Jurnal Infotel
Vol. 16, No. 1, February 2024, pp. 17–31. doi:10.20895/infotel.v16i1.1049

Research Article

# Resampling Strategies and their Influence on Heart Rate Variability Features in Low Sampling Rate Electrocardiogram Data

Muhammad Zakariyah[1,*], Umar Zaky[2], Muhammad Nurjaman[3], Agil Ghani Istikmal[4], and Hafizh Athallah Widianto[5]

[1,3,4]Informatics Program, University Technology of Yogyakarta, Sleman 55285, Indonesia
[2,5]Information System Program, University Technology of Yogyakarta, Sleman 55285, Indonesia

*Corresponding email: muhammad.zakariyah@staff.uty.ac.id

*Received: September 26, 2023; Revised: November 21, 2023; Accepted: November 27, 2023.*

**Abstract:** Heart rate variability (HRV) is a parameter to measure fluctuations in the interval between heartbeats. HRV provides essential insights into the cardiovascular function and autonomic nervous system. Electrocardiograms (ECG) on wearable devices are often recorded at low sampling rates, limiting temporal resolution and information. Resampling is a technique of changing the sampling rate from a high sampling rate to a lower sampling rate and vice versa. This research aims to evaluate the effect of resampling ECG data with a low sampling rate on HRV features. ECG data consists of 50 Hz and 100 Hz sampling rates. Data with a 50 Hz sampling rate is up-sampled up to 100 Hz, while 100 Hz data is down-sampled up to 50 Hz and up-sampled up to 250 Hz using the Fast Fourier Transform Interpolation Method. Upsampling from 50 Hz to 100 Hz shows unsatisfactory results, except for some HRV features such as NN20, pNN20, and CVI. Better results were found when up sampling from 100 Hz up to 250 Hz, with some HRV features showing good concordance values. However, downsampling from 100 Hz up to 50 Hz is unsuitable for HRV feature analysis. To obtain accurate HRV analysis results in all domains, it is highly recommended to use a sampling rate above 100 Hz.

**Keywords:** electrocardiogram, heart rate variability, interpolation, resampling, sampling rate

## 1 Introduction

An electrocardiogram (ECG) is a vital tool in the medical world that is used to record the heart's electrical activity. ECG represents electrical signals produced by contraction and

 ISSN: 2085-3688; e-ISSN: 2460-0997

relaxation of the heart muscle. Processing ECG data has many clinical features that can be extracted, including heart rate variability (HRV). HRV measures fluctuations in the interval between heartbeats in-depth and is an essential indicator in analyzing cardiovascular function and the condition of the autonomic nervous system [1,2].

ECG recording is usually carried out using modern equipment that has a high sampling rate. However, nowadays, ECG recording can be done by using a portable device. These mobile ECGs often record data with a low sampling rate, due to data storage efficiency considerations [3,4]. As a result, ECG data with low sampling rates have limited temporal resolution and may contain missing information [5]. A low sampling rate in HRV analysis can affect the extraction of HRV features and the interpretation of analysis results [6].

Resampling is a method used to overcome challenges in ECG data with a low sampling rate. This technique involves changing the data sampling rate through upsampling or downsampling. Although data resampling has been applied in various signal processing domains [7], the implications of this method for HRV analysis of ECG data with low sampling rates still need to be clarified.

Previous research was conducted by Kwon *et al.* [8] and Mahdiani *et al.* [9]. Both studies used down-sampling techniques to investigate whether ECG data with high sampling rates (1,000 Hz and 5 kHz) could be reduced. The down-sampling results are explained by keeping the information contained in the HRV feature the same. Data with a high sampling rate can only be obtained from devices with unique technology. Based on this, this research intends to use ECG data with a low sampling rate that could be obtained with wearable devices, and then up-sampling is carried out in stages.

The main aim of this study was to investigate the impact of data resampling on HRV features using ECG data with a low sampling rate. This research will analyze how data resampling affects the extraction results of HRV features and the interpretation of the analysis results. Through the results of this analysis, we hope that it can provide an understanding of when and how this method can be used effectively in the context of cardiovascular analysis.

## 2 Research Methods

This research begins with collecting ECG data, which has 50 Hz and 100 Hz sampling rates (original ECG data). The initial ECG data (baseline) is then processed for resampling (up-sampling or down-sampling) using the interpolation method. The following process is to detect the peak (R-Peak) of the original signal and resampled signal for calculating RR-Interval. Calculated RR-Interval produces new data (time-series data). The next step is to extract values of time-series data for the HRV features. Correlation analysis was performed to determine the significance of the original signal (50 Hz and 100 Hz) with the resampled signal using the concordance correlation coefficient (CCC) method. Figure 1 shows the proposed analysis method to compare the original and resampled signal.

### 2.1 Dataset

Electrocardiogram data in this study uses a secondary dataset from the Physionet site [10]. The lowest sampling rates of ECG data used for HRV analysis are 50 Hz and 100 Hz. The ECG data consists of 815 data with a 50 Hz sampling rate [11] and 98 data with a 100 Hz

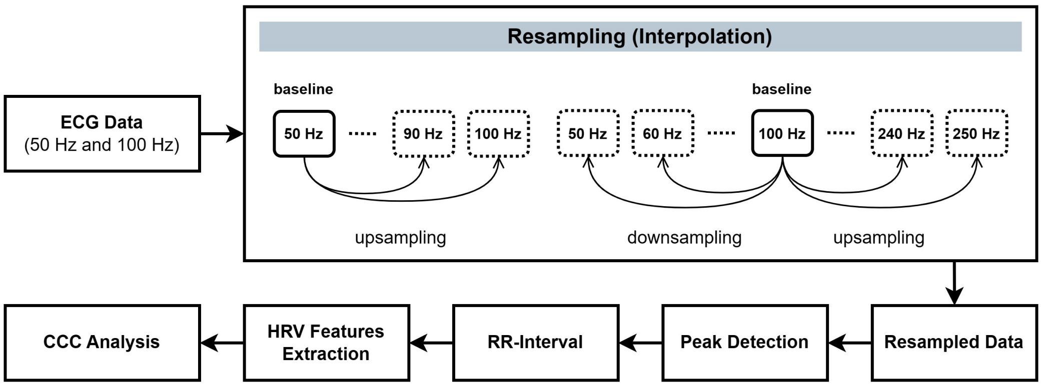


Figure 1: Proposed method.

