A Novel Approach to GSR modelling via ECG labelling

Obaid Tamboli 20198

***Abstract*—This project involves the development of a device equipped with ECG and GSR sensors to monitor sleep quality. The ECG data obtained from the device was processed using the Hearty library to extract various parameters of the ECG waveform and rhythm, such as beats per minute, inter-beat interval, standard deviation of intervals between adjacent beats, and more. Ground truth labels were generated using the available ECG data, and a model was developed for GSR to predict sleep quality and heart rate based on these labels. The results of this project could have significant implications for sleep monitoring and analysis and labelling GSR data.**

Keywords*— *GSR, ECG, Modelling, Sleep Quality

I. PROBLEM FORMULATION

The problem formulation for this project stemmed from the fact that ground truth data for GSR is not readily available, and wearing an ECG for sleep quality monitoring can be intrusive. On the other hand, GSR can be non-intrusive, making it an ideal candidate for monitoring sleep quality. Hence, the project aimed to develop a model that maps ECG data to GSR data, which can be used to predict sleep quality.

To achieve this, the team modulated the modeling approach by first converting ECG data to beats per minute and categorizing the data into labels {0, NA}, {1, very-poor}, {2, good}, and {3, excellent} based on BPM values. These labels were then used to model the variations in GSR data.

It is noteworthy that GSR, which falls under the umbrella term of electrodermal activity (EDA), refers to changes in sweat gland activity that are reflective of the intensity of our emotional state, otherwise known as emotional arousal. Our level of emotional arousal changes in response to the environment we're in, and both positive and negative stimuli can result in an increase in arousal and skin conductance. The GSR signal is, therefore, not representative of the type of emotion, but the intensity of it.

On the other hand, the ECG signal reflects the electrical activity of the heart observed from the strategic points of the human body and represented by quasi-periodic voltage signal. The ECG signal contains essential information about the cardiac pathologies affecting the heart, characterized by five peaks known as fiducial points, which are represented by the letters P, Q, R, S, and T. The QRS complex is the depolarization of the right and left heart ventricles, which is used as a reference point for signal analysis. The diagnosis of the signal relies on the morphology of the waves, as well as the duration of each peak and the segments that make it up.

Therefore, detection of each section of the ECG signal is essential for health professionals in screening, diagnosis, and monitoring of several heart conditions. However, the team in this project aimed to use ECG data to predict GSR data, which can then be used to monitor sleep quality non-intrusively. The development of such a model has significant implications for sleep monitoring and analysis.

II. DATA COLLECTION AND SETUP

A. Connection Setup

To summarize, both the GSR-Groove sensor and the AD8244 ECG sensor were connected to an ESP32 microcontroller. The ground pin and input pin of both sensors were connected in parallel to the ground and 3V pins of the ESP32, respectively. The output of the AD8244 was collected on GPIO 15, while the low pass filter (lo+ev) and high pass filter (lo-ev) outputs were connected to pins 16 and 17, respectively. For the GSR sensor, a GPIO pin was used to read the analog signal and convert it to a digital value. This setup allowed for the collection of both GSR and ECG data for use in a sleep quality monitoring project. Ref Fig.1.

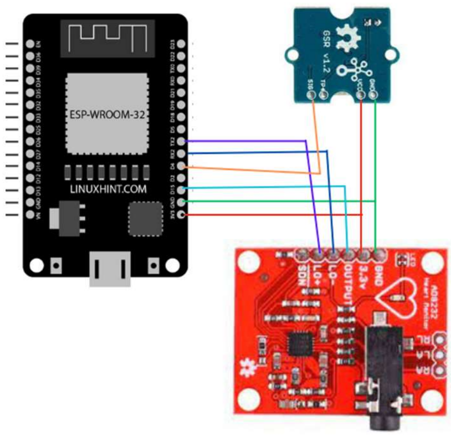

Fig. 1. Connection Diagram ESP-ECG-GSR

B. Data Collection

To further expand on the data collection process, after the GSR and ECG sensors were connected to the ESP32 microcontroller, an Arduino program was written in the Arduino IDE. The program utilized the 'analogRead()' function to read the analog signals from the sensors and convert them to digital values at a sampling rate of 125Hz. These values were then printed to the serial communication (COM) port at a predefined baud rate.

To capture and save the data, a Python program was written using the 'pyserial' library to listen on the serial terminal on the COM port. The Python program was able to receive the data and store it in a comma-separated values (CSV) file for further processing and analysis.

This setup allowed for the collection of GSR and ECG data in real-time, via COM, which was then saved to a file for use in developing and testing the sleep quality monitoring model. The collected data could also be used for further analysis and exploration of the relationship between GSR and ECG signals and their potential applications in other fields.

III. Plots, Results and Discussion

To further process and analyze the ECG data collected from the AD8244 sensor, the Heartpy library was used. This library provides functions for filtering and analyzing ECG data to extract important parameters such as heart rate, interbeat intervals, and various statistical measures such as standard deviation and median absolute deviation.

One common issue with ECG data is the presence of noise due to interference from other sources. The Hearty py library provides filtering functions to remove this noise and improve the accuracy of the data. Once the data was filtered, it was possible to extract various parameters from the ECG waveform and analyze them.

