Article

A Tandem Feature Extraction Approach for Arrhythmia Identification

Javier Tejedor *[ID]**, David G. Marquez, Constantino A. Garcia and Abraham Otero**

Information Technology Department, Escuela Politécnica Superior, Universidad San Pablo-CEU,
CEU Universities, Urbanización Montepríncipe, 28668 Boadilla del Monte, Madrid, Spain;
david.gonzalezmarquez@ceu.es (D.G.M.); constantino.garciama@ceu.es (C.A.G.); aotero@ceu.es (A.O.)
* Correspondence: javier.tejedornoguerales@ceu.es

Abstract: Heart disease is currently the leading cause of death in the world. The electrocardiogram (ECG) is the recording of the electrical activity generated by the heart. Its low cost and simplicity have made it an essential test for monitoring heart disease, especially for the identification of arrhythmias. With the advances in electronic technology, there are nowadays sensors that enable the recording of the ECG during the daily life of the patient and its wireless transmission to healthcare facilities. This type of information has a great potential to detect cardiac diseases in their early stages and to permit early interventions before the patient's health deteriorates. However, to usefully exploit the large volume of information obtained from ambulatory ECG, pattern recognition techniques that are capable of automatically analyzing it are required. Tandem feature extraction techniques have proven to be useful for the processing of physiological parameters such as the electroencephalogram (EEG) and speech. However, to the best of our knowledge, they have never been applied to the ECG. In this paper, the utility of tandem feature extraction for the identification of arrhythmias is studied. The coefficients of a regression using Hermite functions are used to create a feature vector that represents the heartbeat. A multiple-layer perceptron (MLP) is trained using these features and its posterior probability outputs are used to extend the original feature vector. Finally, a Gaussian mixture model (GMM) is trained on the extended feature vectors, which is then used in a GMM-based arrhythmia identification system. This approach has been validated using the MIT-BIH Arrhythmia database. The accuracy of the Gaussian mixture model increased by 15.8% when applied over the extended feature vectors, compared to its application over the original feature vectors, showing the potential of tandem feature extraction for ECG analysis and arrhythmia identification.

Keywords: tandem feature extraction; Hermite function; arrhythmia identification; ECG

1. Introduction

Cardiovascular diseases are the main cause of death in the world [1,2]. Approximately 18 million people died from cardiovascular diseases in 2019, representing 33% of all deaths worldwide. In the near future, it is expected that the proportion of global deaths from heart disease will increase [3], and this increase will be more pronounced in developing countries due to changes in diet and lifestyle derived from the greater purchasing power of their citizens [4].

The electrocardiogram (ECG) is a fundamental test in the clinical routine for the diagnosis and monitoring of cardiovascular diseases. Its low cost, simplicity of usage, non-invasive nature and the simplicity of the instrumentation necessary for its acquisition, make it an ideal candidate for long-term ambulatory monitoring during the patient daily life [5,6]. In the ECG, a lead is a measure of the electrical activity of the heart given by the difference in potential between two points. This difference can be measured between two electrodes (bipolar lead) or between a virtual point and an electrode (monopolar lead). Different leads provide different perspectives of the electrical stimulus, and therefore complementary information.

The QRS complex is the most distinctive element of the heartbeat in an ECG [7]. It corresponds to the ventricular depolarization, which causes contraction of the right and left ventricles. Normal QRS complexes last between approximately 60 and 120 ms. The duration, amplitude and morphology of the QRS complex provide valuable information about the state of the heart and are useful in the diagnosis of cardiac arrhythmias, that is, of conduction abnormalities and other heart disorders (see Figure 1). Nowadays, there are algorithms that are capable of detecting the position of the QRS complexes with an acceptable degree of satisfaction, obtaining a sensitivity of about 99.9% [8]. However, the identification of the heartbeat morphology is still an open problem [9,10].
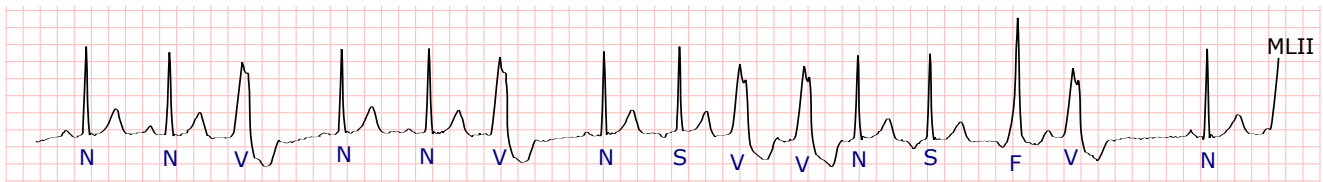

Figure 1. Fragment of an ECG showing different types of heartbeats associated with different QRS morphologies. 'N' stands for normal heartbeat, 'S' for supraventricular heartbeat, 'V' for ventricular heartbeat, and 'F' for fusion heartbeat. (Source: MIT-BIH Arrhythmia Database, recording 208, between 0:17:45 and 0:17:55).

In the last decade, there have been significant advances in electronic technology, including miniaturization of components, increased battery life and decreases in production costs. At the same time, advances in communication technologies have facilitated the wireless transmission of information both in local area networks (such as Wi-Fi, Bluetooth and Zigbee [11]) and over long distances (such as 4G–5G [12]). This has made the ambulatory monitoring of patients during their daily live activities technologically possible and cost effective [13]. However, all this information has a volume too high to be exploited manually by healthcare staff. A normal patient has approximately 100,000 heartbeats per day, and each ECG lead recorded is capturing a different electrical representation of those heartbeats. Hence the need for pattern recognition techniques for the automatic analysis of ECG that are suitable for using in the context of a wearable telemonitoring platform [6,14].

1.1. Related Work

Significant research has been conducted on the usage of pattern recognition techniques in ECG analysis that aim to automatically detect different types of arrhythmias over ECG recordings. The most widely used pattern recognition approaches for years are based on two different stages: feature extraction and classification. The feature extraction stage aims to extract robust (and commonly hand-crafted) features that effectively represent the ECG signal and the classification stage employs those features to carry out heartbeat classification. Topological features derived from persistent homology [15] such as mean, standard deviation, skewness and kurtosis for persistence, birth time, death time, persistence entropy, number of dimensions, sums of persistence, number of layers in the landscapes, number of valleys per layer, and mean, standard deviation, skewness and kurtosis for the number of peaks per layer, along with other feature types such as demographic data and RR intervals have been employed in [16]. Some other works present time-frequency-based features such as the wavelet transform [17–19], wavelet packet entropy [20] and frequency-based features such as fast Fourier transform [19]. Fusion of different feature sets (morphological features based on discrete wavelet transform (DWT), statistical variational features and temporal features) have also been explored [21].

