J Electron Test
DOI 10.1007/s10836-016-5602-4

# Fault Diagnosis in Highly Dependable Medical Wearable Systems

**Cristina C. Oliveira**[1] · **José Machado da Silva**[1]

Received: 6 November 2015 / Accepted: 23 June 2016


**Abstract** High levels of dependability are required to promote the adherence by public and medical communities to wearable medical devices. The study presented herein addresses fault detection and diagnosis in these systems. The main objective resides on correctly classifying the captured physiological signals, in order to distinguish whether the actual cause of a detected anomaly is a wearer health condition or a system functional flaw. Data fusion techniques, namely fuzzy logic, artificial neural networks, decision trees and naive Bayes classifiers are employed to process the captured data to increase the trust levels with which diagnostics are made. Concerning the wearer condition, additional information is provided after classifying the set of signals into normal or abnormal (e.g., arrhythmia, tachycardia and bradycardia). As for the monitoring system, once an abnormal situation is detected in its operation or in the sensors, a set of tests is run to check if actually the wearer shows a degradation of his health condition or if the system is reporting erroneous values. Selected features from the vital signals and from quantities that characterize the system performance serve as inputs to the data fusion algorithms for Patient and System Status diagnosis purposes. The algorithms performance was evaluated based on their sensitivity, specificity and accuracy. Based on these criteria the naive Bayes classifier presented the best performance.

**Keywords** Dependability · Fault detection · Wearable medical systems · Machine learning

## 1 Introduction

The advances on sensors, wireless communications and information technologies have promoted a rapid development of various wellness or disease monitoring systems, which enable extended independent daily living and improve the quality of life. Traditionally, medicine has been practised on an intervention basis (drugs, surgeries, prosthesis, etc.) to treat them. Nowadays, and regardless of the patients' age, the health care community is trying to focus on prevention and wearable monitoring systems have been proposed to meet this task.

Remote health monitoring can be used only if the monitoring device is based on a comfortable, easy to use, and customizable sensing interface. The textile approach to the implementation of sensing elements embedded in clothing items, allows for low cost and long-term monitoring of patients and to easily customize the sensor configuration according to the needs of each individual [27]. By applying this concept it is possible to reduce health care costs, maintaining high quality of care, shift the focus of health care expenditures from treatment to prevention, provide access to health care to a larger number of patients, reduce the length of hospital stays and address the issue of specific requirements for elderly population and/or chronically ill patients.

Because these wearable monitoring systems are to be used for medical purposes, their dependability has to be perfectly controlled. Unfortunately, the complexity and the functional specificities of these systems make the existing dependability techniques, specifically developed for the

Responsible Editors: G. Léger and C. Wegener.

✉ Cristina C. Oliveira
cristina.oliveira@fe.up.pt

[1] INESC TEC, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

aeronautics, space and automotive applications, not totally appropriate for the medical field [17].

To overcome the lack of a dependability model that can be used for the development of complex pervasive medical monitoring devices, a fault tree analysis approach has been developed to identify the main risk of failure. A typical wearable device (hereafter the system) comprises a module to capture the biosignals, including the electrodes and the analogue front-end, a microcontroller, and a radiofrequency module to transmit data to a smartphone or personal computer. In our approach the captured biosignals are received and analysed within a smartphone. A classification algorithm decides whether these are normal or not. If not, it is diagnosed if the wearer shows an abnormal situation or instead the system is faulty. That is, a data abnormality can be due to a wearer irregular state (pathological condition or intense physical activity) or due to a degradation of the system operation. On the other hand, cases can occur where measured data is taken as correct when actually the system is faulty or the measurement procedure is not performed correctly. A flatline ECG is obtained when either the person is dead or the ECG meter is faulty. A correct heart beat rate could be fooled by an oscillating circuit.

Methodologies have been proposed to increase the reliability of medical wearable systems, which address faults detection in sensors, electronic modules, and communication links [2, 11, 12, 15, 18, 31]. Data fusion techniques have also been applied as a means for a combined analysis of several physiological signals to extract additional information on a patient's condition. Kenneth et al. performed the fusion of ECG, blood pressure, saturated oxygen content and respiratory data for achieving improved clinical diagnosis of patients in cardiac care units [14].

Our objective is to achieve the fastest and most efficient way to detect and diagnose deviations occurring in the captured data, having in mind the concern of correctly differentiating errors due to faults in the system, from those due to a change of the person health status. This procedure should be done in ambulatory and able to run in the microcontroller of the data capture module and/or in a smartphone. That is, the overall procedure is divided in different operations, being the data aggregation device (typically a smartphone) the main information processing device, due to the need of high computation capabilities and to minimize power consumption in the wearable devices. These are involved in the dependability process when specific local test operations need to be executed.

A data fusion model for wearable medical systems based on fuzzy logic was presented in [25]. It was shown how fuzzy logic can be explored to correlate data obtained from different sensors in order to obtain status indicators that characterize the operation correctness of a monitoring system or a pathological condition of the wearer. Ideally there should be a database with all possible hardware faults (a fault dictionary) and a database with the clinical data of each wearer (patient being monitored). If clinical data records are not available, or the database of a new patient does not have enough information to build a reliable classifier, the diagnosis could be run based on fuzzy logic, until enough data is gathered to create a robust classifier.

In this paper a database with information regarding the hardware (fault dictionary) and clinical data is introduced and utilized to explore possible data fusion algorithms capable of both system and patient diagnosis. The data fusion algorithms evaluated here are all supervised learning algorithms. Their inputs (features from the monitoring system and the patient vital signs) and outputs (system and patient status) are labelled and this information is used in the training proccess. There are several supervised learning algorithms available, with different processing speeds, memory usage needs and interpretability. Four data processing approaches to fuse data extracted from features, based on Fuzzy Logic (FL), Artificial Neural Networks (ANN), Classification Trees (CT), and Naive Bayes Classifiers (NBC), were evaluated for anomaly detection and classification purposes.

Fuzzy Logic can be used to solve problems of different types and domains, including medicine, as it resembles human reasoning and decision making. It looks into the not precisely formulated relationships and solves uncertainties and ambiguities created by human language where everything cannot be described in precise and discrete terms. ANN offer a number of advantages, such as requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms. However, its "black box" nature does not allow to see the relations between inputs and outputs. CT have a fast prediction speed, small memory usage and the results are easy to interpret. NBC prediction speed and memory usage vary according to the distribution size, but usually perform better than CT and the results are also easy to interpret.

The on-line methodology being proposed is based on a top-down evaluation process that starts with the detection of anomalies in the captured data. This detection triggers an analysis to check if actually the wearer shows a degradation of his health condition or if the system is reporting erroneous values. If it is found that the system is faulty, specific test operations can be executed to find the origin of the fault. To the best of our knowledge, no previous work was published where fusing data from both the data acquisition system and patient vital signs, is explored for simultaneous diagnosis of the patient health condition and the system status.