sampling rate [12] as a baseline. Data with a 50 Hz sampling rate will be up-sampled into 60 Hz, 70 Hz, 80 Hz, 90 Hz, and 100 Hz. For data with a 100 Hz sampling rate as the baseline, two resampling processes will be carried out: upsampling up to 250 Hz and downsampling up to 50 Hz (multiples of 10).

## 2.2 Fast Fourier Transform Interpolation

The fast fourier transform (FFT) interpolation method for resampling data is an approach that leverages the frequency-domain concept to alter the data's sampling rate. The interpolation method with FFT works by transforming the original signal into the frequency domain using FFT, changing the sampling rate by a specific factor, and finally converting the signal back to the time domain using the inverse FFT (IFFT). The goal is to obtain new data with a different sampling rate from the original data. Steps of FFT-based interpolation are FFT transformation, adding zero frequencies (zero padding), IFFT, and resampling result.

### 2.2.1 FFT transformation

First, perform the FFT transformation of the original data to convert it into the frequency domain. The general FFT formula is shown in (1).

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j\frac{2\pi kn}{N}} \tag{1}$$

where $X(k)$ is the frequency coefficient in the frequency domain, $x(n)$ is a data point in the time domain, $k$ is the frequency index, and $N$ is the data length.

### 2.2.2 Adding zero frequencies (zero padding)

Increasing the sampling rate can be done by adding zero values to the frequency coefficients in the frequency domain according to the desired frequency enhancement factor.

#### 2.2.3 Inverse FFT (IFFT)

After adding the necessary zero frequencies, perform the inverse FFT to return the data to the time domain. The IFFT formula can be seen in (2).

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \cdot e^{j\frac{2\pi kn}{N}} \tag{2}$$

where $x(n)$ is the resampled data in the time domain, and $X(k)$ is the altered frequency coefficient from the original data.

#### 2.2.4 Resampling result

The data obtained after IFFT is the resampled data with the altered frequency as needed.

### 2.3 Peak Detection and RR-Interval

This research uses two moving averages (TMA) to detect R-peaks in electrocardiogram (ECG) signals. R-peaks represent the highest points in the QRS complex, indicating the depolarization of the ventricles. The TMA algorithm leverages the concept of moving averages to identify R-peaks by tracking variations in signal amplitude over time [13]. Here are the steps for the TMA algorithm for R-peak detection.

#### 2.3.1 Initialization

Define the window lengths for the fast-moving average (FMA) and the slow-moving average (SMA). The FMA captures short-term changes, while the SMA focuses on longer-term trends in the ECG signal.

#### 2.3.2 Initial averages

Compute the initial FMA and SMA values using the data points within their respective windows.

#### 2.3.3 Iterate through data

Start iterating through the ECG signal data points, moving forward step by step (for each data point).

a. Update FMA: Calculate the FMA value using the current data point and the FMA window.
b. Update SMA: Calculate the SMA value using the current data point and the SMA window.
c. Compare Averages: Check if the FMA crosses over the SMA.
d. If FMA crosses over SMA and the difference exceeds threshold, mark the data point as a potential R-peak.

## 2.4 Heart Rate Variability Features

Heart Rate Variability (HRV) refers to the variation in time between successive heartbeats (RR-Intervals) in an electrocardiogram (ECG) signal. HRV analysis involves quantifying the fluctuations in heart rate over time and exploring the relationship between heart rate and various physiological and environmental factors. Analysis of HRV is conducted through two main approaches: linear analysis (time domain and frequency domain) and nonlinear analysis (Poincaré plot and entropy). Several HRV features provide different insights into heart rate variability. In this study, we used HRV features as shown in Table 1.

Table 1: Heart rate variability features

| **Domain Analysis** | **HRV Features** | **Description** |
|---|---|---|
| Time Domain | MeanRR | Mean RR-interval |
| | SDRR | Standard deviation of RR-intervals |
| | RMSSD | Root mean square of successive differences |
| | CVRR | Coefficient of variation of RR-intervals, percentage of the mean RR-interval |
| | NN20 | The number of pairs of successive RR-intervals differing by more than 20 ms |
| | pNN20 | The percentage of NN20 divided by the total number of RR-intervals |
| Frequency Domain | VLF | Power in the very low-frequency range (0.0033 - 0.04 Hz) |
| | LF | Power in the low-frequency range (0.04 - 0.15 Hz) |
| | HF | Power in the high-frequency range (0.15 - 0.4 Hz) |
| | LF/HF Ratio | The ratio of LF to HF power |
| | CSI | Cardiac sympathetic index |
| | CVI | Cardiac vagal index |
| Poincaré plot | SD1 | Standard deviation of points perpendicular to the line of identity in a Poincaré plot |
| | SD2 | Standard deviation of points along the line of identity in a Poincaré plot |
| | SD1/SD2 Ratio | The ratio of SD1 to SD2 |
| | CVSD | Coefficient of variation of SD1 |
| Entropy | SampEn | Sample Entropy, measure of complexity and irregularity in HRV |

## 2.5 Concordance Correlation Coefficient (CCC)

Lin's concordance correlation coefficient (CCC) measures how well the observed bivariate is valid against standard measurement methods. CCC was used to evaluate the strength of agreement between the two methods [14]. CCC is the product of the Pearson correlation coefficient and the squared difference between measurements. A high coefficient means strong agreement and a linear relationship between methods. The form of the concordance correlation coefficient is defined in Equation 3.

$$\rho_c = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\mu_x - \mu_y)^2} \tag{3}$$

where $\mu_x$ and $\mu_y$ are the means for the two variables, $\sigma_x^2$ and $\sigma_y^2$ are the corresponding variances, and $\rho$ is the correlation coefficient between the two variables. CCC ($\rho_c$) measures precision ($\rho$) and accuracy ($C\beta$), which has values ranging from 0 to $\pm$ 1. Table 2 shows the interpretation of CCC measurement results [15].

