Architecture
Image Studies
AIS Scientific Journal

Vol.6, Issue 4, pp.173-179, 2025
DOI: https://doi.org/10.62754/ais.v6i4.412

https://journals.ap2.pt/index.php/ais/index

# Study on The Identification of Normal/Abnormal Conditions by Applying MFCC Standard Deviation Characteristics of Heart Rate and Respiration Signals

Sundo June[1]

## Abstract

This study investigates a neural network with clustering capabilities designed to identify whether a person is in a normal physiological condition using heart rate and respiration signals. In particular, the study extracts MFCC (Mel-Frequency Cepstral Coefficients) features from the bio-signals such as heart rate and respiration and performs clustering using an Autoencoder and K-Means algorithm. Based on the resulting clusters, a simulated experiment was conducted to determine whether the method could distinguish between healthy individuals and those in abnormal states. Through these experiments, it was confirmed that the classification performance varies depending on the number of clusters (states), which is a key parameter of the K-Means algorithm. Based on further analysis and discussion, the study proposes a method to automatically calculate the optimal number of K-Means states by reflecting the standard deviation characteristics of the MFCC distributions of healthy individuals and patients with heart failure. To validate the proposed approach, simulated experiments were conducted using heart rate and respiration data of healthy individuals and heart failure patients from Physio.Net, applying both training and non-training datasets. As a result, the proposed method successfully estimated the optimal number of states, enabling reliable identification performance of heart rate and respiration signals using the automatically determined number of states.

**Keywords:** *Heat Rate, Respiration , K-Means, State, Standard Deviation.*

## Introduction

Biometric information provides valuable insights into a person's health and condition. In particular, technologies that utilize signal processing and neural methods to assess human health status are being continuously studied[1][2][3][4][5]. These technologies can be used in various healthcare applications, such as caregiving systems for the elderly and wearable health monitoring systems.

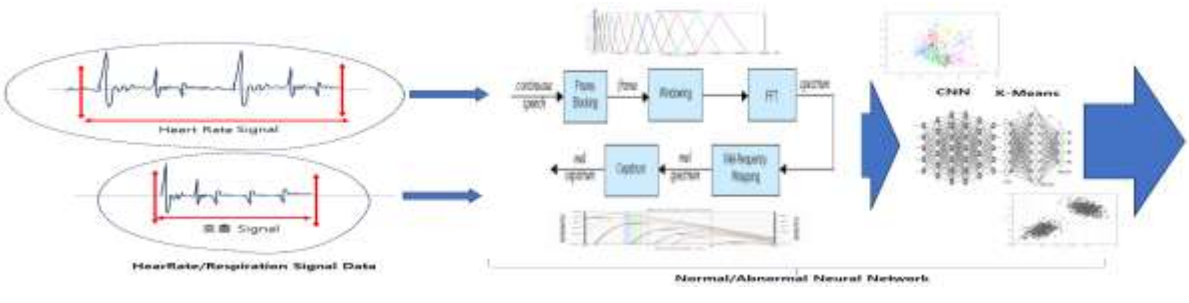


**Figure 1. Heart Rate/Respiration Monitoring and Neural Network-Based System**

This study focuses on a method to classify health status using heart rate and respiration signals by extracting MFCC (Mel-Frequency Cepstral Coefficient) features and applying K-Means clustering[6][7][8]. To achieve this, the heart rate and respiration signals were extracted as time-series waveforms, and their features were derived using MFCC, which applies a Mel-scale filter bank in the frequency domain. Subsequently, an AutoEncoder was used to reduce the dimensionality (latent representation) of the data, and K-Means clustering was performed. The study specifically investigated whether the clusters formed by K-Means could effectively distinguish between normal and abnormal data. Through experiments, it was found that the identification performance varied depending on the

[1] Gyeonggi University of Science and Technology, Gyeonggido, Korea, Email : jsd@gtec.ac.kr, ORCID: https://orcid.org/0009-0009-2443-6144

number of states (clusters). Additionally, statistical analysis of the MFCC values of heart rate and respiration signals showed that signals with larger standard deviations tend to have lower identification accuracy for the same number of states. Based on these findings, a formula was proposed to determine the number of states in accordance with the standard deviation of the biometric signals. Simulations were conducted, and the results demonstrated that the proposed method, which adapts the number of states to the signal characteristics, provides reliable classification performance.

