**ELSEVIER**

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
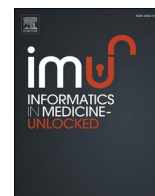
A non-contact infection screening system using medical radar and Linux-embedded FPGA: Implementation and preliminary validation

Cuong V. Nguyen [a], Truong Le Quang [a], Trung Nguyen Vu [b,c], Hoi Le Thi [b], Kinh Nguyen Van [b], Thanh Han Trong [a], Tuan Do Trong [a], Guanghao Sun [d,e,*], Koichiro Ishibashi [d]

[a] School of Electronics and Telecommunications, Hanoi University of Science and Technology, Hanoi, 100000, Viet Nam
[b] National Hospital of Tropical Diseases, Hanoi, Viet Nam
[c] Hanoi Medical University, Hanoi, Vietnam, Hanoi, Viet Nam
[d] Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo, 182-8585, Japan
[e] Center for Neuroscience and Biomedical Engineering, The University of Electro-Communications, Tokyo, 182-8585, Japan

A R T I C L E   I N F O

Keywords:
Infection screening
Embedded system
Field programmable gate array
Digital signal processing
Machine learning

A B S T R A C T

Objectives: In this study, an infection screening system was developed to detect patients suffering from infectious diseases. In addition, the system was also designed to deal with the variability in age and gender, which would affect the accuracy of the detection. Furthermore, to enable a low-cost, non-contact and embedded system, multiple vital signs from a medical radar were measured and all algorithms were implemented on a Field Programmable Gate Array, named PYNQ-Z1.

Methods: The system consisted of two main stages: digital signal processing and data classification. In the former stage, Butterworth filters, with flexible cut-off frequencies depending on age and gender, and a time-domain peak detection algorithm were deployed to compute three vital signs, namely heart rate, respiratory rate, and standard deviation of heart beat-to-beat interval. For the classification problem, two machine learning models, Support Vector Machine and Quadratic Discriminant Analysis, were implemented.

Results: The Student's *t*-test showed that our proposed digital signal processing algorithms coped well with the variability of human cases in age and gender. Meanwhile, the f_1-score of roughly 98.0% represented the high sensitivity and specificity of our proposed machine learning methods.

Conclusion: This study outlines the implementation of an infection screening system, which achieved competent performance. The system might be beneficial for fast screening of infected patients at public health centers in underdeveloped areas, where people have little access to healthcare.

1. Introduction

With the widespread epidemics of influenza [1], the Middle East Respiratory Syndrome (MERS) [2], or the Ebola hemorrhagic fever [3] in recent years, it is imperative to diagnose the presence of infection at mass gathering places. Our previous studies were conducted focusing on the noncontact dengue fever detecting systems based on multiple vital signs using medical radar and thermography. For example, [4–7] developed methods that processed the recorded signal to obtain heart rate (HR) and respiratory rate (RR) by built-in algorithms, namely Autocorrelation, Fast Fourier transform (FFT) and Multiple Signal Classification (MUSIC), on computer software. However, those research studies used built-in algorithms with fixed parameters regardless of the

age and sex of subjects, which could reduce the accuracy of the system. Regarding the task of discriminating patients with infectious diseases, [6] employed the Linear Discriminant Analysis method on a small dataset of 18 samples, which achieved an indifferent accuracy level of 88.9%. Moreover, [8–10] employed sophisticated methods using Neural Networks. However, this approach required high computational complexity to train the Neural Network, partly because it was susceptible to the variability in age and gender.

In this study, several signal processing and machine learning classification algorithms were employed to improve the performance of both tasks. To measure vital signs, a Digital Butterworth filter [11] and a simple time-domain peak detection algorithm were deployed to process the signal acquired from medical radar. To deal with the variability in

HR [12] and RR [13,14] depending on age and gender, instead of the inflexible methods used in Refs. [4–7], filters with flexible cut-off frequencies were designed in this work. For the classification problem, two machine learning models, namely Support Vector Machine (SVM) [15] and Quadratic Discriminant Analysis (QDA) [16] were used, with a far larger dataset of 101 samples, than the 18-sample dataset in Ref. [6], to increase the classification accuracy. Moreover, those algorithms were implemented on a Field Programmable Gate Array (FPGA), since one of the long range objectives of this study is to enable a low-cost, portable, and standalone system, which would be beneficial for public health centers in underdeveloped areas, where people have little access to healthcare, thus they have higher risks of succumbing to infectious diseases. FPGAs have been widely used for multiple signal processing problems, especially filtering [17,18], due to their flexible and reconfigurable structure. Moreover, research focusing on FPGA implementations of machine-learning problems, such as classification, has been thriving in recent years [19–21], since the flexibility, parallelism and energy- efficiency of FPGAs are beneficial for accelerating complicated work. In this study, the FPGA called PYNQ-Z1 was employed to implement all signal processing and classification algorithms.

2. Materials and methods

2.1. The PYNQ-Z1 board

The PYNQ-Z1 board [22] is designed to be used with the PYNQ opensource framework, which is described in Fig. 1. The framework enables embedded programmers to deploy All Programmable System-on-Chip (APSoC) platforms without having to use hardware description languages such as Verilog to design combinational or sequential logic circuits. At the hardware level, programmable logic circuits of the FPGA are presented as libraries called overlays. These overlays are analogous to software libraries and can be accessed through an application programming interface (API). On the other hand, regarding the software level, a high-level programming language, Python, with the Jupyter Notebook design environment [23], as well as the kernels of the Linux operating system, can be deployed for high performance embedded application. The PYNQ-Z1 board combines all of the above elements of both software and hardware levels to simplify and improve APSoC design.

In this study, the capabilities of the PYNQ-Z1 were exploited to enable a low-cost and portable platform. Specifically, the software level of the board enabled the implementations of all algorithms in Python for both the signal processing and classification tasks. Instead of inflexible built-in algorithms on computer software [4–7], flexible digital filters were designed to process the radar signal. The filters were designed with SciPy [24], a Python library for mathematics, science, and engineering. In addition, SVM and QDA classifiers could also supplant complicated

Neural Networks [8–10] deployed on bulky and expensive computers or workstations. The two classifiers were implemented on the board using scikit-learn [25], a framework supporting many machine learning algorithms.

