Combining Rhythmic and Morphological ECG Features for Automatic Detection of Atrial Fibrillation: Local and Global Prediction Models

Gennaro Laudato[1(✉)], Franco Boldi[2(✉)], Angela Rita Colavita[3(✉)],
Giovanni Rosa[1(✉)], Simone Scalabrino[1,4(✉)], Aldo Lazich[5,6(✉)],
and Rocco Oliveto[1,4(✉)]

[1] STAKE Lab, University of Molise, Pesche, IS, Italy
{gennaro.laudato,simone.scalabrino,rocco.oliveto}@unimol.it,
g.rosa@studenti.unimol.it
[2] XEOS., Roncadelle, BS, Italy
franco.boldi@xeos.it
[3] ASREM, Campobasso, CB, Italy
angelarita.colavita@asrem.org
[4] Datasound srl, Pesche, IS, Italy
{rocco,simone}@datasound.it
[5] Ministero della Difesa, Rome, RM, Italy
aldo.lazich@marina.difesa.it
[6] DIAG, University of Rome "La Sapienza", Rome, Italy
lazich@diag.uniroma1.it

Abstract. Atrial fibrillation (AF) is the most common type of heart arrhythmia. AF is highly associated with other cardiovascular diseases, such as heart failure, coronary artery disease and can lead to stroke. Unfortunately, in some cases people with atrial fibrillation have no explicit symptoms and are unaware of their condition until it is discovered during a physical examination. Thus, it is considered a priority to define highly accurate automatic approaches to detect such a pathology in the context of a massive screening.

For this reason, in the recent years several approaches have been defined to automatically detect AF. These approaches are often based on machine learning techniques and—most of them—analyse the heart rhythm to make a prediction. Even if AF can be diagnosed by analysing the rhythm, the analysis of the morphology of a heart beat is also important. Indeed, during an AF events the P wave could be absent and fibrillation waves may appear in its place. This means that the presence of only arrhythmia could be not enough to detect an AF events.

Based on the above consideration we have presented MORPHYTHM, an approach that use machine learning to combine rhythm and morphological features to identify AF events. The results we achieved in an empirical evaluation seems promising. In this paper we present an extension of MORPHYTHM, called LOCAL MORPHYTHM, aiming at further improving the detection accuracy of AF events. An empirical evaluation of LOCAL

MORPHYTHM has shown significantly better results in the classification process with respect to MORPHYTHM, particularly for what concerns the true positives and false negatives.

Keywords: Healthcare · Atrial fibrillation · Decision support system · Machine learning

1   Introduction

During the last few years has occurred a rapid technological evolution in the scientific field of the Internet of Medical Things (IoMT) and Wireless Body Area Network (WBAN). The main demands for these systems can be summarised as follows: (i) reducing the healthcare costs while keeping the quality of the services and (ii) promoting wellness programs to shift the health expenditure from treatment to prevention [3].

All the efforts in this field by the scientific research communities has made it possible to obtain electronic devices of minimal size and wearable [31]. This has created a fertile ground for telemedicine. Telemedicine can be commonly defined as the use of advanced telecommunications technologies for the purpose of supporting many medical activities. In the last years, this industry has grown and most US health institutions and hospitals are currently employing such kind of technology [12].

In this context, medical activities become responsible—beyond the knowledge and clinical skills—in handling an ample amount of data related to the patient health. Thus, appropriate elaboration of the clinical data are required to facilitate the work of experts and promote a policy of welfare. A Decision Support System (DSS) is the key component of an effective telemedicine system. Such a component is basically a layer of software that latently and continuously analyze the acquired data aimed at providing recommendations, even at patient-level, to the medical experts for the identification of a risky situation or for the diagnosis of a specific pathology [38].

In this paper we present an approach that could be integrated in a DSS of a telemedicine system aiming at supporting the identification of atrial fibrillation (AF) episodes through the analysis of ECG. AF is the most common sustained arrhythmia and is associated with significant morbidity and mortality [10]. We decided to focus on AF detection because of the incidence statistics of such a pathology. Indeed, around one third of all ischemic strokes are caused by AF [14] and the early phase of appearance is a particularly high-risk period for the development of stroke [34]. In addition, AF is often asymptomatic. Thus, it is crucial to detect onset episodes of AF with high accuracy to allow a proper intervention of cardiologist [18].

The detection of AF episodes generally involves two electrocardiogram (ECG) sources of information: (i) *beat morphology*, because during an AF episode, it is possible to observe fluctuating wave forms instead of P waves and (ii) *rhythm*, because during an AF episode it is possible to observe an irregularity of heart

rate. The fact that an ECG recording of the episode is a diagnostic criterion can make the process cumbersome, especially if the arrhythmia is paroxysmal and not easily provoked during a recording session. In order to capture the episode, an extended recording time (at least 24 h) through an Holter monitoring is required [2]. These recordings from wearable ECG devices introduce an amount of data which results complicated for the physician to inspect and analyze. This recall the need of semi-automatic approaches to determine onset and duration of AF episodes.

A lot of effort in the research community has been devoted to the definition of methods to automatically detect AF. These are often based on Machine Learning techniques and—most of them—are based only on the analysis of R-R intervals (RRI), *i.e.,* they just exploit the rhythmic source of information. Even if the accuracy of such approach is generally very high in terms of accuracy (more than 95%), the proposed approach still misclassifies *fibrillant* heart beat signals as *non-fibrillant* [41]. This suggests that there is still room for improvement. Especially, our conjecture is that by combining morphological and rhythm features is possible to improve the accuracy of approaches based on just one of the two source of information.

