applied sciences

MDPI

Article

Impact and Classification of Body Stature and Physiological Variability in the Acquisition of Vital Signs Using Continuous Wave Radar

Beatriz Soares [1,2,*,†], Carolina Gouveia [3,†], Daniel Albuquerque [4,5,†] and Pedro Pinho [1,2,†]

[1] Instituto de Telecomunicações, 3810-193 Aveiro, Portugal; ptpinho@ua.pt
[2] Departamento de Eletrónica, Telecomunicações e Informática, Universidade de Aveiro, 3810-193 Aveiro, Portugal
[3] Colab Almascience, Madan Parque, 2829-516 Caparica, Portugal
[4] CISeD, Polytechnic of Viseu, 3504-510 Viseu, Portugal; dfa@ua.pt
[5] ESTGA, University of Aveiro, 3750-127 Águeda, Portugal
* Correspondence: blsoares@ua.pt
† These authors contributed equally to this work.

Abstract: The Bio-Radar system, useful for monitoring patients with infectious diseases and detecting driver drowsiness, has gained popularity in the literature. However, its efficiency across diverse populations considering physiological and body stature variations needs further exploration. This work addresses this gap by applying machine learning (ML) algorithms—Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest—to classify subjects based on gender, age, Body Mass Index (BMI), and Chest Wall Perimeter (CWP). Vital signs were collected from 92 subjects using a Continuous Wave (CW) radar operating at 5.8 GHz. The results showed that the Random Forest algorithm was the most accurate, achieving accuracies of 76.66% for gender, 71.13% for age, 72.52% for BMI, and 74.61% for CWP. This study underscores the importance of considering individual variations when using Bio-Radar, enhancing its efficiency and expanding its potential applications.

Keywords: Continuous Wave radar; dataset; machine learning; body stature variability; physiological variability

Citation: Soares, B.; Gouveia, C.; Albuquerque, D.; Pinho, P. Impact and Classification of Body Stature and Physiological Variability in the Acquisition of Vital Signs Using Continuous Wave Radar. *Appl. Sci.* **2024**, *14*, 921. https://doi.org/10.3390/app14020921

Academic Editor: Atsushi Mase

Received: 19 December 2023
Revised: 16 January 2024
Accepted: 19 January 2024
Published: 22 January 2024

1. Introduction

The Bio-Radar system, a non-contact method for monitoring vital signs, has emerged as a promising technology in the field of healthcare and biomedical research. Unlike traditional contact sensors, which require physical contact with the subject, Bio-Radar operates remotely, using radar signals to detect and monitor physiological parameters such as heart rate and respiration. This non-intrusive method offers several advantages over contact sensors. Firstly, it eliminates the discomfort and potential skin irritation caused by wearable sensors, thereby improving patient comfort and compliance. Secondly, it reduces the risk of cross-infection in clinical settings, which is particularly important in the care of patients with highly infectious diseases.

The potential applications of Bio-Radar are vast and varied. It has been used in a wide range of scenarios, from monitoring patients in hospital settings to tracking the vital signs of drivers to prevent accidents caused by fatigue [1]. In the field of sleep research, Bio-Radar has been utilized to monitor sleep patterns without disturbing the subject's natural sleep environment [2]. Additionally, in emotion recognition, it has been employed to detect physiological changes associated with different emotional states [3].

The growing interest in Bio-Radar is reflected in the increasing number of studies that utilize this technology for capturing vital signs. Table 1 presents a comprehensive review

of such studies, with a specific focus on those employing Bio-Radar in Continuous Wave (CW) mode.

CW Bio-Radar offers advantages such as the ability to provide real-time monitoring and the potential for detecting even minute physiological changes. The studies reviewed in Table 1 have explored these benefits in various contexts, ranging from sleep monitoring to emotion recognition. However, the results presented in this table are based on a very small population (at most 50 subjects) and often neglect individual characteristics.

Table 1. Several studies in the area of Bio-Radar.

Reference	Population	Gender Distinction in Result	Respiratory Rate	Heart Rate	Carrier Frequency [GHz]
[4]	1 S	-	X	X	10
[5]	U	-	X	X	2.4
[6]	12 S	-	X	X	2.4
[7]	2 M	-	X	X	24
[8]	5 M; 5 F	No	-	X	5.8
[9]	10 M	-	X	X	34
[10]	4 S	-	X	X	5.8
[11]	U	-	-	X	5.8
[12]	7 M; 4 F	No	-	X	24.17
[13]	8 M; 2 F	No	-	X	24
[14]	5 M; 1 F	No	X	X	0.915
[15]	2 M; 4 F	No	-	X	2.45
[16]	14 M; 16 F	No	X	X	24
[17]	7 M; 4 F	No	X	X	24
[18]	1 M; 1 F	Yes	X	X	10
[19]	3 M; 3 F	No	X	X	24
[20]	50 NB	-	X	X	24
[21]	14 M; 21 F	No	X	X	24
[22]	1 M; 4 F	No	X	-	2.4
[23]	6 M; 4 F	No	X	X	2.4
[24]	1 M	-	X	X	2.4
[25]	U	-	X	X	2.4
[26]	8 S	-	X	X	24

M: Male, F: Female, S: Subject, U: Undefined, -: There is no information about the topic, X: Whether the topic is studied, NB: Newborn. The study performs tests on subjects; however, the number or gender is not specified.

The impact of human body characteristics and aging on vital signals is significant and multifaceted. Generally, men have larger trunk dimensions than women [27], which may allow for more space to accommodate lungs during inhalation, potentially leading to a higher tidal volume for the same weight [28]. This could result in a greater amplitude in the acquired signal for men compared to women of the same weight. Body Mass Index (BMI) is another factor that can influence the amplitude of the respiratory signal. Studies have shown that respiratory amplitude tends to be lower in individuals with higher BMI values [29]. Additionally, aging impacts respiratory functions. With advancing age, the expansion of the rib cage on inspiration decreases, lung elastic recoil reduces, muscles lose strength, and the respiratory center becomes less sensitive. Consequently, breathing tends to weaken with age, which could be reflected in the acquired signals [30,31]. These factors underline the importance of considering individual physiological and body stature variations when analyzing signals acquired by Bio-Radar.

Considering this, the study of individual physiological and body stature variations in Bio-Radar analysis is crucial for several reasons. Firstly, these variations can affect signal quality, necessitating the adaptation of algorithms for different population characteristics to enhance the accuracy of vital signal extraction. Secondly, distinguishing features like gender and age can improve biometric identification, serving as a preliminary identification layer. Lastly, discerning parameters such as BMI and CWP can provide valuable health and fitness insights, aiding in preventative healthcare and fitness monitoring. In conclusion,

considering individual physiological and body stature variations is essential for the accuracy of Bio-Radar in monitoring vital signals and opens up a range of potential applications in biometrics, fitness, and health.

In response to this need to characterize signals according to the subject's characteristics, and noting that none of the studies presented in Table 1 differentiate between physical characteristics, an innovative study was developed. Accordingly, machine learning (ML) algorithms, including Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest, were utilized to distinguish between different genders, age ranges, Body Mass Index (BMI) categories, and Chest Wall Perimeter (CWP) measurements.

The paper is structured as follows: Section 2 provides a detailed description of the system setup and respiratory signal processing techniques utilized. Section 3 discusses the data records, the features selected for analysis, and the statistical methods employed to identify the most informative features. The results of our ML algorithms are presented in Section 5. Finally, Section 6 concludes the paper and discusses potential future work.

2. Bio-Radar Prototype Description

The Bio-Radar system employed in this research utilizes a Continuous Wave (CW) Doppler radar. As the terminology suggests, this type of radar continuously transmits a sinusoidal carrier wave, which is digitally generated, and captures the signal reflected by the target. Due to the micro-Doppler effect, a phase shift occurs when the subject's chest wall moves towards or away from the radar, resulting in phase modulation in the received signal [32]. The Bio-Radar prototype used in this study, the operating principles of which are illustrated in Figure 1, was developed in a previous work [33].

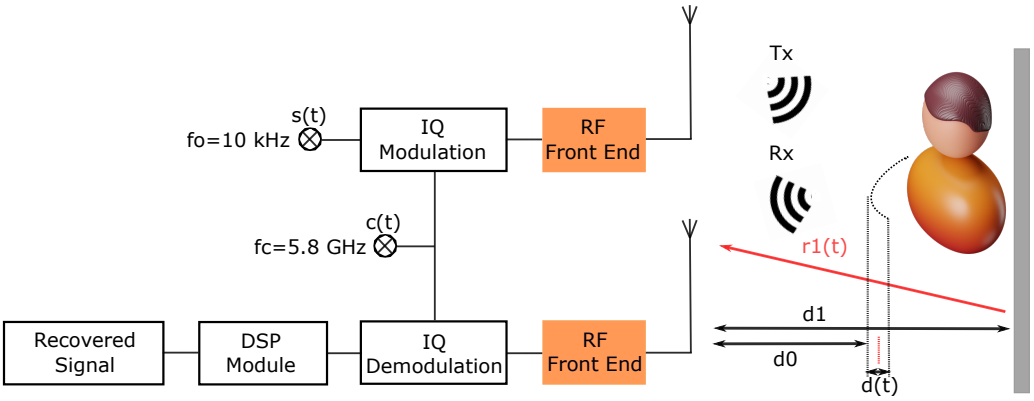


Figure 1. Block diagram of the Bio-Radar system, adapted from [33].