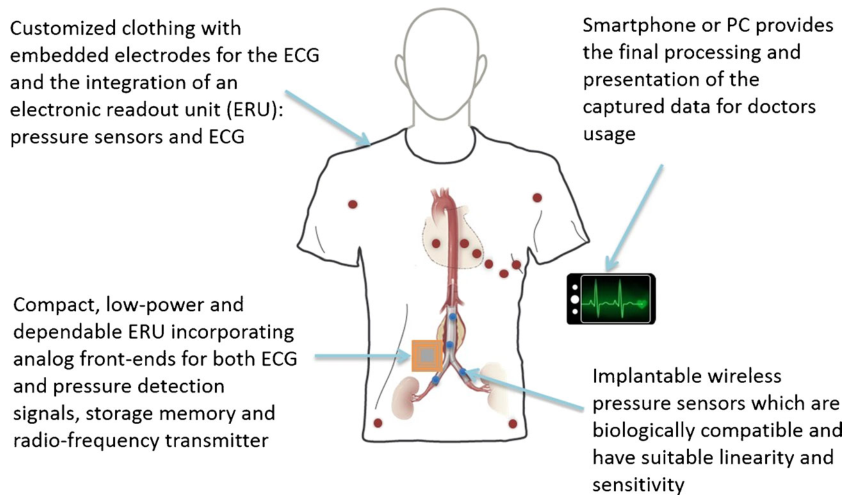


**Fig. 1** Wearable ECG data capture and transmitter module

Next section describes the wearable system being used as a case study. Section 3 presents the dependability strategy being proposed and introduces the mode how the four data fusion algorithms are being explored. Section 4 discusses the obtained results and addresses the main conclusions highlighted in Section 5.

## 2 The SIVIC Monitoring System

The SIVIC[1] system, a combined cardiac and aortic monitoring system under development (Fig. 1), provides the synchronous measurement of a patient's ECG (electrocardiogram) and pressure in the abdominal aortic aneurysm (AAA) sac, in order to obtain a more robust and reliable monitoring. Biologically compatible capacitive pressure sensors, which show suitable linearity and sensitivity [22], are used to capture the intra-sac AAA pressure and detect endoleaks.

The pressure monitoring system relies on an inductive coupling interface to capture the resonant frequencies of a cluster of LC sensors placed on the stent-graft wall [21]. An electronic readout unit (ERU) capable of energizing sensors and capture the pressure data is placed in the patient's chest. This unit provides also the monitoring of a 12-lead ECG using textile dry electrodes [30]. The electronic unit and the electrodes are built in a customized clothing. Data is transmitted to a smartphone for further processing and diagnosis, data display, and eventually can also be transmitted by the smartphone to a healthcare center.

Figure 2 shows the SIVIC 12-lead ECG DAT (Data Acquisition and Transmission Unit) that was developed to be integrated within the T-Shirt and transmit the data to a smartphone. It is a circular board (30 mm ⊘) with an ECG analogue front-end based on the Texas Instruments low-power, 8-channel, 24-bit ADS1298 chip for biopotential measurements and the PAN1740 Bluetooth Low Energy (BLE) module from Panasonic. The internal microcontroller (32-bit ARM Cortex M0) present on the BLE module is used to perform data acquisition, preliminary processing, and communication operations, thus saving the cost of an external microcontroller, the additional PCB area, and power consumption. The DAT is set with a sampling frequency of 250 Hz, which provides a good balance in terms of data accuracy and power consumption. This frequency is adequate to ensure that phase noise does not impair the estimation of the R-wave fiducial point [16].

Wireless ECG monitoring systems with a high number of leads (e.g. 12-lead) are usually designed for clinical usage, being systems with a lower number of acquisition channels (e.g. 1 to 3 channels) commonly used in ambulatory cases [3, 8]. Our system was designed having in mind its use in both clinical and ambulatory scenarios and thus the number of ECG data acquisition channels is reconfigurable between a single lead (1 channel) to twelve leads. Inputs not used to capture ECG signals can be used to acquire other biosignals.

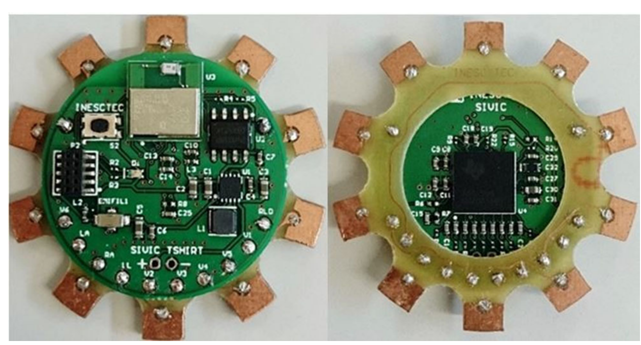

**Fig. 2** Front and back pictures of the SIVIC 12-leads ECG acquisition unit prototype

[1] Portuguese acronym for integrated cardiovascular surveillance system.

Figure 3a shows a prototype of the T-shirt with the cardiac monitoring SIVIC system. It integrates five snap fasteners in the configuration corresponding to the derivations of the triangle of Einthoven and Wilson leads I, II and III and the precordial V1. The T-shirt was fabricated using seamless knitting technology and is composed of two superimposed fabrics, integrating snap fasteners to attach the ECG electrodes. These interconnect to the DAT unit through embedded textile interconnects and a custom designed socket. Five snaps are also available to connect to a external conventional Holter for comparison purposes.

Figure 3b shows ECG signals displayed on a smartphone, captured when a person wears the T-Shirt in a steady position and without any skin preparation. In this case the smartphone receives three leads (LI, LII, V1) and from them calculates four more leads (LIII, aVR, aVL, aVF) [30].

## 3 Dependability Strategy

### 3.1 Failure Mode and Effects Analysis

A failure mode and effects analysis (FMEA) based on the approach made in [7] was carried out during the system's design. This analysis is very important to identify the most problematic components and functions of the system and determine, at the system design phase, which components or blocks should include built-in self-test (BIST) or other type of testing. From this analysis it was concluded that the wearable system blocks with higher severity and probability classification, i.e., those whose failure is more dangerous and probable, are located in the sensing hardware, power supply and microcontroller. Therefore it is necessary to be able to detect faults in the sensors (ECG electrodes and implantable pressure sensors), the sensors conditioning circuits, the battery and the microcontroller.

Concerning testing and BIST for these parts the following features are available. The ADS1298 IC provides an electrodes impedance measurement circuit, as well as internally-generated test stimuli for subsystem verification, which allows testing the entire analogue front-end chain. The electrode-skin impedance measurement circuit enables detecting if the electrodes are connected to the patient or are loose/disconnected.

The signal-to-noise ratio (SNR) of bioelectrical signals is known to be related to the electrode-skin impedance [24]. Since the impedance varies for each person and is affected by other factors like temperature and applied pressure, the electrode-skin impedance is measured when the system is switched on and afterwards it is monitored periodically to establish a normal region for the impedance values, for which the acquired ECG quality is considered acceptable. These values are then used for comparison with the impedance measured during normal operation of the system.

