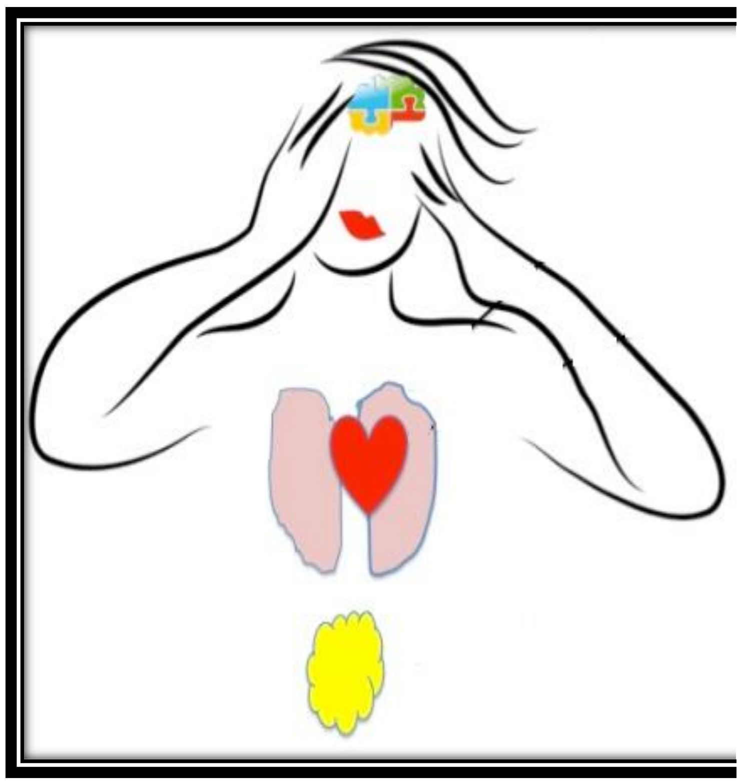
A Low-Cost, Long-Term, Contactless Identification System of Cognitive Stress Using Machine Learning and Physiological Parameters



Shifa Somji
NEWPORT HIGH SCHOOL

Table of Contents

Project Summary

Stress has become prevalent in today's American society. Money, work, and the economy are the most common causes of stress for Americans. In a recent survey, thirty-nine percent of the participants said that their stress had increased over the past year and forty-four percent said that their stress had increased over the last five years. Stress is one of the leading causes of diseases that range from cardiovascular disease to depression to drug abuse. According to American Psychological Association, stress causes fifty-four percent of Americans to fight in their families. Stress costs more than $300 billion each year in health care, missed work, employee turnover, legal expenses, workers' compensation and insurance (Stress in America: Our Health at Risk). While there is a growing understanding of stress and its effects on the overall health of the affected populace, this knowledge does not translate into better prevention or management of stress. The lack of active stress management is centered around available equipment needed to measure and diagnose stress. This equipment often requires blood, urine and saliva samples from the patient and assistance of a trained health care professional. In this project, my objective is to design and implement a low-cost, long-term, contactless stress identification system, which can operate without any professional intervention. My system consists of three parts:

i. Doppler radar module
ii. Custom designed electronic circuit board with NI-USB 6008 DAC
iii. Software Application comprising of Custom Signal Processor and Machine Learning Classifier

I utilize an off-the-shelf K-band (24 GHz) short-range Doppler radar module to track human physical movement, heart rate, and breathing rate. On the electronic circuit board, the output signals from the Doppler radar module are amplified by an op-amp, with a set gain of 11 times, housed inside NI-USB 6008 DAC (data-acquisition-device). After amplification, an ADC, housed inside NI-USB DAC, of 12-bit resolution is utilized to sample the signal at 5 KHz. The sampled signal is transmitted through a USB cable to a personal computer, where a running software application saves and analyzes the captured data. The custom signal processor filters the captured data, through signal processing techniques, to extract heart rate and breathing rate signals of the subject. It then computes various time and frequency domain parameters of mean RR (average heart rate variability), mean SD (standard deviation of RR), square root of the mean of the sum of the squares of differences between adjacent RR intervals (RMSSD), mean NN50 (number of consecutive RR intervals that differ more than 50ms), mean pNN50 (proportion of NN50), mean HR (average heart rate) and mean LF/HF (LF: 0.04 to 0.15 Hz, HF: 0.15 to 0.4 Hz), a common measure of sympatho/vagal balance, as features for machine learning classification.

The computed time and frequency domain parameters, listed above, are used as features into a machine learning classifier to measure human stress in real time. I trained and evaluated the classifier using test dataset collected with myself as a subject. I collected four 150 seconds' datasets. During the initial 150 seconds' interval, no dataset was collected while I relaxed by listening to my favorite music. In recording the first and second sets, I observed myself during my normal state through a pulse oximeter, utilized to establish the ground truth, and through my system, utilized to establish the baseline. In recording the third set, I placed myself under cognitive stress through stress-inducing techniques like the Berg Card Sorting Task (BCST), mental arithmetic, recitation of the alphabets backward, estimation, number series, and analogies. Finally, I utilized the fourth set to collect another dataset of myself experiencing cognitive stress. I used the second and the third datasets to train the classifier and the fourth set to evaluate the classifier. The results show that by integrating physiological parameters collected in a contact-free manner, I was able to achieve 92.9% recall (Naïve Bayes) and 100% recall (SVM) in predicting whether a given subject is under cognitive stress or not. The total cost of parts of my design is under $65. The modular and simple design of the system makes it efficient to measure cognitive stress in a contact-free manner in many settings such as a car dashboard for driver stress and on the back of a workplace chair for work related stress.

Theory

i. Fundamentals of Doppler Radar

According to Doppler theory, a subject with a time varying position will reflect the signal with its phase modulated proportionally to the time-varying target position. For example, a Doppler radar with the chest-wall as the target will receive a signal similar to the transmitted signal but with its phase modulated by the time-varying chest-wall position. If the heartbeat and breathing rate signals are to be monitored, demodulating the phase will then give a signal proportional to the chest-wall position that contains information about movement due to heartbeat and respiration, from which heart and breathing rates can be determined. Based on this principle, a non-contact heartbeat and breathing rate monitor can be envisioned (Rahman, Adams and Ravichandran). The fundamental principle of detecting vital signals using continuous wave Doppler radar is demonstrated in Figure 1 below.

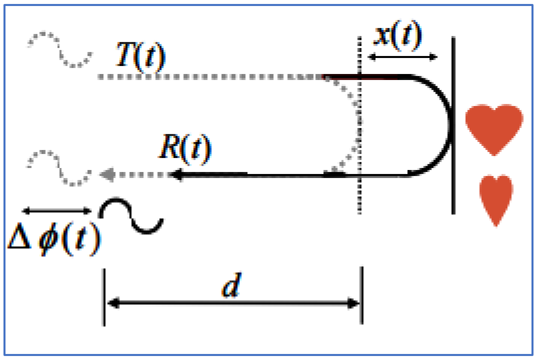


Figure 1: Doppler Effect (http://etd.fcla.edu/UF/UFE0017533/xiao_y.pdf)

The module transmits a single tone $T(t)$ on a carrier frequency f, wavelength $\lambda = c/f$, combined with phase noise $\phi(t)$ from the oscillator, approximated by Equation 1:

$$T(t) = \cos\left(2\pi f t + \phi(t)\right)$$

Equation 1

When the signal $T(t)$ is reflected back by a target, given by $R(t)$, in this case the subject's body generating periodic chest movements due to respiration, given by $x_r(t)$, and heartbeat, given by $x_h(t)$, at a distance d. As a result, the reflected signal received by the radar is approximated by Equation 2:

$$R(t) = \cos\left(2\pi f t - \frac{4\pi d}{\lambda} - \frac{4\pi x(t)}{\lambda} + \phi\left(t - \frac{2d}{c}\right)\right)$$

Equation 2

Where $x(t) = x_r(t) + x_h(t)$, c is the speed of light, λ is the signal's wavelength. More importantly, the information of $x(t)$ is phase modulated in $R(t)$ in addition, to the distance between the human body and the radar, d, and a time delayed version of the phase noise $\phi(t - \frac{2d}{c})$.