### Characteristics of Heart Rate and Respiration

Heart rate signals are biometric signals that represent cardiac activity over time by measuring the heartbeat. They are typically measured using an electrocardiogram (ECG), with key waveform components including the P wave, QRS complex, and T wave. These signals allow for the analysis of heart rhythm, heart rate, arrhythmias, and other cardiac conditions.

Respiratory signals are biometric signals that reflect the activity of the lungs, including inhalation and exhalation. These signals are generally measured using respiratory sensors and are used to analyze parameters such as respiratory rate, breathing patterns, and depth of respiration. Key features used to characterize respiration include the respiratory rate (the number of breaths per minute), as well as patterns related to regularity, depth, and asymmetry of breathing.

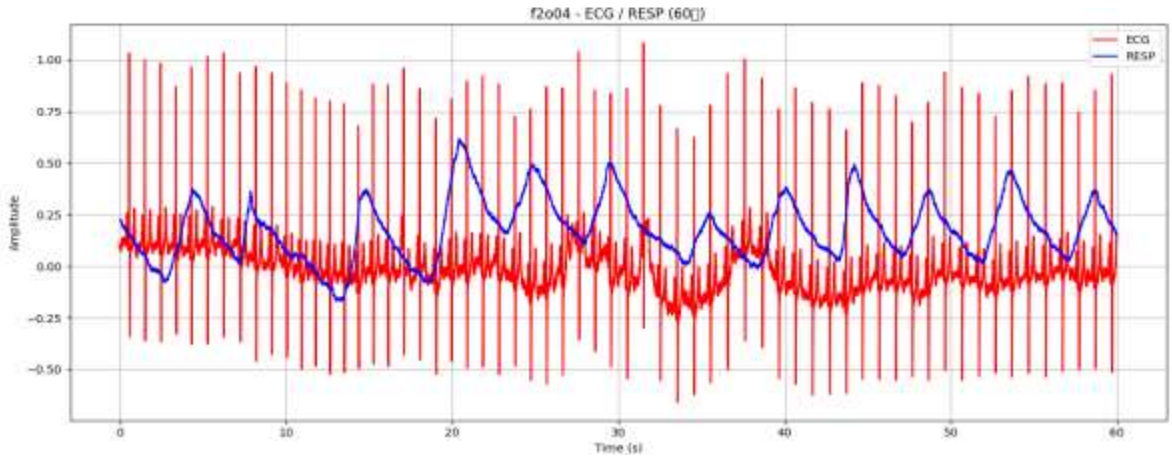


**Figure 2. Heartrate and Respiratory Signals ( Heart Rate : ECG in red, Respiration : RESP in blue).**

### Clustering of Heart Rate and Respiration Signals Using K-Means

**MFCC (Mel-Frequency Cepstral Coefficients)** is a widely used feature extraction algorithm designed to capture the characteristics of audio signals. It is based on the **Mel scale**, which reflects the human auditory perception of sound frequency. MFCC is primarily used in fields such as **speech recognition**, **emotion analysis**, and **biometric authentication**, where accurate representation of audio or physiological signals is essential

$$\mathrm{Mel(f)} = 2595 \cdot \log_{10}(1 + \mathrm{f}/700) \quad (1)$$

$$S_m = \log E_m \quad (2)$$

$$C_n = \sum_{m=1}^{m=M} S_m \cos(n\pi(m - 0.5)/M) \quad (3)$$

An **Autoencoder** is an unsupervised learning model based on neural networks that encodes input data into a low-dimensional **latent space**, and then reconstructs the original data from this representation.