The primary hardware component of the prototype is the Universal Software Radio Peripheral (USRP) B200 [34], which operates within a Continuous Wave (CW) frequency range of 70 MHz–6 GHz. This platform enables the digital programming of Bio-Radar characteristics, such as the power of the transmitted wave, set at 2 dBm, and its sampling frequency. A carrier frequency of 5.8 GHz was selected, falling within the Industrial, Scientific, and Medical (ISM) dedicated band. The system also features two 2×2 patch array antennas, with one designated for Transmission (Tx) and the other for Reception (Rx). The specific operational parameters of the Bio-Radar system utilized in this study are detailed in Table 2, providing a comprehensive overview of the system's technical specifications.

Table 2. Technical parameters of the Bio-Radar system.

Parameter	Value
Operating Frequency	5.8 GHz
Transmit Power	2 dBm
Sampling Rate	100 kHz
Frequncy Offset	10 kHz
Antenna Type	2 × 2 patch array

Signal Processing Algorithm

The Digital Signal Processing (DSP) Module runs an algorithm designed for the rigorous extraction of vital signals information. Figure 2 represents the Bio-Radar system, more specifically the implemented DSP algorithm implemented in the receptor chain [35].

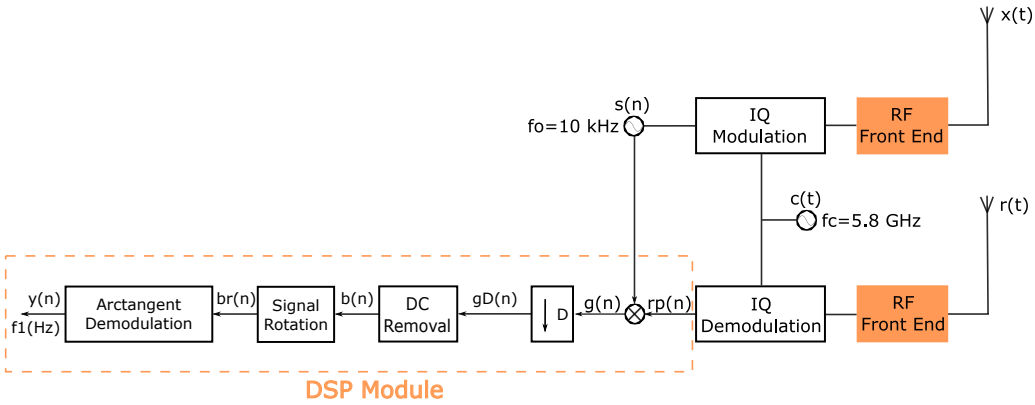


Figure 2. Block diagram of the Bio-Radar system including DSP module, adapted from [35].

The signal received by our system is initially processed through a step called downconversion, which converts it to a baseband signal. This results in a complex signal, referred to as $g(n)$. Following this, the signal undergoes a process known as decimation, which reduces its original sampling rate from 100 kHz to 100 Hz. This reduction produces a new signal, denoted as $g_D(n)$. As illustrated in Figure 3a, the complex baseband signal $g_D(n)$ is represented in the complex plane, typically visualized as an arc. It is crucial to understand that the transmitted signal, which our system sends out, is reflected back not just by the target object, such as the subject's chest wall, but also by other objects in the vicinity.

These other objects, assuming they are stationary, contribute to the signal in a specific way. Their reflections are added to the received signal as DC offsets, which are constant components that shift the center of the arc in the complex plane. This means that the overall signal received and visualized includes not just the vital information from the subject's chest wall but also these additional static reflections from the surrounding environment. Understanding and accounting for these additional components is important for accurately interpreting the data from our Bio-Radar system. Assuming these objects remain stationary, their reflections are perceived as DC offsets added to the center of the arc. Afterwards, the DC offsets are estimated and subsequently removed from the decimated signal, resulting in signal $b(n)$.

Figure 3 serves as a visual guide to the DC offset removal step, which is crucial for accurate signal analysis. The process begins with the received radar signal, which often includes an unwanted DC component due to signal propagation effects. Using the Optimized Cost (OC) method, proposed in a previous study [35], it was calculated the median of the signal (Figure 3a) and establish a fictitious boundary to contain the cost function optimization (Figure 3b). The lowest cost point within this boundary indicates the precise DC offset to be removed. Once this offset is subtracted, the signal median should theoretically align with the origin (Figure 3c). Finally, it is demonstrated the signal

post-correction with an example of a rotated signal for illustrative purposes (Figure 3d). This process ensures that the signal is appropriately centered and ready for subsequent stages of analysis, such as feature extraction and pattern recognition.

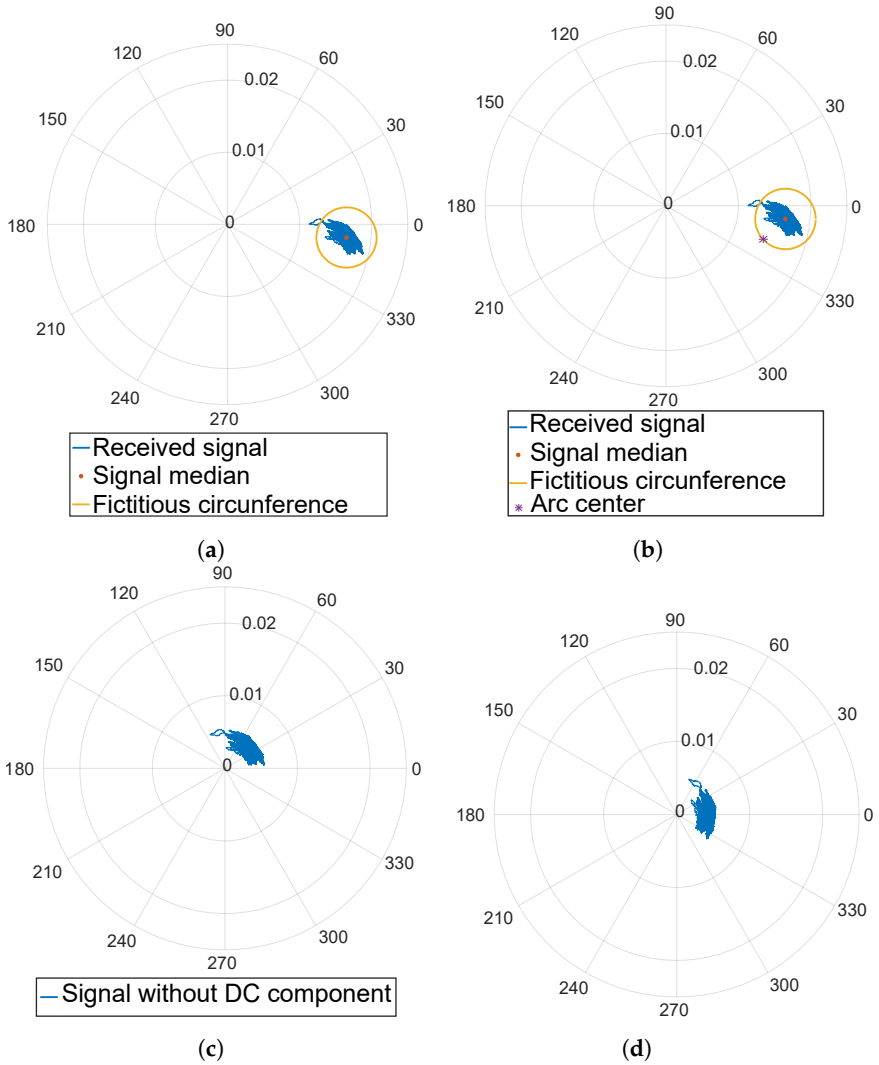


Figure 3. Sequential illustration of DC offset removal process in radar signal using the Optimized Cost Method. (**a**) Original radar signal with DC offset shown by 'Received signal' and the guide for optimization 'Fictitious circumference' with 'Signal median' as its center. (**b**) Identification of 'Lowest cost point' indicating the optimal DC offset values, with 'Arc center' marking the new median post-correction. (**c**) Signal post-DC offset removal, where the median aligns with the origin. (**d**) The signal after offset removal and a rotation to align for further analysis.

In order to avoid wraps caused by the possibility of the signal crossing the 180° value, it is necessary to rotate the signal to oscillate around the 0° angle, resulting in $br(n)$, as shown in Figure 3d. This process not only avoids wraps but also simplifies the signal subsequent analysis. Lastly, to obtain the received vital signs in the time domain, $y(n)$, the arctangent demodulation is applied, resulting in the signal of Figure 4.