2.2. Overview of the embedded screening system

Fig. 2 provides overall information about the system. The medical radar (NJR 4262J, Japan), which has the frequency range from 24.05 to 24.25 GHz transmitted microwave to the person's chest and received the reflected wave, which contained both the cardiac and respiratory frequency components due to the Doppler effect [6]. The signal was monitored for 30 seconds, then it was transferred to the board for processing.

A digital band-pass filter extracted the cardiac frequency component from the signal, then a peak detection algorithm was used to obtain HR and Standard deviation of heart beat-to-beat interval (SDHI). Similarly, the respiratory component was extracted by a digital low-pass filter, then passed to the peak detection algorithm to compute the RR.

From a machine learning perspective, the signal processing stage performs as a feature extraction process [26], which transforms raw data into a reduced set of features containing relevant information of the original. This process, therefore, reduces the complexity of following classification algorithms, which take the small set of relevant features as input. In this study, each digital signal was processed to generate 3 features, which were HR, SDHI and RR. Then, those 3 features were passed to a classification algorithm. With such a small number of features, machine-learning-based algorithms, such as SVM and QDA can perform well, and they should not be inferior to deep-learning complex structures like Neural Network. Therefore, in this study, the SVM and QDA classifiers were chosen and implemented.

The performance time of the system was also evaluated, which consisted of all amounts of time taken for running algorithms on the PYNQ-Z1 board. The performance time was measured after the 30-second recording time by the medical radar. Assessing performance time should be vital for practical purposes, since if negligible, the screening result of the subject could be displayed promptly; thus a real-time platform could be enabled.

2.3. Data acquisition

Machine learning algorithms often require a large dataset for training, depending on the complexity of the problem, the number of classes and input features. The size of each class in the dataset should be similar. In this study, experiments were conducted on the undergraduates at Hanoi University of Science and Technology (HUST), 41 males and 13 females, aged from 18 to 22, who did not have any illnesses over the measurement period, and 54 samples were collected for the healthy class. For the infected group, the 3 vital signs of 47 dengue fever patients, 26 males and 21 females, from 15 to 77 years of age, were measured at the National Hospital of Tropical Diseases (NHTD), Hanoi, Vietnam, from July 31 to August 2, 2017. In total, the dataset has 101 samples of both healthy and infected classes, which is far larger than the 18-samples dataset used in Ref. [6]. This joint research's data acquisition was approved by the Ethics Committee of the NHTD, HUST and the University of Electro-Communications (UEC).

2.4. Digital signal processing

The system consisted of two main stages: Digital signal processing and data classification. The first stage, including two forms of digital filters, band-pass and low-pass filter, and the peak detection algorithm in time domain, was designed to calculate the 3 vital signs HR, SDHI, RR of a person from the radar signal.
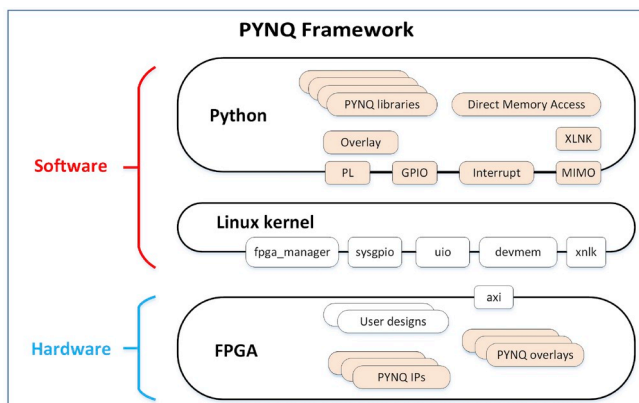

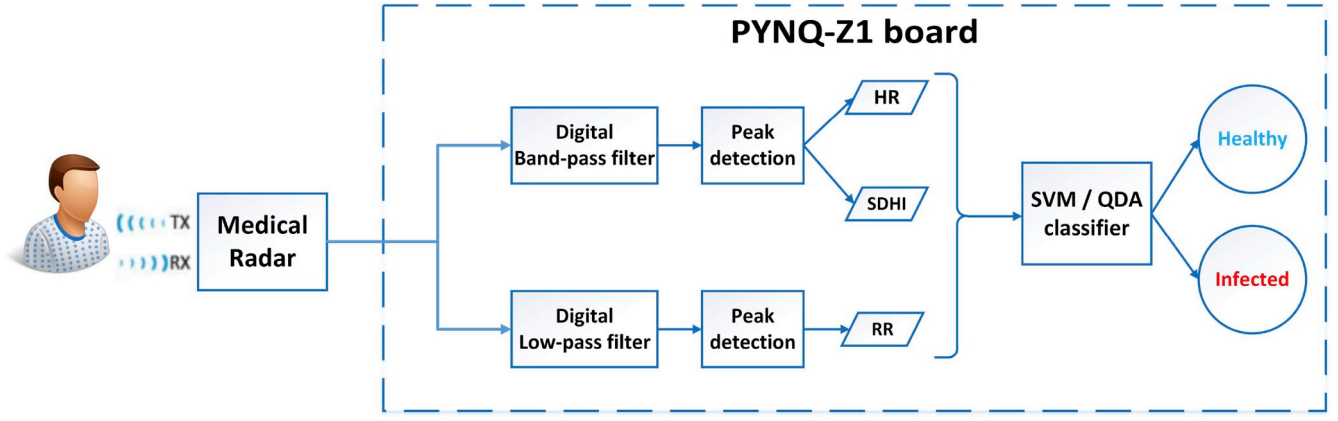**Fig. 1.** Overview of the PYNQ framework.

Fig. 2. Block diagram of the system.

2.4.1. Digital Butterworth filter

The signal recorded by the medical radar contained cardiac, respiratory information, and high-frequency noise [6]. The primary purpose of our filtering method was to eliminate unwanted frequency components. The frequency of the heartbeat was greater than that of the respiration [27] and lower than the noise [6]; therefore, a band-pass filter was applied for calculating the HR. Similarly, a low-pass filter was employed for RR measurement, to remove both the cardiac and noise frequency ranges. The digital filters required a parameter of normalized frequency f_n [28], which was related to the analog frequency f_a [Hz] and the sampling rate $f_s = 100$ [samples/second] in this study by Equation (1). The unit of the normalized frequency f_n is half-cycles per sample (hcps).

$$f_n = 2 \times \frac{f_a}{f_s} \left[\text{hcps} \right] \tag{1}$$

A third-order band-pass and a fifth-order low-pass Butterworth filter were implemented on the PYNQ-Z1 board. By comparing the HR and RR results calculated by different filter orders with the results calculated by reference methods, the third and the fifth orders appeared to produce the most reliable HR and RR outcomes, respectively.

