Non-Invasive Classification of Atrial Flutter Subtypes from F-Wave Morphology Using Dimensionality Reduction and Machine Learning

David Hernando [1], Samuel Ruipérez-Campillo [2,3], Sergio Castrejón [4], Maria Eugenia Martínez [4], José Luis Merino [4], Francisco Castells [1], José Millet [1]

[1] ITACA Institute, Universitat Politecnica de Valencia, Valencia, Spain
[2] Department of Computer Science, ETH Zurich, Zurich, Switzerland
[3] AI Center, ETH Zurich, Zurich, Switzerland
[4] Department of Cardiology, La Paz University Hospital, Madrid, Spain

Abstract

Atrial flutter is a supraventricular reentrant arrhythmia that requires invasive electrical mapping for its correct characterization. The differentiation between the main variants — common and perimitral, both in clockwise or counterclockwise direction — is essential for planning ablation on the macroreentrant circuit. In this work, a non-invasive method based on the morphology of the F-wave of the ECG is proposed. By means of dimensionality reduction on the atrial signals and machine learning classifiers under stratified nested cross-validation, the four AFL subtypes are classified in a cohort of 150 patients. Three lead configurations are evaluated: 12 leads, the 4-lead subset (II, III, aVF, V1) and V1 in isolation. For the four-lead subset, the model obtained AUROCs of 0.84, 0.84, 0.88 and 0.72 for C-CCW, C-CW, PM-CCW and PM-CW respectively, matching or exceeding the full ECG in most subtypes. Notably, V1 in isolation maintained substantial performance, opening the possibility of preliminary subtype classification from a single lead. These results indicate that F-wave morphology, captured through a reduced set of leads and lower-dimensional representations, contains enough discriminative information for AFL subtype classification. Thus, we demonstrate that the F-wave contains morphological information to guide non-invasive ablation planning, laying the groundwork for future diagnostic support systems in the context of cardiac arrhythmias.

1. Introduction

Atrial flutter (AFL) is the second most frequent supraventricular tachyarrhythmia after atrial fibrillation (AF), characterized by rapid, organized and sustained atrial activity, maintained by a macro-reentrant circuit that circumvents anatomical or functional barriers. In the 12-lead electrocardiogram (ECG), it usually manifests through F-waves with a ”sawtooth” morphology, although AFL comprises different subtypes. The most prevalent subtype is typical flutter, dependent on the cavotricuspid isthmus (CTI), whose circuit can rotate counterclockwise or clockwise with inverted F-wave polarities [1]. There are also atypical forms not dependent on the CTI, such as perimitral flutter, whose reentrant circuit surrounds the mitral annulus in the left atrium, with greater electrocardiographic heterogeneity and more complex ablation strategies. Although the initial diagnosis is made with the 12-lead ECG, it has limitations in 2:1 AV conduction ratios [2].

Previous works have demonstrated that the vectorcardiogram (VCG), obtained from the ECG, provides three-dimensional morphological descriptors capable of distinguishing AFL subtypes with AUC values above 0.90 [3,4]. Luongo et al. [5] demonstrated the feasibility of classifying the flutter substrate directly from the ECG using a hybrid approach, although its applicability depends on the extraction of dozens of complex features and the use of 12 leads. Principal component analysis (PCA) is an established technique in ECG signal processing [6] that allows condensing morphological variability into a few orthogonal components, and whose application to atrial arrhythmias was demonstrated by Castells et al. [6]. Thus, our hypothesis is that the morphology of the F-wave itself contains sufficient information to discriminate between AFL subtypes without the need for manual feature extraction. In this context, we further investigate whether this discriminative information can be effectively captured through dimensionality reduction and preserved across clinically relevant lead configurations, with the aim of assessing the extent to which simplified representations of the ECG can support accurate subtype classification.

