# Classification of Persistent Atrial Fibrillation Targets Using Machine Learning on Multipolar Electrograms

Rafael Costa de Almeida[1], Douglas Almonfrey[1], Fabien Squara[2], Vicente Zarzoso[3]

[1] Instituto Federal do Espírito Santo, Vitória, Espírito Santo, Brazil
[2] Hôpital Pasteur, Nice, France
[3] Université Côte d'Azur, CNRS, I3S Laboratory, Sophia Antipolis, France

## Abstract

*Cardiovascular diseases, particularly atrial fibrillation (AF), remain a significant global health burden. Despite advancements in diagnostic and treatment techniques, the long-term success rates of AF ablation procedures remain suboptimal. This is primarily due to the complexity of underlying mechanisms, challenges in accurately identifying arrhythmogenic substrates, and the efficacy heavily relying on physician interpretation, which contributes to variability in procedural outcomes. This work addresses these limitations by leveraging machine learning (ML) for the classification of persistent AF using multi-lead electrograms (EGMs). We investigate logistic regression with handcrafted features, as well as convolutional neural networks (CNNs) and Long Short-Term Memory (LSTM) networks designed to interpret transformed EGMs and capture temporal dependencies, with the aim of enhancing the accuracy of identifying regions suitable for ablation. All three approaches show promise in identifying persistent AF behavior, even in data-limited settings. These results highlight the potential of ML to improve diagnostic precision and support more effective, personalized ablation strategies for persistent AF.*

## 1. Introduction

AF is the most common sustained cardiac arrhythmia, affecting millions of people globally and significantly contributing to stroke, heart failure, and other serious complications [1]. Ablation strategies aim to eliminate arrhythmogenic sources and modify the atrial substrate to prevent AF recurrence. One common strategy involves pulmonary vein isolation (PVI), which focuses on eliminating high-frequency pulmonary vein potentials and creating a bidirectional block to prevent ectopic pulmonary vein activity from triggering AF [2].

Recent advances in artificial intelligence (AI) and machine learning (ML) have improved cardiac medicine, particularly in managing complex cardiac arrhythmias. These technologies are increasingly used to enhance diagnostic accuracy, personalize treatment, and improve procedural outcomes.

ML models can enhance cardiac ablation by accurately identifying arrhythmogenic regions. Algorithms detecting spatio-temporal EGM dispersion have improved success rates in persistent AF, with tailored strategies achieving higher one-year freedom from recurrence (88% vs. 70%) compared to pulmonary vein isolation (PVI) alone, particularly in long-duration AF [3]. In addition, ML-based patient stratification, such as uplift modeling, has identified persistent AF patients benefiting from more extensive ablation (PVI-plus), yielding significantly lower recurrence than standard PVI [4]. AI has also enhanced diagnostic capabilities by objectively analyzing complex electrophysiological data. CNNs have achieved high accuracy (95.0%) in distinguishing rotational activation patterns from intracardiac EGMs when converted to visual image grids via Hilbert transforms, outperforming traditional statistical methods or support vector machines [5]. However, classifying ablation targets from multipolar EGMs remains challenging. This work proposes a pipeline combining signal processing, feature engineering, and ML models to address this gap.

## 2. Methods

### 2.1. Real Data Used for Classification

The dataset used for classification comprises electrograms from 53 patients with persistent AF, collected using the *PentaRay*® multielectrode mapping catheter at Nice Pasteur University. It extends a previous dataset of 16 patients [6], now containing over 10,000 pre-ablation samples. Each sample consists of a 2500 ms window from 10 catheter leads, annotated by physicians during the procedure. Labels distinguish regions relevant for ablation (e.g., CFAE, flutter, scar, or normal). In this work, scar tissue

 ISSN: 2325-887X DOI: 10.22489/CinC.2025.437

is excluded, and all non-normal conditions are grouped as ablation targets, yielding a binary classification problem. For each sample, if any bipole is labeled as an AF ablation target, the entire sample is considered AF.

Extensive preprocessing was required to address duplication artifacts and synchronization issues. In collaboration with physicians and the software provider, only samples collected between the initial tagging and the start of ablation were retained, forming the *Raw Dataset* of 13,888 samples (1,035 ablation targets; 12,853 non ablation targets) from 53 anonymized patients.