Regarding the classification stage, discriminative models that aim to differentiate between the heartbeat classes are widely used (e.g., support vector machine (SVM) [17] and random forest [16,20]). More advanced discriminative models such as convolutional neural networks have been investigated in [18], and fusion of different classifiers (SVM and nearest neighbour) were also considered in [21].

However, in the recent years, the traditional approaches based on those two stages are being replaced by deep neural network-based approaches. These approaches do not need the feature extraction stage, since they are able to carry out classification from the raw ECG signal [22–26]. Transformer models based on attention and encoder-decoder architectures have also been employed [19,27]. However, all these approaches present an important drawback when considered to be used as part of a daily live monitoring solution: they cannot be fully-integrated on a wearable device due to their computational requirements.

1.2. Motivation and Organization of this Paper

The approaches presented in the literature for ECG analysis (and hence for arrhythmia detection) are commonly based on hand-crafted features. For ECG analysis, Hermite functions have shown to be a compact and robust representation in the presence of noise for feature extraction in ECG signal classification systems [28–30]. On the other hand, the tandem approach for feature extraction was firstly presented in early 2000 [31] for speech recognition tasks. This aims to augment the original hand-crafted features with discriminatively-trained features. These features were originally based on a multiple-layer perceptron (MLP) although linear discriminant analysis (LDA) has also been explored later for some other signals different to speech. To do so, the MLP is employed to obtain a posterior probability for each class to be identified. The resulting set of posterior probabilities in the output layer of the MLP is added to the original features to create the so-called tandem approach. For LDA-based features, the projections of the LDA are employed. For speech signals, the tandem approach does typically incorporate logarithm and PCA decorrelation-based transformations to match the speech signal characteristics. This tandem approach has been shown to significantly improve pattern recognition performance with MLP-based features in speech recognition [32–34], speaker verification [35], language identification [35–37] and fiber optic recognition [38,39]; both MLP and LDA-based features proved useful in electroencephalogram (EEG) recognition [40], with the best result obtained from the MLP-based features.

Based on the power of Hermite functions for ECG signal representation as hand-crafted features [28–30], this work explores whether the tandem approach could be useful for the classification of heartbeat morphology. To do so, the proposal uses an augmented feature extraction strategy, which incorporates new features based on the posterior probabilities output by an MLP to the Hermite-based features. Then, the augmented features are fed to a Gaussian mixture model (GMM)-based classification system which carries out the final classification of the heartbeats. It must be noted that the work presented in [30] also employs Hermite functions for ECG signal representation and MLP for classification. However, our work differs from [30] since we propose the use of the MLP within the feature extraction and we base our classification system in Gaussian mixture modelling. The work presented in [41], which employs multiscale principal component analysis for signal preprocessing, statistical features (i.e., mean, average power, standard deviation and mean value ratio) related to the coefficients of the DWT for feature extraction and decision trees for classification also differs from our approach since the signal preprocessing, feature extraction and classification stages are all different. Therefore it can be said that, to the best of our knowledge, this is the first work that employs a tandem approach for feature extraction in arrhythmia identification from ECG. Moreover, due to the low complexity of the GMM approach employed for classification in this work, the system is able to be fully-integrated in a (low-cost) wearable device.

The rest of the paper is organized as follows: Section 2 presents the database used in this work. The novel tandem feature extraction for ECG arrhythmia identification is presented in Section 3. The experimental procedure is presented in Section 4. Section 5 presents the experiments and results, which are discussed in Section 6. Finally, Section 7 concludes the paper.

2. Database

To validate the technique presented in this work, the most referenced database in the literature of arrhythmia identification will be used: the MIT-BIH Arrhythmia Database [42]. The wide variety of patients, the different types of heartbeats and the large number of annotations have fostered the use of this database [28,43–47]. This database contains 48 electrocardiogram recordings obtained from 47 different patients. Each recording consists of two leads among the following: MLII, V1, V2, V3, V4 and V5. The recordings are digitized at a sampling rate of 360 Hz with a resolution of 11 bits. The database should not be considered a representative sample of the population as the records were carefully selected to try to cover the widest variety of cardiac disorders as possible. Each heartbeat was reviewed by at least two cardiologists, being approximately 68% of them considered as normal and the other 32% were divided into 16 types of abnormal heartbeats.

After the publication of the MIT-BIH Arrhythmia Database, the Association for the Advancement of Medical Instrumentation (AAMI) proposed guidelines for evaluating the performance of arrhythmia identification algorithms and this recommended to use a division of heartbeats into five types: normal (N), supraventricular (S), ventricular (V), fusion (F) and indeterminate (Q) heartbeats [48]. This classification has become a de facto standard, and it will be used here to evaluate the results of our novel approach. The original labels of the MIT-BIH Arrhythmia Database were mapped to the five heartbeat labels recommended by AAMI, as in [28]. This mapping is presented in Table 1.

Table 1. Mapping between MIT-BIH heartbeat classes and AAMI heartbeat classes. 'N' stands for normal heartbeat, 'S' for supraventricular heartbeat, 'V' for ventricular heartbeat, 'F' for fusion heartbeat and 'Q' for indeterminate heartbeat.

AAMI Heartbeat Class	MIT-BIH Heartbeat Class
N	Normal heartbeat Left bundle branch block heartbeat Right bundle branch block heartbeat
S	Aberrated atrial premature heartbeat Supraventricular premature heartbeat Atrial premature heartbeat Nodal (junctional) premature heartbeat Nodal (junctional) escape heartbeat Atrial escape heartbeat
V	Ventricular flutter wave Ventricular escape heartbeat Premature ventricular contraction
F	Fusion of ventricular and normal heartbeats
Q	Paced heartbeat Unclassifiable heartbeat