The Programmable Gain Amplifier (PGA) can be tested for high gain variations, e. g. due to components aging and for catastrophic faults, by using an internal square wave stimulus with an amplitude of +/- 2 mV and a frequency of 4 Hz. Class B safety software library routines, which comply with IEC 60730 Class B certification process standard, allow detecting faults in the microcontroller. A power supply monitoring circuit, such as an watchdog circuit, allows detecting battery failures.

A methodology to test and measure the L and C values of the pressure sensors, after measurements of the power and impedance seen from the reader circuit, was developed using a simple data fusion approach [23]. The processing of the implantable pressure sensor signals with other physiological signals like the electrocardiogram (ECG) and arterial blood pressure (ABP) allows obtaining better resolution and decision trust with the acquired information in terms of diagnosability of faults eventually occurring in the stent-graft and the pressure sensors.


**Fig. 3** **a**) A SIVIC 5-leads unit and T-shirt prototype; **b**) Android application screenshot

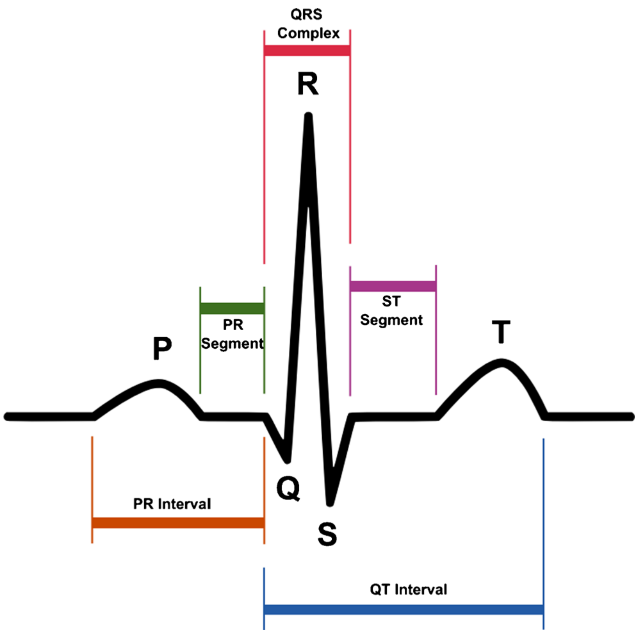


**Fig. 4** Typical ECG signal and its main characterizing waves [9]

An alternative BIST approach that could be used to test the front-end of data acquisition unit was proposed in [28]. The targeted tests are electrode-skin impedance measurement and detection of functional deviations of the signal conditioning circuit.

### 3.2 ECG Feature Extraction

The ECG conveys important hemodynamic information, such as the heart rate (HR). During an ECG cycle three main events take place: the P wave (contraction of the atria), the QRS complex (corresponding to the contraction of left ventricle) and the T wave (relaxation of the ventricles) (see Fig. 4). Their morphologies (amplitude and interval/segment length) will vary according to the person's physiological condition.

The HR is given in beats per minute (bpm) of consecutive R-waves. However, noise contamination such as baseline wander, power line interference, and body activity can corrupt the signal and reduce the clinical value of an ECG recording. Since wearable devices are more prone to perturbations by noise, filtering of the ECG is a necessary pre-processing step to ensure a reduction of the noise components while preserving the QRS complex shape. The Pan-Tompkins algorithm is used for ECG filtering and HR calculation [26]. Other biosignals, such as the blood pressure (BP), can provide important information about the patient's condition. BP is defined by the mean arterial pressure (MAP) and is measured in millimetres of mercury (mmHg). BP is affected by the physical activity of the patient or associated diseases.

The main features of each signal eventually measured by the SIVIC system and the classification of the patient or system condition are presented in Table 1.

### 3.3 Fault Dictionary

As the SIVIC system is still under development not enough data from measurements performed in patients is yet available. To overcome this a database with all possible hardware faults, a fault dictionary, and clinical data of patients being monitored was created.

A diagram of the main SIVIC blocks is presented in Fig. 5. This model includes the vital signs being measured, the sensor model (for the pressure sensors) and the electrode-skin interface model (for the ECG and respiration electrodes), the amplifiers, the ADC, the microcontroller, the BLE communication, and the smartphone. Associated to these are the anomalies that can occur: faults in the system (electrodes, analogue front-end, microcontroller, communication link) and extreme change of the physical activity or pathology in the wearer.

To ensure that the evaluation of the test and diagnosis approaches under analysis is carried out with a large number of real cases, the inputs of this model (ECG and ABP signals) are obtained from the MIT Multiparameter database (MGH/MF) [10, 32]. This database includes ECG signals (leads I, II and V), the arterial blood pressure, and

**Table 1** Data fusion model for the measured signals

| Signals | Features | Classifier |
|---|---|---|
| ECG | HR I<br>HR II<br>HR III<br>⋮ | Normal/Abnormal |
| Blood Pressure | MAP | Hypotensive/Normal/Hypertensive |
| AAA Pressure | Mean Pressure | Normal/Endoleak |
| Electrode-Skin Impedance | Resistance | Electrodes Connected/ Disconnected |
| Square-wave Stimulus | Amplitude and Frequency | PGA Stuck-at/ High Gain / High Frequency variation |

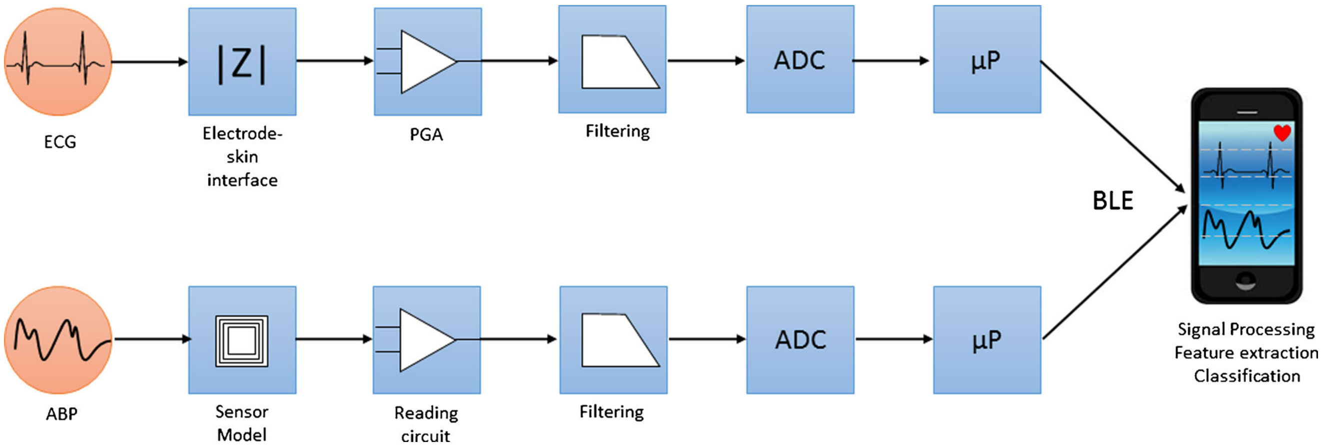


**Fig. 5** Block diagram of the SIVIC system describing the signal path from the sensors to the smartphone

respiratory rate. Since this database includes annotations of the patients' conditions made by doctors and presents also a broad diversity of cases, information is available to validate the diagnosis performed in the smartphone.