Table 2: CCC Interpretation

| **Value of CCC** | **Interpretation** |
|---:|---|
| > 0.99 | Almost Perfect |
| 0.95 to 0.99 | Substantial |
| 0.90 to 0.95 | Moderate |
| < 0.90 | Poor |

# 3 Result

This section discusses electrocardiogram resampling and correlation coefficients.

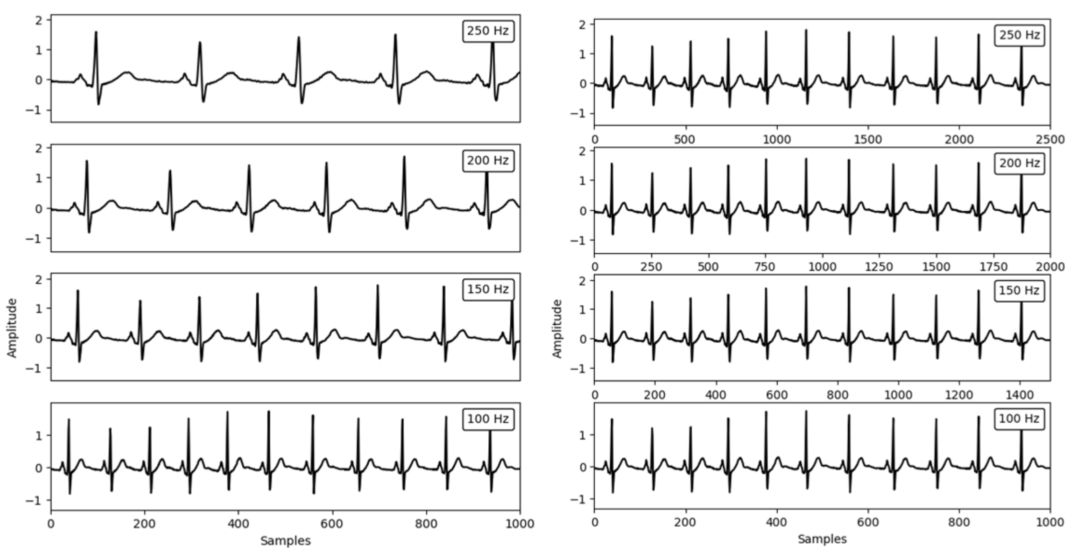


Figure 2: ECG signal from 100 Hz baseline resampled up to 250 Hz: Data with initial 1,000 samples (left), adjustment sample size for up sampled data (right).

## 3.1 Electrocardiogram Resampling

ECG data, which has a baseline 50 Hz sampling rate, is sampled up to 100 Hz (60 Hz, 70 Hz, 80 Hz, 90 Hz, 100 Hz). Meanwhile, ECG data with a 100 Hz baseline is down-sampled up to 50 Hz and up-sampled up to 250 Hz (using multiples of 10). Each baseline signal is resampled using the FFT interpolation method. Figure 2 shows data with 100 Hz baseline up sampled into 150 Hz, 200 Hz, and 250 Hz.

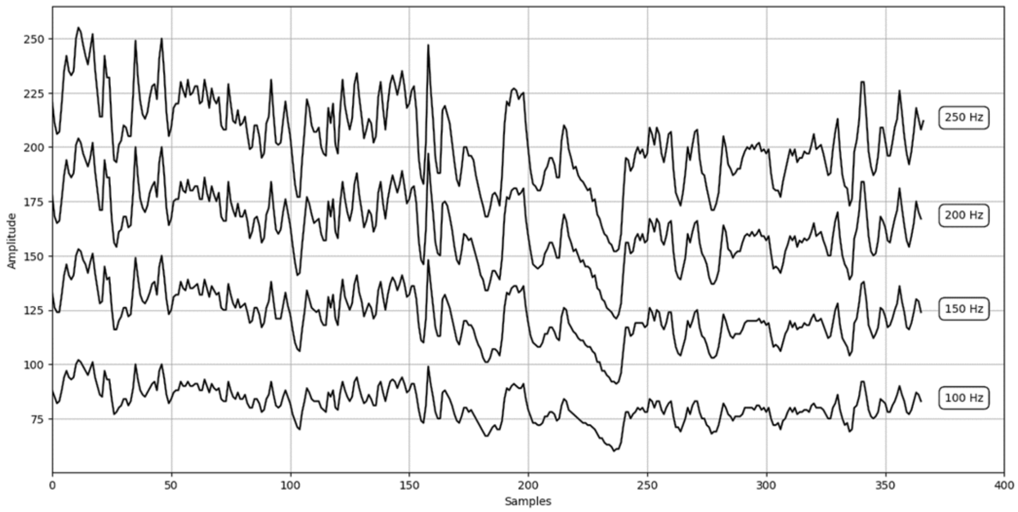


Figure 3: RR-interval plot from 100 Hz (baseline) up to 250 Hz (upsampling).

As shown in Figure 2 (left), data with 100 Hz as a baseline, when up-sampled into 150 Hz, 200 Hz, and 250 Hz, will produce different numbers of R peaks in the first 1,000 samples. However, the difference in the number of samples after the resampling process did not affect overall number of R peaks detected (Figure 2 right). This can also be seen in the RR-Interval plot results in Figure 3.