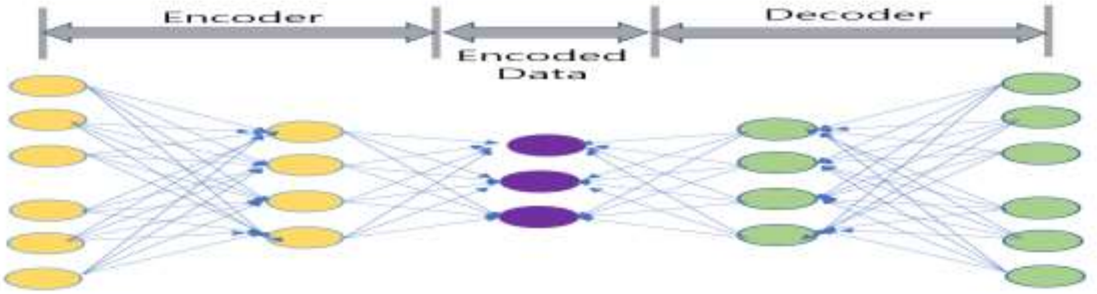


**Figure 3. Autoencoder Neural Architecture**

The **K-Means algorithm** is a clustering method that partitions a given dataset into **K predefined clusters**. Each cluster is formed by assigning data points based on their proximity to a central point called the **centroid**, and the algorithm aims to minimize the distance between the data points and their respective centroids.

$$J = \sum_{k=1}^{K} \sum_{\{x_i' \in C_k\}} ||x_i' - \mu_k||^2 \quad (4)$$

**Clustering Experiment of Heart Rate and Respiratory Data**

The data used in the experiment were standard biometric datasets downloaded from **PhysioNet**. For normal (healthy) data, sleep recordings from the **MIT-BIH Polysomnographic Database** and age-diverse recordings (ranging from young adults to the elderly) from the **Fantasia Database 1.0.0** were utilized. For data representing heart failure symptoms, recordings from the **BIDMC Database** were downloaded and used for training.

MFCC (Mel-Frequency Cepstral Coefficients) features were extracted from these biometric signals. The extracted features were then processed using an **Autoencoder** to obtain latent representations, which were subsequently clustered using the **K-Means** algorithm.