## 3.4 Classification

The proposed anomaly detection algorithms yields two outputs (Patient Status and System Status) that constitute a critical and extremely useful step of the diagnosis process.

The dataset for training, testing and validating the algorithms contains 161 data cases collected during experiments and/or from simulations (in the case of the hardware). The dataset was normalized so that all the input values range from 0 to 1 and randomly divided in two sets: training (70 %) and testing (30 %).

The goal is to compare all the supervised learning classification algorithms presented in this section in terms of performance, complexity and computation load (time and memory requirements), so that the classification can be implemented in the wearable system and achieved in real time.

The inputs for the classifiers are:

– HR: Mean heart rate of the 5 last beats;
– BP: Mean MAP of samples with the same length as the HR;
– ABP: Pressure measured by the sensors implanted inside the aneurysm sac;
– Z: electrode-skin impedance;
– $Amp_{Min}$: Minimum amplitude of the test square wave;
– $Amp_{Max}$: Maximum amplitude of the test square wave;
– Freq: Frequency of the test square wave.

The outputs of the classification algorithms that provide binary classifiers (Fuzzy Logic and ANN) are:

– Patient Status: 1 if ok, 0 if there is a health problem;
– System Status: 1 if ok, 0 in case a fault is detected.

while for the categorical classification algorithms (Decision Trees and Naive Bayes Classifier) the provided output classes are:

– *ok* - both the patient and the system are good;
– *pf* - the patient shows a health problem;
– *sf* - the system is failing;
– and *pf-sf* - the patient shows a health problem and the system is failing.

### *3.4.1 Fuzzy Logic*

Fuzzy Logic enables the creation of a decision making process based on logic and straightforward principles. Its implementation is relatively easy and thus suitable for implementation in a smartphone. The extensive literature in the medical field provides a solid knowledge base for the implementation of a medical decision support system. This technique can be applied to monitoring a patient's vital signs during an invasive surgery [5], support medical decisions in a intensive care unit [4], or cancer diagnosis based on image processing [1, 19].

In our case, FL is used in the data fusion process taking advantage of its probability assignment based on rules. Since the values of the features extracted from the biosignals can be sorted in ranges well defined in the medical literature, the creation of rules is relatively straightforward. Table 2 shows common normal values for the HR and BP, and some pathological cases.

The FL decision process comprises 4 main components: fuzzy rules (knowledge base), fuzzy sets, fuzzy inference engine and defuzzification (Fig. 6) [33]. The inputs of the FL algorithm are the features previously extracted from the measured signals (Table 1). The outputs are the diagnostic results inferred from the observed quantities values and their correlations, i. e, *Patient Status* and the *System Status*, which

**Table 2** Fusion rules for patient condition diagnosis

| Features | Condition | Rule |
|---|---|---|
| HR (ms) | Normal | HR between 60 and 100 bpm |
| | Asystole | No QRS for at least 4 seconds |
| | Extreme Bradycardia | HR lower than 40 bpm for 5 consecutive beats |
| | Extreme Tachycardia | HR higher than 140 bpm for 17 consecutive beats |
| MAP (mmHg) | Normal | 70-105 |
| | Hypotension | $< 70$ |
| | Hypertension | $< 105$ |
| ABP (mmHg) | Normal | Low pressure (∼40) |
| | Endoleak | Sistemic Pressure ($\geq$ MAP) |
| Z (Ω) | Normal | $< 1e^6$ |
| | Disconnected | $> 1e^6$ |
| $\text{Amp}_{Min,Max}$ (mV) | Normal | ± 2mV * PGA Gain |
| | Fault | DC or dif $>$ 25 % from expected |
| Freq (Hz) | Normal | 4 |
| | Fault | 0 or dif $>$ 25 % from expected |

can be normal or anomalous. The outputs are determined based on the input values of the fuzzy sets and the rules assigned for each output. The rules to define the *Patient Status* are based on information found in the literature or provided by a physician, the rules for the *System Status* are defined from the system specifications and the previously performed FMEA analysis.

The fuzzy sets include the HR for each channel, the blood pressure (MAP), and can also include the ABP pressure, the acceleration of the patient's activity, and the electrode-skin contact resistance if these data are available.

In case the impedance values are higher than expected, signalling a potential loose connected electrodes situation, the fuzzy logic system updates the *System Status*.

The trapezoidal curve was chosen for the membership function. This is a function of a vector, x, and depends on four scalar parameters $a$, $b$, $c$, and $d$ (Eq. 1). The parameters $a$ and $d$ locate the "feet" of the trapezoid and the parameters $b$ and $c$ locate the "shoulders".

$$\mu_{trapezoidal}(x) = \begin{cases} 0, & x < a \text{ or } x > d \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \dfrac{d-x}{d-c}, & c \leq x \leq d \end{cases} \quad (1)$$

The BP is divided in three sets: low (hypotension), normal and high (hypertension). Figure 7 displays the membership functions for diastolic (a) and systolic (b) BP sets, respectively.

The HR includes the following sets: bradycardia, normal and tachycardia. The vertical black line in Fig. 8 represents a HR measurement of 130 bpm, which has a membership level of 0.3 in the normal set and a level of 0.8 in the tachycardia set.

The output variables *Patient Status*, *System Status*, and *Global Status* have 2 sets: abnormal (from 0 to 0.5) and normal (from 0.5 to 1). The normal sets from the inputs are assigned to the normal set of the outputs, and the remaining input sets are assigned to the abnormal output set.

### 3.4.2 Artificial Neural Networks

Artificial Neural Networks (ANN) are learning algorithms inspired by biological neural networks. These are typically organized in layers, which are made up of a number of interconnected *nodes* containing an *activation function*. Patterns are presented to the network via the *input layer*, which communicates to one or more *hidden layers* where the actual processing is done after a combination of weighted *connections*. The hidden layers then link to an *output layer* where the answer is collected. Most ANN contain some form of learning rule which modifies the weights of the connections according to the patterns presented to the input.

For anomaly diagnosis purposes, the backpropagation neural network algorithm was used. The backpropagation is a supervised process that occurs with each cycle or 'epoch' (i.e., each time the network is presented with a new input pattern) through a forward activation flow of outputs, and the backwards error propagation of weight adjustments.

