ML PROJECT ON ARRHYTHMIA DETECTION

Heartbeat Classification Fusing Temporal and Morphological Information

via Ensemble of Classifiers

Published in Elsevier Journal

Sidhant Mohanty (230002069) - covering slides 6-11

Piyush Raj (230002051) – covering slides 1-5

PROBLEM DEFINITION

- Our goal is to **automate** arrhythmia detection via ML techniques. However, two main challenges arise.
- First, there's **class imbalance**: about 90% of beats in the MIT-BIH dataset are perfectly normal, and fewer than 10% are abnormal. A classifier could simply call every beat 'normal' and still be over 90% accurate—which isn't helpful.
- Second, we use an **inter-patient split. T**he model trains on ECGs from one group of patients and tests on completely different patients. That mimics real-world situations with brand-new situations.
- We discuss further how we tackled these challenges.

NOVELTY OF THE WORK

The following ideas distinguishes our approach from earlier ECG classifiers:

1. Instead of one “big” SVM that ingests every feature at once, we build **five separate SVMs**, each focused on a different slice of information.

2. We perform **minimal preprocessing**, only removing slow, wandering baselines with median filters — no heavy noise-reduction or wave delineation. After each SVM makes its own predictions, we use three **fusion rules**:

   A. Product rule

   B. Sum rule

   C. Majority vote

MODEL ARCHITECTURE

- Independent One-vs-One SVM per feature.
- Features: R-R Intervals, Wavelets, HOS, LBP, Custom Morphology.
- Combine class probabilities via fusion rule.
- Final decision based on max accumulated probability.
- Elaborated further in the next slide.

1. Beat Segmentation & Preprocessing

– We locate each R-peak (that is the highest spike in the ECG) and extract a fixed window of 180 samples around it.
– We subtract a slowly varying baseline to remove drift.

2. Feature Extraction

– We compute five distinct feature sets (R-R intervals, wavelets, higher-order statistics, LBP, custom morphology).

3. Five One-vs-One SVMs

– For each feature, we train an SVM on every pair of the four beat classes (Normal, SVEB, VEB, Fusion)—that's six binary models per feature.

4. Probability Fusion

– Each SVM outputs class probability estimates. We accumulate those per class and then combine across features using product, sum, or majority vote.

5. Final Decision

– The class with the highest fused score becomes our predicted beat type."

IMPLEMENTATION DETAILS

- Our experiments use the **MIT-BIH Arrhythmia Database**, which contains 48 half-hour ECG recordings sampled at 360 Hz. We follow the standard **inter-patient split**: 22 records in **DS1** for training, and 22 in **DS2** for testing—no patient appears in both sets.
- For preprocessing, we apply two consecutive median filters—200 ms and 600 ms windows—to remove baseline wander, but we skip any additional high-frequency noise filtering to preserve waveform details.
- Each SVM uses an **RBF kernel** (Radial Basis Function), with the kernel width γ set to 1 over the number of features. We tune the penalty parameter **C** via **10-fold cross-validation** on DS1, splitting that set into ten parts, training on nine and validating on one in rotation.
- Finally, we retrain on all of DS1 with the chosen C and evaluate on DS2, mimicking deployment on new patients.

FEATURE EXTRACTION

1. **R-R Intervals (8-dimensional)**
   – Time between consecutive R-peaks: Pre-RR and Post-RR.
   – Local-RR is the average of the last 10 Pre-RRs; Global-RR is the average over 20 minutes.
   – We also normalize each interval by the recording's mean, for a total of eight values.

2. **Wavelets (23-dimensional)**
   – We apply the Daubechies "db1" wavelet at three decomposition levels, yielding 23 coefficients that capture time-frequency structure.

3. **Higher-Order Statistics (10-dimensional)**
   – We split the beat into five equal segments and compute skewness and kurtosis in each—measuring waveform asymmetry and peakiness.

4. **1D-LBP (59-dimensional)**
   – Local Binary Patterns encode simple up/down relationships between neighboring samples, then we build an 8-bit histogram of those codes.