3. Tandem Feature Extraction for Arrhythmia Identification

The tandem feature approach has been tested on the ECG arrhythmia identification system presented in Figure 2, which is based on four different stages: (1) signal preprocessing, where the ECG signals are initially denoised, (2) raw feature extraction, in which a *raw* set of discriminant features is extracted from the denoised ECG signals, (3) augmented feature extraction, in which the raw feature vector is enhanced with the MLP-based features to create the so-called tandem feature vectors, and (4) pattern classification, which involves two different stages itself: training, which trains the Gaussian mixture model for each AAMI heartbeat class from the training tandem feature vectors, and testing, which classifies each heartbeat into one of the predefined AAMI heartbeat classes from the testing tandem feature vectors (see Table 1). These stages are explained in more detail next.
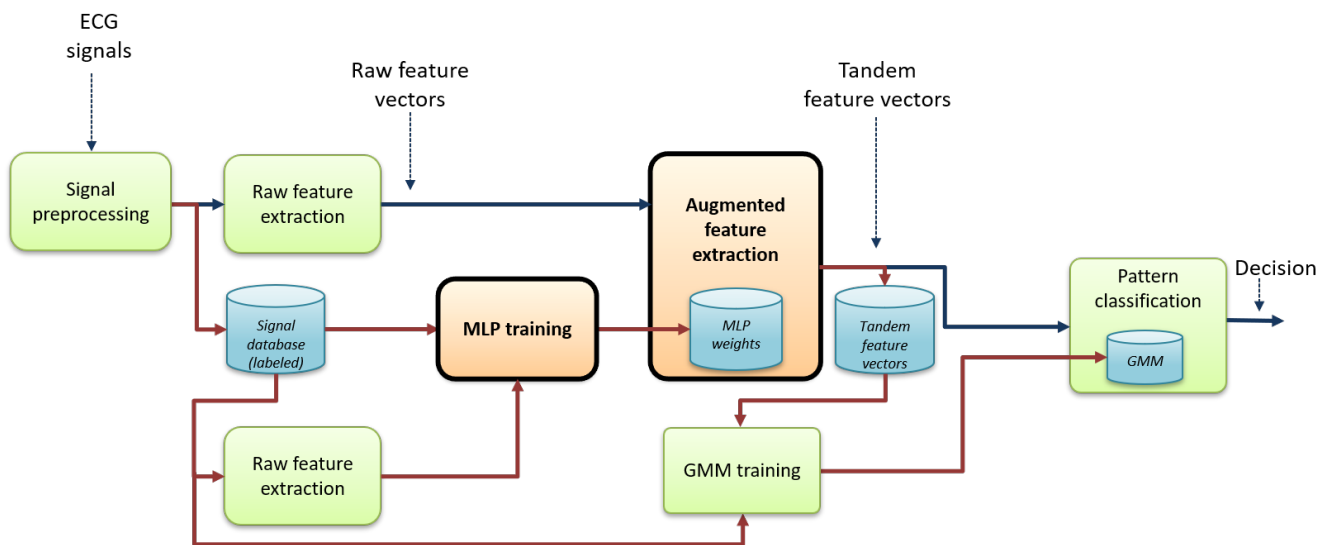
Figure 2. Architecture of the ECG arrhythmia identification system with the modules that convey the main contribution of this paper in bold font.

3.1. Signal Preprocessing

Signal preprocessing aims to filter noise in the signal to allow the feature extraction step to be based on the morphological properties of the heartbeat, without being affected by issues such as baseline drift or high frequency noise. Two filters were applied on the ECG recordings. First, an eight-level Daubechies-based wavelet transform with extremal phase filters of width 4 was applied for baseline drift removal. The result of this filter is used to reconstruct a time series with the baseline drift of the ECG recording, and this drift is subtracted from the original recording. Afterwards, a low-pass Butterworth filter with a cut-off frequency of 40 Hz was applied to eliminate high frequency noise as well as the power line hum. The filtered ECG signals comprise the output of this module and are the input to the raw feature extraction stage.

3.2. Raw Feature Extraction

From the denoised ECG signals, the raw feature extraction aims to obtain the most discriminant information from the various types of heartbeats present in the database. To do so, the Hermite functions were used in this work for raw signal representation. The orthogonal Hermite functions have a shape reminiscent of QRS morphology and include a width parameter that enables an efficient modelling of QRS complexes of different amplitudes. This makes it possible to obtain accurate heartbeat representations with few coefficients. The heartbeat is represented by a feature vector with the coefficients that permit its reconstruction from the combination of the Hermite functions. This representation has been shown to be compact and robust in the presence of noise [28].

From the ECG signals, a 200 ms window was extracted for each heartbeat by considering the samples before and after the actual heartbeat position labelled in the database. Hermite functions tend to zero both in $-\infty$ and ∞. To make Hermite functions converge at window edges, a 100 ms zero segment was added at both sides of the QRS complex so that the resulting window has length a of 400 ms. This window can be represented as Equation (1):

$$x[l] = \sum_{n=0}^{N_h-1} c_n(\sigma)\phi_n[l,\sigma] + e[l],$$ (1)

where l refers the window sample, N_h is the number of Hermite functions, $c_n(\sigma)$ represents the coefficients of the linear combination, $\phi_n[l,\sigma]$ is the n-Hermite discrete function that is obtained by sampling the corresponding Hermite continuous function (i.e., $\phi_n(t,\sigma)$), $e[l]$ is the approximation error between the actual window $x[l]$ and the Hermite representation,

σ is a dilation parameter that relates the width of the Hermite function with the width of the QRS complex and l varies according to Equation (2):

$$l = -\left\lfloor \frac{W \cdot Fs}{2} \right\rfloor, -\left\lfloor \frac{W \cdot Fs}{2} \right\rfloor + 1, ..., \left\lfloor \frac{W \cdot Fs}{2} \right\rfloor, \tag{2}$$

where W is the window size, Fs is the sampling frequency and $\lfloor \rfloor$ represents the floor function.

The Hermite functions $\phi_n[l, \sigma]$ are defined as Equation (3):

$$\phi_n[l, \sigma] = \frac{1}{\sqrt{\sigma 2^n n! \sqrt{\pi}}} e^{\frac{-(l \cdot T_s)^2}{2\sigma^2}} H_n(\alpha), \tag{3}$$

where T_s is the sampling period (i.e., the inverse of the sampling frequency Fs) and α is defined as Equation (4):

$$\alpha = \frac{l \cdot T_s}{\sigma}. \tag{4}$$

The Hermite polynomial $H_n(\alpha)$ in Equation (3) is defined recursively as Equation (5):

$$H_n(x) = 2x H_{n-1}(x) - 2(n-1) H_{n-2}(x), \tag{5}$$

where $H_0(x) = 1$ and $H_1(x) = 2x$.

This Hermite representation enables the representation of the heartbeat contained in each signal window from the N_h coefficients of the linear combination (referred to as $c_n(\sigma)$) from the Hermite functions, and from σ.