2.4.2. Time-domain peak detection algorithm

The next step was to detect all peaks in the filtered signal. A low-complexity algorithm was designed, which found all local maxima in the filtered signal, whose differences with the two adjacent troughs were higher than a threshold. The elaborate procedure was described in Algorithm 1.

Algorithm 1. Time-domain peak detection.

Algorithm 1 Time-domain peak detection

Input: S[0…n-1], the digital signal
Output: P[0…], list of all peaks
1: $P \leftarrow [\ ]$
2: $E \leftarrow$ *All Extreme Points* (S)
3: $max \leftarrow Max(S)$
4: $min \leftarrow Min(S)$
5: $threshold \leftarrow 0.01(max - min)$
6: for each peak in E do
7:     **if** $(peak - adj\_troughs) > threshold$ **then**
8:         $Append(P, peak)$
9:     **end if**
10: **end for**

2.4.3. Heart rate measurement

A person's heart rate normally ranges between 60 and 100 beats per

minute (bmp) [29], while patients suffering from dengue haemorrhagic fever have higher heart rates [30]. In addition, the heart rate range also varies according to this problem, since there is no parameter which can be optimal for all human cases. In this study, to address with the problem, band-pass filters were designed with different cut-off frequency ranges by sex for four age groups, and the heart rate range was expanded to cover the infected case, as shown in Table 1.

With the sampling rate $f_s = 100$ [cycles/sample], the normalized higher and lower cut-off frequencies of the band-pass filter were computed, as shown in Equations (2) and (3). For example, with the heart rate range [50, 140] [bpm], equivalent to [0.83, 2.33] beats per second:

$$f_{Ln} = 2 \times \frac{0.83}{100} = 0.017 \left[\text{hcps} \right] \tag{2}$$

$$f_{Hn} = 2 \times \frac{2.33}{100} = 0.047 \left[\text{hcps} \right] \tag{3}$$

Applying the peak detection algorithm for the filtered signal, the HR can be obtained by:

$$HR = \frac{n_{peaks}}{t_m} \times 60 \left[\text{bpm} \right] \tag{4}$$

In Equation (4), n_{peaks} is the total number of peaks that were detected, and t_m is the measuring time in seconds.

2.4.4. Standard deviation of heart beat-to-beat interval

The standard deviation is a statistical measure that indicates the amount of variation in a dataset. The heart beat-to-beat interval is the time between consecutive heartbeats, which can be measured as the duration between consecutive peaks in the filtered signal. According to Ref. [31], the SDHI of a normal person should be higher than that of an infected patient, due to the phenomenon of respiratory sinus arrhythmia (RSA) [32]. Therefore, the SDHI should be a beneficial feature to improve the accuracy of diagnosis. Statistically, it is calculated by the following formula:

Table 1
Heart rate range and normalized cut-off frequency by age and sex.

Gender	10–29	30–49	50–69	70–99
Male (HR range [bpm])	50–140	45–130	45–125	40–120
Male (Cut-off [hcps])	0.017–0.047	0.015–0.043	0.015–0.042	0.013–0.040
Female (HR range [bpm])	55–140	50–135	40–125	40–120
Female (Cut-off [hcps])	0.018–0.047	0.017–0.045	0.013–0.042	0.013–0.040

$$SDHI = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2} \ \left[s\right] \qquad (5)$$

In Equation (5), $n = (n_{peaks}- 1)$ is the number of heart beat-to-beat intervals, $x_i[s]$ is the *i*th heart beat-to-beat inter-val, which is the time between two consecutive peaks *i*th and $(i+1)^{\text{th}}$, $\mu[s]$ is the mean value of all heart beat-to-beat intervals x_i in the band-pass filtered signal.

2.4.5. Respiratory rate measurement

The respiratory rate range varies significantly according to age, especially from birth to 18 years of age [13]. The range of children and adolescents, from 18 to 30 breaths per minute [BPM], tends to higher than that of adults, which is from 12 to 20 [BPM] [14]. Therefore, the inflexible methods used in Refs. [4–7] should be highly influenced by this variability. In this study, to measure the subjects' respiratory rate more accurately, a low-pass filter was designed with different cut-off frequencies: 30 [BPM] (0.5 breaths per second) if the subject is aged 18 or more, and 20 [BPM] (0.33 breaths per second) otherwise. Similar to the HR calculation method, the low-pass filter was applied with the normalized cut-off frequencies f_{Cn1} and f_{Cn2} obtained by:

$$f_{Cn1} = 2 \times \frac{0.5}{100} = 0.01 \ \left[\text{hcps}\right] \qquad (6)$$

$$f_{Cn1} = 2 \times \frac{0.5}{100} = 0.01 \ \left[\text{hcps}\right] \qquad (7)$$

After the peak detection stage, the RR was computed by Equation (8):

$$RR = \frac{n_{peaks}}{t_m} \times 60 \ \left[\text{BPM}\right] \qquad (8)$$

The frequency responses of the band-pass and low-pass filters, with normalized cut-off frequencies of [0.017, 0.047] and 0.01 [*hcps*], were illustrated in Fig. 3, in solid red and dashed blue curves, respectively. The normalized cut-off frequencies of the two filters were showed in thin dotted vertical lines as well.

2.5. Data classification

The two classification algorithms proposed in this paper were SVM and QDA. This section briefly provides the mathematical concepts of those methods.