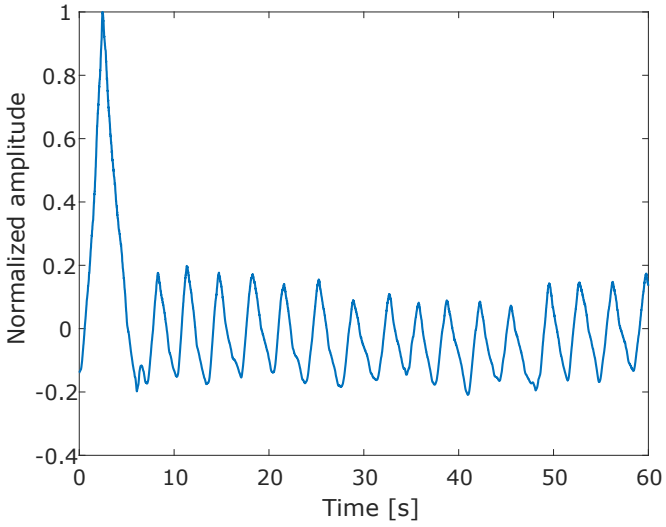
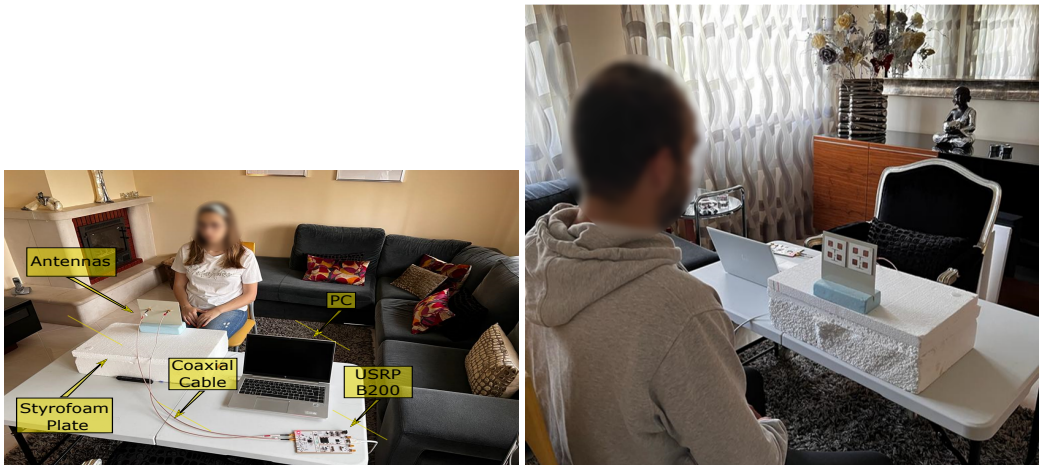


Figure 4. Example of a demodulated signal.

3. Data Records and Procedures

Given the absence of a public dataset encompassing a broad range of subjects' stature variability, it was necessary for us to gather a substantial dataset of vital signs. This study received approval from the Ethics and Deontology Committee of the University of Aveiro, Portugal (No. 29-CED/2021), and was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants. The data collection took place in a single unaltered environment: an 18 square meter room. The setup involved the Bio-Radar system described earlier, positioned on a plastic table. Styrofoam plates were utilized to adjust the height of the antennas to match the stature of the subject being tested. Figure 5 provides an example of the setup with two volunteers—a woman (Figure 5a) and a man (Figure 5b). The data collection involved 92 subjects (46 males and 46 females) aged between 18 and 50 years old with Body Mass Index (BMI) values ranging from 17.5 to 40 kg/m^2 and Chest Wall Perimeter (CWP) measurements between 77 and 129.5 cm.



(**a**) Woman participant. (**b**) Man participant.

Figure 5. Overview of the measurement setup and the system configuration.

4. Features Selection and Analysis

To explore the possibility of classifying subjects based on gender, age, BMI, and CWP, we extracted a comprehensive set of features from the vital signs data. We then conducted both visual and statistical analyses of these features to identify those capable of distinguishing between different class groups. Finally, we utilized the selected group of

relevant features in the machine learning algorithms. Initially, we determined a broad set of 40 features within a 60 s window with a 50% overlap (30 s), as depicted in Figure 6.

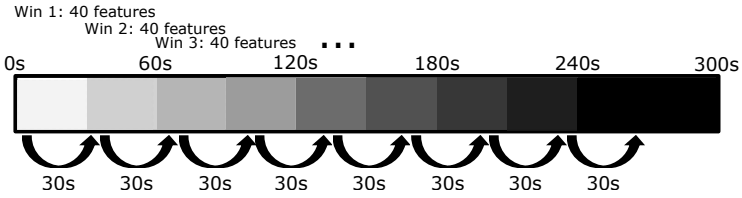


Figure 6. Sliding window of the feature extraction.

4.1. Features Description

The Bio-Radar signal features were selected taking into account the ones typically used in the literature [21,26,36–40], resulting in the following list:

(1) **L:** Amplitude of the respiratory signal;
(2) **med ibi:** Interbreath Interval (IBI) median;
(3) **mean d1 signal resp:** Mean of the respiratory signal first derivative;
(4) **mean Nd1 signal resp:** Mean of the normalized respiratory signal first derivative;
(5) **mean d2 signal resp:** Mean of the respiratory signal second derivative;
(6) **mean Nd2 signal resp:** Mean of the normalized respiratory signal second derivative;
(7) **skewness T inhale:** Skewness of inhalation time;
(8) **med T inhale:** Median of inhalation time;
(9) **iqr T inhale:** Inter-Quartile Range (IQR) of inhalation time;
(10) **mean T inhale:** Mean of inhalation time;
(11) **skewness T exhale:** Skewness of exhalation time;
(12) **med T exhale:** Median of exhalation time;
(13) **iqr T exhale:** IQR of exhalation time;
(14) **mean T inhale:** Mean of exhalation time;
(15) **RR:** Respiratory Rate;
(16) **p01–p06:** Average Power Spectral Density (PSD) energy in the band 0–0.1 Hz;
(17) **p02:** PSD energy in the band 0.1–0.2 Hz;
(18) **p03:** PSD energy in the band 0.2–0.3 Hz;
(19) **p04:** PSD energy in the band 0.3–0.4 Hz;
(20) **p05:** PSD energy in the band 0.4–0.9 Hz;
(21) **p02:** PSD energy in the band 0.9–1.5 Hz;
(22) **med sign:** Signal median;
(23) **mean sign:** Signal mean;
(24) **iqr sign:** Signal IQR;
(25) **iqr ibi:** Interbreath Interval Inter-Quartile Range;
(26) **mean ibi:** IBI mean;
(27) **skewness ibi:** Skewness of the IBI;
(28) **AppEn max:** Signal maximum approximate entropy;
(29) **kurtosis:** Respiratory signal kurtosis;
(30) **SDNN Resp:** Standard Deviation of IBI (SDNN) of the respiratory signal;
(31) **SAM:** Number of samples above the mean value of the waveform;
(32) **RMSSD Resp:** Root Mean Square of Successive Differences (RMSSD) between IBI of the respiratory signal;
(33) **skewness sign:** Skewness of the signal;
(34) **RRV:** Respiratory Rate Variability (RRV);
(35) **RMS:** Quadratic mean value of the signal;
(36) **freq resp:** Respiratory frequency;
(37) **mean peak valley width:** Average of the peaks and valleys width of the respiratory signal in the time domain;

(38) **Ratio PSD:** PSD of the ratio between the PSD of the Low Frequency (0.1–0.4 Hz) and High Frequency (0.4–1.5 Hz);
(39) **var sign:** Variance of the respiratory signal;
(40) **AppEn resp:** Approximate entropy of the respiratory signal.

More details regarding the feature extraction are now provided: Feature 1 is determined through the calculation of arc length. Features 2, 25–27, 30, and 34 are derived from the Interbreath Interval. The Interbreath Interval (IBI) refers to the time interval between individual breaths. The Inter-Quartile Range (IQR) is a statistical measure that represents the spread of the middle 50% of the IBI data, specifically the difference between the 75th and 25th percentiles. This metric is important in our analysis as it provides insight into the variability and regularity of the breathing pattern. The features related to inspiration and expiration, numbers 7–14, are statistical measures extracted from the peaks and troughs of the respiratory signal. The determination of RR involves calculating the ratio between the number of signal peaks and the time period in which they occur, measured in minutes. This calculation yields the RR in Breaths Per Minute (BrPM). Features 16–21 were computed using the Welch method. The Welch method is a spectral estimation technique that divides the signal into overlapping segments, computes a periodogram for each, and averages them to yield a more accurate and less noisy estimation of the PSD (Figure 7). This method was applied to the signal after removing its average. Similarly to the latter ones, features 3–6, 17–19, 22–24, 28–33, 37, 39, and 40 are also statistical measures concerning the entire respiratory signal. Feature 34 is determined in [40], using (1):

$$RRV = \left(100 - \frac{H1}{DC}\right)\%. \quad (1)$$

where $H1$ and DC are the amplitude powers of the first and the zeroth harmonic peaks, respectively.

