile health applications where interpretability, fast execution, and user feedback are priorities. Future work will explore adaptive threshold optimization and validation on additional ECG databases.

REFERENCES

[1] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.

[2] I. Silva, G. B. Moody, and L. Celi, "Improving the quality of ECGs collected using mobile phones: The PhysioNet/Computing in Cardiology Challenge 2011," in *Proc. Computing in Cardiology*, Hangzhou, China, 2011, pp. 273–276.

[3] H. Xia et al., "Computer algorithms for evaluating the quality of ECGs in real time," in *Proc. Computing in Cardiology*, Hangzhou, China, 2011, pp. 369–372.

[4] Q. Li and G. D. Clifford, "Signal quality and data fusion for false alarm reduction in the intensive care unit," *J. Electrocardiol.*, vol. 45, no. 6, pp. 596–603, Nov. 2012.

[5] D. Hayn, B. Jammerbund, and G. Schreier, "Real-time visualization of signal quality during mobile ECG recording," in *Proc. Computing in Cardiology*, Hangzhou, China, 2011, pp. 357–360.

[6] P. Kligfield et al., "Recommendations for the standardization and interpretation of the electrocardiogram," *Circulation*, vol. 115, no. 10, pp. 1306–1324, Mar. 2007.

[7] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.

[8] A. Torres, "psdpar.m - Power spectral density parameters," IBEC, Barcelona, Spain, BSA Course Materials, 2024.

[9] B. E. Moody, "A rule-based method for ECG quality control," in *Proc. Computing in Cardiology*, Hangzhou, China, 2011, pp. 361–363.