The received signal, $R(t)$ is similar to the transmitted signal, $T(t)$, but has a time delay determined by the distance of the target and a phase modulation due to the periodic motion of the target. Periodic chest wall motion can thus be retrieved if the received signal is multiplied by a copy of the transmitted signal. The resulting signal will be free of any information related to the carrier frequency and will preserve the change in phase to the signal corresponding to $x(t)$. A Doppler radar will split $R(t)$ into two baseband signals, $B_i(t)$ and $B_q(t)$, 90^o out of phase with each other, which are given by Equation 3:

$$B_i(t) = \cos\left(\theta + \frac{\pi}{4} + \frac{4\pi x(t)}{\lambda} + \Delta\phi(t)\right)$$

$$B_q(t) = \cos\left(\theta - \frac{\pi}{4} + \frac{4\pi x(t)}{\lambda} + \Delta\phi(t)\right)$$

Equation 3

The portion of interest, in the above baseband signal, therefore, is the phase modulation due to physiological movements $x(t)$ given by $\frac{4\pi x(t)}{\lambda}$. θ contains the target distance d information and $\Delta\phi(t)$ represents residual phase noise. Since θ is generally constant, the baseband signal can be approximated as given by Equation 4:

$$B(t) \approx \frac{4\pi x(t)}{\lambda} + \Delta\phi(t))$$

Equation 4

In order to retrieve the needed chest-wall motion $x(t)$ accurately from the final baseband signal, residual phase noise would need to be carefully filtered in the custom signal processor during the system design phase.

ii. The Electrocardiogram

The electrocardiogram (ECG) is a visualization of the electrical activity in the heart. It is widely considered the golden standard for calculating heart rate and heart rate variability (HRV) and used in numerous medical applications. The wave shapes seen on an ECG represents net electrical pulses observed by each of the ECG leads. Figure 2 below shows an ECG shape with labeled peaks and waves as well as the state of the heart valves during the cycle:

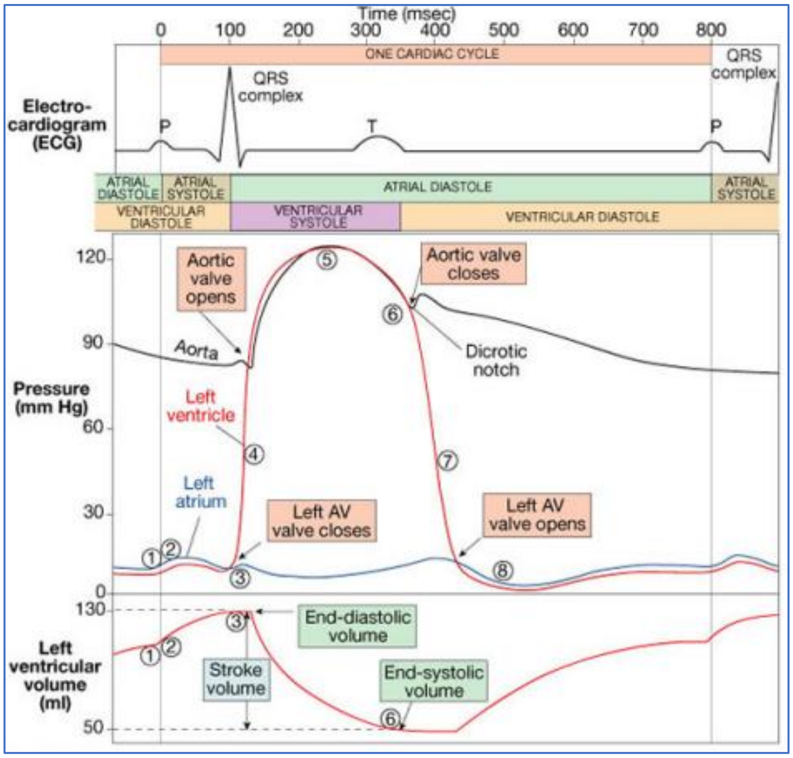


Figure 2: Polarization and Depolarization Electrical Activity in the Heart as seen on an ECG (https://scholar.sun.ac.za/bitstream/handle/10019.1/86649/koegelenberg_application_2014.pdf)

Figure 3 shows the physical and electrical activity of the heart rate to one another.

1. De-oxygenated blood from the body enters the right atrium from the superior and inferior vena cavae as atrial depolarization is started.
2. Once the atrium is completely depolarized, the atrium is full and contracts.
3. Ventricular depolarization then starts and the blood is pushed into the right ventricle. The atria start repolarizing at this point.
4. The tricuspid valve then briefly open to allow the blood through. Once closed, the valve prohibits the blood from flowing back into the atrium. During ventricular depolarization, the ventricles contract pushing the blood out of the pulmonary arteries.
5. The pulmonary valve allows the blood to flow out to the arteries but stops any flow in the opposite direction. The ventricles now repolarize.

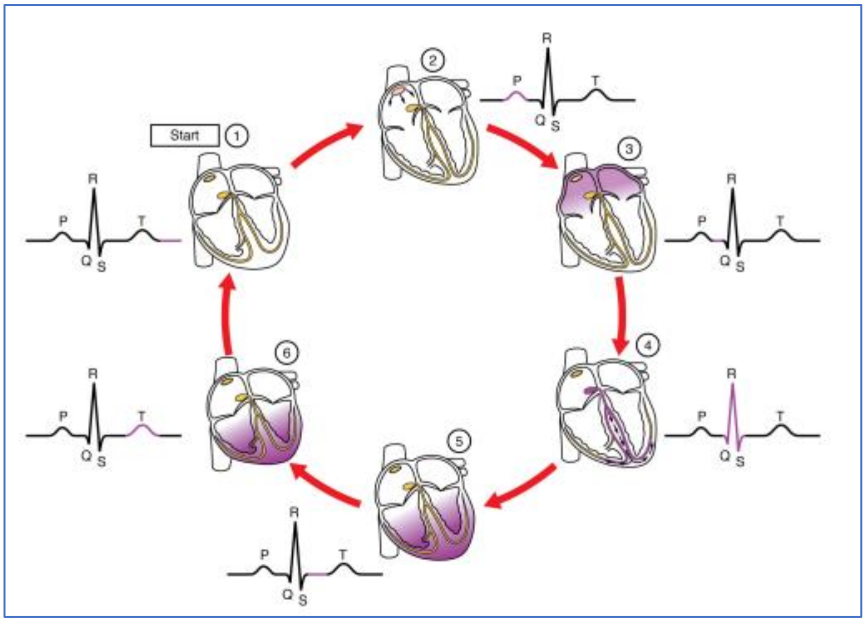


Figure 3: An ECG trace with synchronized timing of heart sounds (https://scholar.sun.ac.za/bitstream/handle/10019.1/86649/koegelenberg_application_2014.pdf)

iii. Human Body in Stress

When the human body is exposed to cognitive stress, relevant part or parts of the body send a signal to the amygdala, an area of the brain that contributes to emotional processing. When the amygdala perceives stress, it instantly sends a distress signal to the hypothalamus.