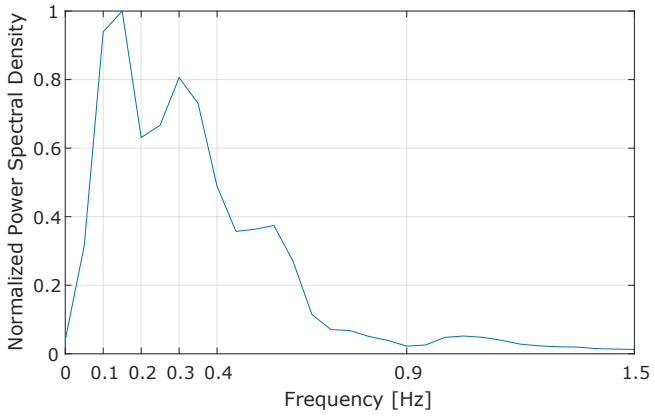


Figure 7. Normalized PSD curve of respiratory signals for a subject.

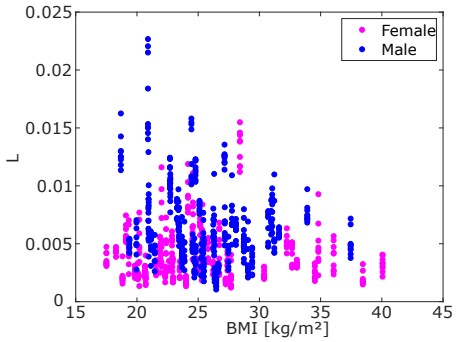
4.2. Features Representation of Physiological and Physical Structure Differences

After extracting an extensive set of 40 features, a preliminary study was carried out to determine whether certain features, such as Respiratory Rate (RR) and Length (L) (features 15 and 1, respectively), contain information that can distinguish between different populations. The literature suggests that differences in respiration between genders can be particularly pronounced in RR and L values. To investigate this, scatter plots were created (shown in Figure 8) for data visualization, and a basic statistical analysis was conducted, the results of which are presented in Table 3. According to this table, men have, on average, a lower RR than women. Conversely, men tend to have a higher L value than women.

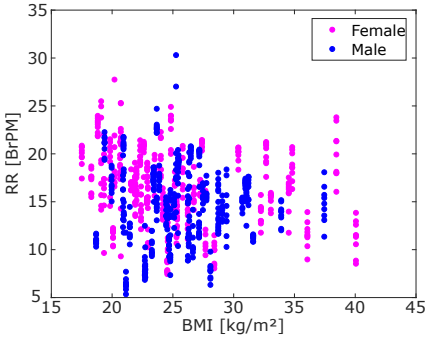
Table 3. Statistical study of the RR and L representation according to gender.

Gender	RR Mean ± STD [BrPM]	RR Median [BrPm]	L Mean ± STD	L Median
Male	14.6252 ± 4.0037	14.7602	0.0064 ± 0.0034	0.0056
Female	16.5741 ± 3.8281	16.6840	0.0040 ± 0.0023	0.0034

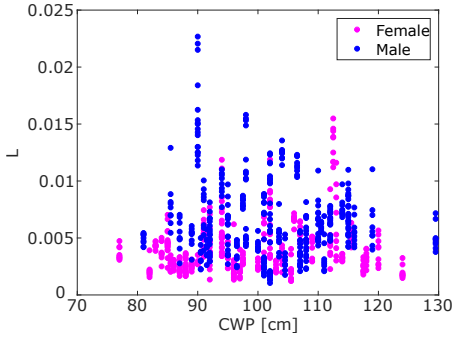
The RR and L features also exhibit distinct distributions in relation to other physical characteristics, such as BMI and CWP. These variations are illustrated in Figure 8a,b for BMI, and Figure 8c,d for CWP. It can be inferred that individuals with higher BMI values tend to have lower L values, as shown in Figure 8a, and also the lowest RR values, as depicted in Figure 8b. Upon analyzing Figure 8d, it is evident that individuals with a smaller CWP have a higher RR. Conversely, the highest L values are associated with individuals having a medium CWP, as demonstrated in Figure 8c. When correlating L with RR, a pattern emerges where higher L values correspond to lower RR values, which is corroborated by the data presented in Figure 8e.



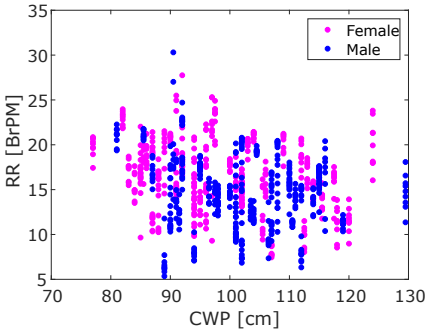
(**a**) Variation of L with BMI.



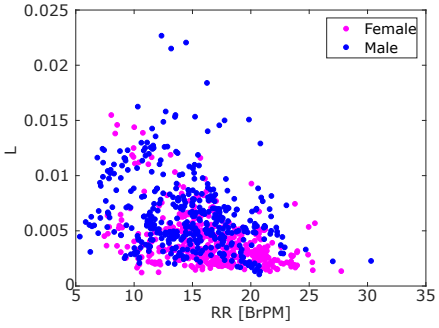
(**b**) Variation of RR with BMI.



(**c**) Variation of L with CWP.



(**d**) Variation of RR with CWP.



(**e**) Variation of L with RR.

Figure 8. Overview of the statistical results.

4.3. Feature Selection

A statistical analysis was conducted on the extracted features to select the most informative ones for use in the machine learning algorithms. This analysis, based on the methodology described in [41], is summarized in Figure 9. As an initial step, a one-way Analysis of Variance (ANOVA) test was applied to determine whether its residuals exhibit a normal distribution.

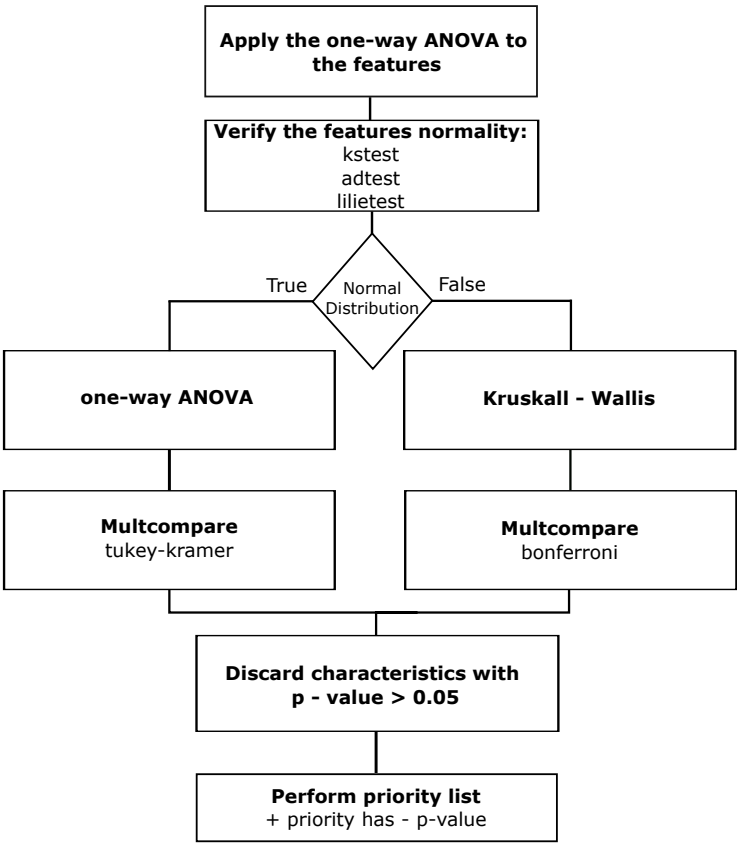


Figure 9. Statistical analysis diagram.

To validate the normal distribution of residuals, we utilized the one-sample Kolmogorov–Smirnov, Anderson–Darling, and Lilliefors tests. The one-sample Kolmogorov–Smirnov test is used to compare the empirical distribution of our sample with a specified theoretical distribution by evaluating the maximum deviation between these distributions. The Anderson–Darling test, while similar in purpose to the Kolmogorov–Smirnov test, places greater emphasis on the tails of the distribution, enhancing sensitivity to tail differences between the empirical and theoretical distributions. Lastly, the Lilliefors test, a variant of the Kolmogorov–Smirnov test, specifically tests for normality by accommodating the estimation of mean and variance from the sample data, unlike the standard Kolmogorov–Smirnov test which assumes these parameters are known. The results indicated that while some features conform to a normal distribution, others do not. Consequently, to determine the *p*-values for each feature, we applied two different tests [42]. These are the Kruskal–Wallis test, which is appropriate when the ANOVA residuals do not exhibit normal distribution, and the ANOVA test, which is suitable when they do [43]. Features with p-values exceeding the threshold of 0.05 were discarded. Subsequently, we constructed a list prioritizing features; those with the lowest *p*-values were given the highest precedence.

This statistical study was conducted for four distinct binary classification tests, each designed to examine the potential for distinguishing subjects based on specific characteristics: gender, age, BMI, and CWP. The objective of these tests is to facilitate binary classification according to predefined criteria for each characteristic.

- **Gender:**
  - Male: 46 people;
  - Female: 46 people.
- **Age:**
  - 18–29 years old: 46 people;
  - 30–50 years old: 46 people.
- **BMI:**
  - 18–25 kg/m^2: 47 people;
  - $\geq$25 kg/m^2: 45 people.
- **CWP:**
  - $\leq$100 cm: 46 people;
  - >100 cm: 46 people.