## 3.2 Correlation Coefficients

A comparison of the values of HRV features among original HRV (50 Hz) and upsampled HRV up to 100 Hz is shown in Table 3. Measurements were carried out using CCC metric and linear measurements (time and frequency) and non-linear measurements (Poincaré plot and sample entropy).

Table 3 shows the result after the up-sampling process from 50 Hz to 60 Hz. From 17 features, only 3 have CCC values above 0.9, namely NN20, pNN20, and CVI. For the remaining upsampling, only the CVI feature shows moderate correlation up to the 80 Hz limit. Upsampling from 50 Hz to 90 Hz or 100 Hz did not show any correlation in all features. The resampling results for data with a 100 Hz baseline are shown in Table 4 and Table 5.

Table 4 shows that when a 100 Hz signal is down-sampled up to 50 Hz, the CVRR feature correlates with each sampling rate (almost perfect at 90 Hz, 80 Hz, 70 Hz, and substantial at 60 Hz and 50 Hz). The CSI, SD1/SD2 Ratio, and CVSD features only show correlation up to 60 Hz with CCC values above 92 % (> 0.92). The NN20 and pNN20 feature correlate up to 70 Hz (CCC value > 0.92). Downsampling results from 100 Hz to 90 Hz show that many features still correlate with CCC values above 92 %, including SDRR, RMSSD CVRR NN20, pNN20 (time domain), CSI and CVI (frequency domain), SD1, SD2, SD1/SD2 Ratio, CVSD (Poincaré plot). Meanwhile, for the Entropy domain (SampEn feature), CCC still does not show any correlation in any downsampling results. Data with

Table 3: CCC result for HRV features rom 50 Hz baseline upsampled up to 100 Hz

| **Features** | **Upsampled coefficients against data sampled at 50 Hz (CCC)** | | | | |
|---|---|---|---|---|---|
| | **60 Hz** | **70 Hz** | **80 Hz** | **90 Hz** | **100 Hz** |
| MeanRR | 0.3243 | 0.1228 | 0.0663 | 0.0431 | 0.0310 |
| SDRR | 0.6645 | 0.4324 | 0.3639 | 0.3644 | 0.4085 |
| RMSSD | 0.5501 | 0.3254 | 0.2779 | 0.2998 | 0.3552 |
| CVRR | 0.7164 | 0.5537 | 0.5286 | 0.5846 | 0.6955 |
| NN20 | **0.9308** | 0.8087 | 0.7414 | 0.6484 | 0.5836 |
| pNN20 | **0.9368** | 0.8271 | 0.7544 | 0.6545 | 0.5782 |
| VLF | 0.0084 | 0.0003 | 0.0002 | 0.0001 | 0.0024 |
| LF | 0.0033 | 0.0002 | 0.0002 | 0.0001 | 0.0014 |
| HF | 0.0028 | 0.0004 | 0.0003 | 0.0002 | 0.0021 |
| LF/HF Ratio | 0.3372 | 0.2769 | 0.2981 | 0.2443 | 0.2153 |
| CSI | 0.8329 | 0.7825 | 0.7535 | 0.7248 | 0.7091 |
| CVI | **0.9663** | **0.9382** | **0.9024** | 0.8675 | 0.8287 |
| SD1 | 0.5501 | 0.3254 | 0.2779 | 0.2998 | 0.3553 |
| SD2 | 0.7182 | 0.5051 | 0.4235 | 0.4020 | 0.4364 |
| SD1/SD2 Ratio | 0.8329 | 0.7825 | 0.7535 | 0.7248 | 0.7091 |
| CVSD | 0.5921 | 0.4231 | 0.4072 | 0.4701 | 0.5743 |
| SampEn | 0.8841 | 0.7304 | 0.6798 | 0.6300 | 0.6869 |

a 100 Hz baseline was also upsampled to 250 Hz with the upsampling results as shown in Table 5.

As shown in Table 5, a comparison of the CCC values from 100 Hz, which upsampled up to 250 Hz, indicates that the CVRR, CVSD, and SampEn features correlate with each sampling rate (except SampEn at 120 Hz). CVRR shows the highest CCC value with a value up to 99.02 % at a 150 Hz sampling rate. The CSI and SD1/SD2 Ratio features correlate at several sampling rates with moderate CCC values (between 90 % and 95 %). Like the results shown in the downsampling process from 100 Hz to 90 Hz, the correlation between 110 Hz and the baseline (100 Hz) is visible in several features; even the SampEn feature also has a moderate correlation (> 0.92).

Scatter plots are used to provide an overview of the distribution of data between the original signal (baseline) and resampled signal. The scatter plot consists of the X-axis, which is the baseline sampling rate, the Y-axis which is the resampling result, the 45-degree line which is the gold standard (faint line), and the deviation line which is the CCC result (clear line). Figure 4 shows a comparison of the data distribution for the CVRR feature from a baseline (50 Hz) when up-sampled to 60 Hz, 80 Hz, and 100 Hz (top image), while the bottom image shows the results of up-sampling from 100 Hz into 110 Hz, 130 Hz, and 150 Hz. The distribution lines between the original 50 Hz signal and the up-sampling results are not in a parallel position. This means that the CCC value from the baseline of 50 Hz has quite a large deviation in the CVRR feature when up-sampled into 60 Hz, 80 Hz, and 100 Hz. Meanwhile, if the original 100 Hz signal is up-sampled into 110 Hz, 130 Hz, and 150 Hz, it shows adjacent/parallel lines. Thus, the up-sampling results shown in Figure 3 have a high level of correlation.