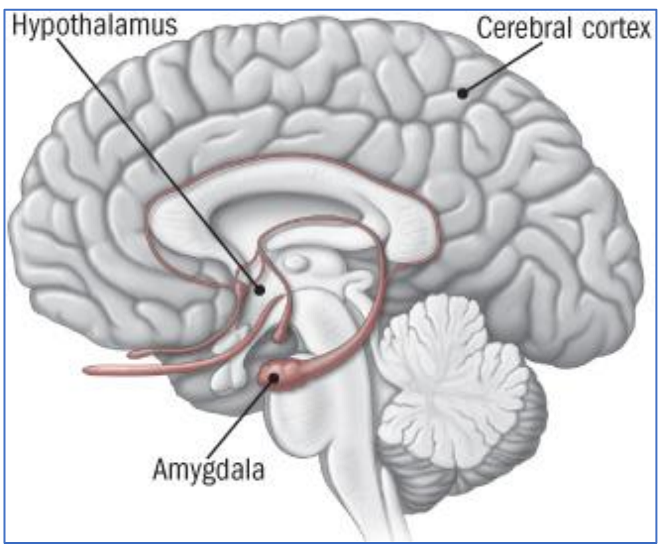


Figure 4: Command Center: illustration of brain showing areas activated by stress (http://www.health.harvard.edu/staying-healthy/understanding-the-stress-response)

The hypothalamus acts as a command center. It communicates with the rest of the body through the autonomic nervous system (ANS), which controls body functions such as breathing, blood pressure, and heartbeat. The ANS has two branches, the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). The SNS triggers the “fight-or-flight” response while the PNS promotes the "rest and digest" response that calms the body down after the danger has passed (ThePaleoMom).

When amygdala sends a distress signal, the hypothalamus engages the SNS by sending signals through the autonomic nerves to the adrenal glands, found above the kidneys. These glands respond by pumping the hormone epinephrine into the bloodstream. As epinephrine circulates through the body, it brings on a number of physiological changes. The heart beats faster than normal, reducing *beat-to-beat differences of the heart (HRV).* Pulse rate and blood pressure go up. The person undergoing these changes also starts to *breathe more rapidly* (Unknown).

After the initial surge of epinephrine subsides, the hypothalamus activates the second component of the stress response system, known as the HPA axis. This network consists of the hypothalamus, the pituitary gland, a protrusion off the bottom of the hypothalamus at the base of the brain, and the adrenal glands. The HPA axis relies on a series of hormonal signals to keep the sympathetic nervous system engaged. If the brain continues to perceive something as dangerous, the hypothalamus releases corticotrophin-releasing hormone (CRH), which travels to the pituitary gland, triggering the release of adrenocorticotropic hormone (ACTH). This hormone travels to the adrenal glands, prompting them to release cortisol. *Cortisol* has a huge range of effects in the body, including controlling metabolism, affecting insulin sensitivity, affecting the immune system, and even controlling blood flow. It has often been referred to as the most important barometer of human stress. Finally, when the danger subsides, the PNS dampens the entire stress response, including the HPA Axis (Unknown).

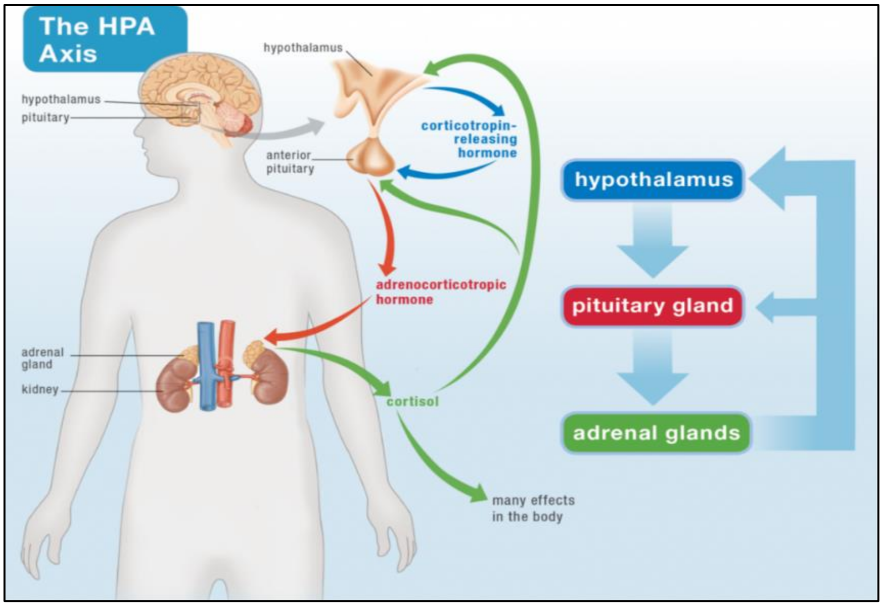


Figure 5: The HPA Axis (https://www.thepaleomom.com/stress-undermines-health)

Stress responses of the human body, therefore, are initiated from the brain via two pathways, i.e. the autonomic nervous system (ANS) and the HPA axis. *Heart rate variability (HRV) and breathing rate (BR) have been demonstrated as accurate measures of ANS activity*. While an accurate measurement of cortisol levels in humans can only be determined through saliva samples, spectral analysis of HRV has shown promise in detecting HPA axis activity. In particular, *the ratio of HRV low frequency (LF: 0.04 to 0.15 Hz) component and HRV high frequency (HF: 0.15 to 0.4 Hz) component has shown a positive correlation to the level of cortisol in the human body* and as a significant measure of HPA axis activity (Sauter). As a general rule, HRV LF component has been correlated to baroreflex (blood pressure maintenance) and the PNS activities in the human body, while, HRV HF power component has been correlated to respiratory sinus arrhythmia (RSA) and activity from the vagus nerve. The relative amounts of LF and HF power, expressed as a ratio, is generally understood as a measure of the balance between sympathetic and parasympathetic nervous system activity (Salai, Vassanyi and Kosa).

iv. Machine Learning

Classification of a dataset is a machine learning problem where a data sample can belong to one of two or more classes. While machine learning algorithms can deal with *more than two* classes, this paper focuses on a binary classification problem where the two possible classes are stress or non-stress samples. In employing machine learning algorithms, training data samples are obtained from a large number of participants. Machine Learning classifiers are trained using these data samples and, once fully trained, are tested with a small set of test data samples to determine accuracy and other related performance metrics (Koegelenberg). Some of the salient performance metrics are as follows:

a. Precision

Precision is the number of True Positives divided by the number of True Positives and False Positives. Precision can be thought of as a measure of a classifiers exactness. A low precision can also indicate a large number of False Positives. Precision is expressed as Equation 5*:*

$$Precision = \frac{t_p}{t_p + f_p}$$

Equation 5

b. Recall

Recall is the number of True Positives divided by the number of True Positives and the number of False Negatives. Recall can be thought of as a measure of a classifiers completeness. A low recall indicates many False Negatives. Recall is expressed as Equation 6*:*

$$Recall = \frac{t_p}{t_p + f_n}$$

Equation 6

c. Confusion Matrix

A clean way to present the prediction results of a classifier is to use a use a confusion matrix. For a binary classification problem, the table has 2 rows and 2 columns. Across the top is the observed class labels and down the side are the predicted class labels. Each cell contains the number of predictions made by the classifier that fall into that cell.