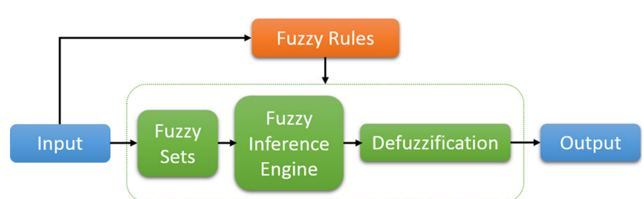


**Fig. 6** Block diagram of the FL algorithm

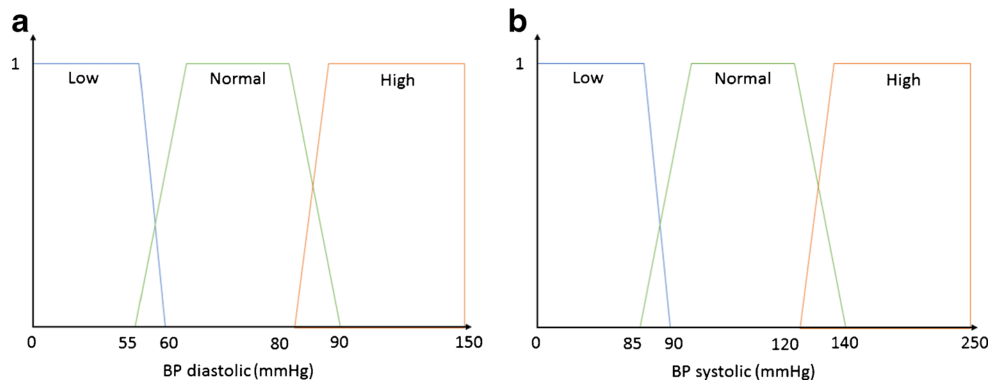


**Fig. 7** Fuzzy sets for the **a**) diastolic BP; **b**) systolic BP

The ANN (see Fig. 9) was trained using the scaled conjugate gradient backpropagation algorithm [20]. The ANN has 7 hidden neurons (the same as the number of inputs) and the training is stopped after 27 epochs. The training dataset was used for building the ANN and the testing dataset was divided in two sub sets: half for validation of the ANN and half for testing.

Figure 10 displays the receiver operating characteristics (ROC) of the ANN for training, validation and test. The ROC is a metric used to check the quality of the classifiers. For each class of a classifier (Class 1: patient status, Class 2: system status) the true positive rate against the false positive rate of the different possible cut-points are calculated. These plots show the trade-off between sensitivity and specificity. The area under the ROC curve (AUC) is a measure of a test accuracy, and also a common metric that can be used to compare different tests.

### *3.4.3 Decision Trees*

Decision trees (DT) are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. A DT important feature is their capability to break down a complex decision-making process into a collection of simpler decisions that are easier to interpret. This algorithm consists of several test nodes and class (or decision) leaves. It classifies an input by executing the tests in the tree beginning at the root, and going down the branches until a leaf is reached which gives the class of the input (or the decision to be taken).

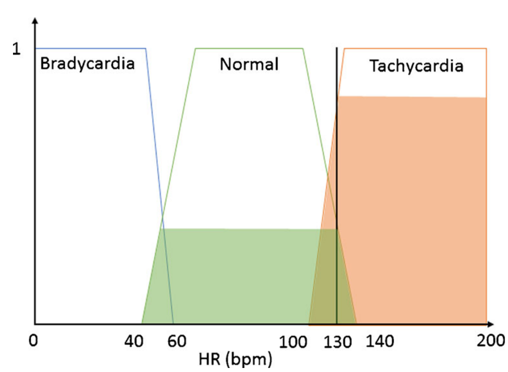


**Fig. 8** Fuzzy sets for the HR

The mathematical formulation is based on [6]. To each leaf a class or even a class probability is assigned. Each of the non-leave nodes represents a split regarding the input space. Such split is represented by a decision $Q$. Given training vectors of features (vital signs and system signals) $x_i \in R^n$, $i = 1, ..., n$ and an output label vector (patient status and system status) $C \in R^k$, the decision tree recursively partitions the space in such a way that the samples with the same labels are grouped together. For each candidate split $\theta = (j, t_m)$ consisting of a feature $j$ and threshold $t_m$, partition the data into branches left $(Q_{left}(\theta))$ or right $(Q_{right}(\theta))$.

$$Q_{left}(\theta) = (x, y)|x_j <= t_m Q_{right}(\theta) = Q \setminus Q_{left}(\theta) \quad (2)$$

A decision tree is built top-down from a root node and involves partitioning the data into subsets that contain instances with similar values (homogeneous). Most DT algorithms use cross entropy to calculate the homogeneity of a sample $m$. If the sample is completely homogeneous the entropy is zero and if the sample is equally divided it has entropy of one.

Using the Matlab decision tree code available in the Statistical Toolbox, a classification decision tree can be generated using the cross-entropy for pruning. The model is constructed using t = classregtree(Inputs,Outputs,'States','HR' 'BP' 'ABP' 'Z' '$Amp_{Min}$' '$Amp_{Max}$' 'Freq') and the tree is optimized using [c,s,n,best] = test(t,'cross',Inputs,C) and tmin = prune(t,'level',best). Figure 11 displays the classification tree for anomaly diagnosis.

### *3.4.4 Naive Bayes Classifier*

The Naive Bayesian Classifier (NBC) is based on Bayes' theorem with independence assumptions between features.

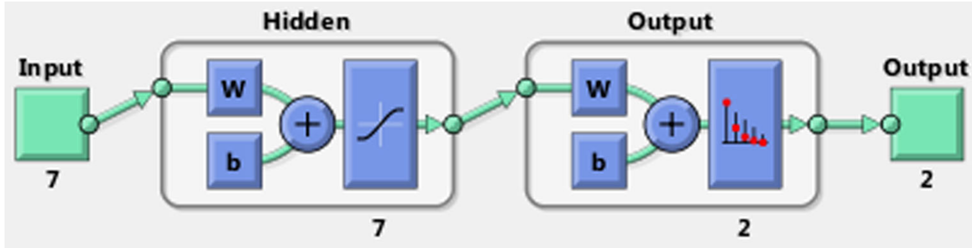


**Fig. 9** ANN architecture

An NBC model is easy to build, with no complicated iterative parameter estimation, making it particularly useful for very large datasets. The NBC relies on a strong hypothesis — the value of any feature is independent of the existence of any other feature. In most of the real life examples the Naive Bayes hypothesis is never satisfied, but the algorithm predicts the classes with good enough accuracy. Despite its simplicity the NBC often does surprisingly well and is widely used because it often outperforms more sophisticated classification methods.