After establishing a priority queue for each binary classification problem, a correlation matrix was constructed to further refine the feature selection [44]. This step helps identify features that contain redundant information, indicated by high correlation values. A feature was deemed highly correlated with others if the absolute value of the correlation exceeded 0.7. If the matrix indicates a correlation value equal to or higher than this threshold, the features in question are evaluated against the priority list. Only the feature with the highest priority is retained for subsequent analysis. Figure 10 graphically illustrates this process for the gender classification test. Figure 10 is a visual representation of the correlation matrix among features where a ball's size and color denote the correlation value, with larger balls (with a diameter equal to the square) indicating a correlation of 1 (red) or −1 (blue). Contrariwise, smaller dot-represented balls represent a correlation near 0. This approach simplifies the analysis of correlations across a larger set of features, providing a more accessible interpretation of the correlation matrix.

The features resulting from this process are those selected for use in the machine learning algorithms to classify the defined binary tests. The resulting features vary for each binary classification test, as presented in Table 4.

Table 4. Priority lists of classification binary test.

Gender	Age	BMI	CWP
L	iqr sign	p05	RR
med ibi	kurtosis	RR	p02
mean d1 signal resp	med T inhale	AppEn resp	p05
RR	L	mean d1 signal resp	p01
med T inhale	SDNN Resp	skewness ibi	skewness sign
p02	AppEn max	med ibi	p04
p05	mean d2 signal resp	RMSSD Resp	med sign
p04	med ibi	med sign	L
p06	p04	mean T inhale	freq resp
p01	SAM	skewness sign	med ibi
med sign	p05	iqr sign	iqr sign
iqr sign			RRV
AppEn max			SDNN Resp
kurtosis			
SDNN Resp			

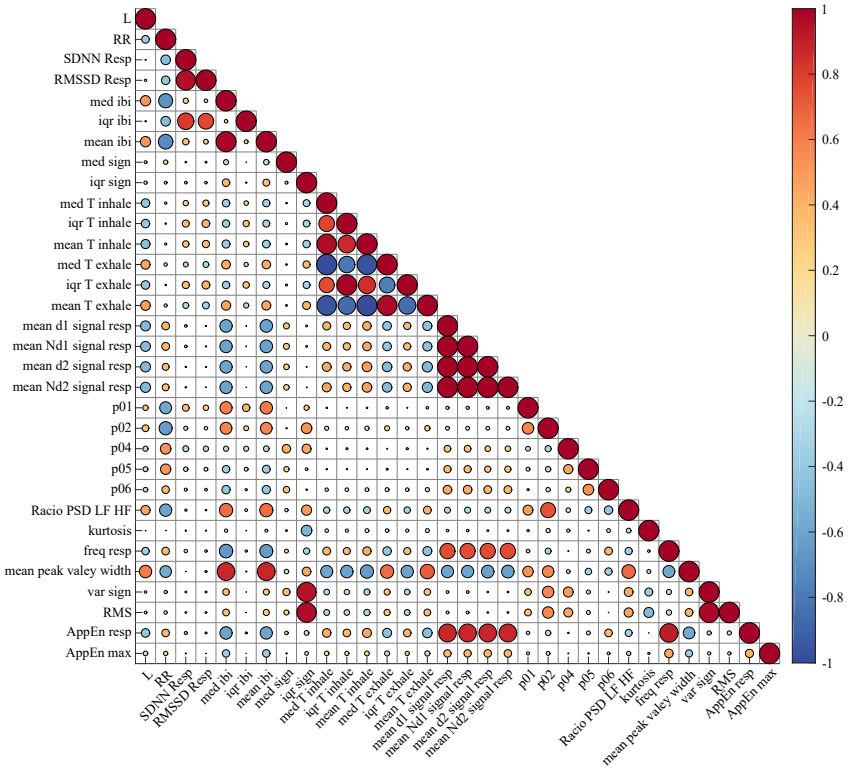
4.3.1. Gender

In relation to the features that effectively categorize the gender test, it is evident that RR and L are among them, aligning with expectations established in Section 4.2, where it was noted that women tend to have a higher average RR, while men have a higher

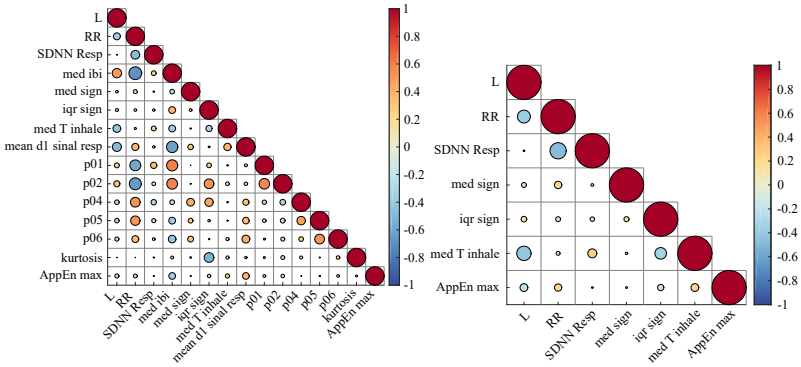
average L. Besides these features, the signal's PSD across various frequency bands also demonstrated significant relevance.

4.3.2. Age

In this binary classification test, the categorization was not characterized by the RR value or any other measure of periodicity. Aging leads to a decreased capacity for thoracic expansion [30], suggesting that the amplitude of the respiratory signal is a significant feature for age categorization. However, since the age range of the participants is relatively narrow and below 50 years, these characteristics may not be as pronounced in this age group. The classification of subjects into different age groups appears to be influenced by features such as respiratory amplitude and the impact of inspiration. This relationship may be attributed to the effects of conditions like osteoporosis, which reduce the thoracic cage's ability to expand during inspiration and compromise diaphragm contraction effectiveness [30].



(**a**) Initial correlation matrix.



(**b**) Using 0.7 of correlation.

(**c**) Using 0.5 of correlation.

Figure 10. Correlation matrices associated for the gender test.

4.3.3. Body Mass Index

In studies reported by [45,46], it was found that a higher percentage of abdominal fat is associated with lower lung volume, which in turn leads to a higher RR. This aligns with our findings, as RR was indeed selected as one of the most important features in this study. RMSSD Resp may also serve as an indicator of this fact. Another consequence of excess fat in the diaphragm region is its potential to hinder chest wall expansion, which could lead to a decrease in respiratory signal amplitude. However, contrary to expectations, this feature did not prove to be representative in our study, as it was not selected for this particular case. Consequently, the inhalation time may increase, a phenomenon that could be indicated by the feature *mean T inhale*.

4.3.4. Chest Wall Perimeter

After reviewing the results presented in Section 4.2, it was anticipated that CWP classification would be influenced by both RR and L, a hypothesis that is supported by the statistical analysis conducted in this section. These results may correlate with the observation that individuals with a higher CWP can accommodate a larger volume of air upon inspiration, thereby reducing the necessity for frequent breaths. Additionally, as shown in Table 4, it is evident that CWP classification may also be influenced by the RRV.

5. Machine Learning Implementation

After extracting signal features, machine learning (ML) algorithms were applied to classify the collected data into gender, as well as into age, BMI, and CWP ranges. For our analysis, we chose to employ Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest algorithms. The decision to utilize these traditional ML algorithms instead of deep learning models was based on several considerations.

Firstly, deep learning models typically require large amounts of data to perform optimally and avoid overfitting [47]. Given the size and complexity of our dataset, SVM, KNN, and Random Forest are more suitable as they can yield robust results with smaller datasets [48]. Secondly, these algorithms offer greater interpretability compared to deep learning models. They allow for an understanding of the importance of different features and their contribution to the final prediction. This aspect is crucial in our study, which aims to understand the impact of physiological and body stature variations on signals acquired by the Bio-Radar [49]. Thirdly, SVM, KNN, and Random Forest are less resource intensive and can be trained more quickly than deep learning models. This is a significant advantage when working with limited computational resources [50]. Lastly, in many scenarios, traditional machine learning algorithms such as SVM, KNN, and Random Forest can achieve performance comparable to those of deep learning algorithms, particularly when the data are not excessively complex or large in volume [51]. Therefore, given the nature of our data and the specific objectives of this research, SVM, KNN, and Random Forest emerged as the most appropriate algorithms for this study.

SVM is a non-parametric classification technique based on statistical learning theory [52]. The main goal of SVM is to find the hyperplane separating the data such that the distance from this hyperplane to each data class is as large as possible. The KNN classifies a new observation by measuring the distance of this observation from its nearest points and then categorizing it in the class to which most of the K points belong. A higher number of K neighbors will imply a smoother classification boundary. On the other hand, a small number of K produces a more flexible boundary with a high variance but low bias [53]. The last classifier, Random Forest, is used in regression and classification problems. A Random Forest consists of a combination of tree predictors. Each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them [54]. The final result of this model is the result returned from the trees with the highest vote, or the average of these.