2.5.1. Quadratic Discriminant Analysis

The QDA algorithm [16] assumes that the dataset follows a Gaussian distribution $X \sim N(\mu_k, \sum_k)$, with each class having different covariance
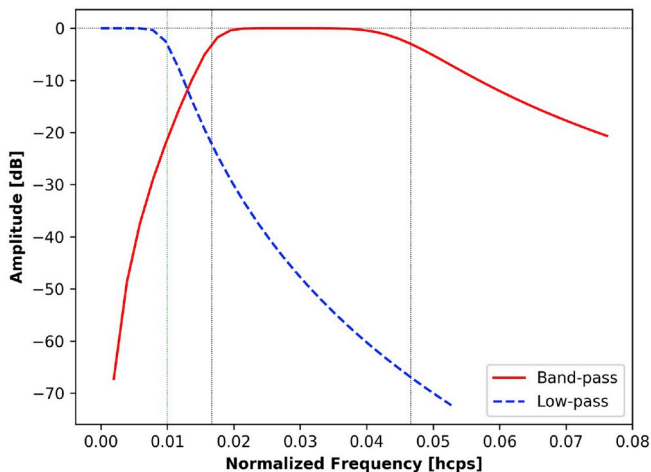

Fig. 3. Frequency responses of Butterworth filters.

matrices, $\sum_1$ and $\sum_{-1}$, as well as different mean values, μ_1 and μ_{-1}. One of the most fundamental concepts of the QDA is the quadratic discriminant function of each class, which is defined for class *k* as:

$$\delta_k\left(x\right) = -\frac{1}{2}\left(x - \mu_k\right)^T \sum_k^{-1} (x - \mu_k) - \frac{1}{2}\log\left|\sum_k\right| + \log\pi_k \qquad (9)$$

The aim of the training phase, as described in Algorithm 2, is to compute all covariance matrices $\sum_k$ with mean values μ_k, and π_k from the labelled dataset, so that the quadratic discriminant function δ_k of an unknown point *x* with respect to class *k* can be calculated. The term π_k is the priori probability of class *k* to the entire dataset.

Algorithm 2. QDA Training phase

Input: Data matrix $X_{(m \times n)}$, containing *m* samples of *n*-dimensional vectors.

Output: Covariance matrices and mean values of all classes: $(\sum_k, \mu_k)$, $k = 1, 2, …, m$

1 Compute the mean of each class:

$$\mu_k = \frac{1}{m_k} \sum x_i, \ \forall x_i \in \text{class·}k$$

m_k is the number of samples of class *k* in the dataset.

2 Calculate the priori probability of each class:

$$\pi_k = \frac{m_k}{m}$$

3 Compute the covariance matrix of each class:

$$\sum_k = \frac{1}{m_k} \sum (x_i - \mu_k)(x_i - \mu_k)^T \ \forall x_i \in \text{ class·}k$$

In the testing phase, given the unknown n-dimensional data *x*, by applying (9) the discriminant functions $\{\delta_1, \delta_{-1}\}$ of *x* were obtained with respect to the two classes. The classification rule assigns the unknown data *x* to the class that has the greater discriminant functions, based on the following classification rule:

$$\bar{y}\left(x\right) = \underset{k}{\text{argmax}}\delta_k(x) \qquad (10)$$

2.5.2. Support Vector Machine

The idea of the SVM algorithm [15] is to find an optimized classifier that maximizes the geometric margins between the decision boundary and the two classes. All points *x* on the decision boundary must satisfy Equation (11):

$$\omega^T x + b = 0 \qquad (11)$$

In (11), (ω, b) are the parameters obtained through the training phase. The geometric margin γ_i of the decision boundary with respect to a sample (x_i, y_i) in the training set is defined in Equation (12):

$$\gamma_i = y_i\left(\left(\frac{\omega}{||\omega||}\right)^T x_i + b\right) \qquad (12)$$

The need to maximize the geometric margin leads to the constrained optimization problem shown in (13):

$$\min_{\gamma, \omega, b} \frac{1}{2}\left|\left|\omega\right|\right|^2, \text{ subject to:} y_i\left(\omega^T x_i + b\right) \geq 1 \ \forall x_i \in X \qquad (13)$$

The problem above can be solved by Lagrange duality [33] and sequential minimal optimization problem (SMO) [34], obtaining the optimal values of (ω, b).

Then, at the testing phase, the SVM classifies an unknown sample x

based on the hypothesis function:

$$h_{\omega,b}(x) = \begin{cases} 1, & if : \omega^T x + b \geq 0 \\ -1, & otherwise \end{cases} \tag{14}$$

The labels 1 and -1 in (14) denote the healthy and infected classes, respectively.

On the other hand, the training phase produces a set of points which are closest to the decision boundary, called the support vectors. The benefit of those support vectors is that when predicting the label of a new point x, following the property of the Lagrange duality:

$$\omega^T x + b = \sum_{i=1}^{m} \alpha_i y_i \langle x, x_i \rangle \tag{15}$$

All of the Lagrange multipliers α_i of all points x_i in the dataset are zero except for support vectors. Hence, many terms in Equation (15) are zero, and the computational cost is significantly reduced.

2.5.3. Hyperparameter optimization

In machine learning, hyperparameters [35] are parameters which are set before the learning process commences. They are different from the parameters of a model generated after the learning process. Different algorithms require different hyperparameters. Hyperparameter optimization, or hyperparameter tuning [35], is the process in which optimal hyperparameters are chosen for a learning algorithm, to produce the best model for a specific problem. There are several methods for optimization, such as grid search and random search [36]. Grid search finds the optimal choice by choosing exhaustively from a set of hyperparameters; thus it guarantees that the optimal hyperparameters in the set will be found. However, its drawback is that when the model needs a number of hyperparameters, the optimizing process can be very time-consuming. Random search, on the other hand, searches the given set randomly, ignoring a certain number of values; thus it is much less time-consuming, but there is no guarantee that the result is the optimal combination of hyperparameters.

The QDA algorithm does not require any hyperparameter, since its training and testing phase can be done sequentially by algebraic operations, while the SVM often has a penalty hyperparameter C of the error term [15]. Because there was only one hyperparameter that needed optimizing, a grid search was employed in this study to find the optimal value of C for the SVM model.

2.6. Assessment methods

Firstly, to assess the accuracy of our proposed approach in measuring vital signs, reference measurements were conducted simultaneously and independently for comparisons. Specifically, the heart rates of all 101 subjects calculated by our method were compared with the ones obtained from a photoplethysmography (PPG) sensor. The PPG signal were processed by a Heart Rate Analysis Toolkit [37] to measure the reference heart rates. For the respiratory rate, the reference values were measured by manually counting how many times the subject's chest rises during the 30-second monitoring period. Three staff separately did the count to obtain a reliable reference.

Our filtering method, with different orders of the band-pass and low-pass Butterworth filters, namely the 2nd, 3rd, 4th, 5th and 6th, was compared with the reference measures. A Student's t-test [39] was employed to find the most appropriate orders. The test showed whether the means of two samples are different from each other, and how significant the differences are, based on statistical computation. In this study, to evaluate the performance of the signal processing stage, the *t*-test was applied for the vital signs obtained by our proposed method and the reference methods.

