Electronic system to monitoring vital signs in pregnancy through Random Forests

Juan Quintanar-Gómez, Daniel Robles-Camarillo, Francisco Rafael Trejo-Macotela, Ocotlán Díaz-Parra

1 Posgrado en Tecnologías de Información y Comunicaciones, Universidad Politécnica de Pachuca
juanquintanargomez@micorreo.upp.edu.mx

Abstract. Currently, the care of women in pregnancy plays an essential role in improving the clinical conditions of women and babies through early detection of hypertensive disorders during pregnancy. For this reason, the clinical guidelines define blood pressure monitoring as a component for the detection of those disorders. In response to this problem, this paper proposes the design of an electronic prototype focused on the acquisition of blood pulses and weight, transmitting the information obtained through Bluetooth to a computer that processes the information obtained through Random Forest, obtaining the patient's blood pressure. As a result, a MAE of 3.215 and 1.55 mmHg is obtained for the systolic and diastolic pressure, respectively. This work is directed to the use of e-Health and artificial intelligence techniques to improve the quality of life of women in the gestational stage or patients who need to monitor their blood pressure during their daily activities.

Keywords: Random Forest, Blood pressure, e-Health.

1 Introduction

One of the diseases with the most presence on a global level is the high blood pressure, a condition caused by a constantly high pressure in the blood vessels, which applies a force against the blood vessel walls when the blood is pumped, causing a more significant effort to the heart during the dilation process and, as a result, it generates complications in the body when it is not detected on time by the individual or medical personnel [1], [2]. These aspects, in the case of women during the pregnancy period, may cause the appearance of hypertensive disorders during pregnancy, such as gestational hypertension, preeclampsia, eclampsia or chronic hypertension with overlapping preeclampsia. These hypertensive disorders, when undetected and untreated in their initial developmental stage, cause damage to both the woman and her baby, from placental hypoperfusion, hypoxia [3], renal dysfunction, birth with a low body mass index [4] and, in the worst case, maternal and fetal death.

Therefore, the clinical guidelines for the care of women during pregnancy propose different methods for the decrease and detection of hypertensive disorders during pregnancy [1], [5]–[7] from eating essential foods for the development of the fetus to visiting the clinical center in order to monitor the gestational period. In addition, the clinical guidelines suggest clinical monitoring of the patient's blood pressure during medical consultation and daily activities [5], [8], becoming an early warning parameter for the detection of hypertensive disorders during pregnancy [5], [8], [9]. Likewise, weight is an essential factor in the detection of hypertensive disorders, identifying a correlation between both factors during the increase or reduction of the patient's weight [10], [11].

In response to this problem, this research proposes the design of an electronic system with the capacity to acquire blood pressure and capture the patient's weight, transmitting the data obtained to a computer, which through the use of Random Forest processes the information and obtains the blood pressure parameters, projecting itself as an additional tool for the detection of blood pressure.
2 State of the art

Vital signs monitoring has positioned itself as a critical practice for the diagnosis of the patient's clinical condition, notifying the user or medical staff through different visualization technologies about the collected parameters. For this reason, the use of Internet of Things (IoT) technologies has been integrated into vital electronic signs monitoring systems, providing them with the ability to communicate the acquired parameters both to the patient and the medical staff responsible for the clinical monitoring.

An example is described by Supriyanti et al. [12] in 2016, which described the use of a mobile application to monitor the blood pressure of patients during the gestational period, focusing the application on data storage purposes. It also establishes in future work that the application can be integrated with algorithms for the prediction of hypertensive disorders.

In 2015, Meng et al. [13] designed a sphygmomanometer with blood pressure acquisition capacity through oscillometric signals, filtering the information obtained through a Butterworth filtering with a central cut-off frequency of 0.8 Hz. It also mentions that the blood pressure acquisition was performed through the Mean Arterial Pressure (MAP) obtained during the filtering stage.

In 2012, Gao et al. [14] developed an electronic system to monitor the blood pressure and cardiovascular hemodynamic parameters of users through the acquisition of oscillometric signals from a pressure transducer.