2. Materials

150 patients with a diagnosis of atrial flutter were analyzed and classified into four subgroups: common clockwise flutter (C–CW), common counterclockwise flutter (C–CCW), perimitral clockwise flutter (PM–CW) and perimitral counterclockwise flutter (PM–CCW). Distribution of the patients in the different gropus is provided in

Table 1. Patient distribution across AFL subtypes.

	C-CCW	C-CW	PM-CCW	PM-CW
N	55	41	29	25

Table 1. All 12-lead electrocardiograms and associated clinical data were provided by Hospital Universitario La Paz (Madrid, Spain). The study was approved by the corresponding ethics committee and all patients provided informed consent, in accordance with the Declaration of Helsinki [7].

3. Methodology

The present work extends the principle proposed by Castells et al. [6] to atrial flutter: the eigenvectors calculated on the F-waves of the ECG define a morphological basis of the signal space, and the projections of each patient onto this basis constitute a compact and interpretable feature vector to feed standard machine learning classifiers. Three lead configurations are evaluated: 12D (12 leads), the clinical subset 4D (II, III, aVF, V1) and V1 in isolation. The choice of the 4D subset responds to its established diagnostic relevance: the inferior leads capture the sawtooth pattern of typical flutter and its direction of rotation, while V1 is critical for discrimination between typical and atypical circuits [8].

For each patient, the atrial activity segment was identified from the 12-lead ECG. During ECG acquisition, adenosine was administered to transiently suppress atrioventricular conduction, thereby isolating atrial activity and removing any ventricular contamination. From each patient, 10 aligned F-wave cycles were selected, with the same number of samples per cycle. The three lead configurations were evaluated independently. In each case, the signals from the selected leads were concatenated into a single vector per patient. Before PCA, row-wise z-score normalization was applied, eliminating global amplitude differences between recordings and forcing the analysis to capture only morphological variability.

The eigenvectors and the projection of each patient onto these vectors were obtained. Unlike non-linear representation methods such as variational autoencoders or deep neural networks, PCA offers complete traceability: each component is a directly interpretable linear combination of the original signal samples, which allows linking the captured morphological variability patterns with the pathophysiology of the reentrant circuit [9]. A reduced representation space was obtained from the ten first principal components (PC1–PC10).The number of principal components was selected by evaluating classification performance across different values on the training folds. Ten components were selected as the best trade-off between discriminative performance and model complexity, limiting the number of input features to reduce the risk of overfitting given the limited cohort size. Thus, the projections of each patient onto this basis — vectors of dimension 10 — constitute the compact representation that directly feeds the machine learning classifiers.

Performance was evaluated using a stratified *nested cross-validation* strategy (external validation k_{ext} = 4, internal hyperparameter optimization k_{int} = 3), chosen to obtain unbiased performance estimates given the limited number of patients. In each outer fold, the inner folds were used to select the optimal hyperparameters via grid search. Then, the final model was retrained on the full outer training set and evaluated on the held-out test set, never seen during the optimization. Four classifiers were trained: Decision Tree, Random Forest (RF), AdaBoostM2 and SVM. All incorporated inverse class frequency weighting to mitigate sample imbalance. The reported metrics included sensitivity, specificity, F1-score, AUROC and confusion matrix. A summary of the methodology is reported in Figure 1.

4. Results

Three lead configurations were evaluated: 12D (12 leads), 4D (II, III, aVF, V1) and V1 in isolation. In all of them, the random forest (RF) offered the most balanced and substantially superior performance across classes in most cases, with sustained sensitivities across the four subtypes and consistently superior AUROC values. The alternative models presented more unstable behaviors: AdaBoostM2 achieved in 12D a sensitivity of 0.83 for C-CCW but reduced PM-CW to 0.20; similarly, SVM obtained in 4D a sensitivity of 0.87 for C-CCW but practically nullified the detection of PM-CW. This pattern of high sensitivity in the majority classes at the expense of collapse in PM-CW motivated the selection of RF as the reference model for the comparative analysis between configurations. The detailed results by class and configuration are presented in Table 2, and the corresponding confusion matrices in Table 3.