Based on the above consideration, in a previous work we have presented MORPHYTHM [23], a new approach based on machine learning techniques where morphological and rhythmic information are fused together. MORPHYTHM showed surprising results, especially for what concerns two vital aspects of the medical classification: increment of true positives and reduction of false negatives.

In this paper we present an extension of MORPHYTHM aiming at further improving its accuracy. We first performed a rigorous feature engineering process in order to identify the features that contribute the most to the prediction of AF events. Then, we experimented most advanced machine learning techniques, including artificial neural network and deep learning techniques. Finally, we integrated in MORPHYTHM the concept of"local" prediction, successfully used in other context [28]. Especially, instead of producing a single prediction model, the new version of MORPHYTHM, called LOCAL MORPHYTHM, automatically build several prediction models based on the characteristics of the ECGs in the training set. In particular, the training set is clustered in order to put together ECGs that exhibits similar characteristics. Then, for each cluster, LOCAL MORPHYTHM builds a prediction model. When a new data point is provided, LOCAL MORPHYTHM first selects the most suitable model based on the characteristics of the new data point, and then it performs the prediction applying the selected model.

The rest of the paper is structured as follows: Sect. 2 provides details on AF and on automatic detectors of AF. Section 3 presents LOCAL MORPHYTHM, our novel approach for AF detection, while Sect. 4 reports the design and the results of the empirical study we conducted to evaluate LOCAL MORPHYTHM. Finally, Sect. 5 concludes the paper and provides suggestions for possible future research directions.

2   Background and Related Work

2.1   Atrial Fibrillation

Normally, the heart contracts and relaxes to a regular beat. In atrial fibrillation, the upper chambers of the heart (the atria) beat irregularly (quiver) instead of beating effectively to move blood into the ventricles[1]. If the ECG recording is available, AF is diagnosed by whenever an irregular heartbeat presents the following characteristics: the absence of P waves (with disorganized electrical activity in their place) and irregular R–R intervals due to irregular conduction of impulses to the ventricles [15].

The prevalence of atrial fibrillation (AF) is increasing all over the world and it is becoming one of the most important clinical issues for industrialised countries [15, 42]. AF is a crucial risk factor for the occurrence of stroke. Beyond stroke, AF can lead also to congestive heart failure. Furthermore, hypertension, diabetes and heart failure are some of the most common comorbidities [24, 39]. In addition, AF presents a sever influence on the global health conditions of individuals who contract it [20].

To produce a diagnosis of AF, a cardiologist checks the clinical history of the patient and the ECG signal, by at least observing a single lead during the revealing of the episode [15]. Unfortunately, AF is often paroxysmal, *i.e.,* there are recurrent episodes that stop on their own in less than seven days [15], and asymptomatic. For these reasons, the screening of such a pathology needs to become a priority.

2.2   Automatic Detection of Atrial Fibrillation

In recent years, the scientific research has provided several works aiming at automatically detect AF episodes. Most of them have shown important results by exploiting only the analysis of heart rhythm, assumed as the observation of the distances between two successive R peaks (RRI, RR intervals) [8, 30, 37, 40]. Indeed, the detection methods based on RRI produce relatively more precise identification of AF since the R-wave peak of QRS complex is the most prominent characteristic feature of an ECG recording and the least susceptible to various kinds of noise [19, 21, 25, 26].

In the work by Hochstadt *et al.* [18], around 18 thousand consecutive RR interval measurements were recorded in 20 patients, including about 12 thousand RR intervals during AF and 6,087 RR intervals during sinus rhythm. The automatic algorithm—based on Lorenz-plot—used by the authors distinguished AF from sinus rhythm with a sensitivity of 100% and specificity of 93.1%.

In the study by Andersen *et al.* [2], a novel approach for AF detection based on Inter Beat Intervals (IBI) extracted from long term electrocardiogram (ECG) recordings is presented. For this purpose, five time-domain features have been extracted from the IBIs and a Support Vector Machine (SVM) has been used for

---

[1] https://bit.ly/3dvrXJX.

classification. The proposed approach has shown a significantly reduced computation time without loss of performance, if compared to a consolidated baseline.

Afdala *et al.* [1] test the ability of simply involving the Shannon entropy in the detection of Atrial Fibrillation episodes. In their research study, they used data from a well-known public data set (Physionet MIT-BIH AFDB) and, as performance, they observed that Shannon entropy has the highest accuracy if a threshold of 0.5 is set.

In the work by Chen *et al.* [5], a new feature extraction method based on RR interval is proposed with the aim at describing an heart rhythm which will be submitted to a classification experiment. As descriptors, they used the robust coefficient of variation (RCV), the distribution shape of RR interval is described with the skewness parameter (SKP), and the complexity of RR interval is described with the Lempel-Ziv complexity (LZC). Finally, the feature vectors have been used as input into the support vector machine (SVM) classifier model to achieve automatic classification and detection of atrial fibrillation. Also in this case, the MIT-BIH atrial fibrillation database was used to verify the data. The final classification results showed a sensitivity of 95.81%, a specificity of % and an accuracy equal to 96.09%.

In the next subsection, the method chosen as baseline—and embedded in MORPHYTHM and consequently in the new approaches proposed in this paper— is described by providing the main ideas and highlighting the computational steps.

The Method Proposed by Zhou *et al.* [41]**.** This section provides details on the method proposed by Zhou *et al.* [41], *i.e.,* our baseline in the evaluation of LOCAL MORPHYTHM. Such an approach consists in the following steps:

– the HR sequence is converted to a symbolic sequence in a fixed interval;
– a probability distribution is constructed from the word sequence which is transformed from the symbolic sequence;
– a coarser version of Shannon entropy is employed to quantify the information size of HR sequence using the probability distribution of word sequence;
– discrimination of the heart beat type (AF or no-AF) using a threshold.