In 2018, Manzano et al. [15] presented the design of an ambulatory monitoring device developed for the diagnosis of hypertension. The prototype developed acquires blood pressure and electrocardiogram data, transmitting the data obtained by low-power Bluetooth version 4.0 to a mobile device.

In 2016, Gutierrez et al. [16] designed a prototype that allows the monitoring of the blood pressure of patients using a Blood Pressure Monitoring Equipment integrated into a Raspberry, which sends the information from the baumanometer to a smartphone via Zigbee protocols, providing an easy interpretation of the patient's blood pressure.

In 2020, Pohl et al. [17] designed an electronic system to acquire and transmit basic patient bio-signals to a mobile device, transmitting the resulting parameters through an HC-05 module.

In 2017, Kario et al. [18] developed an ambulatory and home blood pressure monitoring system, identifying the variability of central and brachial blood pressure during the patient's daily activity. In addition, the device monitors the temperature, air pressure and patient location using a highly sensitive actograph.

In 2016, Adi et al. [19] presented a device to monitor the blood pressure of the patient through a pressure transducer, the information obtained from the transducer is filtered through a double bandpass filtering with a cut-off frequency of 0.278 Hz and 5.837 Hz. The information obtained from the processing stage is transmitted to a Raspberry Pi3 module through a Zigbee module. It also proposes in future works the transmission of information through the use of encryption algorithms, safeguarding confidentiality during the transmission process.

In 2016, Balamurugan et al. [20] developed a low-cost non-invasive blood pressure measurement system that uses the Photoplethysmogram (PPG) technique using reflective infrared sensors, identifying the blood pressure of the user.

In 2019, Turk et al. [21] developed a low-cost microprocessor-based wireless blood pressure and pulse monitoring system, receiving the blood pressure information through amplification and filtering of the signal acquired through the use of the BIOPAC SS19LA blood pressure bracelet and transducer, transmitting the data via a Bluetooth module.

In each of the works described previously, it is possible to identify the modules and sensors responsible for acquiring the user's clinical parameters. Moreover, it is essential to have an algorithm responsible for processing and giving the resulting information to the end-user with a minimum error concerning the parameters obtained by certified clinical devices. In view of this, it is possible to identify the use of Artificial Intelligence techniques as an alternative for processing the user's biological signal and delivering its results through communication technologies or through the display integrated into the device.
Example of this is described in 2018 by Anisimov et al. [22], describing the use of neural networks for the determination of blood pressure coefficients using MATLAB software, previously acquiring the patient's oscillometric signal through a certified clinical device and presenting a mean square error (MSE) of 20.5 mmHg for systolic pressure detection and 29.3 mmHg for diastolic pressure detection.

In 2017, Skorobogatova et al. [23] designed a device to acquire blood pressure parameters using neural networks, obtaining an MSE of 11 mmHg during systolic pressure identification and 26 mmHg for diastolic pressure, obtaining the result of this processing through the device's LCD screen.

In 2019, Sidhu et al. [24] identified the performance of the various oscillometric estimation algorithms such as: the Maximum Amplitude Algorithm (MAA), the Maximum-Minimum Slope Algorithm, the Arterial Lumen Area Algorithm (ALA) and the Pulse Time Transit algorithm. The results described by [24] indicated that the ALA algorithm followed by PTT are relatively accurate in estimating systolic pressure, mean blood pressure and diastolic blood pressure in cardiac patients.

In 2014, Chen et al. [25] reported a predictive analysis of cardiovascular disease using vector support machines (VSM), the nearest K-neighbors (kNN). In addition [25] describes that in future work, KNN algorithms can be used: Random Forests, Neural Networks and Set-based Learning Algorithms in order to compare their predictive capacity.

In 2017, Koohi et al. [26] presented a version of the ALA algorithm for blood pressure estimation, highlighting a 56.7% and 57.3% improvement in mean absolute error (MAE), 98.9% and 64.4% improvement in mean error, 50% and 59% improvement in standard deviation of errors, and up to 57.6% and 55.8% improvement in measurement uncertainty for systolic and diastolic pressures respectively.

In 2016, Kachuee et al. [27] presented a pulse time-of-arrival algorithm for the estimation of systolic, diastolic and MAP blood pressure indices through the processing of vital signals and the extraction of two types of characteristics, which are based on physiological parameters or a complete representation of vital signals.