5. Discussion

The results show that the morphology of the F-wave of the ECG, compressed into the ten first principal components, contains sufficient discriminative information to classify the four AFL subtypes with promising performance. It is noteworthy that the 4D configuration (II, III, aVF, V1) matched or exceeded the full 12 leads in most metrics: sensitivity for C-CCW improved from 0.63 to 0.69, PM-CCW from 0.58 to 0.62, and AUROC values were superior in all classes except PM-CW, where both configurations obtained 0.72.

Importantly, this work does not aim to maximize classification performance (e.g., AUROC), but rather to explore whether a simplified representation of the ECG—based only on F-wave morphology and a reduced subset of leads—can retain enough discriminative information be-

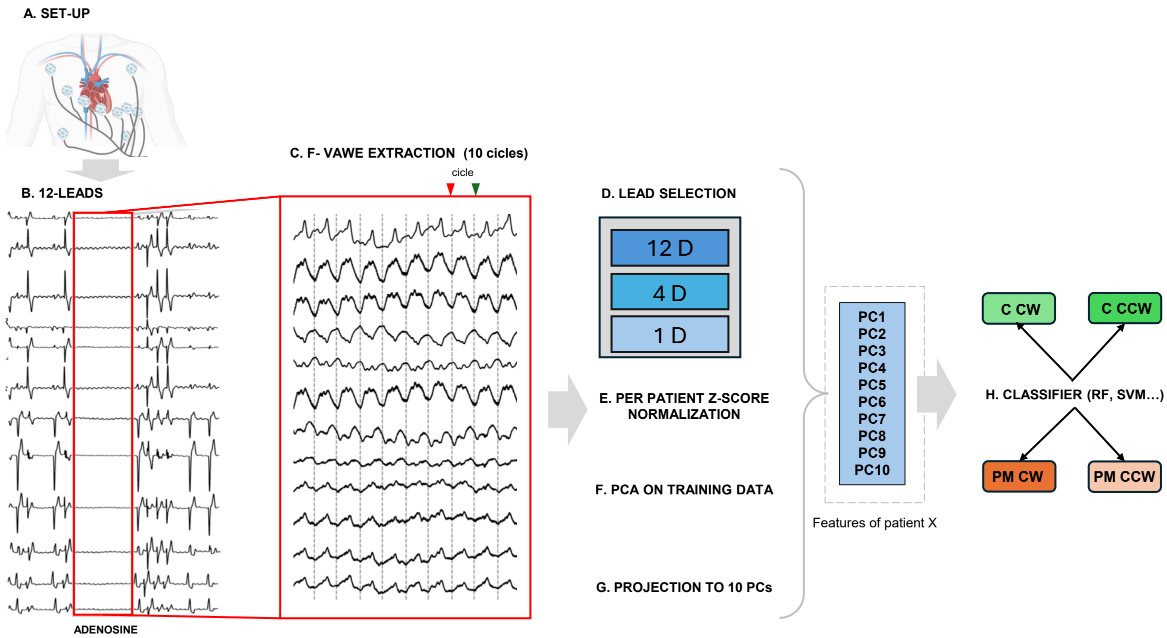

Figure 1. Proposed methodology. A: acquisition with adenosine. B: 12-lead ECG. C: 10 F-wave cycle selection. D: lead configuration selection (12D, 4D, 1D). E: per-patient z-score normalization. F–G: PCA calculation on training data and projection to 10 principal components. H: AFL subtype classification.

tween classes. In this sense, the results should be interpreted as evidence of redundancy reduction and model simplification rather than as a direct competition with more complex approaches.