Secondly, to evaluate the accuracy of the two classification measures, the k-fold cross-validation method [40] was employed. The entire dataset was shuffled randomly and split into k similar-in-size disjoint subsets. (k - 1) subsets were used for training, and the remaining subset was used for testing, which is also called the hold-out cross-validation set. The model made predictions of all samples in the holdout set, and by comparing those with the ground truth class labels, all evaluation metrics such as training accuracy, precision, recall and f_1-score [41] were collected. The process is repeated k times, which means each subset is used 1 time as the hold-out set and ($k - 1$) times for the training stage.

The precision (P) and the recall (R) are defined as:

$$P = \frac{TruePositive}{TruePositive + FalsePositive} \tag{16}$$

$$R = \frac{TruePositive}{TruePositive + FalseNegative} \tag{17}$$

The precision score measures the accuracy of a machine learning model in terms of false positive cases, i.e., the number of infected patients predicted to be healthy. The lower the false positive cases, the higher the precision of the model. Meanwhile, the recall score gives information about false negative cases, which is the number of predictions that a person has an infectious disease but is actually healthy. In the context of disease screening, it appears that ignoring infected patients is far more severe than misdiagnosing healthy people; thus the precision makes more sense than the recall in terms of evaluating a model.

The f_1-score is an overall statistical measure that combines and balances both the precision and recall values; thus it is the most suitable metric for model evaluation.

$$f_1 = 2 \times \frac{PR}{P + R} \tag{18}$$

In this study, 4-fold cross validation (k = 4) was chosen, which means that each subset constitutes 25% of the entire dataset. To make a reliable assessment about the robustness of a classifier, especially when shuffling the dataset before splitting, the 4-fold cross validation was conducted 10 times repeatedly to obtain the average and standard deviation (STD) of evaluation metrics.

However, the cross-validation method was not used to measure the training and testing time of the two classifiers, since this method reduced the size of the dataset used for training. Instead, the entire dataset containing the 3 vital signs of 101 objects was used for training, and only one test case was used for testing.

3. Results

Fig. 4 illustrates the results of the digital signal processing stage and how the three vital signs were computed. In the second subplot, through the band-pass filter, all of the high-frequency noise and respiratory information in the original signal were eliminated, to produce the HR from the extracted cardiac frequency component. The third subplot illustrates the waveform of a reference cardiac signal, monitored by the PPG sensor simultaneously with the radar signal. Meanwhile, the fourth subplot shows the low-pass filtered signal, which contains information to calculate the RR. All red dots represent the peaks detected in the two signals after filtering, by applying Algorithm 1. Then using (4), (5) and (8), with the measuring time $t_m = 30s$, 32 peaks in the cardiac, and 8 peaks in the respiratory filtered signals, the 3 vital signs were obtained: $HR = 64 [bpm]$, $SDHI = 0.21 [s]$, $RR = 16 [BPM]$.

Table 2 and Table 3 show the assessment of the accuracy of our signal processing method, in comparison with the reference methods by the PPG sensor and manual counting for HR and RR, respectively. A *t*-test was conducted, with the null hypothesis that the two results are different from each other. *t*-value and *p*-value are two metrics of the test. The lower the absolute of the *t*-value, the more similarity between the results of the two methods. Meanwhile, if the *p*-value, from 0 to 1, was smaller than a threshold, typically 0.05, the null hypothesis that there is no significant difference between the two methods could be rejected. In
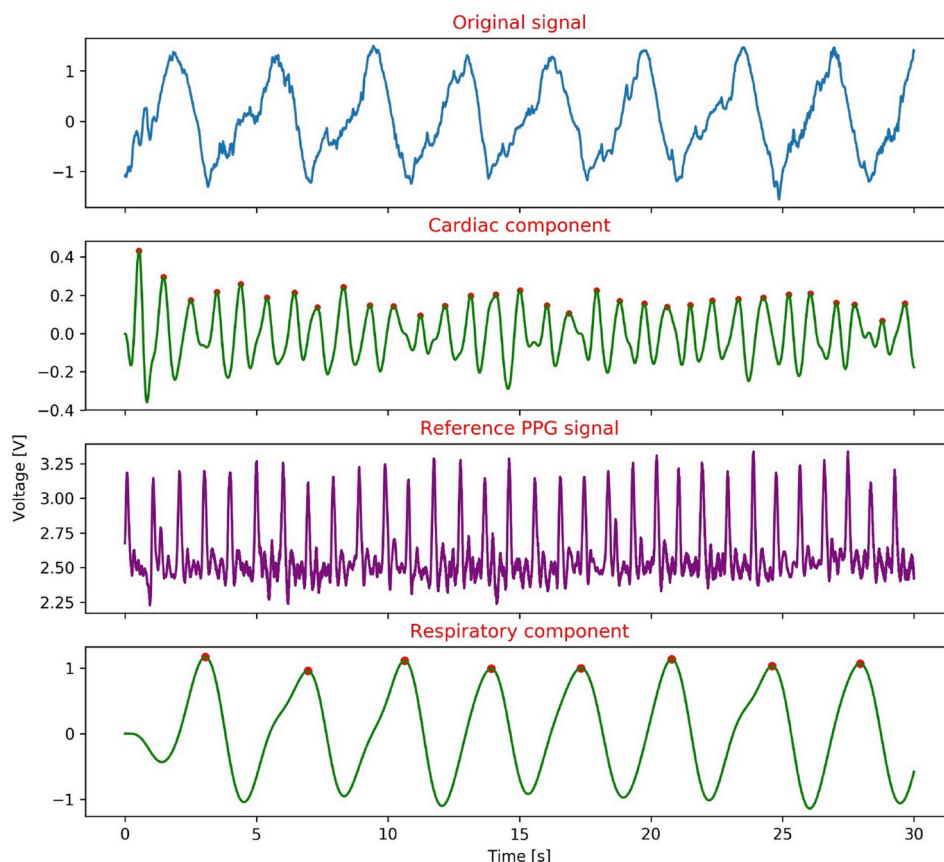
Fig. 4. Original, filtered and reference signals.

Table 2
Assessment of heart rate measurement.