	Positive	***Negative***
Positive	*True Positive*	*False Positive*
Negative	*False Negative*	*True Negative*

Table 1

d. Accuracy

Accuracy measures how well a binary classification test correctly identifies or excludes a condition. That is, the accuracy is the proportion of true results (both true positives and true negatives) among the total number of cases examined. Accuracy is expressed as Equation 7*:*

$$Accuracy = \frac{t_p + t_n}{t_p + f_p + t_n + f_n}$$

Equation 7

e. Naïve Bayes

Naïve Bayes is a machine learning algorithm which employs the Bayes theorem to derive the probability that a given feature vector is associated with a specific class. The algorithm naively assumes that there is independence between every pair of features (Koegelenberg).

f. Support Vector Machines

Support Vector Machines (SVM) is a machine learning algorithm proposed by Cortes and Vapnik. It learns a non-linear function by a combination of linear mappings in high dimensional feature space. Support vectors are especially good in high dimensionality spaces as they are independent of the input space dimensionality (Koegelenberg).

System Design

My system consists of three parts:

i. Doppler radar module
ii. Custom designed electronic circuit board with NI-USB 6008 DAC
iii. Software Application comprising of Custom Signal Processor and Machine Learning Classifier

The Doppler radar module tracks physiological parameters such as breathing rate and heart rate along with unavoidable signals of human physical movements. It is an off the shelf component connected to a custom designed electronic circuit board on a simple USB port. The electronic circuit board houses a data acquisition device which amplifies and digitizes the signal from the radar module and transmits it to a software application running on any smartphone, tablet or PC.

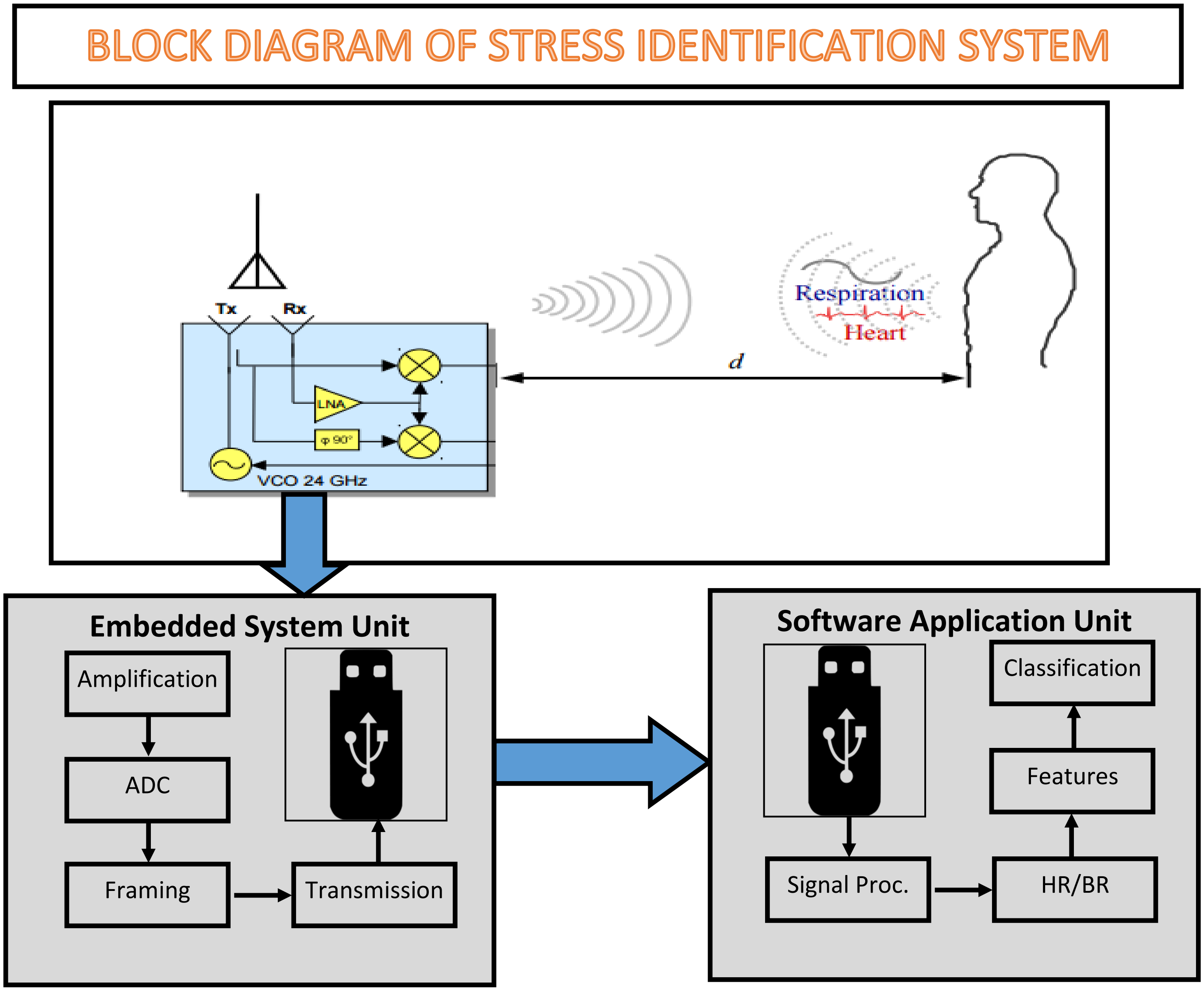
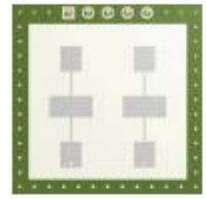


Figure 6: Stress Identification System

i. Doppler Radar Module

Doppler Radar is the primary component of the system. While the concept of contactless signal monitoring system using microwave signals has been around for a long time, a lot of research has been done around choices of radio frequencies, from 1150 MHz to detect vital signs underneath concrete to 27 GHz to improve sensitivity of detecting really faint signals as well as antenna architecture to compensate for phase noise in transmission and receipt of microwaves (Xiao). In this system, since heart and breathing signals are faint, I needed a radar component that can detect vital signs at a distance of 3m with high accuracy. Based on prior research (Rahman, Adams and Ravichandran), I selected a K band (24 GHz) RADAR module with following specifications:

- 24 GHz short range transceiver
- Beam aperture 80°/34°
- 150 MHz typical sweep rate
- High sensitive LNA receiver
- Low cost design
- Compact size: 25mm x 25mm x 6mm



K-LC5 actual size

ii. Electronic Circuit Board

A custom designed electronic circuit board is the next component of the system. While the Doppler radar module can also amplify the baseband physiological signal, the signal needed further amplification before analog to digital conversion by an ADC. Since the aim of the electronic circuit board is to amplify and sample the signal before transmitting the converted digital signal to the software application, I researched and settled on a low-cost bus-powered multifunction DAC (data-acquisition-device) manufactured by National Instruments, NI-USB 6008 DAC, which also provides a simple USB interface. The block diagram of the DAC is as follows:

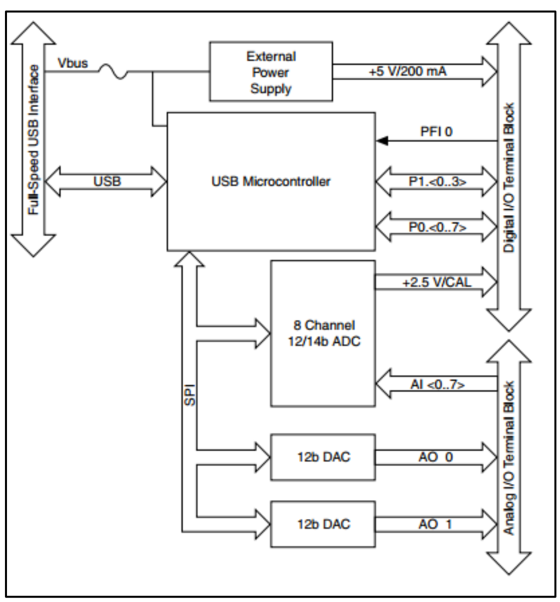


Figure 7: NI USB-6008/6009 Block Diagram (http://www.ni.com/pdf/manuals/371303n.pdf)

The DAC amplifies the signal with a set gain of 11 times. After amplification, an ADC, housed inside NI-USB DAC, of 12-bit resolution is utilized to sample the signal at 5 KHz. The sampled signal is transmitted through a USB cable to a personal computer, where a running software application saves and analyzes the captured data

iii. Software Application

The software application synthesizes the captured data from the previous two components of the system and measures a subject's stress in real time. It comprises of two distinct components, a custom signal processor and a machine learning classifier.

a. Custom Signal Processor:

Before the signal processing pipeline can be invoked, the processor, running as a MATLAB module, surveys all USB devices attached to the system. It opens up a session to the attached NI-USB DAC and sets the sampling rate as well as the length of the sampling window. Once it is able to get the requisite data from the DAC, it saves the data in a text file for further processing.

The first segment of the custom signal processor reopens the previously saved text file and applies a Chebyshev filter to filter out noise and extract heart rate and breathing rate signals of the subject. For extracting breathing rate accurately, I used a type 1 order 2 Chebyshev filter with frequency bounds of 0.1 Hz and 0.5 Hz while for extracting heart rate signal, I settled on a type 1 order 1 Chebyshev filter with frequency bounds of 1 Hz and 2 Hz. Even though this filtering is done in frequency domain, I retain the data in time domain for further processing. The figure below illustrates the frequency response of a fourth-order type 1 Chebyshev low-pass filter with epsilon of 1.

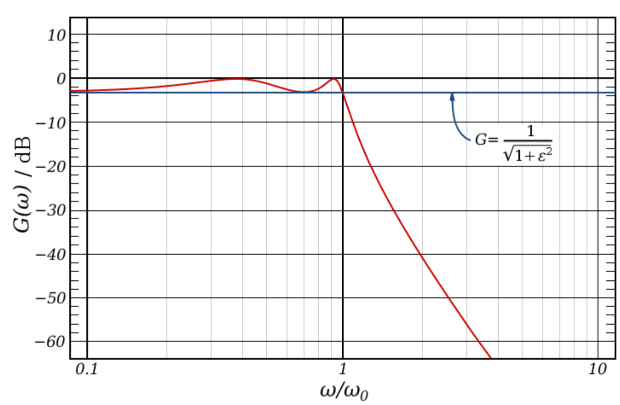


Figure 8: Chebyshev Filter (https://en.wikipedia.org/wiki/Chebyshev_filter)

The conventional measure of heart rate variability is obtained by computing the duration between R wave peaks in the heart rate signal. The next segment of the custom signal processor computes the time series of R-R intervals by detecting R wave peaks in the filtered heart rate signal. It computes duration between subsequent peaks and hands the time-series of R-R intervals to subsequent segments in the pipeline.

The next segment of the custom signal processor computes mean RR (average heart rate variability), mean SD (standard deviation of RR), square root of the mean of the sum of the squares of differences between adjacent RR intervals (RMSSD), mean NN50 (number of consecutive RR intervals that differ more than 50ms), mean pNN50 (proportion of NN50) and mean HR (mean heart rate) as time domain features of heart rate variability (HRV) from the time-series of R-R intervals as features for the machine learning classifier.

The final segment of the custom signal processor computes frequency domain parameters of heart rate variability (HRV). The power spectral density of R-R intervals is estimated by computing a periodogram. The spectrum is divided into three frequency bands; very low frequency (VLF), 0.01-0.04 KHz, low frequency (LF), 0.04-0.15 KHz and high frequency (HF), 0.15-0.4 KHz. The ratio of LF/HF is computed as a frequency domain parameter feature for the machine learning classifier.

b. *Machine Learning Classifier:*

The machine learning classifier, running as a WEKA module, obtains the computed feature set from the custom signal processor through an ARFF file and determines if the subject is experiencing cognitive stress using the following HRV features:

CATEGORY	FEATURES
TIME	Mean RR
	SDRR
	RMSSD
	Mean NN50
	Mean pNN50
	Mean HR
FREQUENCY	LF/HF

Table 2

This system employs two different classification schemes: Naïve Bayes and Support Vector Machines. Both classifiers utilized all the features listed above and their model parameters were obtained using training data. Segmentation between rest and stress states is the final outcome of the classifier.

Figure 9 demonstrates a fully assembled system, attached to a PC:

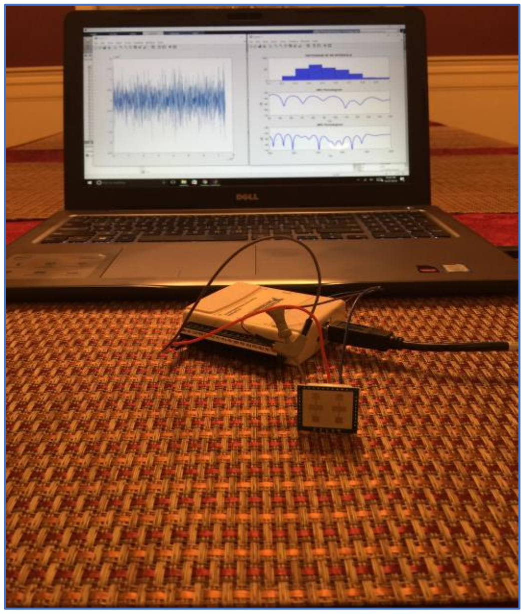


Figure 9: Fully Assembled System with Doppler Radar

System Testing

System Validation Against Ground Truth

In order to validate the results obtained through system testing, I collected ground truth data using a pulse oximeter. Figure 10 below shows the estimated heartbeat waveform obtained using my system, overlaid with the ground truth heartbeat waveform from the pulse oximeter:

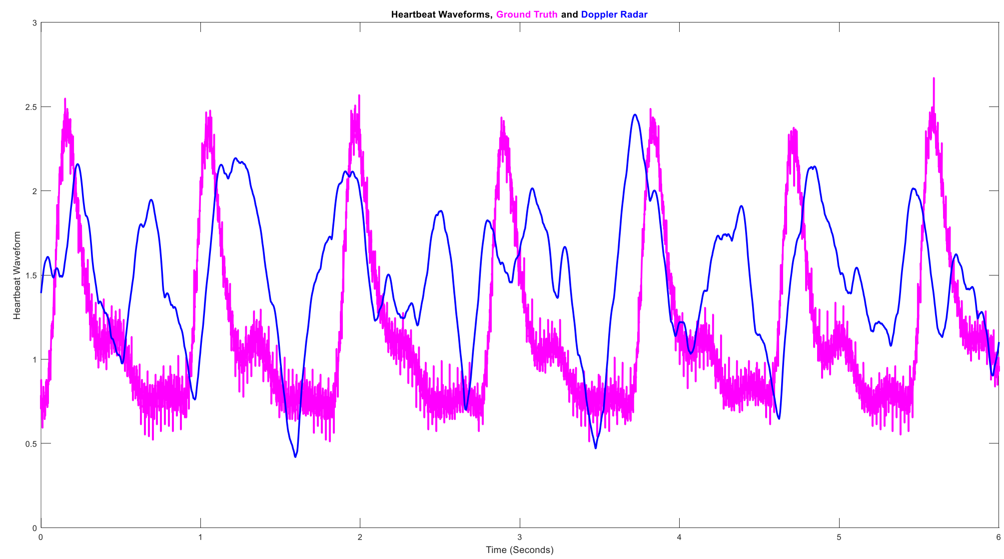


Figure 10: Heartbeat Waveform, Ground Truth and Doppler Radar

In Figure 10, for every cycle in the ground truth heartbeat waveform, I obtain roughly two cycles, corresponding to inhalation and exhalation, of the estimated heartbeat waveform using my system. Hence, the estimated heart rate is half the frequency of the filtered signal. Figure 11 shows the power spectral density of the estimated heartbeat, where the peak corresponds to the dominant frequency, which is the heart rate.

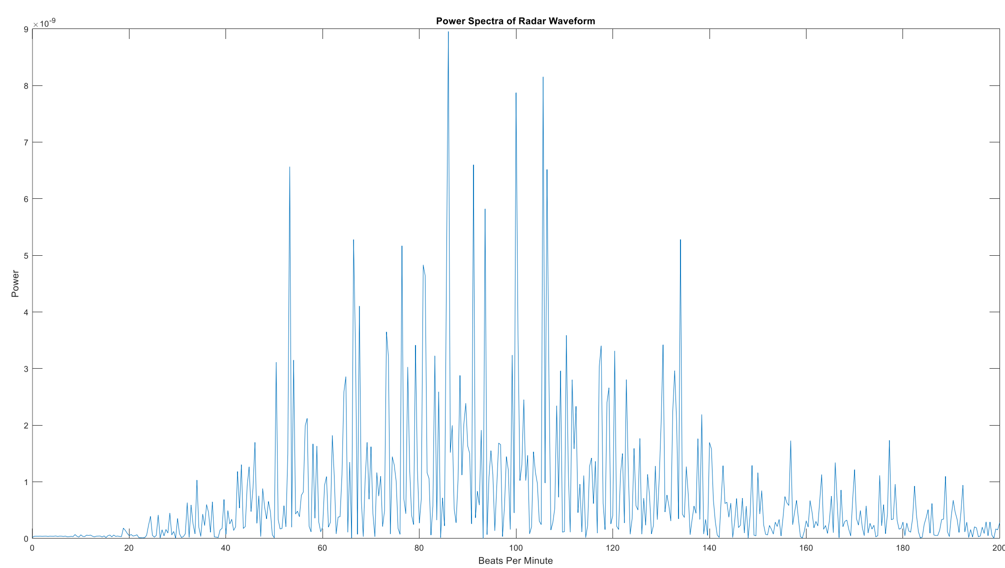


Figure 11: Power Spectral Density, Estimated Heartbeat Waveform

The error of my estimated heart rate across using Mean Absolute Percentage Error (MAPE), is around 10.05%. These results indicate that my system can track heart rate with a reasonable accuracy.

Stress vs. Non-Stress

In order to train the machine learning classifier and determining the final outcome of rest or cognitive stress, I collected time and frequency domain parameters mentioned above. The results obtained are presented in Table 3 below:

	Rest	Stress
Mean RR (s)	**0.550**	**0.527**
Mean SDRR	0.174	0.180
Mean RMSSD	0.174	0.180
Mean NN50	216	219
Mean pNN50 (%)	83	82
Mean HR	**84**	**102**
Mean LF/HF	**2.76**	**2.07**

Table 3: Time and Frequency Domain Parameters Across Rest and Stress Conditions

From the table above, I conclude that Mean RR or HRV decreases and Mean HR increases when I am under cognitive stress. Similarly, in the frequency domain, as expected, Mean LF/HF shows a substantial decline. Figure 11 shows the histogram of RR intervals, LF and HF for myself under rest while Figure 12 shows the histogram of RR intervals, LF and HF for myself under cognitive stress. Figure 11 and Figure 12 highlight the above discussed results; Mean RR and Mean LF/HF decrease under cognitive stress.