The CCC calculation results show that ECG signal resamples on several HRV features have varying impacts, depending on the measured feature and the level of resampling

Table 4: CCC result for HRV features from 100 Hz baseline downsampled up to 50 Hz

| Features | Downsampled coefficients against data sampled at 100 Hz (CCC) | | | | |
|---|---|---|---|---|---|
| | 90 Hz | 80 Hz | 70 Hz | 60 Hz | 50 Hz |
| MeanRR | 0.3413 | 0.1050 | 0.0439 | 0.0221 | 0.0123 |
| SDRR | **0.9389** | 0.7526 | 0.5379 | 0.3453 | 0.2038 |
| RMSSD | **0.9605** | 0.8596 | 0.7056 | 0.4953 | 0.2795 |
| CVRR | **0.9932** | **0.9913** | **0.9941** | **0.9888** | **0.9671** |
| NN20 | **0.9757** | **0.9544** | **0.9202** | 0.7913 | 0.4645 |
| pNN20 | **0.9766** | **0.9560** | **0.9215** | 0.7956 | 0.4752 |
| VLF | 0.8613 | 0.6966 | 0.3721 | 0.2163 | 0.1156 |
| LF | 0.7895 | 0.5675 | 0.4000 | 0.2310 | 0.1439 |
| HF | 0.8304 | 0.5670 | 0.3967 | 0.2376 | 0.1083 |
| LF/HF Ratio | 0.8080 | 0.7606 | 0.7003 | 0.6600 | 0.5321 |
| CSI | **0.9420** | **0.9611** | **0.9569** | **0.9262** | 0.8099 |
| CVI | **0.9573** | 0.8363 | 0.6587 | 0.4645 | 0.2744 |
| SD1 | **0.9605** | 0.8596 | 0.7057 | 0.4954 | 0.2796 |
| SD2 | **0.9267** | 0.7197 | 0.4974 | 0.3177 | 0.1934 |
| SD1/SD2 Ratio | **0.9420** | **0.9611** | **0.9569** | **0.9262** | 0.8099 |
| CVSD | **0.9744** | **0.9715** | **0.9773** | **0.9633** | 0.8741 |
| SampEn | 0.8955 | 0.8621 | 0.8616 | 0.7857 | 0.7941 |

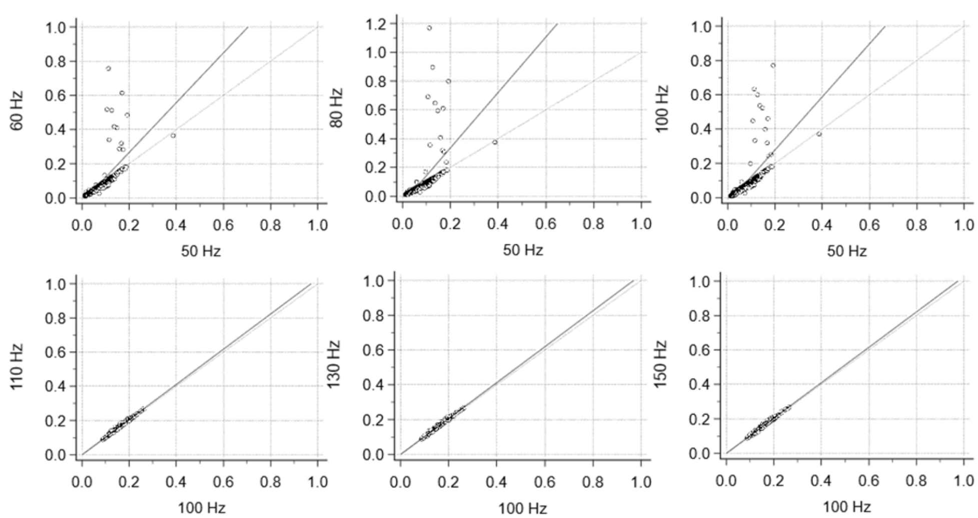


Figure 4: Upsampling scatter plot from 50 Hz baseline (top) and 100 Hz (bottom) for coefficient varians of RR-interval (CVRR) feature.

applied to the original signal. Some features may retain information from the original signal after resampling, while others show significant changes. Therefore, the resampling process should be carefully considered depending on the features to be used in the HRV analysis.