The data resulting from Section 3 are split into training, test, and new subject categories. This dataset is composed of 92 individuals that have nine 1-minute signal windows. Two scenarios regarding this division were examined: in the first, two individuals were designated as new data, and in the second, five individuals were designated as new data. In both scenarios, the remaining participants were used for training and testing, following a 70:30 hold-out strategy. This selection process is illustrated in Figure 11. The same procedure was repeated over twenty iterations. Figure 12 depicts the implemented process. In this study, four different combinations of training–test models were conducted to identify the conditions that would yield the best results:

- Correlation value of 0.7 with 2 people as new data;
- Correlation value of 0.7 with 5 people as new data;

This division was strategically chosen for several reasons, which we elucidate here to provide clarity on our methodology.

Firstly, the division into two cases allowed us to assess the model's adaptability and performance with varying degrees of data variability. Analyzing the model's behavior with a smaller subset (two subjects) and then a larger one (five subjects) enabled us to observe how it responds to an incremental increase in dataset complexity. This approach is particularly relevant for evaluating the scalability and robustness of our model in real-world applications, where the number of subjects and the variability in data can significantly influence performance. Secondly, the choice of these specific subset sizes was partly influenced by the limitations in the available dataset. While ensuring meaningful analysis, we also had to consider the practical constraints of our data pool. The subsets of two and five subjects represent different levels of variability and complexity within these constraints, providing a balanced perspective on the model's generalizability and effectiveness. Furthermore, the incremental analysis, starting with two subjects and extending to five, offered insights into the model's ability to maintain accuracy and reliability as the diversity of the data increased. It was crucial for our study to understand not only how well the model performs with a limited and controlled dataset but also how it scales up when exposed to a broader range of subject data.

The difference in the number of subjects between the two cases might seem modest. However, our results indicate that even this modest increase in the number of subjects provided valuable information about the model's performance under varying conditions. This incremental approach helped us to better understand the nuances of our model's behavior in response to changes in the dataset size and composition. In conclusion, the division of the new dataset into these two specific cases was a deliberate methodological decision aimed at thoroughly evaluating the model's adaptability, scalability, and robustness in handling Bio-radar signal data with varying characteristics.

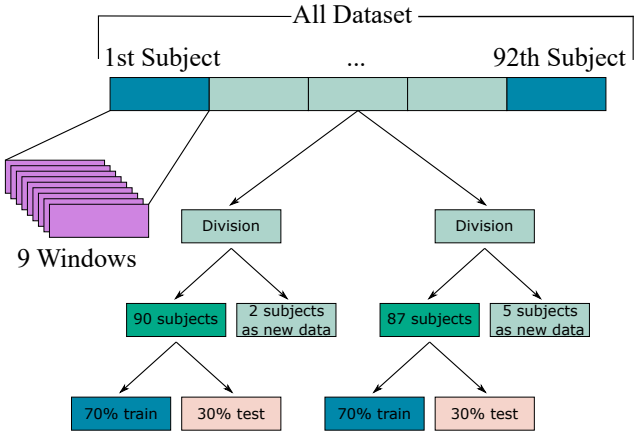


Figure 11. Data splitting into train, test, and new data.

The selection of hyperparameters was carried out for each binary classification test in multiple phases. Initially, hyperparameters were obtained using the default range provided by using functions from MATLAB's Statistics and Machine Learning Toolbox, namely fitsvm, fitknn, and treebagger, through a loop of twenty iterations. The results from this first step were recorded, and the range of hyperparameters was subsequently narrowed. This process was repeated until a single value was reached, indicating that the final models' hyperparameters were those with the highest vote count. The selected hyperparameters for the three machine learning models are presented in Table 5.

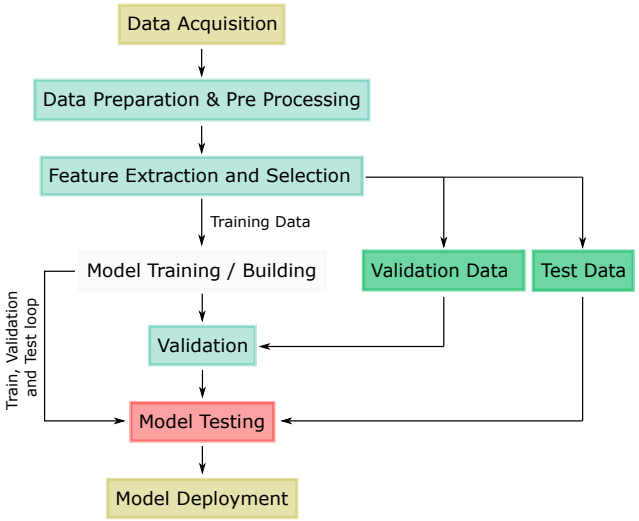


Figure 12. Classification workflow, modified from [55].

Table 5. Selected hyperparameters for each binary test.

ML Algorithm	Hyperparameters	Classification Binary Test			
		Gender	**Age**	**BMI**	**CWP**
SVM	KernelFunction	gaussian	gaussian	gaussian	gaussian
	Standardize	1	1	1	1
	BoxConstraint	12	7	5	19
	KernelScale	1.7	1.3	2.7	1.3
KNN	NumNeighbors	5	4	10	4
	Standardize	1	1	1	1
	Distance	minkowski	minkowski	minkowski	minkowski
	Exponent	-	0.5	0.6165	-
	DistanceWeight	squaredinverse	squaredinverse	squaredinverse	squaredinverse
Random Forest	Method	regression	regression	regression	regression
	OOBPredictorImportance	On	On	On	On
	MinLeafSize	5	5	5	5

Tables 6–9 display the accuracy results for each binary classification test. Each table presents the mean accuracy value and standard deviation (STD) across three testing scenarios: cross-validation (CV), testing using a 30% hold-out strategy, and testing with new data. For the CV process, we employed the K-Fold method with K set to 5. This method involves randomly dividing the dataset into K subsets, each containing approximately the same number of observations. Additionally, the tables include results for two distinct training–test model combinations, labeled as (a) and (b).

5.1. Results

5.1.1. Gender

The results for the gender binary test are in Table 6. Analyzing these results, it is possible to conclude that the results of the Random Forest are better than those of the remaining algorithms.

Comparing the results of the (a) and (b) tests, generally, the (a) test results are better than the (b) test results. Therefore, in the gender classification, an accuracy of 83% was obtained in the cross-validation (CV) method. In the test data, an accuracy of 76.7% was achieved, and an accuracy of 63.3% was achieved in the new data.

Table 6. ML results for the gender test.

Algorithm	Parameters	(a) Acc ± STD [%]	(b) Acc ± STD [%]
SVM	CV	76.0 ± 1.8	76.0 ± 1.7
	Test	**77.9 ± 2.5**	77.9 ± 3.3
	New Data	62.8 ± 20.2	63.1 ± 12.5
KNN	CV	74.2 ± 1.3	74.9 ± 2.1
	Test	**76.0 ± 2.9**	74.9 ± 2.6
	New Data	64.2 ± 26.2	60.8 ± 15.4
Random Forest	CV	83.0 ± 0.4	83.2 ± 0.7
	Test	**76.7 ± 3.0**	76.4 ± 2.7
	New Data	63.3 ± 30.8	61.2 ± 18.4

Acc—Accuracy; CV—Cross-Validation; STD—Standard Deviation. (a)—Correlation = 0.7 and new subjects = 2; (b)—Correlation = 0.7, new subjects = 5.

5.1.2. Age

Upon analyzing the ML results, the Random Forest algorithm once again emerged as the best performer. In comparing the test results, the most favorable outcomes were observed in test (b). This test achieved an accuracy of 80.2% in cross-validation (CV), 71.1% accuracy in the test data, and 55.8% accuracy with new data.

Table 7. ML results for the age test.

Algorithm	Parameters	(a) Acc ± STD [%]	(b) Acc ± STD [%]
SVM	CV	67.3 ± 1.6	67.7 ± 1.7
	Test	**68.5 ± 2.0**	68.1 ± 2.9
	New Data	54.7 ± 19.9	52.1 ± 10.8
KNN	CV	72.7 ± 2.2	73.8 ± 1.7
	Test	**73.5 ± 2.0**	73.4 ± 2.6
	New Data	47.2 ± 21.2	48.1 ± 12.3
Random Forest	CV	80.2 ± 0.5	80.2 ± 0.7
	Test	70.6 ± 3.2	**71.1 ± 3.0**
	New Data	59.4 ± 23.8	55.8 ± 12.7

Acc—Accuracy; CV—Cross-Validation; STD—Standard Deviation. (a)—Correlation = 0.7 and new subjects = 2; (b)—Correlation = 0.7, new subjects = 5.

5.1.3. Body Mass Index

Consistent with the first two tests, the Random Forest algorithm again demonstrated the best results in the BMI classification test. Similar to previous patterns, test (b) stood out with the highest performance. Consequently, for the BMI categorization, the achieved accuracies were 80.6% in cross-validation (CV), 72.5% on the test data, and 56.7% on new data.

Table 8. ML results for the BMI test.