This result confirms that the discriminative information is concentrated in the inferior frontal plane and in V1 [8], and that the remaining eight leads are largely redundant once PCA is applied. The V1 isolated configuration, despite the expected reduction in sensitivity for typical subtypes, maintained notable performance: AUROC of 0.80–0.85 and a sensitivity for PM-CW of 0.40, even higher than that of 4D (0.24), suggesting that V1 captures relevant morphological features of the perimitral circuit independently of the frontal plane leads.

PM-CW was systematically the most difficult class in all configurations, with AUROC of 0.72 for all of them and consistently lower sensitivities. This behavior is not inherent to the model but to the class itself as it presents the greatest intraclass morphological heterogeneity, overlap with other subtypes and the smallest sample size (n=25), which limits the robust estimation of the models [4].

These results can be contextualized with respect to previous work [4], which employed a fundamentally different approach based on VCG-derived metrics combined with clinical variables, achieving AUROC $\sim$0.90 for typical subtypes with PM-CW remaining the most challenging class (AUC 0.72, sensitivity 0.25). However, it is important to emphasise that both approaches pursue different objectives and are not directly comparable in terms of performance alone. While previous work focuses on maximising classification accuracy through multimodal information and engineered features, the present study explores an alternative paradigm: directly analysing raw F-wave morphology, reducing dimensionality, and identifying minimal lead subsets that preserve discriminative information. Despite the methodological differences — the proposed approach relies solely on raw F-wave morphology without VCG reconstruction or clinical variables — and lower sensitivity perfomance, similar AUROC values are achieved, suggesting that the F-wave already encodes a substantial part of the discriminative information. From this perspective, the slightly lower performance observed in some metrics should be interpreted in the context of a significantly simpler and more interpretable pipeline.Incorporating clinical variables or non-linear representations could further improve classification, particularly for PM-CW.

The main limitations of the study are the reduced cohort size and the linear nature of the dimensionality reduction employed. Future works could address non-linear representations such as autoencoders or convolutional networks

Table 2. Multiclass classification results of the RF model on the external test folds for the three evaluated lead configurations. Values reported as [95% CI]. Sens, Spec and F1 use exact Clopper–Pearson intervals. AUROC is reported as a point estimate on aggregated external folds in a one-vs-rest scheme. 12D: 12 leads; 4D: II, III, aVF, V1; V1: single lead.

A. Common Flutter

		C_{CCW}	C_{CW}
Sens	12D	0.63 [0.49, 0.76]	0.65 [0.49, 0.80]
	4D	0.69 [0.55, 0.81]	0.63 [0.47, 0.78]
	V1	0.58 [0.44, 0.71]	0.59 [0.42, 0.74]
Spec	12D	0.77 [0.67, 0.85]	0.87 [0.79, 0.93]
	4D	0.79 [0.69, 0.87]	0.85 [0.77, 0.91]
	V1	0.80 [0.71, 0.88]	0.88 [0.81, 0.94]
F1	12D	0.63 [0.52, 0.73]	0.66 [0.49, 0.83]
	4D	0.67 [0.56, 0.79]	0.63 [0.54, 0.71]
	V1	0.60 [0.45, 0.75]	0.62 [0.59, 0.64]
AUROC	12D	0.82	0.81
	4D	0.84	0.84
	V1	0.80	0.80

B. Perimitral Flutter

		PM_{CCW}	PM_{CW}
Sens	12D	0.58 [0.39, 0.76]	0.32 [0.15, 0.54]
	4D	0.62 [0.42, 0.79]	0.24 [0.09, 0.45]
	V1	0.62 [0.42, 0.79]	0.40 [0.21, 0.61]
Spec	12D	0.88 [0.81, 0.93]	0.89 [0.83, 0.94]
	4D	0.92 [0.85, 0.96]	0.87 [0.80, 0.92]
	V1	0.86 [0.79, 0.92]	0.86 [0.79, 0.92]
F1	12D	0.57 [0.38, 0.75]	0.35 [0.10, 0.59]
	4D	0.63 [0.37, 0.90]	0.26 [0.21, 0.30]
	V1	0.56 [0.28, 0.85]	0.39 [0.24, 0.53]
AUROC	12D	0.84	0.72
	4D	0.88	0.72
	V1	0.85	0.72

that could capture more complex morphological patterns, especially for PM-CW.