Table 5: CCC result for HRV features from 100 Hz baseline upsampled up to 250 Hz

| Features | Downsampled coefficients against data sampled at 100 Hz (CCC) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 110 Hz | 120 Hz | 130 Hz | 140 Hz | 150 Hz | 160 Hz | 170 Hz | 180 Hz | 190 Hz | 200 Hz | 210 Hz |
| MeanRR | 0.4090 | 0.1542 | 0.0800 | 0.0499 | 0.0345 | 0.0256 | 0.0202 | 0.0163 | 0.0137 | 0.0117 | 0.0101 |
| SDRR | **0.9132** | 0.7758 | 0.6416 | 0.5243 | 0.4459 | 0.3635 | 0.3137 | 0.2689 | 0.2392 | 0.2131 | 0.1880 |
| RMSSD | **0.9077** | 0.8192 | 0.7325 | 0.6348 | 0.5785 | 0.4758 | 0.4340 | 0.3809 | 0.3566 | 0.3244 | 0.2818 |
| CVRR | **0.9896** | **0.9887** | **0.9885** | **0.9872** | **0.9902** | **0.9827** | **0.9841** | **0.9818** | **0.9848** | **0.9850** | **0.9798** |
| NN20 | **0.9347** | **0.9030** | 0.8503 | 0.7746 | 0.6974 | 0.5821 | 0.5068 | 0.4251 | 0.3626 | 0.2968 | 0.2409 |
| pNN20 | **0.9377** | **0.9053** | 0.8540 | 0.7792 | 0.7015 | 0.5880 | 0.5125 | 0.4321 | 0.3698 | 0.3037 | 0.2480 |
| VLF | 0.8403 | 0.7042 | 0.6273 | 0.4724 | 0.3698 | 0.2916 | 0.2324 | 0.1949 | 0.1586 | 0.1288 | 0.1069 |
| LF | 0.8743 | 0.7178 | 0.5358 | 0.3930 | 0.2813 | 0.2045 | 0.1559 | 0.1304 | 0.1078 | 0.0928 | 0.0809 |
| HF | 0.7683 | 0.5065 | 0.3439 | 0.2765 | 0.2502 | 0.2097 | 0.1730 | 0.1495 | 0.1246 | 0.1068 | 0.0900 |
| LF/HF Ratio | 0.8803 | 0.7591 | 0.6232 | 0.4088 | 0.3894 | 0.3255 | 0.2571 | 0.2390 | 0.2062 | 0.1945 | 0.1725 |
| CSI | **0.9385** | **0.9330** | **0.9502** | **0.9230** | **0.9114** | 0.8971 | **0.9064** | **0.9175** | **0.9182** | **0.9237** | 0.8863 |
| CVI | **0.9267** | 0.8279 | 0.7339 | 0.6318 | 0.5624 | 0.4746 | 0.4281 | 0.3754 | 0.3480 | 0.3146 | 0.2802 |
| SD1 | **0.9077** | 0.8192 | 0.7326 | 0.6348 | 0.5785 | 0.4759 | 0.4340 | 0.3809 | 0.3567 | 0.3244 | 0.2819 |
| SD2 | **0.9163** | 0.7671 | 0.6222 | 0.5014 | 0.4182 | 0.3411 | 0.2907 | 0.2482 | 0.2177 | 0.1932 | 0.1710 |
| SD1/SD2 Ratio | **0.9385** | **0.9330** | **0.9502** | **0.9230** | **0.9114** | 0.8971 | **0.9064** | **0.9175** | **0.9182** | **0.9237** | 0.8863 |
| CVSD | **0.9638** | **0.9645** | **0.9695** | **0.9627** | **0.9701** | **0.9460** | **0.9547** | **0.9493** | **0.9594** | **0.9577** | **0.9398** |
| SampEn | **0.9246** | 0.8997 | **0.9029** | **0.9202** | **0.9337** | **0.9292** | **0.9202** | **0.9323** | **0.9198** | **0.9359** | **0.9278** |
| **Features** | **Downsampled coefficients against data sampled at 100 Hz (CCC) (Contd.)** | | | | | | | | | | |
| | **220 Hz** | **230 Hz** | **240 Hz** | **250 Hz** | | | | | | | |
| MeanRR | 0.0089 | 0.0079 | 0.0071 | 0.0065 | | | | | | | |
| SDRR | 0.1704 | 0.1545 | 0.1422 | 0.1318 | | | | | | | |
| RMSSD | 0.2605 | 0.2395 | 0.2277 | 0.2140 | | | | | | | |
| CVRR | **0.9790** | **0.9779** | **0.9805** | **0.9806** | | | | | | | |
| NN20 | 0.2018 | 0.1676 | 0.1379 | 0.1135 | | | | | | | |
| pNN20 | 0.2089 | 0.1745 | 0.1450 | 0.1202 | | | | | | | |
| VLF | 0.0891 | 0.0804 | 0.0674 | 0.0563 | | | | | | | |
| LF | 0.0687 | 0.0592 | 0.0494 | 0.0427 | | | | | | | |
| HF | 0.0716 | 0.0608 | 0.0484 | 0.0414 | | | | | | | |
| LF/HF Ratio | 0.1543 | 0.1379 | 0.1230 | 0.0876 | | | | | | | |
| CSI | 0.8868 | 0.8882 | **0.9149** | **0.9047** | | | | | | | |
| CVI | 0.2600 | 0.2404 | 0.2259 | 0.2119 | | | | | | | |
| SD1 | 0.2605 | 0.2395 | 0.2277 | 0.2141 | | | | | | | |
| SD2 | 0.1544 | 0.1397 | 0.1277 | 0.1181 | | | | | | | |
| SD1/SD2 Ratio | 0.8868 | 0.8882 | **0.9149** | **0.9047** | | | | | | | |
| CVSD | **0.9399** | **0.9397** | **0.9494** | **0.9485** | | | | | | | |
| SampEn | **0.9432** | **0.9276** | **0.9320** | **0.9318** | | | | | | | |

# 4 Discussion

This research aims to determine the effect of resampling on HRV features in ECG data. Upsampling is carried out on ECG data from a low sampling rate to a higher sampling rate, while downsampling is used from a high sampling rate to a lower sampling rate. There are 17 HRV features used in this study, consisting of MeanRR, SDRR, RMSSD, CVRR, NN20, and pNN20 (time domain), VLF, LF, HF, LF/HF Ratio, CSI, and CVI (frequency domain), SD1, SD2, SD1/SD2 ratio, and CVSD (Poincaré plot), and sampEn (entropy).

An up-sampling process of ECG data with a 50 Hz baseline was done. Only a few research studies discuss up-sampling techniques for ECG data, especially for HRV analysis. The baseline signal, with a 50 Hz sampling rate up-sampled up to 100 Hz, does not show significant concordance values for most of the HRV features. The results of this research show that of the 17 HRV features used, only three features show the suitability of the up-sampling process. These results were only acceptable in up-sampling from 50 Hz to 60 Hz via the NN20, pNN20, and CVI features.

NN20 in HRV indicates the number of RR-interval pairs with a difference of more than 20 milliseconds (ms). This feature reflects a faster response to external stimuli or changes in physical activity. Like NN20, the pNN20 feature indicates how significant the heart rate fluctuations are from all observed heart rates in percentage. Elevated NN20 and pNN20 may indicate the more dynamic response of the body to external stimuli or changing conditions. The cardiac vagal index (CVI) reflects the degree to which parasympathetic (vagal) activity contributes to heart rate variability. A higher CVI value indicates domination in parasympathetic activity, which is usually considered an indicator of balance and a better level of relaxation.