Step 1: Converting the HR Sequence. The first step of the method regards the generation of a symbolic dynamic starting from the analysis of a sequence of heart beat (hr_n). Especially, the authors encode the information included in hr_n to a sequence of fewer symbols, where each symbol aims at representing an instantaneous state of heart beating. The mapping function is the following:

$$sy_n = \begin{cases} 63, & \text{if } n \text{ hr} \geq 315 \\ \lfloor hr_n \rfloor, & \text{other cases} \end{cases}$$

where $[\lfloor \cdot \rfloor]$ represents a floor operator.

Step 2: Building the Symbolic Sequence. The authors apply a 3-symbols template in order to explore the entropic properties of the symbolic series sy_n. Thus, to examine the chaotic behavior, the word value can then be calculated as:

$$wv_n = (sy_{n-2} \times 2^1 2) + (sy_{n-1} \times 2^6) + sy_n$$

Step 3: Computing the Entropy. The authors define a coarser version of Shannon entropy $H''(A)$ to quantitatively calculate the information size of wv_n. In this study, the dynamic A comprises of 127 consecutive word elements from wv_{n-126} to wv_n, as proposed in the function below:

$$H''(A) = -\frac{k}{Nlog_2N} \sum_{i=1}^{k} p_i log_2 p_i$$

where N and k are total number of the elements and characteristic elements in space A, respectively.

Step 4: Classification. Based on the obtained entropy value, a final beat-to-beat classification (*fibrillant* or *not-fibrillant*) is presented by applying a threshold discrimination. The optimal threshold was empirically identified at 0.639.

3   The Proposed Approach: An Overview

This section describes LOCAL MORPHYTHM, an evolution of the approach recently proposed by Laudato *et al.* [23], called MORPHYTHM. LOCAL MORPHYTHM is able—given a heart beat signal—to classify it as fibrillating or not fibrillating.

As well as MORPHYTHM, LOCAL MORPHYTHM uses supervised machine learning techniques[2] to combine rhythmic and morphological features extracted from an ECG and predict whether or not a heart beat is fibrillating or not fibrillating. However, in LOCAL MORPHYTHM, (i) a rigorous feature engineering process and (ii) a local prediction strategy have been adopted in order to identify respectively the features that best contribute to the prediction of AF episodes and to evaluate if a local approach may be preferred instead of a global one.

3.1   Pre-processing

Before extracting features, the ECG data have to be pre-processed according to [33] and [6]. The main steps involved in this phase are: (i) the detrend of the ECG signal, (ii) the application of a filtering stage (where a low and high pass filters have been applied to get rid of baseline wander and discard high frequency noise, respectively) and (iii) the normalization of the samples.

Once executed the previous steps, the Pan-Tompkins [33] QRS-detection method has been applied with the aim at segmenting the ECG in heart beat signals. In this work, as *heart beat*, it is intended the signal included between two successive R peaks. Such an interpretation is very suitable for AF detection, because it highlights the atrial activity.

---

[2] In the LOCAL MORPHYTHM evaluation, we experimented several supervised machine learning technique.

3.2   Feature Extraction and Selection

As the name suggests, LOCAL MORPHYTHM embeds features extracted from both the heart rhythm and the morphology of a heart beat. Rhythmic features are based on one or more heart beats and they aim at capturing aspects that mostly regard the regularity of the heart beat signal. Zhou *et al.* [41] state that the detection methods based on RRI are more useful to produce a precise and accurate identification of AF because the R-wave peak of the QRS complex is the most prominent characteristic feature of an ECG recording. Such a characteristic is less subject to noise [26].

Even if the acquisition of rhythmic features can be very reliable, such features can only help detecting arrhythmia, which is just one of the possible signs of AF. Thus, morphological features are necessary to detect anomalies in the shape of a single heart beat signal and could be particularly useful to corroborate the warnings raised by analysing the rhythm.

Thus, similarly to MORPHYTHM, also in LOCAL MORPHYTHM we consider both rhythm and morphological features. Especially, we consider the same set of features used in MORPHYTHM [23]:

– *Rhythmic Features*: we used two features based on the observation of a single heart beat signal, *i.e.,* Heart Beat Length (HBL) and Heart Beat Discrete Length (HBDL), and two additional rhythmic features that consider the information of a sequence of consecutive heart beats, *i.e.,* Heart Beat Regularity (HBR) and Entropy, as defined by [41]. HBL represents how long a single heart beat signal lasts. HBDL is a classification of the heart beat signal in three classes, based on its length: a beat is (i) *short* if it takes less than $0.5\,\mathrm{s}$, (ii) *long* if it takes more than $1.2\,\mathrm{s}$, and (iii) *regular* otherwise. HBR is based on HBDL. It considers a rhythmic pattern of 10 consecutive discrete heart beats lengths. Once obtained the pattern, we compute HBR simply counting the number of regular heart beats.
– *Morphological Features*: given a sequence of samples provided for a heart beat signal, we computed several features: (i) the Mean Signal Intensity (MSI), (ii) the Signal Intensity Variance (SIV), (iii) the Signal Intensity Entropy (SIE). MSI, SIV and SIE are features obtained by measuring respectively the mean, the variance and the entropy [29] of all the samples acquired in a heart beat signal. To try to enrich the knowledge of classifiers, we also used the segmented version of these last features: we divided proportionally the heart beat signal in 10 segment and for each portion we evaluated the MSI, SIV and SIE. Finally, we included the features obtained by (i) the application of the Fast Fourier Transform on 32 points and (ii) the estimation of the coefficients of the Auto-Regressive model of order 16.