Metrics	2nd	3rd	4th	5th	6th	Auto-Correlation	FFT	MUSIC
t-value	0.21	0.11	0.13	0.22	0.29	0.21	0.19	0.16
p-value	0.16	0.03	0.11	0.18	0.23	0.19	0.15	0.10

Table 3
Assessment of respiratory rate measurement.

Metrics	2nd	3rd	4th	5th	6th	Auto-Correlation	FFT	MUSIC
t-value	0.59	0.31	0.24	0.09	−0.46	0.27	0.22	0.18
p-value	0.38	0.25	0.19	0.03	0.36	0.21	0.18	0.14

Table 2, the 3rd-order filter produced the best performance for heart rate measurement, with t-value = 0.11 and p-value = 0.03. Similarly, Table 3 indicates that the 5th, with t-value = 0.09 and p-value = 0.03, was the most appropriate order for calculating the respiratory rate.

In addition, three previous methods to measure vital signs from the digital radar signal, namely Auto-Correlation [4], FFT [5,6] and MUSIC [7], were also evaluated by the t-test with the reference method. Those measures were operated on our dataset of 101 subjects. Tables 2 and 3 show that the performance of our proposed filtering method should be superior to the three methods, since [6] employed those three as rigid

built-in algorithms, with fixed parameters, which were more susceptible to the variability of human cases.

Table 4 compares the 3 signal processing phases in terms of the average amount of time taken to obtain the three vital signs from a signal recorded in 30 seconds by the medical radar. It is evident that the peak-detecting process accounted for by far the highest amount, approximately 80% of the total. Meanwhile, since the calculation of HR and SDHI require two same steps of band-pass filtering and peak detecting, the gross processing time of both the two features was reduced to approximately 35.87 ms. In total, the signal processing stage on the board took about 70 ms.

Fig. 5 visualizes pairs of the 3 vital signs HR, SDHI, and RR calculated from the signal processing stage, of all healthy and infected people included in our research study. It is clear that the HR and RR of ordinary people appeared to be lower than those of the unhealthy, while the converse can be seen in the SDHI.

The grid search was used for tuning the hyperparameter C of the SVM
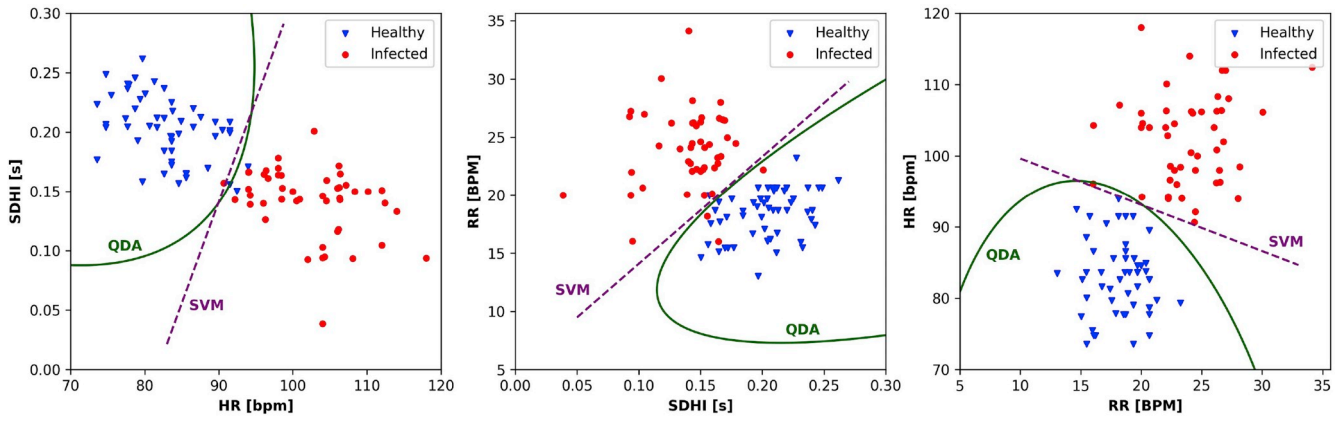
Table 4
Digital signal processing time (ms) on the board.

Vital sign	Filtering	Peak detecting	Calculating	Total
HR	5.7	29.6	0.07	35.37
SDHI	5.7	29.6	0.5	35.80
RR	4.7	29.6	0.01	33.61

Fig. 5. Data visualization.

Table 5
Assessment of the two classification algorithms.

Model	Training time (ms)	Training accuracy	Predicting time (ms)	Precision	Recall	f_1-score
QDA	4.63	98.5%	1.63	98.1%	98.0%	98.1%
STD	0.1%	0.1%	0.09	0.013%	0.012%	0.014%
SVM	403	98.8%	1.04	98.0%	97.8%	97.9%
STD	0.12%	0.1%	0.06	0.6%	0.6%	0.6%

model, ranging from 1 to 100, with the evaluating metrics f_1-score. The results showed that the value C = 45 seems to be optimal, with which the SVM model achieved the highest f_1-score of 97.9%.

The solid green parabolas and dashed purple lines represent the 2-D decision boundaries of the QDA and SVM algorithms, respectively. The decision boundary of the linear-kernel SVM [42] is a straight line, while that of the QDA algorithm is a parabolic curve. It can be seen that the two classifiers distinguished the 2 classes quite separately, and subsequent results quantitatively show that both the SVM and the QDA measures achieved high precision, recall, and f_1-scores.

Table 5 compares the performance of the SVM and the QDA models. In the training phase, the SVM model generated higher accuracy than the QDA. In contrast, all the metrics with regard to the testing phase of the QDA were slightly higher.

The average predicting time of the QDA algorithm for one test case was approximately 1.63 ms, significantly higher than the 1.04 ms value of SVM. These results could be attributed to (9) and (15), which show that the QDA required more complicated operations with the covariance matrices, such as determinant and inverse, than solely dot product and addition of the SVM. In contrast, the SVM required 403 ms for training the model, nearly 90 times as high as the training time of the QDA. It was probably because the SVM includes an iterative procedure to solve the constrained optimization problem, whereas training the QDA classifier only requires a fixed number of steps for basic matrix operations. In practical purposes, all machine learning models can be trained prior to testing; thus the predicting time appears to outweigh the training time. Therefore, on balanced, the performance of the SVM model appeared to be superior.

4. Discussion

4.1. Main findings

In Tables 2 and 3, the small absolute *t*-value and *p*-value, around 0.10 and 0.03 respectively, show that our proposed methods that used flexible cut-off frequencies of Butterworth filters coped well with the variability in vital signs depending on age and sex. Therefore, the measurements of heart rate and respiratory rate were more accurate.