For a given value of σ, the Hermite functions form an orthogonal basis, as shown in Equation (6):

$$\int_{-\infty}^{\infty} \phi_n(l) \phi_m(l) = \delta_{nm}. \tag{6}$$

It must be noted that Equation (6) holds if the discrete Hermite function $\phi_n[l, \sigma]$ is close enough to zero on both the edges and outside the analysis window. For the edges of each analysis window, $\phi_n[l, \sigma]$ is at most $1/10$ of its maximum value within the window, as defined in Equation (7):

$$|\phi_n[-l_0, \sigma]| = |\phi_n[l_0, \sigma]| < \frac{1}{10} \max_{l \in [-l_0, l_0]} |\phi_n[l_0, \sigma]|, \tag{7}$$

where $-l_0$ and l_0 refer to the first and last window samples, respectively. Moreover, we also impose that the value of $\phi_n[l, \sigma]$ is smaller outside the analysis window than in the edge of the analysis window, as shown in Equation (8):

$$|\phi_n[l, \sigma]| \le |\phi_n[l_0, \sigma]|, \forall |l| > l_0. \tag{8}$$

For a certain value of σ, the linear combination coefficients $c_n(\sigma)$ are computed by minimizing the summed squared error given by Equation (9):

$$\sum_l (e[l])^2 = \sum_l \left(x[l] - \sum_{n=0}^{N_h - 1} c_n(\sigma) \phi_n[l, \sigma] \right)^2, \tag{9}$$

in which the squared error is approximated following $c_n(\sigma) = x[l] \cdot \phi_n[l, \sigma]$. For a predefined window size and for a fixed number of Hermite functions, it is possible to calculate theoretical limits for the value of σ. Through an incremental iterative process, different values of σ are tested, starting at 0 and going up to the theoretical maximum, until the one that minimizes the error is found. The average values of σ for $N \in [1, 30]$ are from 14 ms to 21 ms.

Then, a raw feature vector x_{rf} is stored for each heartbeat, which consists of the N_h numerical values of the $c_n(\sigma)$ Hermite representation of the corresponding heartbeat plus the σ

value. This process is carried out per each ECG lead available; since our system employs two different ECG leads, $N_r = 2(N_h + 1)$-dimensional raw feature vectors x_{rf} comprise the output of this module and are given to the augmented feature extraction module.

3.3. Augmented Feature Extraction

This module takes the raw feature vectors x_{rf} as input and produces tandem feature vectors x_{tf} as output. An MLP is employed to add the feature-level augmented information to each heartbeat in the ECG arrhythmia identification system. The MLP consists of three layers, as shown in Figure 3: an input layer with N_r raw feature vector values, a hidden layer, whose number of units was selected based on preliminary experiments, and an output layer, which employs the softmax activation function, with a number of units equal to the number of heartbeat classes (five in our case).
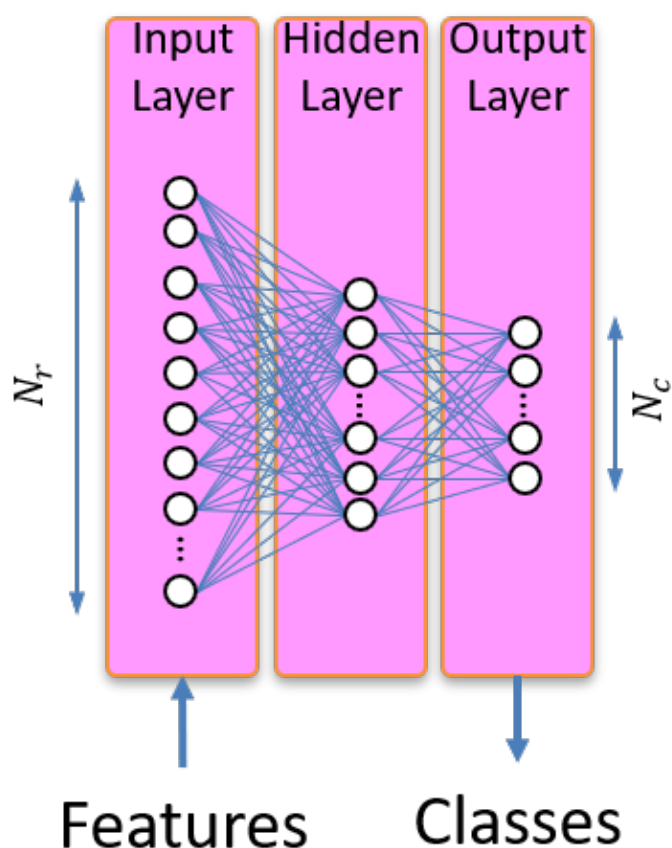

Figure 3. Architecture of the MLP employed in the augmented feature extraction module. N_r is the number of features that are used to represent each heartbeat in the raw feature extraction and N_c is the number of AAMI heartbeat classes to recognize.

The MLP models are trained by the MLP training module in Figure 2. The standard back-propagation algorithm [49] is employed to learn the MLP weights (i.e., connections between the units of the input and hidden layers and connections between the units of the hidden and output layers, as shown in Figure 3) so that the classification error in the training data is minimized. Henceforth, the set of weights learned are used then to obtain the posterior probability vectors.

The augmented feature extraction consists of two different stages, which are applied to each of the N_r-dimensional raw feature vectors x_{rf}, as presented next.

3.3.1. Posterior Probability Vector Computation

From the raw feature vectors x_{rf} and employing the weights computed during MLP training, the MLP calculates a posterior probability for each class to be recognized. This process is similar to the use of the MLP for classification in which each raw feature vector is assigned the class with the highest posterior probability. However, instead of making a class decision for each raw feature vector, the MLP generates one posterior probability per class, as shown in Figure 3. These posterior probability values are then used as new features, hence building a set of N_c-dimensional posterior probability vectors, being N_c the number of different AAMI heartbeat classes.

3.3.2. Tandem Feature Vector Construction

This stage concatenates the original N_r-dimensional raw feature vectors x_{rf} (those generated by the raw feature extraction module) and the N_c-dimensional posterior probability vectors computed by the MLP. Therefore, $(N_r + N_c)$-dimensional tandem feature vectors x_{tf} are built, which are then used in the pattern classification system.

The ICSI QuickNet toolkit [50], which was originally developed for the tandem approach in speech recognition tasks, provides different tools for developing signal processing systems based on MLP strategies. Here, we have used the ICSI QuickNet toolkit with the default parameter values for MLP training, posterior probability vector computation and tandem feature vector construction.

3.4. Pattern Classification

Gaussian mixture modelling has a widespread usage within pattern classification tasks (e.g., speech recognition [51], image recognition [52], video recognition [53], etc.). For ECG arrhythmia identification, GMMs are a suitable tool because: (1) GMMs can be trained from a limited amount of data [54], as it occurs for some heartbeat types present in the MIT-BIH Arrhythmia Database; (2) GMMs provide a simple strategy for classification, making it suitable for embedding the ECG arrhythmia identification system in a wearable device that aims to continuously monitor heart activity; and (3) GMMs can represent a large class of sample distributions (e.g., those corresponding to the training and testing data).

Therefore, for the GMM λ_k, being k one of the five heartbeat classes, the probability that a certain feature vector x_{tf} belongs to the class represented by that model λ_k can be obtained. We will denote this probability as $p(x_{tf}|\lambda_k)$.

3.4.1. Training

From a subset of the heartbeats for a certain class k, which comprises the training subset, the training stage estimates the parameters (i.e., mean and covariance values) of each GMM λ_k from the tandem feature vectors of that subset. To do so, the Expectation-Maximization algorithm [55], which makes use of a maximum likelihood criterion, is employed. This training stage is needed just once, so that the classification stage employs the set of trained GMMs. For the sake of simplicity, a single component for each GMM has been used to train each model.