For each heart beat signal we extract a total number of 76 different features (eight rhythm features and 68 morphological features). In order to select the most appropriate features for the detection of AF events we used the Weka *Info-GainAttributeEval* as Attribute Evaluator and *Ranker* as Search Method. The

Table 1. Features ranking using Information Gain.

Rank	InfoGain	Attribute	Type
1	0.86	Entropy from Zhou *et al.* [41]	Rhythmic
2	0.20	Entropy from the rhythmic pattern	Rhythmic
3	0.18	Heart beat absolute length	Rhythmic
4	0.14	Coeff. no. 10 from AR model	Morphological
5	0.13	Coeff. no. 11 from AR model	Morphological
6	0.12	Coeff. no. 7 from AR model	Morphological
7	0.11	Coeff. no. 12 from AR model	Morphological
8	0.11	Coeff. no. 1 from AR model	Morphological
9	0.11	Coeff. no. 8 from AR model	Morphological
10	0.11	Coeff. no. 9 from AR model	Morphological
11	0.11	Coeff. no. 3 from AR model	Morphological
12	0.10	Coeff. no. 6 from AR model	Morphological
13	0.10	Coeff. no. 2 from AR model	Morphological
14	0.10	Coeff. no. 4 from AR model	Morphological
15	0.10	Coeff. no. 3 from FFT model	Morphological
16	0.10	Coeff. no. 31 from FFT model	Morphological
...	...	...	...
75	0.03	Entropy of Sample Amplitudes	Morphological
76	0.01	Length discrete class	Rhythmic

former basically evaluates the worth of an attribute by measuring the information gain with respect to the class, while the latter ranks attributes by their individuals evaluations.

The feature selection process has been conducted on the MIT-BIH AF Database [16], a commonly used benchmark which contains recordings of 25 patients. Each recording in the data set lasts 10 h and contains two ECG signals sampled at 250 samples per second (12-bit resolution). Due to the embedding of morphology descriptors, the feature selection process has been performed on the $AFDB_2$, i.e., the AFDB without records 00735 and 03665 because, for such records, only information on the rhythm is available [16] and without 04936 and 05091 because—as others have shown [25]—the records 04936 and 05091 include many incorrect manual AF annotations.

The outcome of the features selection process is reported in Table 1. From the analysis of the results achieved, we observe that:

– rhythmic information in AF episodes detection represent the main contribution in terms of information gain;
– morphological features of an ECG can provide a contribution in terms of information gain for the automatic classification of heart beats. Specifically,
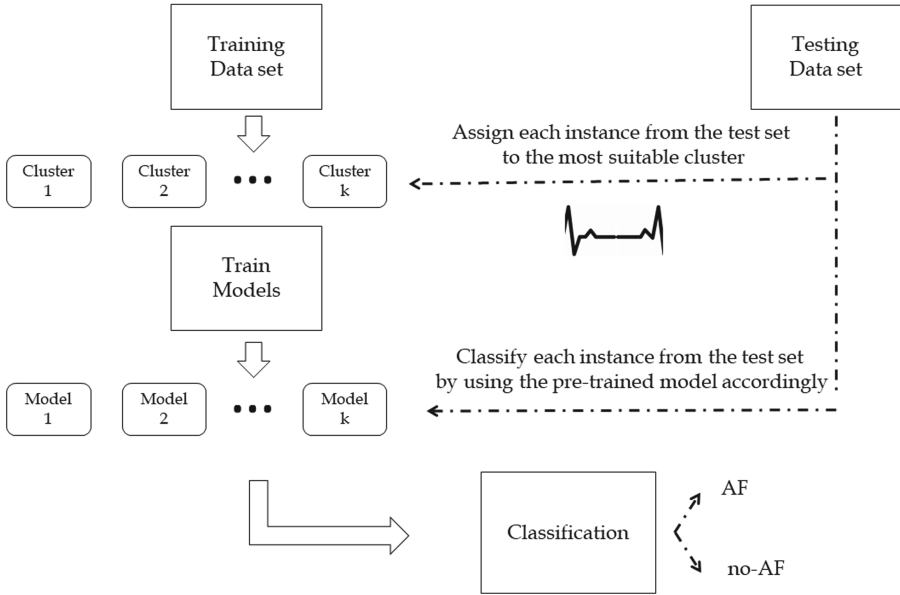
Fig. 1. Workflow of LOCAL MORPHYTHM.

these features refer to the middle and the last part of the signal, where the fibrillating rhythm appears and where the P-wave can exhibit its changes.

By selecting a fixed threshold of 0.12, we obtain a selection of a group of six features containing a balanced number of morphological and rhythmic features. Thus, we decided to incorporate in LOCAL MORPHYTHM the first six features reported in Table 1.

3.3   Making the Prediction

The main difference between LOCAL MORPHYTHM and MORPHYTHM regards the way as the prediction is performed. In MORPHYTHM, as in any canonical approach based on supervised machine learning techniques, a training set is used to build a (global) prediction model. Such a model is used on all the new data points where a prediction is required. Especially, when a new heart beat signal is provided, MORPHYTHM first computes the features on this new heart beat signal and then uses the prediction model to determine whether or not the heart beat is fibrillating or not fibrillating.

However, the heart beat signals in the training set could be quite different each other. The heterogeneity of the training set might negatively impact the accuracy of the prediction model [28]. In order to mitigate such a problem, in LOCAL MORPHYTHM we integrated a local prediction strategy [28].