Table 5 shows the negligible difference between the SVM and the

QDA models regarding precision, recall, and f_1-score. Overall, the two classifiers distinguished the healthy and infected classes quite accurately, as they both achieved f_1-scores of about 98.0%, significantly higher than the accuracy of 88.9% of the Linear Discriminant Analysis in Ref. [6]. Those models were also not inferior to the complicated Neural Networks used in Refs. [8,10], which achieved an accuracy of 97.1% and 98.0%. Additionally, low standard deviations of all the metrics indicate that the two classifiers are robust and less susceptible to overfitting [43]. Moreover, the far larger dataset of 101 samples contributed to the robustness of our approach, compared with the dataset of 18 samples in Ref. [6]. Furthermore, mitigating the effects of variability in human cases in the preceding signal processing stage also benefited the performance of the two models.

4.2. Limitations

The total performance time on the PYNQ-Z1 board was nearly 72 ms, including 70 ms of signal processing and about 2 ms to make a binary prediction, which was negligible compared with the 30s for monitoring the signal via medical radar. On one hand, the longer the measuring time, the better the estimates of vital signs, as both the HR and RR in (4) and (8) are susceptible to how long the radar signal is monitored. On the other hand, 30 seconds is rather long for practical applications. Furthermore, the dataset of 101 subjects with only 3 features per sample, as well as the narrow age range of the healthy group, might be other limitations of our study.

5. Conclusion

In this research, we developed the embedded infection screening system, exploiting the capabilities of the PYNQ-Z1 board. Specifically, algorithms for digital signal processing and data classification were implemented. The flexible parameters of Butterworth filters mitigate the effects of the variability of human cases, which improved the accuracy of measuring vital signs. Two machine learning models, SVM and QDA, were also proposed for classification, to increase the sensitivity and specificity of the system. The results show that the proposed algorithms achieved good performance.

The low-cost embedded system has considerable potential for

practical purposes. In future work, we intend to measure additional vital signs such as blood pressure and oxygen saturation and increase the number of people from different age groups involved in our research, as well as reduce the monitoring time of the system, while maintaining the competent and robust performance.

Funding

Author contributions

Conceptualization, C.V.N., K.I. and G.S.; methodology, C.V.N., T.L. Q., T.H.T. and T.D.T.; software, C.V.N., G.S. and T.L.Q; validation, G.S. and K.I.; formal analysis, C.V.N.; investigation, T.L.Q., H.L.T., K.N.V. and T.N.V.; resources, G.S. and K.I.; data curation, G.S., K.I., H.L.T., K.N. V. and T.N.V.; writing–original draft preparation, C.V.N. and T.L.Q.; writing–review and editing, T.D.T., T.H.T., G.S. and K.I.; visualization, C.V.N.; supervision, T.D.T., T.H.T. and G.S.; project administration, K.I.; funding acquisition, G.S. and K.I.

Ethical statement

This study was approved by the Ethics Committee of the National Hospital of Tropical Diseases (Vietnam), Hanoi University of Science and Technology (Vietnam) and The University of Electro-Communications (Japan).

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

References

[1] Nishiura Hiroshi, Kamiya Kazuko. Fever screening during the influenza (h1n1-2009) pandemic at narita international airport, Japan. BMC Infect Dis 2011;11(1): 111.

[2] MATSUYAMA Shutoku. Diagnosis of middle east respiratory syndrome coronavirus in Japan. J Vet Epidemiol 06 2016;20:S18–9.

[3] Trevor R Shoemaker, Stephen Balinandi, Tumusiime Alex, Nyakarahuka Luke, Lutwama Julius, Mbidde Edward, Kofman Aaron, Klena John D, Ströher Ute, Rollin Pierre E, Nichol Stuart T. Impact of enhanced viral haemorrhagic fever surveillance on outbreak detection and response in Uganda. Lancet Infect Dis apr 2018;18(4):373–5.

[4] Sun Guanghao, Matsui Takemi. Rapid and stable measurement of respiratory rate from Doppler radar signals using time domain autocorr-elation model. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE; 2015. p. 5985–8.

[5] Sun G, Vinh NQ, Matsuoka A, Miyata K, Chen C, Ueda A, Kim S, Hakozaki Y, Abe S, Takei O, Matsui T. Design an easy to-use infection screening system for non-contact monitoring of vitalsigns to prevent the spread of pandemic diseases. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society; Aug 2014. p. 4811–4.

[6] Yang Xiaofeng, Ishibashi Koichiro, Negishi Toshiaki, Kirimoto Tetsuo, Sun Guanghao. Short time and contactless virus infection screeni-ng system with discriminate function using doppl-er radar. In: 12th international conference on bio-inspired computing: theories and application, vol 11; 2017. p. 263–73.

[7] Negishi Toshiaki, Sun Guanghao, Sato Shohei, Liu He, Matsui Takemi, Abe Shigeto, Nishimura Hidekazu, Tetsuo Kimimoto. Infection screen-ing system using thermography and ccd camera with good stability and swiftness for non-contact vital-signs measurement by feature matching and music algorithm. In: The 41st annual international coference of the IEEE engineering in medicine and biology society (EMBC). IEEE; 2019.

[8] Sun Guanghao, Matsui Takemi, Hakozaki Yukiya, Abe Shigeto. An infectious disease/fever screening radar system which stratifies higher-risk patients within ten seconds using a neural network and the fuzzy grouping method. J Infect 2014; 70:12.

[9] Yao Y, Sun G, Kirimoto T, Matsui T, Schiek M. Online state space filtering of biosignals using neural network-augmented kalman filter. In: 2017 10th biomedical engineering international conference (BMEiCON); Aug 2017. p. 1–5.

[10] Yang Xiaofeng, Kumagai Koki, Sun Guanghao, Ishibashi Koichiro, Nguyen Vu Trung, Van Kinh Nguyen, et al. Dengue fever screening using vit-al signs by contactless microwave radar and ma-chine learning. In: 2019 IEEE sensors applications symposium (SAS). IEEE; 2019. p. 1–6.

[11] Selesnick Ivan, Burrus Charles. Generalized digital butterworth filter design. In: IEEE transactio-ns on signal processing. vol 46; 05 1998.