In 2016, Li et al. [28] reported the use of a retropropagation algorithm to calculate the weight of hidden layer nodes in a neural network, highlighting a standard deviation between blood signals and actual signals of 4.4797 mmHg, complying with the American National Standards of the Association for the Advancement of Medical Instrumentation.

In 2018, Tan et al. [29] presented a non-invasive method of continuous blood pressure measurement based on a neural network, integrating a genetic algorithm for blood pressure determination. The author [29] reports that the neural network results are more accurate than regression models.

In 2018, Lu et al. [30] proposed a method of blood pressure estimation based on a neural network model algorithm that calculates the Pulse Width Transit Time (PWTT) through data analysis, identifying an error of less than five mmHg and a standard deviation of less than nine mmHg.

3 System overview

The developed system consists of the following elements: hardware, information acquisition, modeling and evaluation. These are described below:

Hardware

The proposed electronic system provides the ability to acquire and transmit the blood pulses of the patient. As shown in the block diagram in Figure 1, the microcontroller responsible for managing the processes of information acquisition and transmission is the PIC18F4550 microcontroller, which is powered by a 4.2 V lithium battery. The prototype integrates LEDs in order to visualize, by turning on the LEDs, the task to be performed. The information regarding the pulses is acquired using the pressure transducer integrated into the device, which is connected through a hemodynamic bracelet and the motors in charge of the inflation and deflation of the bracelet. Likewise, the patient's weight is acquired by manually capturing this parameter in the computer. The prototype adds a Bluetooth HC-05 transmission module to transmit information to the computer.
Figure 1. Block diagram of the electronic device

Blood pulse acquisition process

The process of acquiring the blood pulses is performed by acquiring the voltage from the pressure transducer. In this case the microcontroller acquires the values of the transducer through the AN1 pin. As mentioned in the clinical guides [31], to acquire the blood pulses it is necessary to manage the pressure exerted by the cuff, this is managed through the air pump and the solenoid valve connected to the microcontroller through the L293D module. Therefore, the use of Pulse Width Modulation (PWM) from the microcontroller is implemented to set the operating period of the inflation and deflation motors. The signal generated from PWM is established at a frequency of 1 kHz and, during the inflation process, the duty cycle of the inflation engine is set at 90% of its load cycle and the deflation engine is configured with a 100% load cycle, preventing air from escaping into the cuff until a set threshold value is reached, configuring the inflation process in the cuff until it reaches a pressure value of 200 mmHg [32]. When the established threshold value is reached, the PWM configuration of the inflation engine is deactivated, and the deflation process is executed in order to decrease the pressure exerted on the individual's wrist, allowing the identification of blood pulses within a 60-second interval (Figure 2).

During the process of acquiring the blood pulses, the prototype filters the signal obtained in a frequency band of 0.5 to 5 Hz [33], identifying the oscillations coming from the blood pulses. The prototype configuration identifies the MAP using the Maximum Amplitude Algorithm (MAA) [33].

Figure 2. Blood pulses detected during the acquisition process.

Through identification of the MAP, the systolic and diastolic blood pressure is identified through supervised learning techniques, allowing through a training phase the learning of patterns and during the test phase the performance of the model against the desired result is identified [34]. As a result, the use of a decision tree was established to predict the user's blood pressure.
Understanding the data

The data collection was obtained from the Kaggle website that refers to a dataset of patient samples focused on the detection of cardiovascular problems [35]. Thus, a sample of 70,000 units is obtained from the dataset. The dataset consists of 12 variables, selecting 3 variables only, which are:

- Weight
- Systolic Pressure
- Diastolic Pressure

As shown in Figure 3, the information related to diastolic and systolic pressure has many outliers; outliers above 500 mmHg are values higher than the pressure levels detected by aneroid or electronic sphygmomanometers. These values are therefore removed during data cleansing. Information on individual weight is added to reports provided by [10], [11], which show a correlation between patient weight and blood pressure. In this case, the weight variable presents numerous outliers. Therefore, these values are imputed.