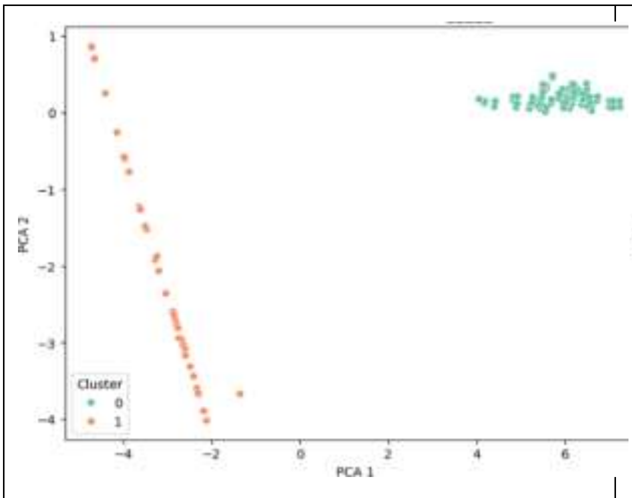


**Figure 3. State = 2, Heartbeat Clustering, PCA1 (1st MFCC order ), PCA2 (2nd MFCC order ).**

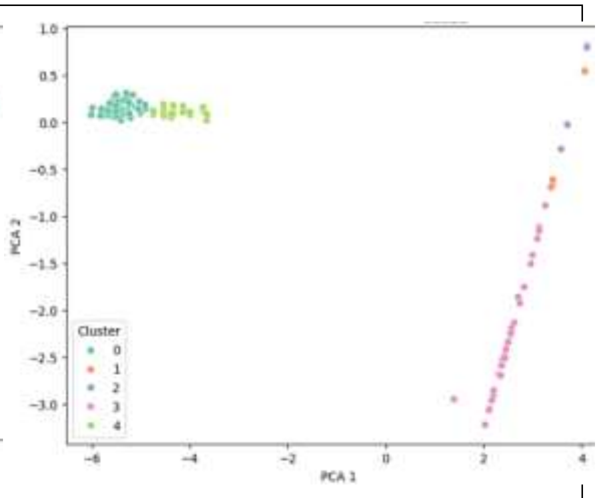


**Figure 4. State = 5, breathing Clustering, PCA1 (1st MFCC order ), PCA2 (2nd MFCC order ).**

**Table 1. Identification Accuracy of Heartbeat and Respiration Using K-Means Clustering.**

| **Number of state** | **Heartbeat Identification** | | **Breathing Identification** | | **Clustering of Respiratory Signals** |
|---|---|---|---|---|---|
| | Training Data | Untrained Data | Training Data | Untrained Data | |
| 2 | **100%** | 100% | 61.6% | 53.7% | 1(normal), 0(Heart Failure) |
| 3 | 100% | 100% | 73.3% | 58.93% | 3 (normal). 0,2 (Heart Failure) |
| 4 | 100% | 100% | 100% | 99.8% | 3(normal), 0,1,2(Heart Failure) |
| 5 | 98.8% | 98% | 100% | 100% | 3(normal), 0,1,2(Heart Failure) |

| | | | | | |
|---|---|---|---|---|---|
| 6 | 80 | 76.8 | 100% | 100% | 3,4(normal), 0,1,2(Heart Failure) |

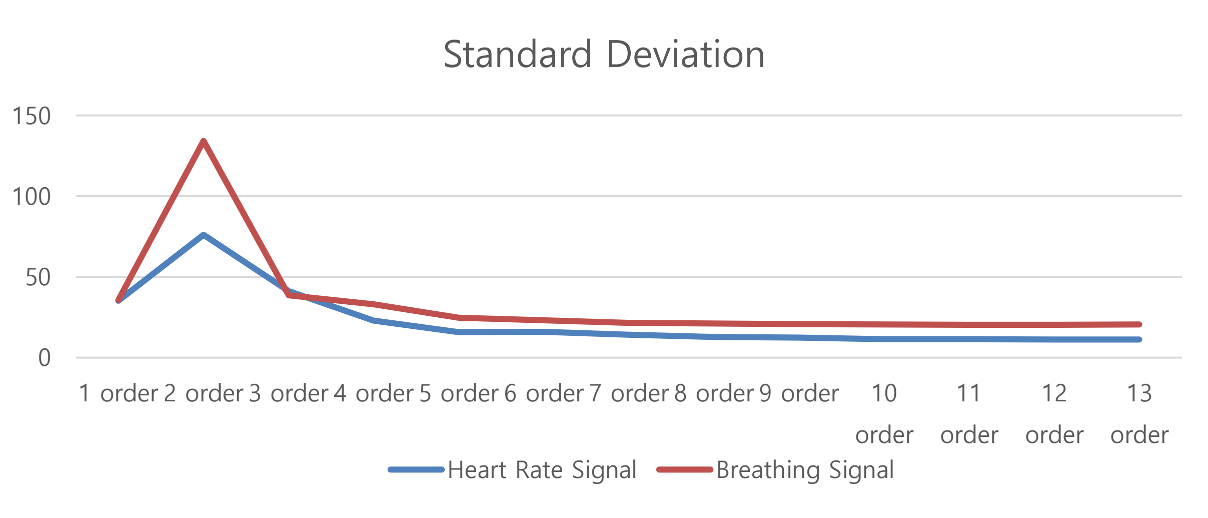


**Figure 5. Comparison of the Standard Deviations of Heart Rate and Respiratory Signal.**

### Analysis and Consideration of Experimental Results

As shown in **Table 1**, in the case of heart rate (HR) signals, using more than two states was sufficient to distinguish between normal HR and HR associated with heart failure. However, when the number of states exceeded six, the classification accuracy decreased.

For respiration signals, a sufficient classification performance was achieved when the number of states was four or more. In contrast, when the number of states was fewer than four, it was difficult to distinguish between normal and heart failure conditions due to clustering limitations. Furthermore, as illustrated in **Figure 5**, statistical analysis of each coefficient in the 13-dimensional MFCC data for both HR and respiration revealed that respiration MFCC data exhibited a larger standard deviation compared to HR MFCC data. This indicates that respiration MFCC data are more widely distributed than HR MFCC data. Consequently, the number of K-means clustering states for respiration MFCC data should be set higher than that for HR MFCC data.