Bayes theorem provides a way of calculating the posterior probability, $P(C_k|x)$, from the prior probability of possible classes (system and patient status) $C_k$ ($P(C_k)$), the evidence of the features (vital signs and system signals) $x = (x_1, x_2, ..., x_n)$ to be classified ($P(x)$), and the class conditional probability $P(x|C_k)$ (Eq. 3).

$$P(C_k|x) = \frac{P(x|C_k)P(C_k)}{P(x)} \tag{3}$$

NBC assumes that the effect of the value of a feature $x_n$ on a given class $C_k$ is independent of the values of other features. This assumption is called class conditional independence. Based on this, the likelihood can be decomposed to a product of terms:

$$P(C_k|x) = P(C_k)\Pi_{i=1}^{n} P(x_i|C_k) \tag{4}$$

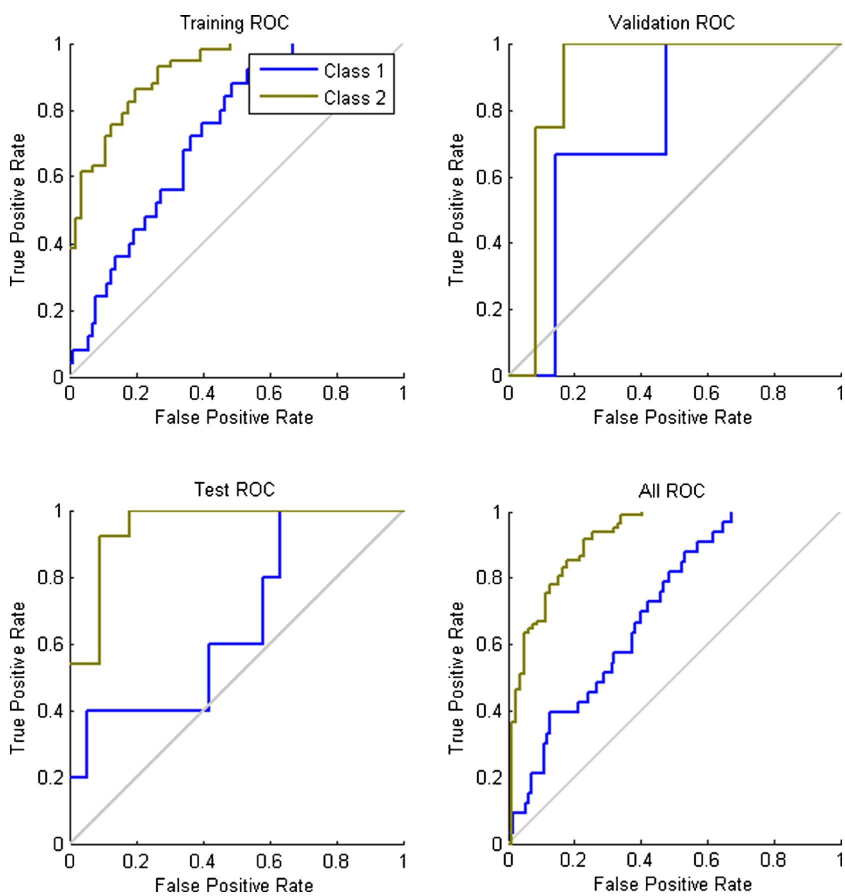


**Fig. 10** Receiver operating characteristic for the ANN

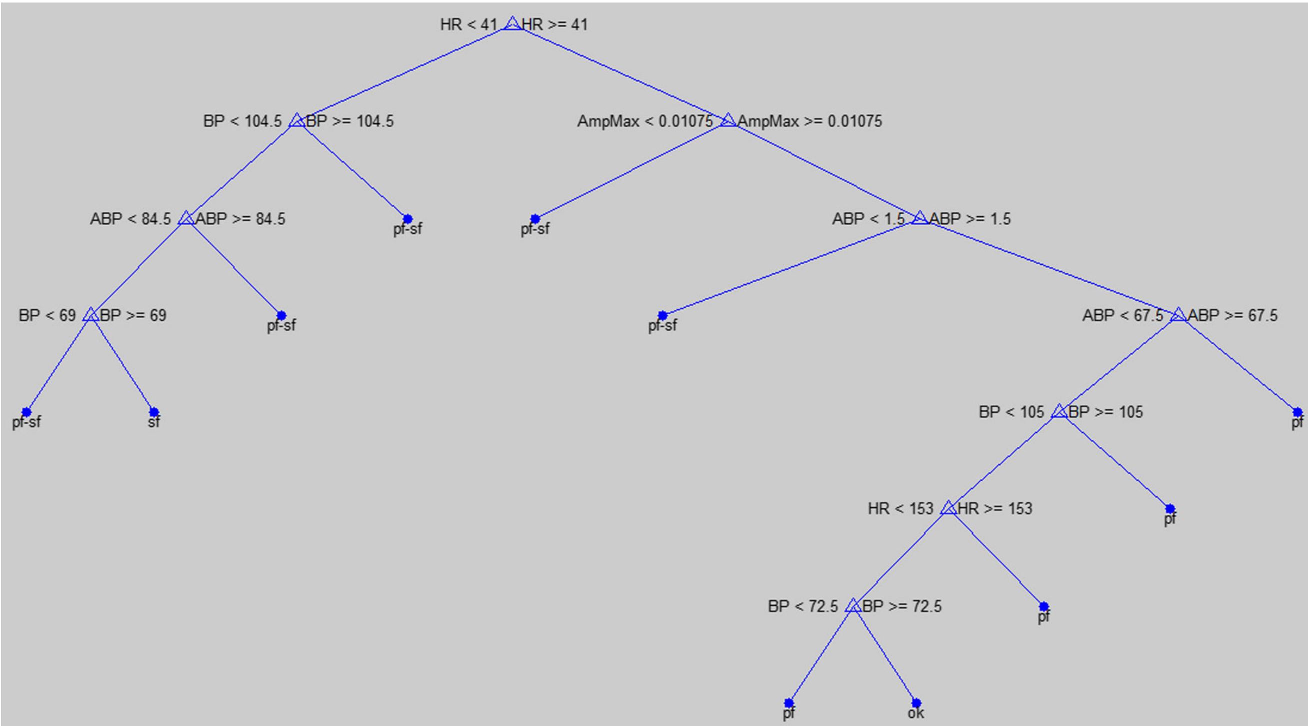


**Fig. 11** Classification decision tree

Using Bayes' rule above it is possible to classify a new set of features (classifier inputs) $x_n$ into a class (classifier outputs) $C_k$ that achieves the highest posterior probability. Naive Bayes can be modelled in several different ways including normal, lognormal, gamma and Poisson density functions. For this work the normal (Gaussian) distribution was assumed:

$$P(x_i|C_k) = \frac{1}{\sqrt{2\pi\sigma_{ik}^2}} e^{-\frac{(x-\mu_{ik})^2}{2\sigma_{ik}^2}} \tag{5}$$

To train this classifier the mean ($\mu$) and variance ($\sigma^2$) of each feature in the training set was calculated. These values are then used in the testing set.

## 4 Results and Discussion

The training set of the database was used to compute the classification models of all four classification algorithms. The test set of the database was used to predict the output values and compare them to the actual outputs. In order to evaluate the performance of the algorithms a confusion matrix was calculated for all the predicted outputs.

A confusion matrix is a table whose rows provide the predicted values for each class (Hypothesized classes) and the columns the true values for each class (True classes). These values allow to calculate the sensibility, specificity and accuracy of the Hypothesized classes. Table 3 shows an example of a confusion matrix containing the values of True Positives(TP), True Negatives (TN), False Positives (FP) and False Negatives (FN) for two classes.