HRV features can provide different information about autonomic nervous activity, both in sympathetic and parasympathetic nerve activity [16]. Based on the results shown in this study, up-sampling is not recommended for ECG data with a baseline of 50 Hz. A sampling rate of 50 Hz can still be used to measure ECG signals without reducing the accuracy of HRV features in the time domain, but the R-peak waveform experiences slight distortion [9].

Different results were shown in the upsampling process carried out on data with a baseline of 100 Hz. ECG data with a 100 Hz baseline, upsampled up to 250 Hz, shows more features that remain correlated, such as CVRR, CSI, SD1/SD2 Ratio, CVSD, and SampEn. CVRR is used to evaluate parasympathetic nerve function [10]. CVRR shows how much variability in the heart rate is about the average heart rate. A decrease in the value of the CVRR feature indicates dysfunction of the parasympathetic nervous system [17].

Several other features, namely CSI and SD1/SD2 Ratio. CSI is a spectral measure that measures how related or aligned the high and low-frequency components are in heart rate variability. A higher CSI value indicates more consistent and balanced nervous activities [18, 19]. SD1/SD2 ratio in HRV compares SD1 and SD2 from the Poincaré plot. The high SD1/SD2 ratio indicates that parasympathetic activity predominates, whereas the lower ratio indicates the predominance of the sympathetic nervous system.

CVSD reflects variations in high-frequency components of the heart rate estimated from short-term measurements. High CVSD values usually indicate more dynamic and responsive autonomic nervous system activity, while lower values indicate a more regulated balance between the sympathetic and parasympathetic nervous systems [19]. Sample Entropy (SampEn) is an analysis method for measuring the complexity and irregularity of HRV data [20]. A higher sampEn value indicates a more irregular and complex HRV. In contrast, lower SampEn values indicate a greater degree of regularity and consistency in the RR-Interval.

Upsampling of ECG data with a 100 Hz sampling rate shows better results compared to ECG data, which has a 50 Hz sampling rate. However, there are still many other features that do not show significant CCC values, mainly if up-sampling is carried out at more than 110 Hz. For HRV analysis that focuses on parasympathetic nervous system activity (CVRR, SD1/SD2 Ratio), the up-sampling process on data with a sampling rate of 100 Hz may be carried out down to 250 Hz. Information resulting from the upsampling process on these features can still be maintained.

The down-sampling process is carried out only on data with a sampling rate of 100 Hz. Downsampling to 50 Hz shows similar results with upsampling up to 250 Hz. HRV features such as CVRR, CSI, CVSD, and SD1/SD2 Ratio still show acceptable concordance values (CCC > 0.9). However, it should be noted that downsampling from 100 Hz for these four features is only able to retain information down to 60 Hz. Thus, downsampling to 50 Hz is not recommended because it will eliminate information from the HRV features that have been extracted.

R-peak detection on ECG data with a high sampling rate provides acceptable results if the down-sampling process is carried out [21, 22]. However, this research shows that although R-Peak detection does not have a significant difference, HRV feature extraction on down-sampling data gives different results. Research regarding the effect of downsampling on HRV analysis was also carried out by Kwon [8]. Downsampling results from 1,000 Hz to 50 Hz proved unacceptable for HRV analysis in the time domain or the frequency domain. At a sampling rate of 50 Hz, the differences between RMSSD and HF features tend to have high random error values.

The results of this study show concordance values for several HRV parameters. Upsampling was done from 100 Hz to 250 Hz in several domains. However, if up-sampling is carried out from 50 Hz, HRV feature analysis can be carried out on NN20, pNN20, and CVI features only at 60 Hz. However, up-sampling ECG data with a baseline of 50 Hz has not given any contribution to HRV analysis.

Therefore, to analyze HRV features in all domains, a minimum sampling rate of above 100 Hz is recommended. The results of this study also support the latest systematic review regarding the use of HRV in the field of clinical care. Most studies use a sampling rate as low as 250 Hz to enable analysis of HRV in various domains [23].

Further research can be carried out by processing the ECG signal using other resampling techniques. Signal processing using different methods (other than interpolation) is expected to provide better results and has the opportunity so that the ECG data obtained is no longer dependent on the sampling rate and device for data acquisition.

# 5 Conclusion

This research evaluates the effect of ECG data resampling on HRV features. Upsampling from 50 Hz to higher levels did not produce significant concordance values, except for the NN20, pNN20, and CVI features (in a 60 Hz sampling rate). Upsampling from 100 Hz to 250 Hz produced better results, with some HRV features still having significant concordance values. Downsampling from 100 Hz to 50 Hz is less satisfactory in HRV feature analysis. Based on the results of this study, it is recommended to use ECG data with a sampling rate above 100 Hz to produce a more comprehensive HRV analysis in all domains.

# Acknowledgments

We want to extend our heartfelt appreciation to the Ministry of Education, Culture, Research, and Technology for providing financial support for this research through the Penelitian Dosen Pemula Program. Our gratitude extends to our research team for their dedication and collaborative effort. We also appreciate the University Technology of Yogyakarta with constructive feedback and engaged in meaningful discussions that helped refine our ideas.

# References

[1] B. W. Johnston, R. Barrett-Jolley, A. Krige, and I. D. Welters, "Heart rate variability: Measurement and emerging use in critical care medicine," *J Intensive Care Soc*, vol. 21, no. 2, pp. 148–157, May 2020, doi: 10.1177/1751143719853744.

[2] A. T. Ginty, T. E. Kraynak, J. P. Fisher, and P. J. Gianaros, "Cardiovascular and autonomic reactivity to psychological stress: Neurophysiological substrates and links to cardiovascular disease," *Autonomic Neuroscience*, vol. 207, pp. 2–9, Nov. 2017, doi: 10.1016/j.autneu.2017.03.003.