In addition, as observed in the HR classification performance, an excessive number of states can lead to the separation of data points that should belong to the same state, thereby reducing classification accuracy. Therefore, it was confirmed that the number of states should be set to an appropriate level.

### Proposed Method for Estimating the Number of States in K-Means Clustering

When applying K-Means clustering, the number of states (clusters) must be predetermined. We propose a method to automatically determine the number of states based on the characteristics of the input, as shown in Equation (5). The next step involves calculating a vector of standard deviations for each order of the input signal's MFCC features, and then computing a distance D through an inverse operation. Ultimately, as expressed in Equations (6) to (8), the distance D is obtained. This distance D, which can be represented as in Equation (9), indicates the degree of uncorrelation among the MFCC data of each order and serves as a criterion for deciding whether additional clusters should be added. A threshold value is set as shown in Equation (10). When this threshold is exceeded, the distance between clusters is considered significant, implying low correlation and thus justifying the addition of new clusters, effectively increasing the number of states. This threshold is determined experimentally by verifying which value yields the optimal number of states. Finally, as described in Equation (11), the degree of standard deviation is used to assess the statistical correlation between the current cluster and potential new clusters, managing the number of states by adding clusters when necessary.

$$R(\tau) = \frac{1}{K}\sum_{k=1}^{K} (X' - \sigma^k)\cdot(X'^T - \sigma^k) \qquad (5)$$

$$D = 1 - R(r\ \ ) \qquad (6)$$

$$D_{sym} = ( D + D^T )/2 \quad (7)$$

$$D_{sym} = D( i,i ) \ for\ all\ i \quad (8)$$

$$\mathrm{dist}(A,B) = \frac{1}{|A||B|} \sum_{i\in A} \sum_{i\in B} d(i,j) \quad (9)$$

$$C_k = \{c_k(|i-j|) | dist( i,j ) \leq threshold\} \quad (10)$$

$$n_{clusters} = number(\{C_1, C_2 \ldots., C_k \}) \quad (11)$$

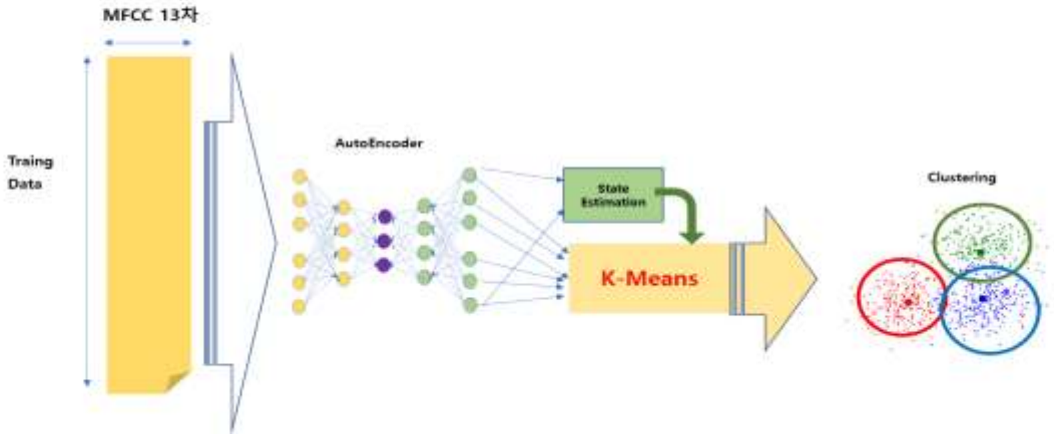


**Figure 6. Overall System Architecture for Cluster Identification of Heart Rate and Respiratory**

## Experiment

Using the proposed method, the experiment described in chapter 6 was conducted. The results showed that the number of states for the heart rate signal was calculated as 2, while the number of states for the respiratory signal was calculated as 4. In this experiment, the threshold was set to 0.05.