The library also provides functions for visualizing the ECG data, allowing for easier interpretation and analysis of the data. This visualization included plotting the ECG waveform, as well as other measures such as the Poincare plot, which helps to identify patterns in the interbeat intervals.

Overall, the Heartpy library was an important tool in processing and analyzing the ECG data collected from the AD8244 sensor. It allowed for the extraction of important parameters and filtering of noisy data, as well as visualization of the data to aid in analysis and interpretation.

The results of the ECG data analysis using the Heartypy library are as follows:

- Beats per minute (BPM): 72.443814
- Interbeat interval (IBI): 828.228070
- Standard deviation if intervals between adjacent beats (SDNN): 93.389554
- Standard deviation of successive differences between adjacent R-R intervals (SDSD): 46.946975
- Root mean square of successive differences between adjacent R-R intervals (RMSSD): 75.153177
- Proportion of differences between R-R intervals greater than 20ms (pNN20): 0.718310
- Proportion of differences between R-R intervals greater than 50ms (pNN50): 0.492958
- Median absolute deviation (MAD): 73.333333
- SD1: 53.052717
- SD2: 121.651510
- S: 20275.660166
- SD1/SD2: 0.436104
- Breathing rate: 0.166667

These parameters provide information about the heart rate variability, which can be used to evaluate the quality of sleep. The BPM, IBI, SDNN, SDSD, and RMSSD are measures of the variation in time between successive heartbeats, while pNN20 and pNN50 are measures of the proportion of time between successive heartbeats that varies by more than 20ms and 50ms, respectively. MAD is a measure of the variability of the data, while SD1 and SD2 are measures of the dispersion of the data in different directions. The S parameter is a measure of the total scatter of the data. The breathing rate provides additional information about the physiological state during sleep.