[3] A. Bansal and R. Joshi, "Portable out-of-hospital electrocardiography: A review of current technologies," *J Arrhythm*, vol. 34, no. 2, pp. 129–138, Apr. 2018, doi: 10.1002/joa3.12035.

[4] A. R. J. Mitchell and P. Le Page, "Living with the handheld ECG," *BMJ Innov*, vol. 1, no. 2, pp. 46–48, Apr. 2015, doi: 10.1136/bmjinnov-2014-000029.

[5] N.-T. Bui and G. Byun, "The comparison features of ECG signal with different sampling frequencies and filter methods for real-time measurement," *Symmetry*, vol. 13, no. 8, p. 1461, Aug. 2021, doi: 10.3390/sym13081461.

[6] J. S. Burma, A. P. Lapointe, A. Soroush, I. K. Oni, J. D. Smirl, and J. F. Dunn, "Insufficient sampling frequencies skew heart rate variability estimates: Implications for extracting heart rate metrics from neuroimaging and physiological data," *J Biomed Inform*, vol. 123, p. 103934, Nov. 2021, doi: 10.1016/j.jbi.2021.103934.

[7] P. Augustyniak, "Adaptive sampling of the electrocardiogram based on generalized perceptual features," *Sensors*, vol. 20, no. 2, p. 373, Jan. 2020, doi: 10.3390/s20020373.

[8] O. Kwon, J. Jeong, H. B. Kim, I. H. Kwon, S. Y. Park, J. E. Kim, and Y. Choi, "Electrocardiogram sampling frequency range acceptable for heart rate variability analysis," *Healthc Inform Res*, vol. 24, no. 3, p. 198, 2018, doi: 10.4258/hir.2018.24.3.198.

[9] S. Mahdiani, V. Jeyhani, M. Peltokangas, and A. Vehkaoja, "Is 50 Hz high enough ECG sampling frequency for accurate HRV analysis?," in *Annual International Conference*

*of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Aug. 2015, pp. 5948–5951. doi: 10.1109/EMBC.2015.7319746.

[10] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C-K. Peng and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[11] N. Kim, A. Krasner, C. Kosinski, M. Wininger, M. Qadri, Z,. Kappus, S. Danish, W. Craelius, "Trending autoregulatory indices during treatment for traumatic brain injury," *J Clin Monit Comput*, vol. 30, no. 6, pp. 821–831, Dec. 2016, doi: 10.1007/s10877-015-9779-3.

[12] T. Penzel, G. Moody, R. Mark, A. Goldberger, and J. Peter, "The Aanea-ECG database," *Comput Cardiol*, vol. 27, pp. 255–258, 2000.

[13] M. Elgendi, M. Jonkman, and F. D. Boer, "Frequency bands effects on QRS detection," in *Proceedings of the Third International Conference on Bio-inspired Systems and Signal Processing, SciTePress - Science and and Technology Publications*, 2010, pp. 428–431. doi: 10.5220/0002742704280431.

[14] T. R. Vetter and P. Schober, "Agreement analysis: What he said, she said versus you said," *Anesth Analg*, vol. 126, no. 6, pp. 2123–2128, Jun. 2018, doi: 10.1213/ANE.0000000000002924.

[15] H. Akoglu, "User's guide to correlation coefficients," *Turk J Emerg Med*, vol. 18, no. 3, pp. 91–93, Sep. 2018, doi: 10.1016/j.tjem.2018.08.001.

[16] R. Tiwari, R. Kumar, S. Malik, T. Raj, and P. Kumar, "Analysis of heart rate variability and implication of different factors on heart rate variability," *Curr Cardiol Rev*, vol. 17, no. 5, Sep. 2021, doi: 10.2174/1573403X16999201231203854.

[17] T. Miyagi, M. Yamazato, T. Nakamura, T. Tokashiki, Y. Namihira, K. Kokuba, S. Ishihara, H. Sakima, Y. Ohya, "Power spectral analysis of heart rate variability is useful as a screening tool for detecting sympathetic and parasympathetic nervous dysfunctions in Parkinson's disease," *BMC Neurol*, vol. 22, no. 1, p. 339, Sep. 2022, doi: 10.1186/s12883-022-02872-2.

[18] J. Jeppesen, S. Beniczky, P. Johansen, P. Sidenius, and A. Fuglsang-Frederiksen, "Using Lorenz plot and cardiac sympathetic index of heart rate variability for detecting seizures for patients with epilepsy," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Aug. 2014, pp. 4563–4566. doi: 10.1109/EMBC.2014.6944639.

[19] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Front Public Health*, vol. 5, Sep. 2017, doi: 10.3389/fpubh.2017.00258.

[20] V. Alcan, "Sample entropy analysis of heart rate variability in RR interval detection," *Mühendislik Bilimleri ve Tasarım Dergisi*, vol. 8, no. 3, pp. 783–790, Sep. 2020, doi: 10.21923/jesd.745275.

[21] Y. Nishikawa, S. Izumi, Y. Yano, H. Kawaguchi, and M. Yoshimoto, "Sampling rate reduction for wearable heart rate variability monitoring," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, IEEE, May 2018, pp. 1–5. doi: 10.1109/ISCAS.2018.8351558.

[22] E. Ajdaraga and M. Gusev, "Analysis of sampling frequency and resolution in ECG signals," in *2017 25th Telecommunication Forum (TELFOR)*, IEEE, Nov. 2017, pp. 1–4. doi: 10.1109/TELFOR.2017.8249438.

[23] S. N. Karmali, A. Sciusco, S. M. May, and G. L. Ackland, "Heart rate variability in critical care medicine: a systematic review," *Intensive Care Med Exp.*, vol. 5, no. 1, p. 33, Dec. 2017, doi: 10.1186/s40635-017-0146-1.