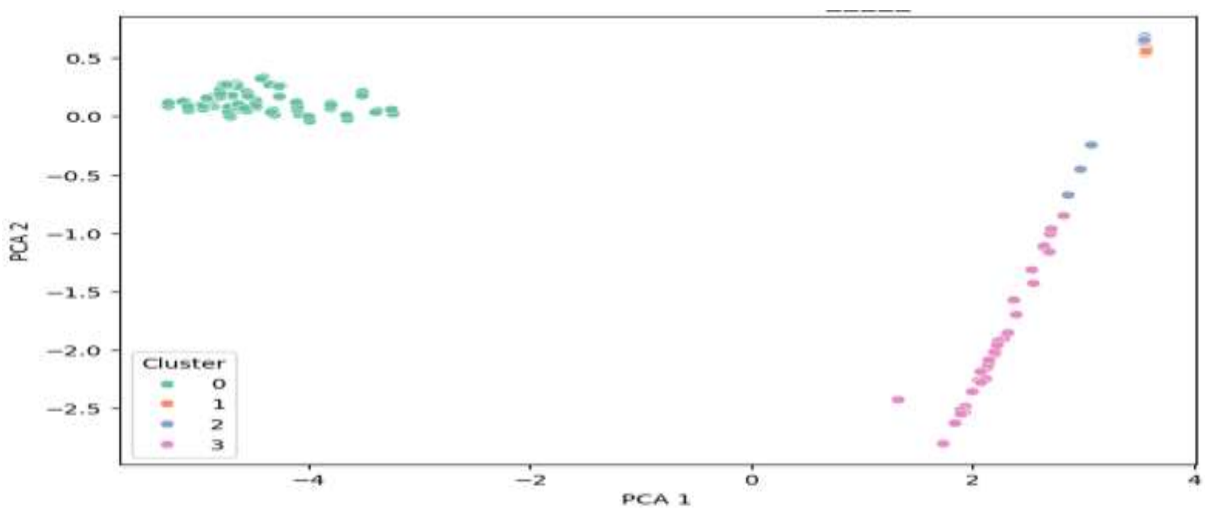


**Figure 7. Case of “State = 2, Heart Rate ”, 100% for Training Data, 100%**

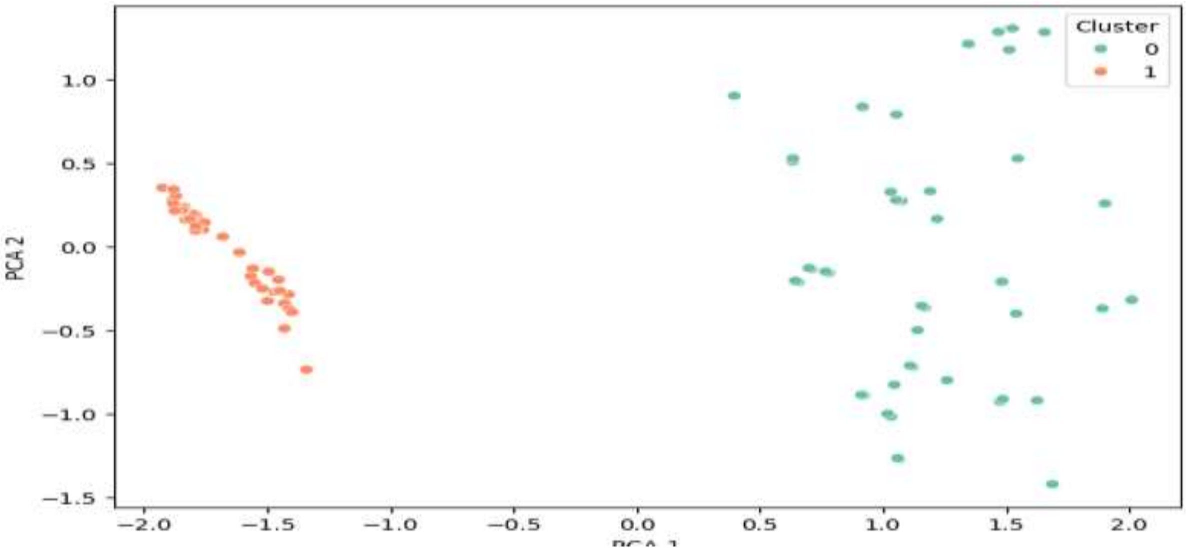


**Figure 8. Case of "State = 4, Respiration" , 100% for Training Data, 99.8% for Untraining Data.**

## Conclusion

This study investigated a method to distinguish between normal and heart failure conditions using heart rate and respiratory signals. Specifically, when clustering with K-Means, the experiments confirmed that the classification accuracy varies according to the number of clusters, which corresponds to the characteristics of the biometric signals. Based on these insights, a method was proposed to calculate the number of states (clusters) for K-Means clustering by computing a distance measure derived from the standard deviation of the MFCC features processed through an Autoencoder.

When experiments were conducted using the proposed method, it successfully determined an appropriate number of states that reflected the characteristics of the MFCC standard deviation in heart rate and respiratory signals, achieving reliable classification accuracy through K-Means clustering. The results of this study can be applied in caregiver systems for elderly or vulnerable individuals. Furthermore, this research is expected to be extended and applied to other classification algorithms, including K-NN enhanced with K-Means learning.

## References


[1] Osathitporn, P., Sawadwuthikul, G., Thuwajit, P., Ueafuea, K., Mateepithaktham, T., Kunaseth, N., … Wilaiprasitporn, T. RRWaveNet: A Compact End-to-End Multi-Scale Residual CNN for Robust PPG Respiratory Rate Estimation. arXiv preprint, 2022. DOI : https://arxiv.org/abs/2208.08672