6. Conclusions

The results support the hypothesis that the morphology of the F-wave of the ECG contains sufficient discriminative information for the non-invasive classification of the four AFL subtypes, with AUROC values of up to 0.88. This information can be effectively captured in simplified, lower-dimensional representations. The four-lead configuration (II, III, aVF, V1) matched or exceeded the full 12-lead setup, and V1 in isolation maintained substantial performance, opening the possibility of a preliminary subtype classification from a single lead. These findings suggest that accurate classification may be achievable using reduced lead configurations, with potential implications for simplified and ambulatory ECG analysis. However, PM-CW remains the main challenge due to its greater morphological heterogeneity and limited sample representation.

Table 3. Confusion matrices of the RF model on the external test folds for the three lead configurations. Rows: true class; columns: predicted class. 12D: 12 leads; 4D: II, III, aVF, V1; V1: single lead.

Config.	True / Pred.	C_{CCW}	C_{CW}	PM_{CCW}	PM_{CW}
12D	C_{CCW}	35	5	8	7
	C_{CW}	9	27	2	3
	PM_{CCW}	7	2	17	3
	PM_{CW}	6	7	4	8
4D	C_{CCW}	38	5	3	9
	C_{CW}	7	26	3	5
	PM_{CCW}	3	6	18	2
	PM_{CW}	10	5	4	6
V1	C_{CCW}	32	5	7	11
	C_{CW}	6	24	6	5
	PM_{CCW}	6	4	18	1
	PM_{CW}	7	4	4	10

References

[1] Varma N, Chomos SK. Clockwise and counterclockwise flutter: mirror-image circuits? HeartRhythm Case Reports 2022;.

[2] Saclova L, Nemcova A, Smisek R, Smital L, Vitek M, Ronzhina M. Reliable p wave detection in pathological ecg signals. Scientific Reports 2022;12(1):6589.

[3] Ruipérez-Campillo S, Castrejón S, Martínez M, Cervigón R, Meste O, Merino JL, Millet J, Castells F. Non-invasive characterisation of macroreentrant atrial tachycardia types from a vectorcardiographic approach with the slow conduction region as a cornerstone. Computer methods and programs in biomedicine 2021;200:105932.

[4] Hernando D, Ruipérez-Campillo S, Castrejón S, Martínez ME, Merino JL, Castells F, Millet J. Modelo multimodal personalizado para la clasificación no invasiva del flutter auricular en base al vectorcardiograma. In Actas XLIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2025). Zaragoza, Spain, 2025; .

[5] Luongo G, Vacanti G, Nitzke V, Nairn D, Nagel C, Kabiri D, Almeida TP, Soriano DC, Rivolta MW, Ng GA, et al. Hybrid machine learning to localize atrial flutter substrates using the surface 12-lead electrocardiogram. EP Europace 2022;24(7):1186–1194.

[6] Castells F, Laguna P, Sörnmo L, Bollmann A, Roig JM. Principal component analysis in ecg signal processing. EURASIP Journal on Advances in Signal Processing 2007; 2007(1):074580.

[7] Manzini JL. Declaración de helsinki: principios éticos para la investigación médica sobre sujetos humanos. Acta bioethica 2000;6(2):321–334.

[8] Cosio FG. Atrial flutter, typical and atypical: a review. Arrhythmia electrophysiology review 2017;6(2):55.

[9] Ruipérez-Campillo S, Reiss M, Ramírez E, Cebrián A, Millet J, Castells F. Clustering and machine learning framework for medical time series classification. Biocybernetics and Biomedical Engineering 2024;44(3):521–533.

Address for correspondence:

David Hernando
davidhernandopr@gmail.com