LOCAL MORPHYTHM first clusters the training set into homogeneous sets of heart beat signals. Then, it builds for each cluster a specific prediction model
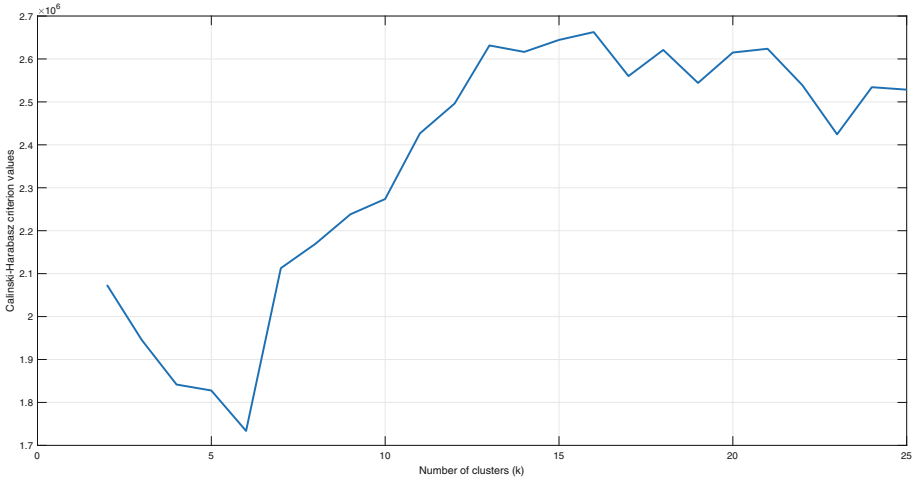
Fig. 2. Results of the Calinski-Harabasz score in order to determine the best value of k for the k-means clustering algorithm. The higher the value of the score the higher the overall quality of the clustering.

using a supervised machine learning technique. In this way, LOCAL MORPHYTHM does not have just one global prediction model, but it has a set of prediction models that are particularly suitable for specific heart beat signals.

When a new heart beat signal is provided, LOCAL MORPHYTHM first computes the features on this new heart beat signal and then it identifies the cluster of heart beat signals more similar to the new heart beat signal. Once identified such a cluster, LOCAL MORPHYTHM uses the model associated to the identified cluster of heart beat signals to noindent predict whether or not the new heart beat is fibrillating or not fibrillating. The workflow of LOCAL MORPHYTHM is depicted in Fig. 1.

In order to cluster the training set, we have exploited the k-means clustering algorithm [27]. This method follows a simple way to classify a given data set through a certain number of clusters fixed *a priori*. The main idea is to define k centroids, one for each cluster. The main steps are described below:

– Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
– Assign each object to the group that has the closest centroid.
– When all objects have been assigned, recalculate the positions of the K centroids.
– Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

We have determined the optimal value of k using the Variance Ratio Criterion (also known as Calinski-Harabasz score) [4]. Especially, we have performed the

clustering of the heart beats by using different values of k from 1 to 25. For each cluster we have computed the Calinski-Harabasz score in order to determine the value of k that determines the clustering with the highest score. The plot in Figure shows that the highest Calinski-Harabasz value occurs with $k = 16$. This number has also been confirmed by involving the Silhouette method [36], an alternative method for the identification of the best k value (Fig. 2).

4   Empirical Evaluation

This section reports the empirical evaluation we conducted to evaluate the accuracy of LOCAL MORPHYTHM.

4.1   Design of the Study

The goal of this study is to evaluate the accuracy of LOCAL MORPHYTHM is classifying AF events in a patient. The perspective is both (i) of a researcher who wants to understand if a local prediction strategy to combine rhythmic and morphological features is worthwhile for detecting AF events, and (ii) of a practitioner who wants to use the most accurate and precise approach in a telemedicine application for the detection of AF events. Thus, the study is steered by the following research question:

> *To what extent, a local prediction model—based on the combination of rhythmic and morphological information—improves the automatic detection of AF episodes?*

The context of this study is represented by the MIT-BIH AF Database [16], and specifically the $AFDB_2$, *i.e.*, the AFDB without records 00735 and 03665 because, for such records, only information on the rhythm is available [16]. Also, records 04936 and 05091 were excluded due to many incorrect manual AF annotations [25].

In the context of our study, we also experimented a large set of machine learning techniques. Indeed, for the classification performances, we have involved in our experiments—beyond the Random Forest [17], J48 [35], Logistic [9], AdaBoost M1 [13] and RepTree [11] already used by Laudato *et al.* to evaluate MORPHYTHM [23]—Neural Networks [22], Multi Layer Perceptron [32], JRip [7] and SGD (which implements stochastic gradient descent for learning various linear models)[3].

As validation technique, we have chosen the Leave One Person Out Cross Validation (L1PO-CV). L1PO-CV means that one person at a time is left out from the training set, so that the training set contains no data specific to the individual who is being tested (the classifier was not tuned with the test data of that person). This is possible since each data segment is associated with an anonymous label corresponding to an individual.

---

[3] https://weka.sourceforge.io/doc.stable-3-8/weka/classifiers/functions/SGD.html.

To answer our research questions we compared:

– True Positives (TP), *i.e.,*, the number of instances classified as fibrillating by the approach and that were actually fibrillating;
– True Negatives (TN), *i.e.,*, the number of instances classified as not fibrillating by the approach and that were actually not fibrillating;
– False Positives (FP), *i.e.,*, the number of instances classified as fibrillating by the approach and that were actually not fibrillating;
– False Negatives (FN), *i.e.,*, the number of instances classified as not fibrillating by the approach and that were actually fibrillating.