[2] Lahr, P., Carling, C., Nauer, J., McGrath, R., & Grier, J. W. Supervised Machine Learning to Examine Factors Associated with Respiratory Sinus Arrhythmias and Ectopic Heart Beats in Adults: A Pilot Study, Hearts, 5(3), 275–287, 2024. DOI : https://www.mdpi.com/2673-3846/5/3/20

[3] Ren, Y., Syrnyk, B., & Avadhanam, N. Dual Attention Network for Heart Rate and Respiratory Rate Estimation, arXiv preprint, 2021. DOI : https://arxiv.org/abs/2111.00390

[4] Akamatsu, Y., Umematsu, T., & Imaoka, H. CalibrationPhys: Self-supervised Video-based Heart and Respiratory Rate Measurements by Calibrating Between Multiple Cameras, arXiv preprint, Oct 2023. (Self supervised dual camera HR/RR). DOI : https://arxiv.org/abs/2310.15043

[5] Md Nazmul Islam Shuzan , Moajjem Hossain ,Muhammad E. H. , Murugappan Murugappan , Enamul Hoque ,Mohamed Arslane Ayari Amith Khandakar, "Machine Learning-Based Respiration Rate and Blood Oxygen Saturation Estimation Using Photoplethysmogram Signals," Biosensors, 10(2):167, 2022. (GPR + feature selection to estimate RR & $SpO_2$ from PPG). DOI : https://pmc.ncbi.nlm.nih.gov/articles/PMC9952751/

[6] Mingu Kang, Siho Shin, Jaehyo Jung, Youn Tae Kim. "Stress Classification Using K-means Clustering and Heart Rate Variability from Electrocardiogram." International Journal of Biology and Biomedical Engineering, 2020, 14: 251–254. DOI : https://www.researchgate.net/publication/348577372_Stress_Classification_Using_K-means_Clustering_and_Heart_Rate_Variability_from_Electrocardiogram

[7] Mohammad Adiban et al. "Statistical feature embedding for heart sound classification." arXiv (2019). DOI : https://arxiv.org/abs/1904.11914

[8] Kun Hsi Tsai et al. "Blind Monaural Source Separation on Heart and Lung Sounds Based on Periodic Coded Deep Autoencoder." arXiv (2020). DOI : https://ieeexplore.ieee.org/document/9167389.