3.4.2. Classification

Once the models have been trained, classification is conducted on a fully independent data subset, the so-called testing subset. The classification stage finds the class represented by the model $\hat{c}$ with the maximum posterior probability. Hence, for a given input tandem feature vector x_{tf} the Bayes' rule is applied as Equation (10):

$$\hat{c} = \underset{k}{argmax} \frac{p(x_{tf}|\lambda_k)p(\lambda_k)}{p(x_{tf})} = \underset{k}{argmax}\{p(x_{tf}|\lambda_k)\}, \tag{10}$$

where we have considered a uniform prior probability for each class.

4. Experimental Procedure

4.1. Evaluation Metrics

The main metric used to test the system was the classification accuracy, which was computed as Equation (11):

$$Accuracy(\%) = \frac{100 * Correct}{N},\tag{11}$$

where *Correct* is the number of correctly classified testing heartbeats and N represents the total number of testing heartbeats.

We also presented the confusion matrix showing the number of testing heartbeats for a given class that were classified as any of the considered AAMI classes, along with the sensitivity and specificity values for each class, which were defined as Equations (12) and (13), respectively:

$$Sensitivity(\%) = \frac{TP_k}{TP_k + FN_k},\tag{12}$$

where TP_k is the number of true positive testing heartbeats for class k (i.e., heartbeats of class k that are correctly classified by the system) and FN_k is the number of false negative testing heartbeats for class k (i.e., heartbeats of class k that are incorrectly classified by the system). It must be noted that sensitivity metric coincided with the intra-class accuracy. The specificity was calculated as:

$$Specificity(\%) = \frac{TN_k}{TN_k + FP_k},\tag{13}$$

where TN_k is the number of true negative testing heartbeats for class k (i.e., heartbeats of all the classes except k that are classified by the system as any class except k) and FP_k is the number of false positive testing heartbeats for class k (i.e., heartbeats that were incorrectly classified by the system as belonging to class k).

4.2. System Configuration

Regarding feature extraction, $N_h = 30$ Hermite functions were used since they showed to optimally represent the vast majority of the heartbeats according to both the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) [28]. This meant that we used a $N_r = 62$-dimensional raw feature vectors, since two leads were used, and each lead provides 30 Hermite coefficients and the σ parameter. Then, the feature vector was augmented to $N_r + N_c = 67$-dimensional tandem feature vectors by adding the five posterior probabilities calculated by the MLP, according to the AAMI heartbeat classes. The MLP training and posterior probability computation employed a hidden layer with 100 units.

4.3. Evaluation Strategy

Experiments were carried out following the training/testing data division presented in [56] (see Table 2). It must be noted that the work presented in [56] did not employ the paced recordings in the MIT-BIH Arrhythmia database. Since our work does employ those recordings (no recording from the database was excluded), we assigned two of these recordings to the training data (102 and 217) and the other two (104 and 107) to the testing data. Both the training and testing data sets were made up of 24 different recordings.

Table 2. Training/testing MIT-BIH Arrhythmia database division.

Data Set	Recording ID
Training	101, 106, 107, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 217, 220, 223, 230
Testing	100, 102, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 217, 219, 221, 222, 228, 231, 232, 233, 234

The number of heartbeats that belonged to each AAMI class for both training and testing data can be found in Table 3. The training data were employed both for GMM and MLP training and the testing data were employed for GMM testing.

It must be noted that, although the same training data were employed to train the MLP used in the augmented feature extraction and to train the GMMs (from the training augmented feature vectors), this did not introduce any over-fitting issues given that the final validation was carried out on the testing data, which were not used in the training of the GMMs nor in the MLP.

Table 3. Number of heartbeats for each AAMI class for training and testing data.

Data Set	N	S	V	F	Q	Total
Training	46,177	976	4426	415	3894	55,888
Testing	44,209	2050	3282	388	4149	54,078

5. Results

5.1. MLP-Based Experiments

An initial set of experiments was carried out to show the potential effectiveness of using the MLP-based features in the classification stage. These experiments employed the N_r-dimensional raw feature vectors as input for the MLP and carries out an MLP-based classification on testing data. For classification, the class with the highest posterior probability was assigned to each heartbeat so that the performance could be evaluated. Results are presented in Table 4 and they showed that the sensitivity (i.e., intra-class) MLP performance was, in general, above chance (i.e., higher than 20%). The only exception was the *S* class, for which there were limited training data and it integrated the highest number of heartbeat morphological classes (see Table 1). This may have dramatically reduced the performance for that class due to both data scarcity and blurred model. The other class results obtained with the MLP classification provided optimism towards the utility of the MLP-based features in our GMM-based classification system to improve its performance.

Table 4. MLP-based classification from the raw feature extraction for testing data.

	N	S	V	F	Q
Sensitivity	67.2%	14.0%	79.1%	53.4%	50.3%
Specificity	35.6%	96.2%	98.5%	99.6%	96.0%
Accuracy			64.5%		

5.2. ECG Arrhythmia Identification System

The results for the raw feature extraction, which was considered as the baseline in this work, and the augmented feature extraction, are presented in Table 5. It must be noted that the raw feature extraction experiments employed the raw feature vectors for both GMM training (using the training recordings of Table 2) and classification (using the testing recordings of Table 2), hence matching the feature type as in the augmented feature

extraction approach. Results showed that the augmented feature extraction approach significantly outperformed the performance obtained with the raw feature extraction, with a 15.8% improvement in the accuracy.

Table 5. ECG arrhythmia identification results with the raw feature extraction alone (Raw) and the augmented feature extraction module (Tandem) with the best statistically significant results for each metric in bold font. Confidence bands for a 95% interval confidence are also presented. 'Feat. extr.' stands for feature extraction, 'Se.' for sensitivity, 'Spe.' for specificity and 'Acc.' for accuracy.

Feat. Extr.	Metric	N	S	V	F	Q
Raw	Se.	$63.1 \pm 0.4\%$	$22.5 \pm 1.8\%$	$94.5 \pm 0.8\%$	$1.3 \pm 1.1\%$	$3.2 \pm 0.5\%$
	Spe.	$30.7 \pm 0.4\%$	$96.9 \pm 0.8\%$	$99.6 \pm 0.2\%$	$99.2 \pm 0.9\%$	$92.5 \pm 0.8\%$
	Acc.			$58.4 \pm 0.4\%$		
Tandem	Se.	$78.6 \pm 0.4\%$	$9.9 \pm 1.3\%$	$91.7 \pm 0.9\%$	$26.0 \pm 4.4\%$	$50.0 \pm 1.5\%$
	Spe.	$44.6 \pm 0.5\%$	$96.3 \pm 0.8\%$	$99.4 \pm 0.3\%$	$99.5 \pm 0.7\%$	$96.0 \pm 0.6\%$
	Acc.			$74.2 \pm 0.4\%$		

The confusion matrix for the raw feature extraction is presented in Table 6 and that of the augmented feature extraction is presented in Table 7.