The sensibility, specificity and accuracy can then be calculated after the following equations:

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

$$Sensibility = \frac{TP}{TP + FN} \tag{7}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

Table 4 presents the accuracy values for all the tested algorithms. The NBC presents the best accuracy and the DT provides the lowest score. The accuracy might not be

**Table 3** Confusion Matrix

| | | True Class | |
|---|---|---|---|
| | | P | N |
| Hypothesized class | P | TP | FP |
| | N | FN | TN |

**Table 4** Accuracy values for all the classification algorithms

| Algorithm | FL | ANN | DT | NBC |
|---|---|---|---|---|
| Accuracy | 85.7 % | 87.5 % | 77.6 % | 95.0 % |

a reliable metric for the real performance of a classifier, because it will yield misleading results if the data set is unbalanced (that is, when the number of samples in different classes vary greatly). For this work the dataset used had balanced inputs, i.e. half of the input values correspond to a fault in the patient/system and half of the values are ok. However, since the combined inputs will affect the outputs (Patient and System Status) differently, the output classes are unbalanced (there are more faults than non faulty situations).

In order to better evaluate the ability of the classifiers to correctly predict the Patient and System Status, the sensitivity and specificity of each output class was calculated for all the algorithms. The results are presented in Table 5 for FL and ANN, and Table 6 for DT and NBC, respectively.

The FL is better at predicting the 'System Status', with high values of sensitivity and specificity. As for the 'Patient Status' the algorithm is able to avoid the false positives (specificity) but sensitivity (true positive rate) is very low, making this algorithm unfit for patient diagnosis. A possible way to improve the algorithm would be to add more features from the ECG or other vital signs in order to correctly detect a health condition.

The data fuzzy model is flexible, in the sense that further inputs can be added to the system with extra information regarding the patient and the system. For instance, behaviour identification sensors like accelerometers can be added to monitor the patient activity. If motion is detected at the same time the ECG signal is degraded, the system can determine the degradation of the biosignal as temporary and not related with any fault from the electronics or the electrodes. A body temperature sensor allows verifying if a moderately accelerated HR is due to a fever situation (the heart rate increases on average 8.5 beats per minute for a 1 degree C increase in body temperature [13]).

**Table 5** Results for the FL and ANN classifications

| FL | | |
|---|---|---|
| Class | 'Patient Status' | 'System Status' |
| Sensitivity | 32.6 % | 85.3 % |
| Specificity | 93.3 % | 79.1 % |
| ANN | | |
| Class | 'Patient Status' | 'System Status' |
| Sensitivity | 33.3 % | 80 % |
| Specificity | 100 % | 100 % |

**Table 6** Results for the DT and NBC classifications

| DT | | | | |
|---|---|---|---|---|
| Class | 'ok' | 'pa' | 'sa' | 'psa' |
| Sensitivity | 62.5 % | 78.5 % | 84.6 % | 77.1 % |
| Specificity | 38.5 % | 88.6 % | 52.4 % | 82.5 % |
| NBC | | | | |
| Class | 'ok' | 'pa' | 'sa' | 'psa' |
| Sensitivity | 100 % | 92.1 % | 100 % | 96.6 % |
| Specificity | 76.9 % | 100 % | 81.0 % | 98.3 % |

The ANN exhibits slightly better results than the FL, but also fails to deliver an acceptable 'Patient Status' prediction.

For the classification algorithms with categorical outputs the sensitivity and specificity are displayed for all the possible status: both the patient and the system are normal (ok), patient anomaly (pa), system anomaly (sa), and both the patient and the system are anomalous (psa).

The DT algorithm presents good scores for detection of a patient health problem (pa) and simultaneous patient health problem and system anomaly (psa) cases. The DT performs poorly predicting when detecting the 'ok' case. For the system anomaly (sa) situation the specificity value is not good enough for our application, because almost half of the anomalous system cases are not being classified as such. Also the DT contains some decision branches that don't make sense in terms of the system features and patient's vital signs. These results might be improved by adding more data to the training set and increasing the number of input variables.

The NBC has acceptable sensitivity and specificity values for all the output cases. This algorithm is able to correctly predict and distinguish a patient health problem and/or a system fault/failure.

The next step of this work is to implement this algorithm in the smartphone connected to the wearable monitoring system to make the diagnosis in real time.

The results presented in this paper could be improved by adding more input features (signals from the hardware and the patient) and/or by a combination of classifiers. Also, since a high level of dependability is required for wearable devices, the system could be improved by guard-banding the predictions and let the classifiers indicate whenever they are not able to diagnose with confidence. This error moderation approach has been proposed in the field of circuit testing using classifiers [29].

## 5 Conclusion

Diverse wearable medical monitoring systems are being made available, all with different architectures, components, characteristics, and designs. It is common sense and well accepted that high levels of reliability, security, safety, availability and maintainability are required. Such high levels of dependability are difficult to achieve due to the complexity of these monitoring systems, which have different blocks and functional layers (sensors, data acquisition front-end, software, networks, etc.) and the fact that the performance of the captured data actually convey information on both the system behaviour and the wearer physical activity and health condition, making it mandatory to correctly interpret the detected anomalies. Four algorithms based on data fusion, for patient and system diagnosis applied to a wearable medical system case, were presented. It is shown how the algorithms perform and can be explored to correlate data obtained from different sensors, in order to obtain status indicators that characterize the operation correctness of the monitoring system and the health condition of the wearer. The Naive Bayes Classifier algorithm presented the best performance in terms of accuracy, sensitivity and specificity. To overcome memory and processing time issues in the smartphone, the training or change (for example, to include more inputs or update the database) of this classifier can be done in a PC, being the probabilities uploaded afterwords to the smartphone for the on-line classification and diagnosis.

**Acknowledgments** This work has been financed by the ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project SIVIC PTDC/EEI-ELC/1838/2012 (FCOMP-01-0124-FEDER-028937), and grant contract SFRH/BD/81476/2011 (first author).