As locations are tagged by the operator while the physician moves the catheter, the tagged location may not always represent the exact sample pointed out by the physician. Although this had no clinical impact, it can affect model training, as the underlying electrical behavior of the tagged region may not accurately reflect the intended target. To address this issue, randomized samples were reviewed and reclassified by a physician. This yielded the *Curated Dataset*, with 430 samples (112 ablation targets; 318 non ablation targets), providing the most reliable labels. DL models are trained on the larger *Raw Dataset*, due to the limited number of samples on the curated version. Final evaluation relies on the *Curated Dataset*.

## 2.2. Signal Processing Operations

Transformations are used to extract salient features. This section describes the transformation techniques applied. Figure 1 summarizes the pipeline by showcasing two representative transformation chains. First showing the raw signal, then the signal after applying the Teager–Kaiser (TK) operator with decimation and cutoff normalization, and finally the same with squared pulses.

**TK operator:** This energy-based transformation enhances local transients by emphasizing instantaneous energy, which helps reveal abrupt events in EGMs [7].

**Decimation:** Decimation is a downsampling technique that effectively decreases the temporal resolution while preserving the overall structure and shape of the signal.

**Cutoff Normalization:** To suppress low-amplitude/no-contact noise while retaining clinically relevant peak information, first, we use minimal amplitude and autocorrelation thresholds to determine if a lead contains an EGM signal or pure noise due to a lack of contact between the catheter and tissue. For leads that pass this first check, we detect peaks and their boundaries within leads using thresholds and rescale signal amplitude within each peak window to [0,1], leaving the inter-peak baseline near zero. This keeps peak timing, width, and fragmentation while removing nuisance variability.

**Squared Signal:** This transformation highlights each peak while simplifying the signal. Each peak is converted into a 0/1 pulse whose width equals the peak duration. This makes the representation less sensitive to absolute amplitude and simplifies downstream modeling. This transformation is used with normalized TK transformed signals since it relies on positive signals with a maximum amplitude of 1, where peaks occur.

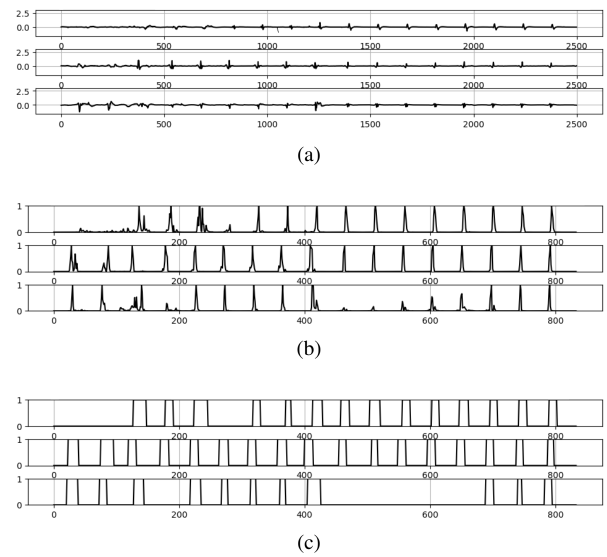


Figure 1: Example of Signal Transformations with a) Raw signal; b) TK + Decimation + Cutoff Normalization; c) TK + Decimation + Cutoff Normalization + Squared Signal

## 2.3. Feature Engineering

Feature engineering is used for simpler models and is comprised of two primary components: the extraction of statistical features from each bipolar EGM and pairwise comparison between neighboring leads.

**Statistical Measures:** To characterize the signal's distribution, central tendency, and variability a group of features are used: textcolorredmean, median, mode, standard deviation, mean absolute deviation, coefficient of variation, interquartile range, percentiles (5%, 25%, 50%, 75%, 90%, 95%), skewness, kurtosis, jarque-bera statistic, number of peaks, mean peak distance, and standard deviation of peak distances.

**Seasonal Decomposition:** Seasonal decomposition isolates trend, seasonal, and residual components of each signal using a convolution filter. The seasonal component serves as input for calculating dynamic time warping (DTW) distances and cross-correlation metrics between leads. DTW provides robust inter-lead distance measures, while cross-correlation identifies time lags, highlighting potential AF synchrony or delay patterns.

**Pairwise Signal Comparison:** All possible pairs among the ten seasonal components are analyzed to extract maximum cross-correlation values and corresponding lags. Minimum and maximum lag values across pairs are retained to assess temporal alignment variability. Pairwise feature differences are computed as normalized percent differences to emphasize magnitude discrepancies, supporting the detection of AF ablation targets.