Algorithm	Parameters	(a) Acc ± STD [%]	(b) Acc ± STD [%]
SVM	CV	71.3 ± 2.2	72.7 ± 2.2
	Test	**71.4 ± 2.9**	71.0 ± 2.4
	New Data	56.7 ± 23.7	55.2 ± 14.5
KNN	CV	71.3 ± 1.5	71.3 ± 1.5
	Test	72.3 ± 2.8	**72.5 ± 2.2**
	New Data	55.0 ± 21.7	51.7 ± 12.5
Random Forest	CV	80.8 ± 0.8	80.6 ± 0.5
	Test	70.6 ± 3.1	**72.5 ± 1.4**
	New Data	59.2 ± 23.0	56.7 ± 11.3

Acc—Accuracy; CV—Cross-Validation; STD—Standard Deviation. (a)—Correlation = 0.7 and new subjects = 2; (b)—Correlation = 0.7, new subjects = 5.

5.1.4. Chest Wall Perimeter

Mirroring the pattern observed in the other tests, the Random Forest algorithm again delivers the best results. In this algorithm, test (a) distinguishes itself from the rest, achieving the following accuracy values: 82.1% in cross-validation (CV), 74.6% on the test data, and 66.1% on new data.

Table 9. ML results for the CWP test.

Algorithm	Parameters	(a) Acc ± STD [%]	(b) Acc ± STD [%]
SVM	CV	76.4 ± 2.0	75.3 ± 2.2
	Test	76.4 ± 2.3	**76.8 ± 2.2**
	New Data	59.2 ± 18.9	57.1 ± 16.3
KNN	CV	75.3 ± 1.5	75.4 ± 1.8
	Test	**75.8 ± 3.1**	75.6 ± 2.8
	New Data	51.7 ± 17.9	54.0 ± 11.1
Random Forest	CV	82.1 ± 0.5	82.2 ± 0.5
	Test	**74.6 ± 2.2**	73.3 ± 2.7
	New Data	66.1 ± 25.4	61.6 ± 15.8

Acc—Accuracy; CV—Cross-Validation; STD—Standard Deviation. (a)—Correlation = 0.7 and new subjects = 2; (b)—Correlation = 0.7, new subjects = 5.

5.2. Results Discussion

In the same way as the other tests, the Random Forest algorithm achieves the best results. In the Random Forest algorithm, test (a) stands out from the remaining tests. This test achieves the following accuracy values: 82.1% on the CV, 74.6% on the test data, and 66.1% on the new data. The results indicate that test (a) performed better in characterizing gender and Chest Wall Perimeter (CWP), while test (b) excelled in age and Body Mass Index (BMI) characterization. A parallel study examined the impact of reducing correlation to 0.5, but this did not enhance the results.

When applied to new data, the results were lower than those obtained in cross-validation and with test subjects, which is expected as these subjects were previously unknown to the algorithm. CWP was the easiest to classify, achieving the highest accuracy of 66.11% with new subjects. This could be due to the Bio-Radar system's direct measurement of chest wall displacement in individuals with higher CWP. Conversely, age was the hardest to classify, with the lowest accuracy of 55.78% for new subjects. This could be because certain characteristics become more pronounced with age.

The respiratory signal features from both tests suggest the possibility of distinguishing between different classification groups in all binary tests. The accuracy results from cross-validation and the hold-out test are mostly similar, indicating the model's robustness.

Adjusting the number of new subjects showed that better results correspond to a larger number of people for testing and training. However, the values achieved with test (b) did not significantly differ from those obtained with test (a), consistently remaining above 50% in all tests. These findings indicate that the dataset of 90 individuals may not be fully representative of the overall population. Nonetheless, the results achieved with the test population, reaching around 70%, suggest the possibility of successfully distinguishing the classes within this specific population.

6. Conclusions

In conclusion, this study presents preliminary results demonstrating the potential of using machine learning algorithms in analyzing physiological and body stature variations in signals acquired by Bio-Radar. With the application of three ML algorithms, the following accuracy values were achieved in test data: 76.66% for gender, 71.13% for age, 72.52% for BMI, and 74.61% for CWP. CWP is the easiest test to classify, followed by gender, then BMI, and finally, age. While the results may not yet reach high levels of accuracy, they provide valuable insights and a promising foundation for future research.

The application of these algorithms has enabled us to distinguish between different genders, age ranges, BMI categories, and CWP measurements. These findings suggest that physiological and body stature variations can indeed be detected by the Bio-Radar, which is a significant step forward in the field of non-contact vital sign monitoring.

For future work, the models will be refined and more data gathered. It is anticipated that the accuracy of predictions will improve. These preliminary results are encouraging and provide a solid foundation for further exploration and development in this area.

Author Contributions: Conceptualization, D.A. and B.S.; methodology, D.A. and B.S.; software, D.A. and B.S.; validation, B.S., D.A., C.G. and P.P.; formal analysis, B.S.; investigation, B.S., D.A., C.G. and P.P.; data curation, B.S.; writing—original draft preparation, B.S.; writing—review and editing, B.S., D.A., C.G. and P.P.; supervision, D.A., C.G. and P.P.; funding acquisition, B.S., D.A., C.G. and P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020-UIDP/50008/2020.

Institutional Review Board Statement: This study received approval from the Ethics and Deontology Committee of the University of Aveiro, Portugal (No. 29-CED/2021), and was conducted in accordance with the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jung, J.; Lim, S.; Kim, B.K.; Lee, S. CNN-based driver monitoring using millimeter-wave radar sensor. *IEEE Sens. Lett.* **2021**, *5*, 3500404. [CrossRef]
2. Kagawa, M.; Suzumura, K.; Matsui, T. Sleep stage classification by non-contact vital signs indices using Doppler radar sensors. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 4913–4916.
3. Gouveia, C.; Tomé, A.; Barros, F.; Soares, S.C.; Vieira, J.; Pinho, P. Study on the usage feasibility of continuous-wave radar for emotion recognition. *Biomed. Signal Process. Control* **2020**, *58*, 101835. [CrossRef]
4. Lubecke, O.B.; Ong, P.W.; Lubecke, V.M. 10 GHz Doppler radar sensing of respiration and heart movement. In Proceedings of the Proceedings of the IEEE 28th Annual Northeast Bio Engineering Conference, Philadelphia, PA, USA, 21 April 2002; pp. 55–56.