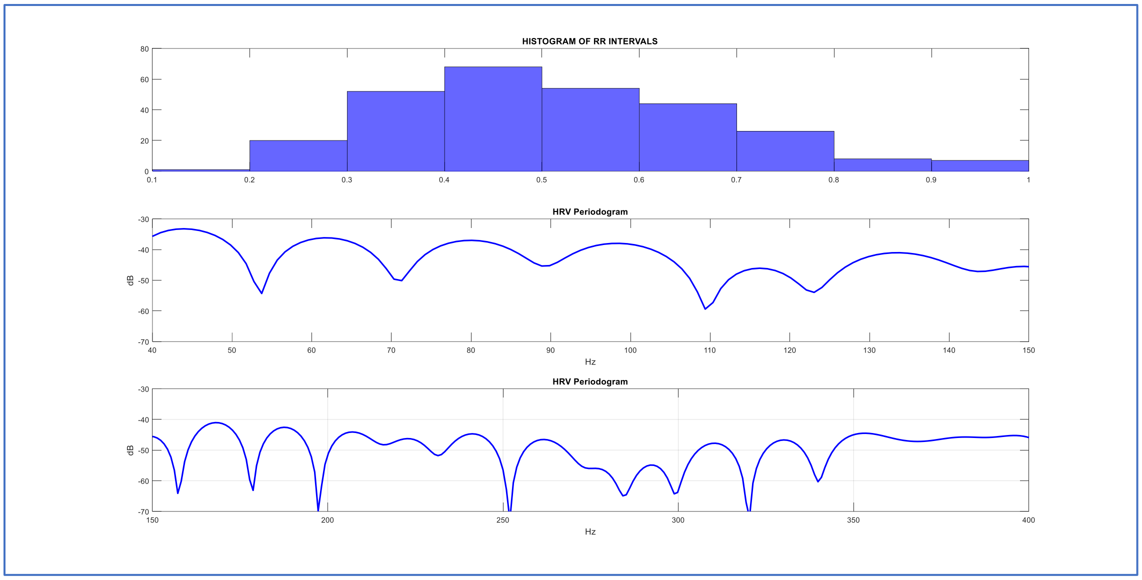


Figure 12: Myself Under Rest

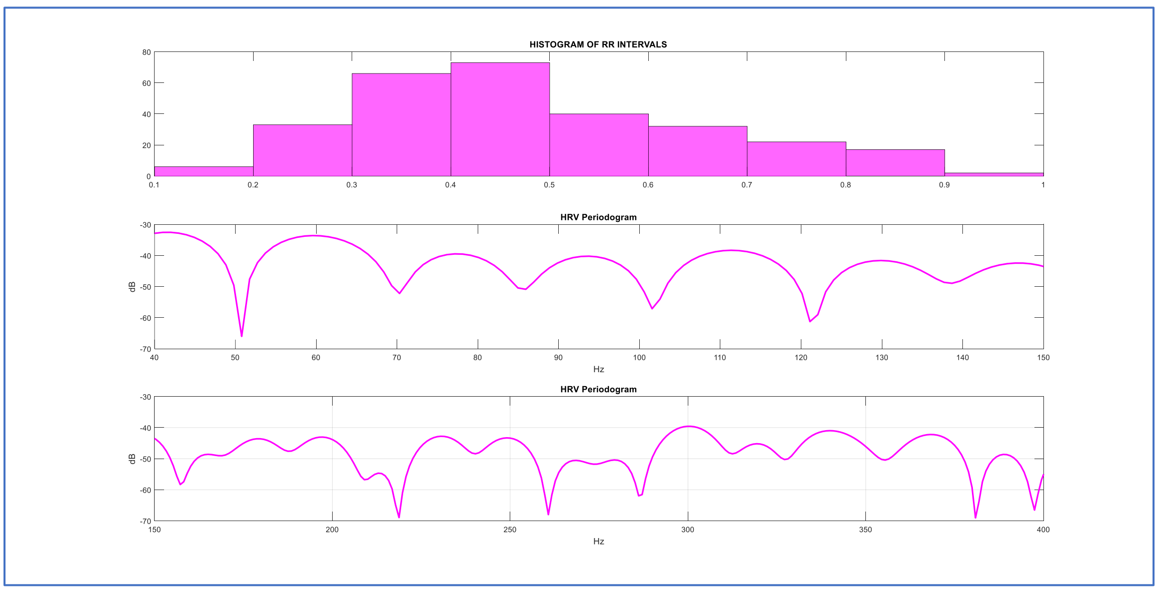


Figure 13: Myself Under Cognitive Stress

I trained the machine learning classifier using two different classification schemes: Naïve Bayes and Support Vector Machines. I evaluated the two classification schemes in terms of precision, recall and accuracy. From the results in Table 4, I observe that SVM outperforms Naïve Bayes and is a better scheme to diagnose stress.

Features	Classifier	Precision (%)	Recall (%)	Accuracy (%)
RR, SD, RMSSD, NN50, pNN50, LF/HF	Naïve Bayes	93.9	92.9	92.9
RR, SD, RMSSD, NN50, pNN50, LF/HF	SVM	100	100	100

Table 4: Performance Metrics of Machine Learning Classifiers

Results

I have successfully designed and implemented a low-cost, long-term, contact-free stress identification system which can operate without any professional intervention. This system tracks an individual's physical body movements, heart beat and breathing rate and determines if a person is undergoing cognitive stress. It shows great promise for continuous, passive and contact-free monitoring of stress in real-world settings. The estimated costs of the component for the system are shown in Table 5.

Description	Cost
Doppler Radar Module	$10
NI-USB 6008 DAC	$49
Miscellaneous (resistors, capacitors, etc.)	$5
Total Cost	**$63**

Table 5

Future Work

Remote sensing opportunities offered by laser-based (Laser Diode Vibrocardiography) or image based (Photoplethusmography) techniques are extremely promising and can be combined with Doppler radar techniques, presented in this paper, to create a multi-modal non-contact stress detection system (Lorenzo). Coupled with the software application, running on a smartphone, this multi-modal system can be deployed at numerous places to assist folks with a history of stress disorders. I would expect a significantly higher accuracy from such a system.

Further, as we know that stress parameters highlighted in this paper, vary greatly for an individual and can change day to day and even hour to hour. Hence, stress score modeling can be significantly enhanced by computing personalized thresholds for each individual. Rather than recording just a single test data from the subject, the software application would need to run in a continuous loop and record subject data. This highly confidential data must be encrypted and secured in a cloud. Numerous further medical insights can be obtained of subjects using this scheme.

Acknowledgments

I would like to thank my mentors, Dr. Brockenbrough, Dr. Ali Ali Yousuf and Mrs. Ravichandran for providing valuable feedback and many suggestions to improve my project. I am also grateful and would like to thank my parents for their dedication throughout this project.

Bibliography

1. Gupta, Otkrist, Daniel McDuff and Ramesh Raskar. "Real-Time Physiological Measurement and Visualization Using a Synchronized Multi-Camera System." *IEEE Explore* (2016).

2. Kim, Dongsoo, et al. "Application and Limitation of Frequency Domain, LF/HF Component in Heart Rate Variability as an Acute Stress Inde." *Proceedings of the International Conference on Biomedical Engineering and Systems* . Prague: International Conference on Biomedical Engineering and Systems , 2014. 128.

3. Koegelenberg, Suretha. *Application of Laser Doppler Vibrocardiography for Human Heart Auscultation*. Matieland: University of Stellenbosch, 2014.

4. Lohman, B, et al. "A Digital Processor For Doppler Radar Sensing Of Vital Signs." 2015.

5. Lorenzo, Scalise. "Non Contact Heart Monitoring." Mills, Richard. *Advances in Electrocardiograms - Methods and Analysis*. Rijeka: InTech Europe, 2012.

6. Lu, Guohua, et al. "Contact-free Measurement of Heart Rate Variability via a Microwave Sensor." *Sensors* (2009).

7. McDuff, Daniel, Sarah Gontarek and Rosalind Picard. *Improvements in Remote Cardio-Pulmonary Measurement Using a Five Band Digital Camera*. Cambridge: MIT Media Lab, 2015.

8. —. *Remote Measurement of Cognitive Stress via Heart Rate Variability*. Cambridge: MIT Media Lab Consortium, 2015.

9. Mostov, K, E Lipsten and R Boutchko. "Medical applications of shortwave FM radar: Remote monitoring of cardiac and respiratory motion." *Medical Physics* March 2010: 1332.

10. Muaremi, Amir, Bert Arnrich and Gerhard Troster. *Towards Measuring Stress with Smartphones and Werable Devices During Workday and Sleep*. Zurich: Springerlink.com, 2013.

11. Peng, Rong-Chao, et al. "Extraction of Heart Rate Variability from Smartphone Photoplethysmograms." *Computational and Mathematical Methods in Medicine* 2015: 11.

12. Plewa, Luke. "iStress: Stress Classification from Heart Rate Variability." *CSC 520* (2015).

13. Rahman, Tauhidur, et al. "DoppleSleep: A Contactless Unobtrusive Sleep Sensing System Using Short-Range Doppler Radar." *American Computing Machinery* (2015).

14. Salai, Mario, Istvan Vassanyi and Istvan Kosa. "Stress Detection Using Low Cost Heart Rate Sensors." *Journal of Healthcare Engineering* 2016: 13.

15. Sauter, Steven. *Stress and Heart Rate Variability*. Los Gatos: SweetWater Health, 2007.

16. *Stress in America: Our Health at Risk*. Washington, D.C.: American Psychological Association, 2012.

17. Taelman, J, et al. "Influence of Mental Stress on Heart Rate and Heart Rate Variability." *IFMBE*. Springer Link, 2008. 1366-1369.

18. ThePaleoMom. *How Stress Undermines Health*. 28 January 2015. <https://www.thepaleomom.com/stress-undermines-health/>.

19. Unknown. *Understanding the stress response*. Cambridge: Harvard Health Publications, 2011.

20. Xiao, Yanming. *Low-Power Radar System For Remote Detection Of Heartbeat And Respiration Using Double-Sideband Transmission And Frequency-Tuning Technique*. Gainesville, 2007.

Appendix

The Appendix details MATLAB modules I employed in the Software Application.