**Complexity and Recurrence-Based Features:** Additional features are derived using Recurrence Quantification Analysis (RQA) and entropy metrics. RQA captures dynamic patterns via features such as determinism, laminarity, recurrence rate, entropy, trapping time, and various line-based statistics. Entropy features such as shannon entropy, conditional entropy, dispersion, phase entropy, and slope entropy quantify signal complexity.

## 2.4. Data Augmentation and Balancing Techniques

To improve robustness and address class imbalance, the dataset was expanded through augmentation and balanced with resampling strategies.

**Augmentation:** Two data augmentation strategies are used. *Lead rotation* reorders the 10-lead EGMs using the Pentaray catheter's symmetry, preserving inter-lead relationships while producing up to 10 distinct spatial configurations. *Time reversal* inverts the temporal sequence of all leads, improving temporal generalization.

**Balancing:** Class imbalance is mitigated through oversampling and undersampling, guided by a configurable Class Imbalance Ratio (CIR) ranging from $0.5$ (undersampling) to $1.5$ (oversampling), enabling flexible rebalancing strategies.

## 2.5. Machine Learning Models

Three distinct machine learning model architectures are investigated for the binary classification: logistic regression, a CNN, and an LSTM. These models are selected based on their ability to capture spatiotemporal patterns and provide interpretable baselines. The CNN and LSTM models have a small number of parameters due to the limited amount of samples available.

- **Logistic Regression**: This model leverages handcrafted features extracted from EGMs instead of raw signal data. Its simplicity and interpretability make it a valuable approach in clinical settings where understanding the model's decision-making process is crucial.
- **CNN**: This compact CNN architecture processes transformed EGM time-series, progressively reducing temporal resolution while extracting local spatial patterns. The model is comprised of two convolutional blocks. The first block, a 2D convolution with a kernel size of $(100 \times 2)$, stride $(2 \times 2)$, and ReLU activation, followed by max pooling, batch normalization, and dropout. The second block uses a kernel size of $(25 \times 1)$ with stride $(2 \times 1)$ and otherwise identical structure. A final dense layer with sigmoid activation produces the binary classification output.
- **LSTM**: This recurrent model is designed to capture temporal dependencies in the EGM sequences. The model applies layer normalization to the input, then employs two stacked bidirectional LSTM layers, the first with 6 hidden units and returning the entire sequence, and the second with 2 hidden units that outputs only the final state. Both layers use recurrent dropout. A dense projection layer with 64 units and ReLU activation expands the representation before applying dropout for regularization. A final dense layer with sigmoid activation produces the binary classification output.

## 2.6. Training and Evaluation Method

We evaluate with stratified 3-fold cross-validation using patient-wise exclusive splits to avoid leakage from correlated signals of the same patient. Stratification preserves AF/non-AF ablation target label proportions in each fold, ensuring comparable class balance.

Models are evaluated using F1-score as the primary metric. Additionally, accuracy, precision, and recall are also reported to provide a comprehensive view of the model's performance. We implement a comprehensive hyperparameter search based on the Tree-structured Parzen Estimator (TPE) [8]. The search jointly optimized training hyperparameters such as batch size, learning rate, but also, class imbalance ratio (CIR), transformation pipeline choices and chaining, feature-engineering options, data augmentation strategies, and model architecture variants. Each configuration was evaluated with the average test F1-score as the optimization objective.

# 3. Results

From the hyperparameter search process, two input pipelines were ultimately considered (Table 1) as they were associated with top models according to F1-Score.

**Method A:** Applies the TK operator, decimation, and cutoff normalization, producing a normalized, positive signal used for feature extraction (Figure 1b). All feature engineering techniques described above are applied to this transformed signal, resulting in 29 features designed to capture key temporal and amplitude-based characteristics.

**Method B:** Extends Method A by appending the Squared Signal step, converting each peak into a 0/1 pulse that preserves timing/width while reducing amplitude sensitivity (Figure 1c). No feature engineering is applied to this representation.

Table 2 presents the results for the top three model architectures evaluated in this work.