Table 6. Confusion matrix of the ECG arrhythmia identification system with the raw feature extraction. The number of heartbeats that are classified as any of the considered classes is shown in each cell. The values between brackets represent the number of heartbeats that belong to the real class. Sensitivity and specificity percentage values are also provided for each class.

			Recognized	Class			
			N	S	V	F	Q
Real class	N	[44,209]	27,891	2817	7109	6381	11
	S	[2050]	475	462	976	137	0
	V	[3282]	145	5	3101	23	8
	F	[388]	261	1	121	5	0
	Q	[4149]	1761	34	2222	1	131
	Sensitivity	[54,078]	63.1%	22.5%	94.5%	1.3%	3.2%
	Specificity	[54,078]	30.7%	96.9%	99.6%	99.2%	92.5%

Table 7. Confusion matrix of the ECG arrhythmia identification system with the augmented feature extraction. The number of heartbeats that are classified as any of the considered classes is shown in each cell. The values between brackets represent the number of heartbeats that belong to the real class. Sensitivity and specificity percentage values are also provided for each class.

			Recognized	Class			
			N	S	V	F	Q
Real class	N	[44,209]	34,733	3231	5234	967	44
	S	[2050]	1366	203	432	49	0
	V	[3282]	183	11	3008	21	59
	F	[388]	175	2	110	101	0
	Q	[4149]	505	55	1514	1	2074
	Sensitivity	[54,078]	78.6%	9.9%	91.7%	26.0%	50.0%
	Specificity	[54,078]	44.6%	96.3%	99.4%	99.5%	96%

6. Discussion

The results of Table 5 show a 15.8% improvement in the accuracy of the augmented feature extraction when compared with the raw feature extraction. This clearly indicates that the MLP is able to produce robust features that are suitable for GMM-based classification by providing complementary information to that of present in the raw feature vector. This supports our hypothesis that the usage of tandem feature extraction can be useful for ECG analysis, in the same way that it has proven its usefulness in the analysis of other physiological signals such as EEG and speech [32,40].

However, when we also consider the intra-class contribution (see Tables 5–7), we can see that not all heartbeat types improve their performance with the incorporation of the tandem features. Although for the 'N', 'F' and 'Q' classes the augmented feature set improves the corresponding intra-class accuracy, for the 'S' and 'V' classes, the opposite occurs. We consider that this may be due to the fact that the 'S' class is the one that integrates a large variety of heartbeat morphology types (see Table 1), which may produce a less robust MLP-based features and a more blurred GMM. This is confirmed by the fact that this class obtained the worst intra-class accuracy in the MLP-based experiments (see Table 4). Furthermore, we must consider the fact that the tandem feature vector includes only features extracted from the QRS morphology. The QRS is the most significant feature of the heartbeat and the most relevant for the identification of arrhythmias, with the possible exception of the 'S' heartbeats. These heartbeats originate in the upper chambers of the heart and usually present a similar propagation through the ventricles as normal ones, whereas the QRS only captures information related to propagation through the ventricles. Therefore, it is often not possible to distinguish between 'S' and 'N' heartbeats using only the QRS morphology. This is consistent with the results in Table 7, which show that most 'S' heartbeats have been classified as 'N' heartbeats, probably because their QRS morphology was similar to that of an 'N' heartbeat. Given the extreme difficulty of reliably identifying and extracting the electrical information of the propagation of the heartbeat through the atria (the P wave), this is often palliated through the incorporation of information related to the distance between each heartbeat and the previous ones [28,30,56]. It is likely that having incorporated this type of information into the tandem feature vector, better performance could be obtained for this heartbeat type.

In Tables 6 and 7, it can be clearly seen that the classes for which limited training data are available due to the lower prevalence of those heartbeat types (i.e., 'S', 'F' and 'Q') obtain the worst performance, and they are mainly confused with the class for which more training data are available (i.e., 'N'). Something similar happens with the MLP results (see Table 4).

It should be noted that the classification of most of the 'F' heartbeats as 'N'/'V' heartbeats is expected, since the former type of heartbeat happens when supraventricular and ventricular impulses concur, hence producing a hybrid complex. In addition, the 'Q' heartbeat in the raw feature extraction (see Table 6) is confused with the 'V' heartbeat, which could be due to the complexity of both paced and unclassifiable heartbeats when using a less robust feature extraction approach.

7. Conclusions and Future Work

This paper has evaluated whether the tandem feature extraction approach is useful for ECG arrhythmia identification. To do so, tandem features have been integrated within a tandem feature extraction approach for a GMM-based arrhythmia identification system. While the use of tandem feature extraction is common in other application domains and it has previously been used in the analysis of other physiological signals such as EEG and speech, to the best of our knowledge it has never been applied to ECG analysis for arrhythmia identification. Our approach consists of adding the posterior probabilities from an MLP as features to the feature vector representing each of the heartbeats. To represent the morphology of each heartbeat we have used the coefficients of a regression based on Hermite functions. Our results have shown that the augmented feature extraction

significantly outperforms the results of the Hermite representation (15.8% improvement in accuracy), and the heartbeat types for which more training data are available benefit more from this approach.

This result suggests that our approach could benefit from the use of data augmentation techniques to handle class imbalance [57,58]. The introduction of features related to the distance between heartbeats should also be considered to be able to better distinguish the 'S' heartbeats. Finally, the use of other types of classifiers, such as those derived from deep learning techniques, could improve the performance of the arrhythmia identification system, although at the cost of higher computational requirements.