5. Ichapurapu, R.; Jain, S.; John, G.; Lie, D.Y.; Banister, R.; Griswold, J. A 2.4 GHz non-contact biosensor system for continuous vital-signs monitoring. In Proceedings of the 2009 IEEE 10th Annual Wireless and Microwave Technology Conference, Clearwater, FL, USA, 20–21 April 2009; pp. 1–3.
6. Massagram, W.; Lubecke, V.M.; HØst-Madsen, A.; Boric-Lubecke, O. Assessment of Heart Rate Variability and Respiratory Sinus Arrhythmia via Doppler Radar. *IEEE Trans. Microw. Theory Tech.* **2009**, *57*, 2542–2549. [CrossRef]
7. Tan, H.; Qiao, D.; Li, Y. Non-contact heart rate tracking using Doppler radar. In Proceedings of the 2012 International Conference on Systems and Informatics, Yantai, China, 19–20 May 2012; pp. 1711–1714.
8. Hu, W.; Zhao, Z.; Wang, Y.; Zhang, H.; Lin, F. Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature Doppler radar sensor. *IEEE Trans. Biomed. Eng.* **2013**, *61*, 725–735. [CrossRef] [PubMed]
9. Kuutti, J.; Paukkunen, M.; Aalto, M.; Eskelinen, P.; Sepponen, R.E. Evaluation of a Doppler radar sensor system for vital signs detection and activity monitoring in a radio-frequency shielded room. *Measurement* **2015**, *68*, 135–142. [CrossRef]
10. Tu, J.; Lin, J. Fast acquisition of heart rate in noncontact vital sign radar measurement using time-window-variation technique. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 112–122. [CrossRef]
11. Li, M.; Lin, J. Wavelet-transform-based data-length-variation technique for fast heart rate detection using 5.8 GHz CW Doppler radar. *IEEE Trans. Microw. Theory Tech.* **2018**, *66*, 568–576. [CrossRef]
12. Will, C.; Shi, K.; Schellenberger, S.; Steigleder, T.; Michler, F.; Fuchs, J.; Weigel, R.; Ostgathe, C.; Koelpin, A. Radar-based heart sound detection. *Sci. Rep.* **2018**, *8*, 11551. [CrossRef]
13. Yamamoto, K.; Toyoda, K.; Ohtsuki, T. Spectrogram-based non-contact RRI estimation by accurate peak detection algorithm. *IEEE Access* **2018**, *6*, 60369–60379. [CrossRef]
14. Park, J.H.; Jeong, Y.J.; Lee, G.E.; Oh, J.T.; Yang, J.R. 915 MHz continuous wave Doppler radar sensor for detection of vital signs. *Electronics* **2019**, *8*, 561. [CrossRef]
15. Kim, J.Y.; Park, J.H.; Jang, S.Y.; Yang, J.R. Peak detection algorithm for vital sign detection using Doppler radar sensors. *Sensors* **2019**, *19*, 1575. [CrossRef] [PubMed]
16. Schellenberger, S.; Shi, K.; Steigleder, T.; Malessa, A.; Michler, F.; Hameyer, L.; Neumann, N.; Lurz, F.; Weigel, R.; Ostgathe, C.; et al. A dataset of clinically recorded radar vital signs with synchronised reference sensor signals. *Sci. Data* **2020**, *7*, 291. [CrossRef] [PubMed]
17. Shi, K.; Schellenberger, S.; Will, C.; Steigleder, T.; Michler, F.; Fuchs, J.; Weigel, R.; Ostgathe, C.; Koelpin, A. A dataset of radar-recorded heart sounds and vital signs including synchronised reference sensor signals. *Sci. Data* **2020**, *7*, 50. [CrossRef] [PubMed]
18. Kebe, M.; Gadhafi, R.; Mohammad, B.; S.; uleanu, M.; Saleh, H.; Al-Qutayri, M. Human vital signs detection methods and potential using radars: A review. *Sensors* **2020**, *20*, 1454. [CrossRef] [PubMed]
19. Xia, W.; Li, Y.; Dong, S. Radar-Based High-Accuracy Cardiac Activity Sensing. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 4003213. [CrossRef]
20. Diewald, A.R.; Landwehr, J.; Tatarinov, D.; Cola, P.D.; Watgen, C.; Mica, C.; Lu-Dac, M.; Larsen, P.; Gomez, O.; Goniva, T. RF based child occupation detection in the vehicle interior. In Proceedings of the 2016 17th International Radar Symposium (IRS), Krakow, Poland, 10–12 May 2016; pp. 1–4.
21. Anishchenko, L. Challenges and potential solutions of psychophysiological state monitoring with bio-radar technology. *Diagnostics* **2018**, *8*, 73. [CrossRef] [PubMed]
22. Gao, Q.; Zhang, L.; Yan, J.; Zhao, H.; Ding, C.; Hong, H.; Zhu, X. Non-contact emotion recognition via CW Doppler radar. In Proceedings of the 2018 Asia-Pacific Microwave Conference (APMC), Kyoto, Japan, 6–9 November 2018; pp. 1468–1470.
23. Liang, Q.; Xu, L.; Bao, N.; Qi, L.; Shi, J.; Yang, Y.; Yao, Y. Research on non-contact monitoring system for human physiological signal and body movement. *Biosensors* **2019**, *9*, 58. [CrossRef]
24. Zhang, L.; Xiong, J.; Zhao, H.; Hong, H.; Zhu, X.; Li, C. Sleep stages classification by CW Doppler radar using bagged trees algorithm. In Proceedings of the 2017 IEEE Radar Conference (RadarConf), Seattle, WA, USA, 8–12 May 2017; pp. 788–791.
25. Boric-Lubecke, O.; Lubecke, V.M.; Host-Madsen, A.; Samardzija, D.; Cheung, K. Doppler radar sensing of multiple subjects in single and multiple antenna systems. In Proceedings of the TELSIKS 2005–2005 uth International Conference on Telecommunication in ModernSatellite, Cable and Broadcasting Services, Nis, Serbia, 28–30 September 2005; Volume 1, pp. 7–11.
26. Rahman, T.; Adams, A.T.; Ravich, R.V.; Zhang, M.; Patel, S.N.; Kientz, J.A.; Choudhury, T. DoppleSleep: A Contactless Unobtrusive Sleep Sensing System Using Short-Range Doppler Radar. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Osaka, Japan, 7–11 September 2015; Volume 10, p. 1145.
27. Lomauro, A.; Aliverti, A. Sex differences in respiratory function. *Breathe* **2018**, *14*, 131–140. [CrossRef]
28. Romei, M.; Mauro, A.L.; D'angelo, M.G.; Turconi, A.C.; Bresolin, N.; Pedotti, A.; Aliverti, A. Effects of gender and posture on thoraco-abdominal kinematics during quiet breathing in healthy adults. *Respir. Physiol. Neurobiol.* **2010**, *172*, 184–191. [CrossRef]
29. Bhatti, U.; Laghari, Z.A.; Syed, B.M. Effect of body mass index on respiratory parameters: A cross-sectional analytical study. *Pak. J. Med. Sci.* **2019**, *35*, 1724. [CrossRef]
30. Sharma, G.; Goodwin, J. Effect of aging on respiratory system physiology and immunology. *Clin. Interv. Aging* **2006**, *1*, 253. [CrossRef] [PubMed]
31. Rossi, A.; Ganassini, A.; Tantucci, C.; Grassi, V. Aging and the respiratory system. *Aging Clin. Exp. Res.* **1996**, *8*, 143–161. [CrossRef] [PubMed]

32. Boric-Lubecke, O.; Lubecke, V.M.; Droitcour, A.D.; Park, B.K.; Singh, A. *Doppler Radar Physiological Sensing*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
33. Gouveia, C. Bio-Radar. Master's Thesis, Universidade de Aveiro, Aveiro, Portugal, 2017.
34. Ettus Research. UB200 Kit. Available online: https://www.ettus.com/all-products/ub200-kit/ (accessed on 16 November 2023).
35. Gouveia, C.; Albuquerque, D.; Vieira, J.; Pinho, P. Dynamic digital signal processing algorithm for vital signs extraction in continuous-wave radars. *Remote Sens.* **2021**, *13*, 4079. [CrossRef]
36. Anishchenko, L.; Bochkarev, M.; Korostovtseva, L.; Sviryaev, Y.; Bugaev, A. Remote Limb Movement Analysis During Sleep by Means of Bioradar. In Proceedings of the 2020 IEEE MTT-S International Microwave Biomedical Conference (IMBioC), Toulouse, France, 14–17 December 2020; pp. 1–3.
37. Anishchenko, L.; Rutskova, E. Estimation of rat's sleep-wake cycle using a bio-radar. In Proceedings of the 2017 International Conference on Electromagnetics in Advanced Applications (ICEAA), Verona, Italy, 11–15 September 2017; pp. 468–471.
38. Lin, F.; Zhuang, Y.; Song, C.; Wang, A.; Li, Y.; Gu, C.; Li, C.; Xu, W. SleepSense: A noncontact and cost-effective sleep monitoring system. *IEEE Trans. Biomed. Circuits Syst.* **2016**, *11*, 189–202. [CrossRef] [PubMed]
39. Anishchenko, L.; Turetzkaya, A. Improved Non-Contact Mental Stress Detection via Bioradar. In Proceedings of the 2020 International Conference on Biomedical Innovations and Applications (BIA), Varna, Bulgaria, 24–27 September 2020; pp. 21–24.
40. Matar, G.; Lina, J.M.; Carrier, J.; Kaddoum, G. Unobtrusive sleep monitoring using cardiac, breathing and movements activities: An exhaustive review. *IEEE Access* **2018**, *6*, 45129–45152. [CrossRef]
41. Justo, I. Automatic Audio Signal Analysis for the Detection of Anomalies in Calls. Master's Thesis, Universidade de Aveiro, Aveiro, Portugal, 2017.
42. Beers, B. P-Value Definition. 2024. Available online: https://www.investopedia.com/terms/p/p-value.asp (accessed on 22 November 2023).
43. Hahs-Vaughn, D.L.; Lomax, R.G. *Statistical Concepts: A Second Course*; Routledge: London, UK, 2020.
44. Arnold, T.; Kane, M.; Lewis, B.W. *A Computational Approach to Statistical Learning*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2019.
45. Sutherland, T.J.; McLachlan, C.R.; Sears, M.R.; Poulton, R.; Hancox, R.J. The relationship between body fat and respiratory function in young adults. *Eur. Respir. J.* **2016**, *48*, 734–747. [CrossRef]
46. Ray, C.S.; Sue, D.Y.; Bray, G.; Hansen, J.E.; Wasserman, K. Effects of obesity on respiratory function. *Am. Rev. Respir. Dis.* **1983**, *128*, 501–506. [CrossRef]
47. Anthimopoulos, M.; Christodoulidis, S.; Ebner, L.; Christe, A.; Mougiakakou, S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. *IEEE Trans. Med. Imaging* **2016**, *35*, 1207–1216. [CrossRef]
48. Koutsoukas, A.; Monaghan, K.J.; Li, X.; Huan, J. Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data. *J. Cheminf.* **2017**, *9*, 42. [CrossRef]
49. Huang, S.; Cai, N.; Pacheco, P.P.; Narrandes, S.; Wang, Y.; Xu, W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genom. Proteom.* **2018**, *15*, 41–51.
50. Lo, Y.C.; Rensi, S.E.; Torng, W.; Altman, R.B. Machine learning in chemoinformatics and drug discovery. *Drug Discov. Today* **2018**, *23*, 1538–1546. [CrossRef] [PubMed]
51. Knights, D.; Costello, E.K.; Knight, R. Supervised classification of human microbiota. *FEMS Microbiol. Rev.* **2011**, *35*, 343–359. [CrossRef] [PubMed]
52. Vapnik, V.N. Conclusion: What is Important in Learning Theory? In *The Nature of Statistical Learning Theory*; Springer: Berlin/Heidelberg, Germany, 1995; pp. 167–175.
53. Brunton, S.L.; Kutz, J.N. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*; Cambridge University Press: Cambridge, UK, 2022.
54. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
55. Pinto, G.; Carvalho, J.M.; Barros, F.; Soares, S.C.; Pinho, A.J.; Brás, S. Multimodal emotion evaluation: A physiological model for cost-effective emotion classification. *Sensors* **2020**, *20*, 3510. [CrossRef]