5. **Custom Morphology (4-dimensional)**
   – We find the max/min amplitude points within four sub-windows around the R-peak and measure their Euclidean distance to the R-peak. Each feature highlights a different aspect of heartbeats—rhythm, spectral content, shape statistics, local texture, and geometric morphology

RESULTS & INSIGHTS

To judge performance we use the **jκ index**, which is a linear combination of:

- **j-score**: focuses on sensitivity for the two key arrhythmias—Supraventricular Ectopic Beats (SVEB) and Ventricular Ectopic Beats (VEB).
- **κ (Cohen's kappa)**: accounts for chance agreement in imbalanced data. It is a more robust measure of a model's performance in imbalanced datasets.

Single-feature SVMs:

- R-R intervals alone gave us the highest single descriptor jκ of about 0.439.
- Wavelets, HOS, LBP, and custom morphology trailed behind.
  Ensemble results:
- The combination of R-R, Wavelet, HOS, and Morphology with the **product rule** achieved jκ = 0.773.

That's more than a 10% improvement over the prior best in the literature.

We also saw that the product fusion rule outperformed both sum and majority-vote, since it heavily penalizes any feature model with low confidence, leading to a tighter consensus

RESULTS AND INSIGHTS (CONTD.)

Here is our final performance on the DS2 test set, summarized in a **classification report** and a **confusion matrix**. Let's break it down:

- **Normal (N) Beats**
  – **Precision 0.99**: Of all beats predicted "normal," 99% really were normal.
  – **Recall 1.00**: We correctly caught all actual normal beats.
  – **F1-score 1.00**: The harmonic mean of precision and recall, showing perfect performance.
  – **Support 10,756**: The total number of normal beats in DS2.

- **SVEB (Supraventricular Ectopic Beats)**
  – Precision & recall are both 0.00 over 41 beats—our model missed every SVEB. This reflects the very low sample count and extreme imbalance.

- **VEB (Ventricular Ectopic Beats)**
  – Precision 0.79 and recall 0.44 over 43 beats, indicating decent precision but that we missed about 56% of actual VEBs.

- **Accuracy 0.9935 (99.35%)**: High, but misleading on its own because of imbalance.

- **Macro average** (equal-weight average across classes) F1 is 0.52, showing the impact of poor SVEB/VEB detection.

- **Weighted average** (accounts for class sizes) F1 is 0.99, dominated by the normal class. The **confusion matrix** below confirms these numbers—showing the raw counts of correct vs. incorrect predictions for each class. This detailed breakdown highlights that while our model excels on the majority class, further work is needed on rare arrhythmias."

CHALLENGES FACED

Throughout this project we encountered a few challenges:

Extreme Class Imbalance
– With roughly 90% normal beats and under 10% abnormal, our model could "cheat" by always predicting normal unless we applied class weights and used metrics beyond plain accuracy.

Inter-patient Generalization
– Splitting by patient prevented data leakage but reduced the number of abnormal examples available for training.

Noisy Raw Signals
– By avoiding heavy noise filtering, we preserved morphology but forced our feature extractors to handle real-world noise. Some texture-based features (LBP) were especially sensitive.

Hyperparameter Tuning at Scale
– Tuning **five** separate SVM pipelines with 10-fold cross-validation each was computationally intensive and required careful scripting.

- **Fusion Rule Selection**
  – Choosing between product, sum, and majority-vote rules meant running full experiments for each and comparing performance on jκ."

PERFORMANCE METRICS

Classification Report:

	precision	recall	f1-score	support
N	0.99	1.00	1.00	10756
SVEB	0.00	0.00	0.00	41
VEB	0.79	0.44	0.57	43
accuracy			0.99	10840
macro avg	0.60	0.48	0.52	10840
weighted avg	0.99	0.99	0.99	10840

Confusion Matrix:

```
[[10751     0     5]
 [   41     0     0]
 [   24     0    19]]
```

Overall Accuracy: 0.9935

CONCLUSION & FUTURE WORK

- We demonstrated that an **ensemble of feature-specific SVMs** can substantially outperform a single SVM trained on all features simultaneously.
- Our best configuration—combining R-R intervals, wavelets, HOS, and custom morphology with the product fusion rule—yielded a **jκ index of 0.773**, over 10% higher than prior work.
- We kept preprocessing minimal and used only **Lead II**, making our approach practical for real-time or resource-limited settings.

Looking ahead, we plan to:

1. Incorporate **multiple ECG leads** to capture richer morphological information.
2. Explore **advanced fusion frameworks** like Dempster–Shafer theory, which can dynamically weight each model's output based on its confidence.
3. Use **data augmentation** or synthetic beat generation to bolster the rare SVEB and VEB classes.