Author Contributions: Conceptualization, J.T., A.O. and D.G.M.; methodology, J.T., D.G.M. and C.A.G.; software, J.T. and D.G.M.; validation: J.T., C.A.G. and A.O.; formal analysis, J.T., D.G.M. and C.A.G.; investigation, J.T., D.G.M. and C.A.G.; resources, D.G.M. and A.O.; data curation, D.G.M. and C.A.G.; writing—original draft preparation: J.T. and A.O.; writing—review and editing, J.T., D.G.M., C.A.G. and A.O.; visualization: J.T., D.G.M. and C.A.G.; supervision, J.T. and A.O.; project administration, A.O.; funding acquisition, A.O. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data employed in this manuscript can be found at https://www.physionet.org/content/mitdb/1.0.0/ accessed on 18 April 2021.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Roth, G.A.; Mensah, G.A.; Johnson, C.O.; Addolorato, G.; Ammirati, E.; Baddour, L.M.; Barengo, N.C.; Beaton, A.Z.; Benjamin, E.J.; Benziger, C.P.; et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. *J. Am. Coll. Cardiol.* **2020**, *76*, 2982–3021. [CrossRef] [PubMed]
2. Roth, G.A.; Mensah, G.A.; Fuster, V. The Global Burden of Cardiovascular Diseases and Risks: A Compass for Global Action. *J. Am. Coll. Cardiol.* **2020**, *76*, 2980–2981. [CrossRef]
3. IDTechEX. *Cardiovascular Disease 2020–2030: Trends, Technologies & Outlook*; Technical Report; IDTechEX: Boston, MA, USA, 2021.
4. Chen, W.W.; Gao, R.L.; Liu, L.S.; Zhu, M.L.; Wang, W.; Wang, Y.J.; Wu, Z.S.; Li, H.J.; Gu, D.F.; Yang, Y.J. China cardiovascular diseases report 2015: A summary. *J. Geriatr. Cardiol.* **2017**, *14*, 1–10. [PubMed]
5. Villegas, A.; McEneaney, D.; Escalona, O. Arm-ECG wireless sensor system for wearable long-term surveillance of heart arrhythmias. *Electronics* **2019**, *8*, 1300. [CrossRef]
6. Jung, J.; Shin, S.; Kang, M.; Kang, K.H.; Kim, Y.T. Development of Wearable Wireless Electrocardiogram Detection System using Bluetooth Low Energy. *Electronics* **2021**, *10*, 608. [CrossRef]
7. Yochum, M.; Renaud, C.; Jacquir, S. Automatic detection of P, QRS and T patterns in 12 leads ECG signal based on CWT. *Biomed. Signal Process. Control* **2016**, *25*, 46–52. [CrossRef]
8. Modak, S.; Taha, L.Y.; Abdel-Raheem, E. A Novel Method of QRS Detection Using Time and Amplitude Thresholds with Statistical False Peak Elimination. *IEEE Access* **2021**, *9*, 46079–46092. [CrossRef]
9. Sinha, N.; Das, A. Discrimination of Life-Threatening Arrhythmias Using Singular Value, Harmonic Phase Distribution, and Dynamic Time Warping of ECG Signals. *IEEE Trans. Instrum. Measur.* **2020**, *70*, 2504508.
10. Laudato, G.; Picariello, F.; Scalabrino, S.; Tudosa, I.; Vito, L.D.; Oliveto, R. Morphological Classification of Heartbeats in Compressed ECG. In Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies, Vienna, Austria, 11–13 February 2021; pp. 386–393.
11. Danbatta, S.J.; Varol, A. Comparison of Zigbee, Z-Wave, Wi-Fi, and bluetooth wireless technologies used in home automation. In Proceedings of the International Symposium on Digital Forensics and Security (ISDFS), Barcelos, Portugal, 10–12 June 2019; pp. 1–5.
12. Tantalaki, N.; Souravlas, S.; Roumeliotis, M. A review on big data real-time stream processing and its scheduling techniques. *Int. J. Parallel Emerg. Distrib. Syst.* **2020**, *35*, 571–601. [CrossRef]
13. Takabayashi, K.; Tanaka, H.; Sakakibara, K. Toward an Advanced Human Monitoring System Based on a Smart Body Area Network for Industry Use. *Electronics* **2021**, *10*, 688. [CrossRef]
14. Hagiwara, Y.; Fujita, H.; Oh, S.L.; Tan, J.H.; San Tan, R.; Ciaccio, E.J.; Acharya, U.R. Computer-aided diagnosis of atrial fibrillation based on ECG signals: A review. *Inf. Sci.* **2018**, *467*, 99–114. [CrossRef]
15. Zomorodian, A.; Carlsson, G. Computing persistent homology. *Discret. Comput. Geom.* **2005**, *33*, 249–274. [CrossRef]

16. Ignacio, P.S.; Bulauan, J.A.; Manzanares, J.R. A Topology Informed Random Forest Classifier for ECG Classification. In Proceedings of the Computing in Cardiology, Rimini, Italy, 13–16 September 2020; p. 297.

17. Liu, Y.; Dong, L.; Zhang, B.; Xin, Y.; Geng, L. Real Time ECG Classification System Based on DWT and SVM. In Proceedings of the International Conference on Integrated Circuits, Technologies and Applications, Nanjing, China, 23–25 November 2020; pp. 155–156.

18. Wang, T.; Lu, C.; Sun, Y.; Yang, M.; Liu, C.; Ou, C. Automatic ECG Classification Using Continuous Wavelet Transform and Convolutional Neural Network. *Entropy* **2021**, *2021*, 119. [CrossRef] [PubMed]

19. Guan, J.; Wang, W.; Feng, P.; Wang, X.; Wang, W. Low-dimensional denoising embedding transformer for ECG classification. *arXiv* **2021**, arXiv:2103.17099.

20. Li, T.; Zhou, M. ECG Classification Using Wavelet Packet Entropy and Random Forests. *Entropy* **2016**, *2016*, 285. [CrossRef]

21. Golrizkhatami, Z.; Acan, A. ECG classification using three-level fusion of different feature descriptors. *Exp. Syst. Appl.* **2018**, *114*, 54–64. [CrossRef]

22. Xu, S.S.; Mak, M.W.; Cheung, C.C. Towards End-to-End ECG Classification with Raw Signal Extraction and Deep Neural Networks. *J. Biomed. Health Inf.* **2015**, *14*, 1574–1584. [CrossRef] [PubMed]

23. Avanzato, R.; Beritelli, F. Automatic ECG Diagnosis Using Convolutional Neural Network. *J. Electron.* **2020**, *9*, 951. [CrossRef]

24. Rana, A.; Kim, K.K. A Novel Spiking Neural Network for ECG signal Classification. *J. Sens. Sci. Technol.* **2021**, *30*, 20–24. [CrossRef]

25. Wasimuddin, M.; Elleithy, K.; Abuzneid, A.; Faezipour, M.; Abuzaghleh, O. Multiclass ECG Signal Analysis Using Global Average-Based 2-D Convolutional Neural Network Modeling. *J. Electron.* **2021**, *10*, 170. [CrossRef]

26. Wang, J.; Qiao, X.; Liu, C.; Wang, X.; Liu, Y.Y.; Yao, L.; Zhang, H. Automated ECG classification using a non-local convolutional block attention module. *Comput. Methods Progr. Biomed.* **2021**, *203*, 106006. [CrossRef]

27. Yan, G.; Liang, S.; Zhang, Y.; Liu, F. Fusing Transformer Model with Temporal Features for ECG Heartbeat Classification. In Proceedings of the 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 18–21 November 2019; pp. 898–905.