## References

1. Abbod MF, von Keyserlingk DG, Linkens DA, Mahfouf M (2001) Survey of utilisation of fuzzy technology in Medicine and Healthcare. Fuzzy Sets Syst 120(2):331–349
2. Alemzadeh H, Jin Z, Kalbarczyk Z, Iyer R (2011) An embedded reconfigurable architecture for patient-specific multi-paramater medical monitoring. In: 2011 Annual international conference of the IEEE engineering in medicine and biology society,EMBC, pp 1896–1900
3. Amour N, Hersi A, Alajlan N, Bazi Y, AlHichri H (2015) Implementation of a Mobile Health System for Monitoring ECG signals. Working Paper, Academy of Science and Engineering, USA
4. Bates JHT, Young MP (2003) Applying fuzzy logic to medical decision making in the intensive care unit. Am J Respir Crit Care Med 167(7):948–952
5. Becker K, Thull B, Ksmacher-Leidinger H, Stemmer J, Rau G, Kalff G, Zimmermann HJ (1997) Design and validation of an intelligent patient monitoring and alarm system based on a fuzzy logic process model. Artif Intell Med 11(1):33–53
6. Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and regression trees. Wadsworth, Belmont
7. Cinque M, Coronato A, Testa A (2012) Dependable services for mobile health monitoring systems. Int J Ambient Comput Intell 4(1):1–15
8. Delano MK, Sodini CG (2013) A long-term wearable electrocardiogram measurement system. In: 2013 IEEE International conference on body sensor networks (BSN), pp 1–6
9. Elgendi M, Eskofier B, Dokos S, Abbott D (2014) Revisiting QRS detection methodologies for portable, wearable, Battery-Operated, and wireless ECG systems. PLoS ONE 9(1):e84,018
10. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE (2000) Physiobank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215–e220
11. Hao S (2013) A self-calibration method for an implantable displacement sensor. In: Microwaves, communications, antennas and electronics systems (COMCAS), 2013 IEEE international conference on, pp 1–5
12. Hovakeemian Y, Naik K, Nayak A (2011) A survey on dependability in body area networks. In: 2011 5Th international symposium on medical information communication technology (ISMICT), pp 10–14
13. Karjalainen JVM (1986) Fever and cardiac rhythm. Arch Intern Med 146(6):1169–71
14. Kenneth E, Rajendra AU, Kannathal N, Lim CM (2005) Data fusion of multimodal cardiovascular signals. In: Engineering in medicine and biology society, 2005. IEEE-EMBS 2005. 27th annual international conference of the, pp 4689–4692
15. Khan M, Kerkhoff H (2011) Soc mixed-signal dependability enhancement: A strategy from design to end-of-life. In: 2011 IEEE International symposium on defect and fault tolerance in VLSI and nanotechnology systems (DFT), pp 374–381
16. Kligfield P, Gettes LS, Bailey JJ, Childers R, Deal BJ, Hancock EW, Herpen VM, Kors JA, Macfarlane P, Mirvis DM, Pahlm O, Rautaharju P, Wagner GS (2007) Recommendations for the Standardization and Interpretation of the Electrocardiogram Part I: The Electrocardiogram and Its Technology: A Scientific Statement From the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society Endorsed by the International Society for Computerized Electrocardiology. Circulation 115(10):1306–1324
17. Le Floch F, Bernard S, Bontorin G, Soulier F, Cathebras G (2011) Global strategy to guaranty dependability of electrical medical implanted devices. In: 2011 5Th international IEEE/EMBS conference on neural engineering (NER), pp 515–518
18. Lesbros G, Sawan M (2006) Wireless implantable sensing device for in-vivo monitoring of electrode-tissues contact variation. In: Computer architecture for machine perception and sensing, 2006. CAMP 2006. International workshop on, pp 67–68
19. Mahfouf M, Abbod MF, Linkens DA (2001) A survey of fuzzy logic monitoring and control utilisation in medicine. Artif Intell Med 21(13):27–42
20. Møller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw 6(4):525–533
21. Oliveira C, Almeida N, Machado da Silva J (2013) A stent-graft endoleakage monitor: Telemetry system based on inductive-coupling transmission for implantable pressure sensors. In: Bioengineering (ENBENG), 2013 IEEE 3rd portuguese meeting in, pp 1–4
22. Oliveira C, Sepúlveda A, Almeida N, Wardle B, Machado da Silva J, Rocha L (2015) Implantable flexible pressure measurement system based on inductive coupling. IEEE Trans Biomed Eng 62(2):680–687

23. Oliveira C, Machado da Silva J (2012) Fault detection system for a stent-graft endoleakage monitor. In: Mixed-signals, sensors and systems test workshop (IMS3TW), 2012 18th international, pp 17–21
24. Oliveira C, Machado da Silva J, Trindade I, Martins F (2014) Characterization of the electrode-skin impedance of textile electrodes. In: 2014 Conference on design of circuits and integrated circuits (DCIS), pp 1–6
25. Oliveira CC, Silva JMd (2015) A fuzzy logic approach for highly dependable medical wearable systems. In: Mixed-signal testing workshop (IMSTW), 2015 20th international, pp 1–5
26. Pan J, Tompkins WJ (1985) A Real-Time QRS Detection Algorithm. IEEE Trans Biomed Eng BME 32(3):230–236. doi:10.1109/TBME.1985.325532
27. Paradiso R, Loriga G, Taccini N (2005) A wearable health care system based on knitted integrated sensors. IEEE Trans Inf Technol Biomed 9(3):337–344
28. Salazar A, da Silva JM, Correia M, Mendes BJ (2014) Built-in self-testing methodology and infrastructure for an emg monitoring sensor module. IARIA Int J Adv Syst Meas 7(1-2):68–79
29. Stratigopoulos HG, Makris Y (2008) Error moderation in low-cost machine-learning-based analog/rf testing. IEEE Trans Comput Aided Des Integr Circuits Syst 27(2):339–351. doi:10.1109/TCAD.2007.907232
30. Trindade I, Martins F, Dias R, Oliveira C, Machado da Silva J (2015) Novel textile systems for the continuous monitoring of vital signals: design and characterization. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 3743–3746
31. Wan J, Kerkhoff H (2011) Boosted gain programmable opamp with embedded gain monitor for dependable SoCs. In: Soc design conference (ISOCC), 2011 international, pp 294–297
32. Welch J, Ford P, Teplick R, Rubsamen R (1991) The massachusetts general hospital-marquette foundation hemodynamic and electrocardiographic database – comprehensive collection of critical care waveforms. J Clin Monit 7(1):96–97
33. Zadeh LA (1988) Fuzzy Logic. Computer 21(4):83–93

**Cristina C. Oliveira** received her M.Sc. in Biomedical Engineering from University of Minho in 2010. During her studies she participated in the ERASMUS programme and conducted her masters dissertation in RWTH Aachen (Germany) with the title "Modeling of the Influence of Electrolyte Imbalances on Electrocardiogram (ECG) and Bioimpedance Spectroscopy (BIS) Measurements". From December 2010 until January 2012 she worked in the SenseCardioHealth Project at FEUP. Currently she is a researcher at INESC TEC and is a PhD candidate of the Biomedical Engineering doctoral program at FEUP. Her research interests include wireless telemetry, sensor and data fusion, wearable systems and fault detection.

**José Machado da Silva** graduated and obtained his PhD, both in Electrical and Computer Engineering from the Faculty of Engineering of the University of Porto, Portugal, in 1984 and 1998, respectively. He joined the University of Porto in 1984 as a lecturer, and INESC Porto in 1991 as a researcher. He is currently Associate Professor at UPorto and Project Leader at INESC TEC, with teaching and research responsibilities in design and test of electronic circuits and systems. His research interests include analog, mixed-signal, and RF circuits, MEMS and sensors, wireless networks, and analog and digital signal processing.