In the context of telemedicine a high number of TP is desirable, because it indicates the number of AF episodes correctly detected. Also, it is desirable to have an approach that does not lose any AF episode: thus, keeping the number of FN low is very important.

4.2   Analysis of the Results

Table 2 compares the prediction accuracy, in terms of TP, TN, FP, and FN, achieved by Local Morphythm, Morphythm, and the approach proposed by Zhou *et al.* [41], the most accurate approach in the literature for the detection of AF events.

From the analysis of the results emerges that for both the approaches Morphythm and Local Morphythm the best overall accuracy is achieved when SGD is used as machine learning techniques.

Using such a technique, Local Morphythm is able to achieve the best results in terms of both TP and FN. Specifically, Local Morphythm is able to identify 8,340 TP more than the baseline (approach by Zhou *et al.*) and 1,114 TP more than Morphythm. Also, Local Morphythm is able to retrieve less FN with respect to both the baseline and Morphythm, *i.e.,* −5,533 and −569, respectively.

However, the approach proposed by Zhou *et al.* [41] is still the best in terms of TN and FP. Specifically, Local Morphythm and Morphythm generate 6,052 and 6,064 FP more than the approach by Zhou, respectively. In terms of TN, instead Local Morphythm and Morphythm retrieves less TN as compared to the baseline, *i.e.,* −8,859 and −8,326, respectively.

By looking at the results achieved at patient level, *i.e.,* by considering a single recording, we observe that Local Morphythm sensibly outperforms—in terms of every metrics—both the baseline and Morphythm for 5 out of 21 recordings (around 24%). Examples of such an improvement is reported in Table 3, where it is possible to observe the classification performances of Local Morphythm with respect to the baseline and Morphythm.

In addition, if we focus the attention on just TP and FN, Local Morphythm outperforms both the other approaches baselines in 8 out of 21 recordings (around 38% of the data set).

For the remaining recordings, the value of all the evaluation metrics are almost balanced, in the sense that no significant improvement can be observed.

Table 2. Comparison of LOCAL MORPHYTHM with MORPHYTHM (with the same features selection strategy used in LOCAL MORPHYTHM) and the approach proposed by Zhou *et al.* [41]. In boldface the best results achieved by these methods.

Approach	TP	TN	FP	FN
Zhou *et al.* on AFDB$_2$	457,001	**554,247**	**15,513**	12,966
MORPHYTHM—Random Forest	459,211	534,822	34,489	11,205
LOCAL MORPHYTHM—Random Forest	458,980	534,824	34,501	11,422
MORPHYTHM—J48	449,471	512,763	54,209	23,284
LOCAL MORPHYTHM—J48	446,947	513,453	55,259	24,068
MORPHYTHM—Logistic	463,730	545,621	22,184	8,192
LOCAL MORPHYTHM—Logistic	464,623	545,624	22,003	7,477
MORPHYTHM—AdaBoost M1	461,635	549,572	16,188	12,332
LOCAL MORPHYTHM—AdaBoost M1	461,214	547,589	18,287	12,637
MORPHYTHM—RepTree	451,962	522,829	42,931	22,005
LOCAL MORPHYTHM—RepTree	452,231	522,819	42,899	21,778
MORPHYTHM—3-layers LSTM NN	462,730	545,621	22,484	8,892
LOCAL MORPHYTHM—3-layers LSTM NN	460,076	546,799	23,081	9,771
MORPHYTHM—3-layers Conv. NN	461,319	546,032	23,260	9,116
LOCAL MORPHYTHM—3-layers Conv. NN	459,660	546,020	23,695	10,352
MORPHYTHM—MultiLayer Perceptron	457,595	544,031	26,964	11,137
LOCAL MORPHYTHM—MultiLayer Perceptron	457,606	544,017	26,992	11,112
MORPHYTHM—JRip	452,966	522,840	42,121	21,800
LOCAL MORPHYTHM—JRip	451,599	523,296	42,571	22,261
MORPHYTHM—SGD	464,227	545,921	21,577	8,002
LOCAL MORPHYTHM—SGD	**465,341**	545,388	21,565	**7,433**

Table 3. Example of records on which LOCAL MORPHYTHM outperforms both MORPHYTHM and the approach by Zhou *et al.* [41] in terms of all the considered evaluation metrics.

Record	Interval	TP	TN	FP	FN
04043	Zhou *et al.* [41]	8,690	44,299	3,063	5,862
	Best MORPHYTHM– Logistic	9,608	43,565	3,797	4,944
	LOCAL MORPHYTHM– AdaBoost M1	**10,090**	**44,991**	**2,371**	**4,462**
06426	Zhou *et al.* [41]	52,104	815	1,229	1,006
	Best MORPHYTHM– Logistic	52,633	629	1,415	477
	LOCAL MORPHYTHM– SGD	**52,576**	**901**	**1,143**	**534**