28. Márquez, D.G.; Otero, A.; García, C.A.; Presedo, J. A study on the representation of QRS complexes with the optimum number of Hermite functions. *Biomed. Signal Process. Control* **2015**, *22*, 11–18. [CrossRef]

29. Vulaj, Z.; Draganic, A.; Brajovic, M.; Orovic, I. A tool for ECG signal analysis using standard and optimized Hermite transform. In Proceedings of the 6th Mediterranean Conference on Embedded Computing, Bar, Montenegro, 11–15 June 2017; pp. 1–4.

30. Ebrahimzadeh, A.; Ahmadi, M.; Safarnejad, M. Classification of ECG signals using Hermite functions and MLP neural networks. *J. AI Data Min.* **2016**, *4*, 55–65.

31. Hermansky, H.; Ellis, D.P.W.; Sharma, S. Tandem connectionist feature extraction for conventional HMM systems. In Proceedings of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, Turkey, 5–9 June 2000; pp. 1635–1638.

32. Zhu, Q.; Chen, B.; Morgan, N.; Stolcke, S. On using MLP in LVCSR. In Proceedings of the International Conference on Speech and Language Processing, Jeju, Korea, 4–8 October 2004; pp. 921–924.

33. Faria, A. *An Investigation of Tandem MLP Features for ASR*; Technical Report; Intenational Computer Science Institute (ICSI): Berkely, CA, USA, 2007.

34. Lal, P. Cross-Lingual Automatic Speech Recognition Using Tandem Features. Ph.D. Thesis, University of Edinburgh, Edinburgh, UK, 2011.

35. Li, M.; Liu, W. Speaker Verification and Spoken Language Identification using a Generalized I-vector Framework with Phonetic Tokenizations and Tandem Features. In Proceedings of the 15th Annual Conference of the International Speech Communication Association, Singapore, 14–18 September 2014; pp. 1120–1124.

36. Wang, H.; Leung, C.C.; Lee, T.; Ma, B.; Li, H. Shifted-delta MLP features for spoken language recognition. *IEEE Signal Process. Lett.* **2013**, *20*, 15–18. [CrossRef]

37. D'Haro, L.F.; Cordoba, R.; Salamea, C.; Echeverry, J.D. Extended phone log-likelihood ratio features and acoustic-based i-vectors for language recognition. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 5342–5346.

38. Tejedor, J.; Macias-Guarasa, J.; Martins, H.F.; Piote, D.; Pastor-Graells, J.; Martin-Lopez, S.; Corredera, P.; Gonzalez-Herraez, M. A Novel Fiber Optic Based Surveillance System for Prevention of Pipeline Integrity Threats. *Sensors* **2017**, *17*, 355. [CrossRef] [PubMed]

39. Tejedor, J.; Macias-Guarasa, J.; Martins, H.F.; Martin-Lopez, S.; Gonzalez-Herraez, M. A Contextual GMM-HMM Smart Fiber Optic Surveillance System for Pipeline Integrity Threat Detection. *J. Lightwave Technol.* **2019**, *37*, 4514–4522. [CrossRef]

40. Ting, C.M.; King, S.; Salleh, S.H.; Ariff, A.K. Discriminative tandem features for HMM-based EEG classification. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan, 3–7 July 2013; pp. 3957–3960.

41. Alickovic, E.; Subasi, A. Medical Decision Support System for Diagnosis of Heart Arrhythmia using DWT and Random Forests Classifier. *J. Med. Syst.* **2016**, *40*, 108. [CrossRef] [PubMed]

42. Moody, G.B.; Mark, R.G. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. Med. Biol. Mag.* **2001**, *20*, 45–50. [CrossRef] [PubMed]

43. Da, S.; Luz, E.J.; Schwartz, W.R.; Cámara-Chávez, G.; Menotti, D. ECG-based heartbeat classification for arrhythmia detection: A survey. *Comput. Methods Progr. Biomed.* **2016**, *127*, 144–164.

44. Anwar, S.M.; Gul, M.; Majid, M.; Alnowam, M. Arrhythmia Classification of ECG Signals Using Hybrid Features. *Comput. Math. Methods Med.* **2018**, *2018*, 1380348. [CrossRef] [PubMed]

45. Alfaras, M.; Soriano, M.C.; Ortín, S. A Fast Machine Learning Model for ECG-Based Heartbeat Classification and Arrhythmia Detection. *Front. Phys.* **2019**, *7*, 103. [CrossRef]

46. Kachuee, M.; Fazeli, S.; Sarrafzadeh, M. ECG Heartbeat Classification: A Deep Transferable Representation. In Proceedings of the IEEE International Conference on Healthcare Informatics, New York, NY, USA, 4–7 June 2018; pp. 443–444.

47. Das, M.K.; Ari, S. ECG Beats Classification Using Mixture of Features. *Int. Sch. Res. Not.* **2014**, *2014*, 178436. [CrossRef] [PubMed]

48. AAMI. *Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms*; Technical Report; American National Standards Institute (AAMI): Arlington, VA, USA, 1998.

49. Bishop, C.M. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 1995.

50. Johnson, D. ICSI Quicknet Software Package. 2004. Available online: http://www.icsi.berkeley.edu/Speech/qn.html (accessed on 19 April 2021).

51. Zhang, Y.; Alder, M.; Togneri, R. Using Gaussian mixture modeling in speech recognition. In Proceedings of the ICASSP, Adelaide, SA, Australia, 19–22 April 1994; pp. 613–616.

52. Permuter, H.; Francos, J.; Jermyn, I. A study of Gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognit.* **2006**, *39*, 695–706. [CrossRef]

53. Yu, T.; Zhang, C.; Cohen, M.; Ru, Y.; Wu, Y. Monocular Video Foreground/Background Segmentation by Tracking Spatial-Color Gaussian Mixture Models. In Proceedings of the Workshop on Motion and Video Computing, Austin, TX, USA, 23–24 February 2007; pp. 1–5.

54. Kim, Y.; Jeong, S.; Kim, D.; López, T.S. An efficient scheme of target classification and information fusion in wireless sensor networks. *Pers. Ubiquitous Comput.* **2009**, *13*, 499–508. [CrossRef]

55. Xuan, G.; Zhang, W.; Chai, P. EM algorithms of Gaussian mixture model and hidden Markov model. In Proceedings of the International Conference on Image Processing, Thessaloniki, Greece, 7–10 October 2001; pp. 145–148.

56. De Chazal, P.; O'Dwyer, M.; Reilly, R.B. Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1196–1206. [CrossRef] [PubMed]

57. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]

58. Shaker, A.M.; Tantawi, M.; Shedeed, H.A.; Tolba, M.F. Generalization of Convolutional Neural Networks for ECG Classification Using Generative Adversarial Networks. *IEEE Access* **2020**, *8*, 35592–35605. [CrossRef]