Equation (1) is also established to obtain an approximate value for the MAP [36] of the data set in order to integrate the coefficient acquired into the set of values to be evaluated, where SP is the systolic pressure and DP is the diastolic pressure.

\[
MAP = \frac{SP + (2DP)}{3}
\]  

Data modeling

Random Forest is a model based on the combination of decision trees (DT) to reveal the relationship between inputs and outputs in order to generate a robust model for both classification and regression problems. In DT, having a set of attributes as input predictors, the objective function (discrete or continuous) can be predicted by placing a series of questions from the root node to the leaf nodes of the DT [37]. Predictor variables can be composed of numerical or categorical variables where numerical variables comprise real numbers, while categorical variables take values from a finite set. In the case of a numerical DT, the condition refers to a range. In regression trees, the fundamental idea is such that the training of the DT is the minimization of the MAE and the MSE [38].
Random Forest builds a set of decision trees, usually trained with the "bagging" method, which averages many models in order to reduce variation and involves training each decision tree on a different data sample. In Random Forest, the DT are constructed using the following instructions:

Random Forest begins by selecting randomly "k" samples from a total of "M" characteristics, where k < m. Then, the selected "k" characteristics are used to find the root node. In the next step, the secondary nodes are calculated using the same approach as in the previous step. The previous steps are performed until the tree is formed with a root node and leaf nodes. This cycle will depend on the depth established by Random Forest. Finally, the previous steps are repeated to create the "n" trees that make up the Random Forest.

During the prediction process it is necessary to test each tree created randomly using a test subset where each DT provides a different prediction (result) for the same characteristic selected from the subset. When obtaining each of the predictions of the DT that compose the Random Forest, the "votes" will be calculated, that is, in the case of the Random Forests used for classification, the number of trees that reported the same coefficient during the test stage is calculated. By identifying the largest number of "votes", Random Forest returns the coefficient obtained as the predicted target. In the case of regression Random Forests, the prediction received is the average of the "votes" given by Random Forest [37].

**Results**

The described prototype is placed on the user's wrist when it is required to acquire the information regarding the patient's blood pressure, as shown in Figure 4. The electronic device is responsible for acquiring the patient's MAP, transmitting the information obtained to the computer. Then, the transmitted information is combined with the patient's weight coefficient, which is previously captured in the computer. The information is processed through Random Forest in order to obtain the patient's diastolic and systolic pressure. Finally, the resulting information is stored in a csv file in order to provide a record of the blood pressure.

![Figure 1. Representation of the use of the electronic prototype in patients.](image)

The evaluation of the model was carried out through the MAE and MSE. These are determined by equations (2) and (3) respectively.

\[
MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \quad (2)
\]

\[
MSE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n} \quad (3)
\]

The performance of the model used is described in Table 1, obtaining an MAE and MSE of 3.215 and 43.133 in the systolic pressure identification test phase and an MAE and MSE of 1.55 and 13.467 in the diastolic pressure identification test phase respectively.
Table 1. Results of the test phase of Random Forest.

<table>
<thead>
<tr>
<th></th>
<th>Systolic Pressure</th>
<th>Diastolic Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Absolute Error</td>
<td>3.215</td>
<td>1.550</td>
</tr>
<tr>
<td>Mean Squared Error</td>
<td>43.133</td>
<td>13.467</td>
</tr>
</tbody>
</table>

4 Conclusions

Through the findings for the detection of blood pressure [33], it is possible to identify the use of different techniques for the identification of this parameter. The results reported have shown that the use of supervised learning algorithms provides acceptable performance, presenting improvements in blood pressure detection.

Based on these considerations, the combination of the MAA algorithm in the device designed together with the integration of Random Forest for blood pressure identification was proposed. It was obtaining an MAA of 3.215 and 1.55 for the detection of systolic and diastolic pressure respectively from the data set used, as well as obtaining an SSM of 43.133 and 13.467 for the detection of systolic and diastolic pressure respectively. In future work, the algorithm can be embedded in the device. Also, the device can integrate an additional communication component that allows the transmission of information over long distances in order to transmit the obtained coefficients to medical personnel autonomously.

Acknowledgements

We thank the Polytechnic University of Pachuca (UPPachuca), Kaggle web services datasets and the National Council for Science and Technology (CONACYT) for the economic support provided for this research.

References