[12] Umetani Ken, Singer Donald H, McCraty Rollin, Atkinson Mike. Twenty-four hour time do-main heart rate variability and heart rate: re-lations to age and gender over nine decades. J Am Coll Cardiol 1998;31(3):593–601.

[13] Fleming Susannah, Thompson Matthew, Stevens Richard, Heneghan Carl, Plüddemann Annette, Maconochie Ian, Tarassenko Lionel, Mant David. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. The Lancet 2011;377(9770):1011–8.

[14] Huang Ming-Chun, Xu Wenyao, Liu Jason, Samy Lauren, Vajid Amir, Alshurafa Nabil, Sarrafzadeh Majid. Inconspicuous on-bed respiratory rate monitoring. In: Proceedings of the 6th international conference on PErvasive technologies related to assistive environments. ACM; 2013. p. 18.

[15] Bishop Christopher M. Pattern recognition and machine learning. Information Science and St-atistics; 01 2006.

[16] Tharwat Alaa. Linear vs. quadratic discriminant analysis classifier: a tutorial. Int J Appl Pattern Recognit 01 2016;3(145).

[17] German-Sallo Zoltan. Signal processing using fpga structures. Procedia Technol 2014;12:112–8.

[18] Jinalkumari K Dhobi, Shukla DYB, Bhatt DKR. Fpga implementation of fir filter using various algorithms: a retrospective. Int J Res Comput Sci 2014;4(2):19–24.

[19] Ahmad Shawahna, M Sait Sadiq, El-Maleh Aiman. Fpga-based accelerators of deep learning networks for learning and classification: a review. IEEE Access 2018;7: 7823–59.

[20] Liu Xu, Ounifi Hibat Allah, Gherbi Abdelouahed, Lemieux Yves, Li Wubin. A hybrid gpu-fpga-based computing platform for machine learning. Procedia Comput Sci 2018;141:104–11.

[21] Ling Andrew, Anderson Jason. The role of fpgas in deep learning. In: FPGA 2017-Proce-dings of the 2017 ACM/SIGDA international symposium on field-programmable gate arrays; 2017.

[22] Pynq - python productivity for zynq (Archived on 04/02/2019), http://webcitation.org/77K5UFLoz.

[23] Kluyver Thomas, Ragan-Kelley Benjamin, Pérez Fernando, Granger Brian, Bussonnier Matthias, Frederic Jonathan, Kelley Kyle, Hamrick Jessica, Grout Jason, Corlay Sylvain, Ivanov Paul, Avila Damián, Abdalla Safia, Willing Carol. Jupyter notebooks – a publishing format for reproducible computational workflows. In: Loizides F, Schmidt B, editors. Positioning and power in academic publishing: players, agents and agendas. IOS Press; 2016. p. 87–90.

[24] Jones Eric, Oliphant Travis, Peterson Pearu, et al. SciPy: open source scientific tools for Python. 2001 [Online; accessed <today>].

[25] Buitinck Lars, Louppe Gilles, Blondel Mathieu, Pedregosa Fabian, Mueller Andreas, Grisel Olivier, Vlad Niculae, Peter Prettenhofer, Gramfort Alexandre, Grobler Jaques, Layton Robert, VanderPlas Jake, Joly Arnaud, Holt Brian, Varoquaux Gaël. API design for machine learning software: experiences from the scikit-learn project. In: ECML PKDD workshop: languages for data mining and machine learning; 2013. p. 108–22.

[26] Guyon Isabelle, Gunn Steve, Nikravesh Masoud, Zadeh Lofti A. Feature extraction: foundat-ions and applications, vol 207. Springer; 2008.

[27] Garde Ainara, Walter Karlen, Ansermino J Mark, Dumont Guy A. Estimating respiratory and heart rates from the correntropy spectral density of the photoplethysmogram. PLoS One 2014;9(1):e86427.

[28] Carlson Gordon E. Signal and linear system analysis. NJ: John Wiley Hoboken; 1998.

[29] Jose Anthony D, Collison D. The normal range and determinants of the intrinsic heart rate in man. Cardiovasc Res 1970;4(2):160–7.

[30] Robert III Carter, Hinojosa-Laborde Carmen, Convertino Victor A. Heart rate variability in patients being treated for dengue viral infection: new insights from mathematical correction of heart rate. Front Physiol 2014;5:46.

[31] Yang Xiaofeng, Ishibashi Koichiro, Le Hoi, Vu Trung Nguyen, Nguyen Van Kinh, Sun Guanghao. Dengue fever detecting system using peak-detection of data from contactless Doppler radar 07 2018;2018:542–5.

[32] Yasuma Fumihiko, Hayano Jun-ichiro. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? Chest 2004;125(2):683–90.

[33] Radu Boǎč, Grad Sorin-Mihai, Wanka Gert. On strong and total Lagrange duality for convex optimization problems. J Math Anal Appl 01 2008;337:1315–25.

[34] Platt John. Sequential minimal optimization: a fast algorithm for training support vector mach-ines. In: Advances in Kernel MethodsSupport Vector Learning, vol 208; 07 1998.

[35] Probst Philipp, Bischl Bernd, Boulesteix Anne-Laure. Tunability: importance of hyperpara-meters of machine learning algorithms. 2018. arXiv preprint arXiv: 1802.09596.

[36] James Bergstra, Bengio Yoshua. Random search for hyperparameter optimization. J Mach Learn Res 2012;13(Feb):281–305.

[37] Van Gent P, Farah H, van Nes N, van Arem B. Heart rate analysis for human factors: development and validation of an open source toolkit for noisy naturalistic heart rate data. In: *Proceedings of the 6th HUMANIST 11 conferen-ce*; 2018. p. 173–8.

[39] Graeme D Ruxton. The unequal variance t-test is an underused alternative to student's t-test and the mann–whitney u test. Behav Ecol 2006;17(4):688–90.

[40] Kohavi Ron, et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Ijcai, vol 14. Montreal: Canada; 1995. p. 1137–45.

[41] Goutte Cyril, Gaussier Eric. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In: European confere-nce on information retrieval. Springer; 2005. p. 345–59.

[42] Patle A, Chouhan DS. In: International conference on advances in techno-logy and engineering (ICATE). Jan; 2013. p. 1–9.

[43] Douglas M Hawkins. The problem of overfitting. J Chem Inf Comput Sci 2004;44 (1):1–12.