The only recording with abnormal classification performances is the recording 08378 where LOCAL MORPHYTHM presents a significant loss in terms of TP and FN with respect to the other two approaches. This suggests that on this particular recording the local prediction strategy is not worthwhile because very
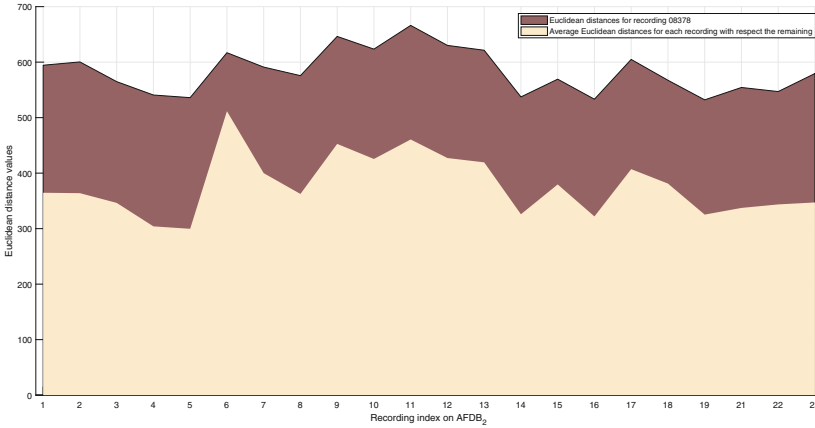
Fig. 3. Average distance between a generic recording i and all the other recordings but 08378 compared to the distance between recording i and 08378.

likely such a recording exhibits characteristics that are quite different from the other recordings in the data set.

In order to validate such a conjecture we compare the average distance between each recording and all the others but 08378 and the distance between each recording and recording 08378. In order to compute the distance between two recordings we considered them as mono dimensional vectors (by selecting the first ECG channel available for each recording) and then compute the Euclidean distance between the two vectors.

The analysis is depicted in Fig. 3. As we can see, the distance between the recording 08378 and a generic recording i is much higher that the average distance between the recording i and all the other recordings but 08378. Such a result confirms our conjecture that the recording 08378 is quite different from the others recording; thus, in this specific case, the local prediction strategy does not provide any benefits as compared to the other two approaches.

Once this recording is excluded from the data set, the classification accuracy of LOCAL MORPHYTHM improves even more. Indeed, LOCAL MORPHYTHM—especially when using the Logistic and the SGD algorithms—avoids a loss of around 1,5 thousands heart beats classified as TP and FN.

5   Conclusion and Future Work

In this paper we presented an extended version of the approach proposed by Laudato *et al.* [23], named MORPHYTHM, where rhythmic and morphological features are combined together in order to improve the classification accuracy of AF episodes. The new approach, called LOCAL MORPHYTHM integrates a more rigorous feature engineering process as compared to MORPHYTHM and more advanced machine learning techniques, including artificial neural networks.

We also extended MORPHYTHM by integrating in the approach the strategy of "local" prediction, successfully used in other contexts [28]. Especially, instead of producing a single prediction model, LOCAL MORPHYTHM automatically builds several prediction models based on the characteristics of the ECGs in the training set. In particular, the training set is clustered in order to put together ECGs that exhibits similar characteristics. Then, for each cluster, LOCAL MORPHYTHM builds a prediction model. When a new data point is provided, LOCAL MORPHYTHM first selects the most suitable model based on the characteristics of the new data point, and then it performs the prediction applying the selected model.

An experimentation conducted on the MIT-BIH AF Database [16] indicates that LOCAL MORPHYTHM is able to increase the TP and reduce the FN as compared to MORPHYTHM and the approach by Zhou *et al.* [41], one of the best approaches in the literature for the detection of AF episodes. Future work will be devoted on the one hand on the replication of the experimentation on other data sets in order to corroborate the results achieved on the MIT-BIH AF Database and on the other hand on the application of a local prediction technique in the context of automatic detection of other types of arrhythmia.

References

1. Afdala, A., Nuryani, N., Nugroho, A.S.: Automatic detection of atrial fibrillation using basic Shannon entropy of RR interval feature. J. Phys.: Conf. Ser. **795**, 012038 (2017)
2. Andersen, R.S., Peimankar, A., Puthusserypady, S.: A deep learning approach for real-time detection of atrial fibrillation. Expert Syst. Appl. **115**, 465–473 (2019)
3. Balestrieri, E., et al.: The architecture of an innovative smart T-shirt based on the internet of medical things paradigm. In: 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA), pp. 1–6. IEEE (2019)
4. Caliński, T., Harabasz, J.: A dendrite method for cluster analysis. Commun. Stat.-Theory Methods **3**(1), 1–27 (1974)
5. Chen, Z., Li, J., Li, Z., Peng, Y., Gao, X.: Automatic detection and classification of atrial fibrillation using RR intervals and multi-eigenvalue. Sheng wu yi xue gong cheng xue za zhi= J. Biomed. Eng.= Shengwu yixue gongchengxue zazhi **35**(4), 550–556 (2018)
6. Clifford, G.D., Azuaje, F., McSharry, P., et al.: Advanced Methods and Tools for ECG Data Analysis. Artech House, Boston (2006)
7. Cohen, W.W.: Fast effective rule induction. In: Machine Learning Proceedings 1995, pp. 115–123. Elsevier (1995)
8. Colloca, R., Johnson, A.E., Mainardi, L., Clifford, G.D.: A support vector machine approach for reliable detection of atrial fibrillation events. In: Computing in Cardiology 2013, pp. 1047–1050. IEEE (2013)
9. Cramer, J.S.: The origins of logistic regression (2002)
10. Dash, S., Chon, K., Lu, S., Raeder, E.: Automatic real time detection of atrial fibrillation. Ann. Biomed. Eng. **37**(9), 1701–1709 (2009). https://doi.org/10.1007/s10439-009-9740-z
11. Devasena, C.L.: Comparative analysis of random forest, REP tree and J48 classifiers for credit risk prediction. Int. J. Comput. Appl. (2014). 0975-8887