Table 1: Summary of Input Methods for Machine Learning

| Method Name | Signal Processing Steps | Feature Engineering |
|---|---|---|
| A | TK Operator → Decimation → Cutoff Normalization | Yes |
| B | TK Operator → Decimation → Cutoff Normalization → Squared Signal | No |

Table 2: Classification Results on Persistent AF Targets

| Model | Params | Recall (%) | Precision (%) | F1 (%) |
|---|---|---|---|---|
| CNN | 272 | 68.8 | 38.4 | 49.2 |
| LSTM | 1,461 | 55.5 | **40.5** | 46.8 |
| Logistic Regression | **29** | **85.9** | 40.2 | **54.7** |

## 4. Discussion

Despite the simplicity of logistic regression, it achieved the highest F1-score, demonstrating the utility of tailored feature extraction and careful regularization in limited data settings. These findings suggest that interpretable models with engineered features can match or even exceed more complex neural architectures when identifying persistent AF behavior from EGMs. Comparing the two other methods, the CNN's relatively higher recall indicates greater sensitivity to AF patterns, whereas the LSTM's precision suggests better discrimination at the cost of recall. Parameter counts remain modest across models (29-1,461), appropriate for the limited dataset size and patient-wise splits. Limitations include label noise from procedural tagging, class imbalance, and the absence of external validation.

## 5. Conclusion

This work presented ML approaches for the classification of ablation targets using multipolar EGMs. By comparing logistic regression with handcrafted features against CNN and LSTM architectures, we demonstrated that interpretable models, combined with careful preprocessing and feature engineering, can achieve stronger performance than DL models despite limited data availability. These findings support the feasibility of integrating ML-driven classification into ablation workflows, potentially assisting physicians in identifying arrhythmogenic regions more objectively and consistently. Future work will expand the dataset, incorporate prospective validation, and add explainability analyses to strengthen clinical interpretability and adoption.

## Acknowledgments

This work has been supported in part by the French government, through the 3IA Cote d'Azur Investments in the project managed by the National Research Agency (ANR) with the reference number ANR-23-IACL-0001. V. Zarzoso holds the "IAblation" 3IA Chair. Support by the CAPES-COFECUB program through project "INTERACTION" (Ma-985-23) is also acknowledged.

## References

[1] Brundel BJJM, Ai X, Hills MT, Kuipers MF, Lip GYH, de Groot NMS. Atrial fibrillation. Nat Rev Dis Primers April 2022;8(1):21.

[2] Wellens HJJ. Pulmonary vein ablation in atrial fibrillation: Hype or hope? Circulation November 2000;102(21):2562–2564. ISSN 1524-4539.

[3] Deisenhofer I, Albenque JP, Busch S, Gitenay E, Mountantonakis SE, Roux A, Horvilleur J, Bakouboula B, Oza S. Artificial intelligence for individualized treatment of persistent atrial fibrillation: a randomized controlled trial. Nature Medicine February 2025;ISSN 1546-170X.

[4] Sato T, Sotomi Y, Hikoso S, Kitamura T, Nakatani D, Okada K, Dohi T, Sunaga A, Kida H, Matsuoka Y, Tanaka N, Watanabe T, Makino N, Egami Y, Oka T. Uplift modeling to identify patients who require extensive catheter ablation procedures among patients with persistent atrial fibrillation. Scientific Reports February 2024;14(1). ISSN 2045-2322.

[5] Alhusseini MI, Abuzaid F, Rogers AJ, Zaman JA, Baykaner T, Clopton P, Bailis P, Zaharia M, Wang PJ, Rappel WJ, Narayan SM. Machine learning to classify intracardiac electrical patterns during atrial fibrillation: Machine learning of atrial fibrillation. Circulation Arrhythmia and Electrophysiology August 2020;13(8). ISSN 1941-3084.

[6] Ghrissi A, Almonfrey D, de Almeida RC, Squara F, Montagnat J, Zarzoso V. Data augmentation for automatic identification of spatiotemporal dispersion electrograms in persistent atrial fibrillation ablation using machine learning. In 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society. 2020; 406–409.

[7] Holambe R, Deshpande M. Nonlinear Measurement and Modeling Using Teager Energy Operator. ISBN 978-1-4614-1504-6, 02 2012; 45–59.

[8] Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for hyper-parameter optimization. In Proceedings of the 25th International Conference on Neural Information Processing Systems, NIPS'11. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781618395993, 2011; 2546–2554.

Address for correspondence:

Rafael Costa de Almeida.
Avenida Vitória, 1729, Jucutuquara, Vitória, Brazil.
contact@rafaelcostadealmeida.com