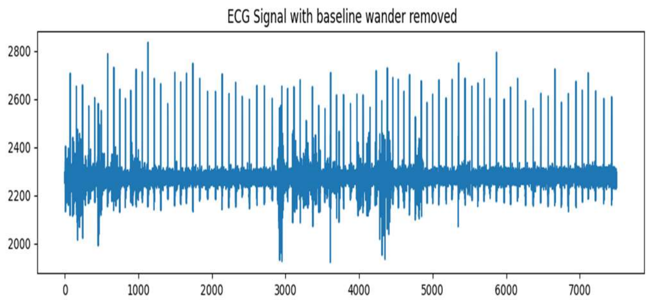

Fig. 2. Filtered ECG Data for 1 sec interval

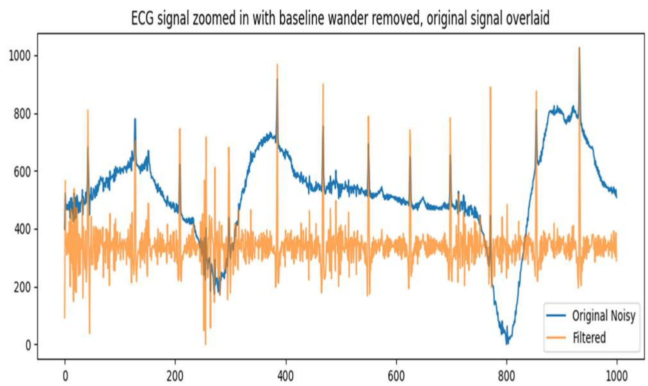

Fig. 3. Filtered ECG Signal overlaid on the Original Noisy Data

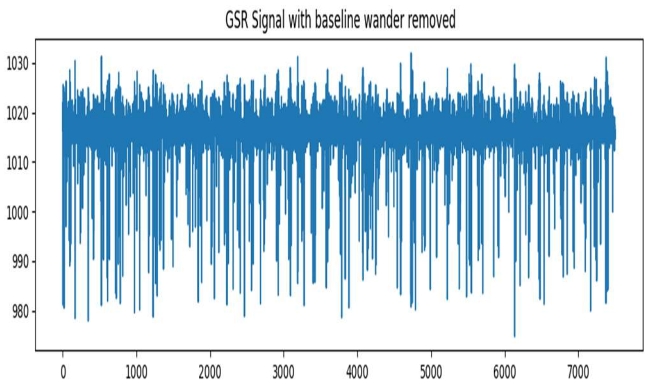

Fig. 4. Filtered GSR Data for 1 sec Interval

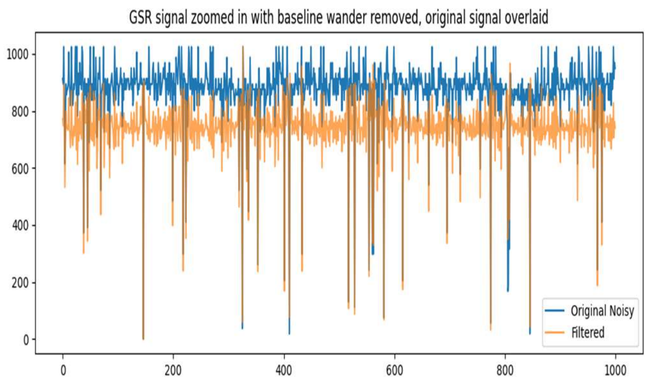

Fig. 5. Filtered GSR Signal overlaid on the original Noisy Data

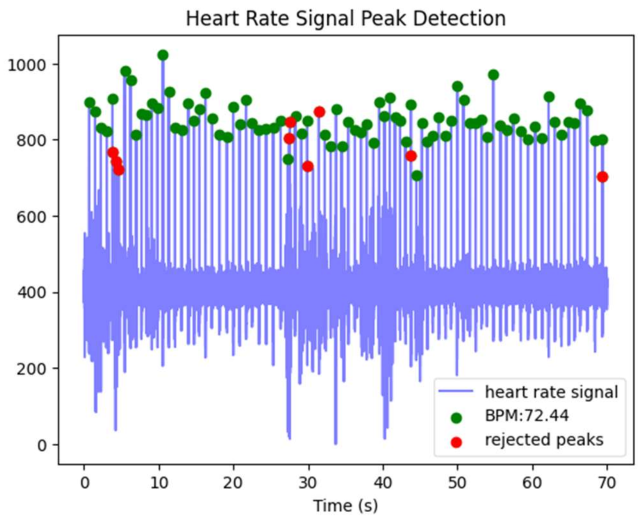

Fig. 6. Beats per Minute detection plot with Heartpy with irrevelant data identified.

The identification of R-peaks is crucial as they represent the depolarization of the ventricles, which is used as a reference point for signal analysis. During the peak identification process, some peaks may be detected incorrectly due to various reasons such as noise, baseline drift, or other artifacts. These false peaks can affect the accuracy of the ECG analysis and can lead to wrong conclusions. Hence, it is essential to identify and remove these false peaks to obtain reliable results.

The plot showing the accepted and rejected peaks helps to visualize the peak identification process and enables us to identify the false peaks. The accepted peaks represent the reliable peaks that are accurately detected, while the rejected peaks represent the false peaks that are detected incorrectly.

By examining the plot, we can identify the false peaks and remove them, ensuring that only the accurate and reliable peaks are used for further analysis. This helps to improve the accuracy and reliability of the ECG analysis, and enables us to obtain accurate results.

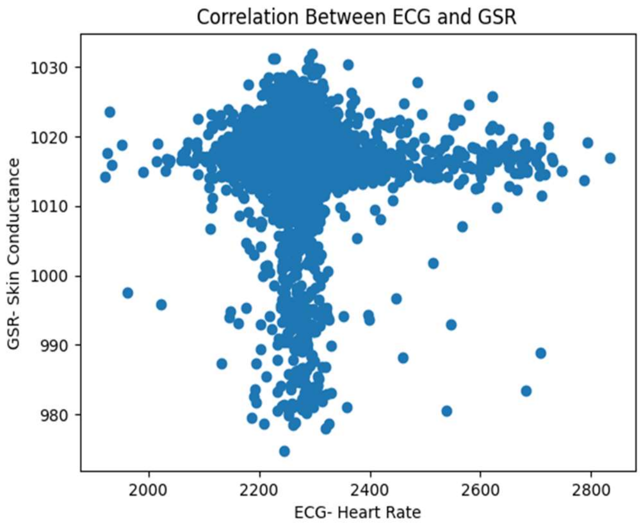

Fig. 7. Scatter Plot between ECG and GSR

From the correlation plot between ECG and GSR, we can infer that there is a correlation between heart rate and skin conductance. As the heart rate increases, the skin conductance also tends to increase and vice versa. This is because the galvanic skin response (GSR) is reflective of the intensity of emotional arousal, and our emotional state changes in response to the environment we are in, which affects our heart rate.

We can also observe that at the peak value of the ECG signal, the GSR is at its maximum, which indicates a high level of emotional arousal. On the other hand, during the diastole phase, the GSR is at its lower bounds, which indicates a lower level of emotional arousal.

These observations can be utilized for sleep quality prediction by building a model that maps ECG to GSR and categorizes the GSR values based on the heart rate. We can use the labels {0, NA}, {1, very-poor}, {2, good}, {3 excellent} to model the GSR variations. By doing so, we can predict the sleep quality of an individual based on their ECG and GSR data.

To build a model for GSR ground truth labels, we can use machine learning algorithms such as regression, decision trees, or neural networks to train the model. We can also use feature engineering techniques to extract useful features from the ECG and GSR data, such as peak detection and spectral analysis, to improve the accuracy of the model. Once the model is trained, it can be used to predict the sleep quality of individuals and provide insights into their emotional state during sleep.

References

[1] https://github.com/obaidtambo/IOT

[2] https://github.com/paulvangentcom/heartrate_analysis_python.

[3] Khizra Saleem, Imran Sarwar Bajwa, Nadeem Sarwar, Waheed Anwar, Amna Ashraf, "IoT Healthcare: Design of Smart and Cost-Effective Sleep Quality Monitoring System", *Journal of Sensors*, vol. 2020, Article ID 8882378, 17 pages, 2020. https://doi.org/10.1155/2020/8882378