12. Elliott, T., Yopes, M.C.: Direct-to-consumer telemedicine. J. Allergy Clin. Immunol. Pract. **7**(8), 2546–2552 (2019)
13. Freund, Y., Schapire, R.E.: A desicion-theoretic generalization of on-line learning and an application to boosting. In: Vitányi, P. (ed.) EuroCOLT 1995. LNCS, vol. 904, pp. 23–37. Springer, Heidelberg (1995). https://doi.org/10.1007/3-540-59119-2_166
14. Friberg, L., Rosenqvist, M., Lindgren, A., Terént, A., Norrving, B., Asplund, K.: High prevalence of atrial fibrillation among patients with ischemic stroke. Stroke **45**(9), 2599–2605 (2014)
15. Fuster, V., et al.: ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American college of cardiology/american heart association task force on practice guidelines and the european society of cardiology committee for practice guidelines (writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation): developed in collaboration with the European heart rhythm association and the heart rhythm society. Circulation **114**(7), e257–e354 (2006)
16. Goldberger, A.L., et al.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)
17. Ho, T.K.: The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Intell. **20**(8), 832–844 (1998)
18. Hochstadt, A., Chorin, E., Viskin, S., Schwartz, A.L., Lubman, N., Rosso, R.: Continuous heart rate monitoring for automatic detection of atrial fibrillation with novel bio-sensing technology. J. Electrocardiol. **52**, 23–27 (2019)
19. Huang, C., Ye, S., Chen, H., Li, D., He, F., Tu, Y.: A novel method for detection of the transition between atrial fibrillation and sinus rhythm. IEEE Trans. Biomed. Eng. **58**(4), 1113–1119 (2010)
20. Kleyko, D., Osipov, E., Wiklund, U.: A comprehensive study of complexity and performance of automatic detection of atrial fibrillation: classification of long ECG recordings based on the PhysioNet computing in cardiology challenge 2017. Biomed. Phys. Eng. Express **6**(2), 025010 (2020)
21. Lake, D.E., Moorman, J.R.: Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. Am. J. Physiol.-Heart Circ. Physiol. **300**(1), H319–H325 (2011)
22. Lang, S., Bravo-Marquez, F., Beckham, C., Hall, M., Frank, E.: WekaDeeplearning4j : a deep learning package for Weka based on deeplearning4j. Knowl.-Based Syst. **178**, 48–50 (2019)
23. Laudato, G., et al.: Combining rhythmic and morphological ECG features for automatic detection of atrial fibrillation. In: HEALTHINF, pp. 156–165 (2020)
24. de Heuzey, J.Y., et al.: Cost of care distribution in atrial fibrillation patients: the COCAF study. Am. Heart J. **147**(1), 121–126 (2004)
25. Lee, J., Nam, Y., McManus, D.D., Chon, K.H.: Time-varying coherence function for atrial fibrillation detection. IEEE Trans. Biomed. Eng. **60**(10), 2783–2793 (2013)
26. Lian, J., Wang, L., Muessig, D.: A simple method to detect atrial fibrillation using RR intervals. Am. J. Cardiol. **107**(10), 1494–1497 (2011)
27. MacQueen, J., et al.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Oakland, CA, USA, vol. 1, pp. 281–297 (1967)
28. Menzies, T., et al.: Local versus global lessons for defect prediction and effort estimation. IEEE Trans. Softw. Eng. **39**(6), 822–834 (2012)

29. Moddemeijer, R.: On estimation of entropy and mutual information of continuous distributions. Sig. Process. **16**(3), 233–248 (1989)
30. Mohebbi, M., Ghassemian, H.: Detection of atrial fibrillation episodes using SVM. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 177–180. IEEE (2008)
31. Niknejad, N., Ismail, W.B., Mardani, A., Liao, H., Ghani, I.: A comprehensive overview of smart wearables: the state of the art literature, recent advances, and future challenges. Eng. Appl. Artif. Intell. **90**, 103529 (2020)
32. Pal, S.K., Mitra, S.: Multilayer perceptron, fuzzy sets, classification (1992)
33. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. **3**, 230–236 (1985)
34. Procter, N.E., Stewart, S., Horowitz, J.D.: New-onset atrial fibrillation and thromboembolic risk: cardiovascular syzygy? Heart Rhythm **13**(6), 1355–1361 (2016)
35. Quinlan, J.R.: C4.5: Programs for Machine Learning. Elsevier, Amsterdam (2014)
36. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**, 53–65 (1987)
37. Sepulveda-Suescun, J.P., Murillo-Escobar, J., Urda-Benitez, R.D., Orrego-Metaute, D.A., Orozco-Duque, A.: Atrial fibrillation detection through heart rate variability using a machine learning approach and Poincare plot features. CLAIB 2016. IP, vol. 60, pp. 565–568. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-4086-3_142
38. Sim, I., et al.: Clinical decision support systems for the practice of evidence-based medicine. J. Am. Med. Inform. Assoc. **8**(6), 527–534 (2001)
39. Stewart, S., Murphy, N., Walker, A., McGuire, A., McMurray, J.: Cost of an emerging epidemic: an economic analysis of atrial fibrillation in the UK. Heart **90**(3), 286–292 (2004)
40. Yuan, C., Yan, Y., Zhou, L., Bai, J., Wang, L.: Automated atrial fibrillation detection based on deep learning network. In: 2016 IEEE International Conference on Information and Automation (ICIA), pp. 1159–1164. IEEE (2016)
41. Zhou, X., Ding, H., Wu, W., Zhang, Y.: A real-time atrial fibrillation detection algorithm based on the instantaneous state of heart rate. PloS One **10**(9), e0136544 (2015)
42. Zoni-Berisso, M., Lercari, F., Carazza, T., Domenicucci, S.: Epidemiology of atrial fibrillation: European perspective. Clin. Epidemiol. **6**, 213 (2014)