READ_DAQ

```
% ACKNOWLEDGMENT: most of this module is derived from an existing implementation
from my mentor

close all;
clear all;
fs = 5000;
refreshrate = 0.25;

% Finds any running DAQ objects and stop them
% creates a session
d = daq.getDevices;
s = daq.createSession('ni');
s.addAnalogInputChannel('Dev1',1,'Voltage');

% sets the sampling rate and length of sampling window
s.Rate = fs;
s.DurationInSeconds = 150;

% set the correct channel to read from
ch = s.Channels(1);

% associate the window with the file id
global window fid
window = zeros(1, 700);

% creates a text file that stores the data
filename = 'Subject7_DopplerModule_Normal_1.txt';
fid = fopen(filename,'w+');

% saves data from the analog to digital converter
lh = s.addlistener('DataAvailable',@plotData);
s.startBackground();
s.wait();
delete(lh);
fclose(fid);
```

COMPUTE_HR_BR

```
% ACKNOWLEDGMENT: cheb_filt, used in the module, to apply a Chebyshev filter to
the raw signal is derived from an existing implementation from my mentor

%opens text file with data from analog to digital converter
filename = 'Subject7_DopplerModule_Normal_1.txt';
features_index = 1;

% Based on filename, determine if this is Stress Data and whether this is
% Ground or Radar data
StressOrNormal = ~isempty(strfind(filename, 'Stress'));
GroundOrRadar = ~isempty(strfind(filename, 'Ground'));

% open and read the file
fid = fopen(filename);
data = fread(fid,'float');
fclose(fid);

% sets sampling rate
fs = 5000;

% calculate the heart rate and plot the data
if GroundOrRadar
    % filter any noise from the data
    hr_filt_mag_data = cheb_filt(1, 1, 2, fs, data);

    % plot the ground truth data
    figure;
    x = (0.0002: 0.0002: 150);
    plot(x, hr_filt_mag_data, 'm')
    xlabel('Time (Seconds)'); % x-axis label
    ylabel('Ground Truth Heart Waveform'); % y-axis label

    % compute the periodogram of the ground truth data
    [pxx,freq] = periodogram(hr_filt_mag_data, [], length(hr_filt_mag_data), fs);

    % find the max frequency at which HR occurs
    [pmax,lmax] = max(pxx);
    f0 = 60*freq(lmax);

    % plot the periodogram
    figure;
    plot(60*freq, pxx);
    xlim([0 200]);
    xlabel('Beats Per Minute'); % x-axis label
    ylabel('Power'); % y-axis label
    title('Power Spectra of Ground Truth Waveform')

    % display the heart rate
    disp('Groud Truth Heart Rate is: '); disp(f0);
```

```
else
    % uses a chebyshev filter of order 2 and frequency bounds of 0.1 and 0.5 Hz
    to find the breathing rate
    br_filt_mag_data = cheb_filt(2, 0.1, 0.5, fs, data);

      % uses a chebyshev filter of order 1 and frequency bounds of 1 and 2 Hz to
      find the heart rate
      hr_filt_mag_data = cheb_filt(1, 1, 2, fs, data);

      % plot the RADAR data
      figure;
      x = (0.0002: 0.0002: 150);
      plot(x, 10^4 * hr_filt_mag_data, 'b');
      xlabel('Time (Seconds)') % x-axis label
      ylabel('Radar Heart Waveform') % y-axis label

      % compute the periodogram for the radar data
      [pxx,freq] = periodogram(hr_filt_mag_data, [],
      length(hr_filt_mag_data), fs);

      % find the max frequency at which HR occurs
      [pmax,lmax] = max(pxx);
      f0 = 60*freq(lmax);

      % plot the periodogram
      plot(60*freq, pxx);
      xlim([0 200]);
      xlabel('Beats Per Minute'); % x-axis label
      ylabel('Power'); % y-axis label
      title('Power Spectra of Radar Waveform')

      % display the heart rate
      disp('Radar Heart Rate is: '); disp(f0);
end
```

COMPUTE TIME AND FREQUENCY DOMAIN PARAMETERS

```
%% PEAKS OF HEART RATE SIGNAL

% calculate RR peaks
% 1.5 for ground truth MinPeakProminence
% 0.00005 for radar MinPeakProminence

if GroundOrRadar
    [peakValues, indexes] = findpeaks(hr_filt_mag_data, 5000, 'MinPeakDistance',
    0.1, 'MinPeakProminence', 1.5);
else
    [peakValues, indexes] = findpeaks(hr_filt_mag_data, 5000, 'MinPeakDistance',
    0.1, 'MinPeakProminence', 0.00005);
end

peakInterval = diff(indexes);

% remove outliers
peakInterval = peakInterval((peakInterval < (median(peakInterval) + 0.5)));
peakInterval = peakInterval((peakInterval > (median(peakInterval) - 0.5)));

% plot the HISTOGRAM of RR INTERVALS
subplot(3,1,1);
histogram(peakInterval);
title('HISTOGRAM OF RR INTERVALS')

%% CALCULATING AVERAGE TIME BETWEEN RR INTERVALS
mean_rr = mean(peakInterval)*1000;
disp('Average time between rr intervals (milliseconds) under normal conditions
is: '); disp(mean_rr);

%% MIN RR INTERVAL
min_rr = min(peakInterval)*1000;
disp('Minimum time between rr intervals (milliseconds) under normal conditions
is: '); disp(min_rr);

%% MAX RR INTERVAL
max_rr = max(peakInterval)*1000;
disp('Maximum time between rr intervals (milliseconds) under normal conditions
is: '); disp(max_rr);

%% STANDARD DEVIATION OF RR INTERVALS
standard_dev_rr = std(peakInterval)*1000;
disp('Standard deviation (milliseconds) in normal conditions is: ');
disp(standard_dev_rr);

%% NUMBER OF RR INTERVALS THAT DIFFER FOR MORE THAN 50 MS
nn50 = 0;
for w = 1:length(peakInterval) - 1
    if(abs(peakInterval(w) - peakInterval(w+1)) > 0.05)
        nn50 = nn50 + 1;
```

```
    end
end

disp('Number of rr intervals that differ for more than 50 ms under normal
conditions is: '); disp(nn50);

%% NN50 DIVIDED BY TOTAL NUMBER OF INTERVALS

pnn50 = 100*(nn50 / length(peakInterval));
disp('Percent NN50 divided by total number of intervals under normal conditions
is: '); disp(pnn50);

%% ROOT-MEAN-SQUARE SUCCESSIVE DIFFERENCE OF INTERVALS

rmssd = sqrt(sum(((mean(peakInterval)- peakInterval).^2))/length(peakInterval-
1))*1000;
disp('Root Mean-Square of Successive RR intervals under normal conditions is: ');
disp(rmssd);

%% FREQUENCY DOMAIN ANALYSIS OF INTERVALS
[pxx,freq] = periodogram(peakInterval, [], 5120, 5000);

TP = 1000*bandpower(pxx, freq, [0 400], 'psd');
VLF = 1000*bandpower(pxx, freq, [0 40], 'psd');
LF = 1000*bandpower(pxx, freq, [40 150], 'psd');
HF = 1000*bandpower(pxx, freq, [150 400], 'psd');

LFHF = LF/HF;

if StressOrNormal

    disp('VLF is: '); disp(VLF);
    disp('LF is: '); disp(LF);
    disp('HF is: '); disp(HF);
    disp('LF/HF is: '); disp(LF/HF);
    set(get(gca,'child'),'FaceColor','m');

    subplot(3,1,2);
    plot(freq, 10*log10(pxx), 'm', 'LineWidth', 2)
    xlim([40 150])
    ylim([-70 -30])
    xlabel('Hz')
    ylabel('dB')
    title('HRV Periodogram')

    subplot(3,1,3);
    plot(freq, 10*log10(pxx), 'm', 'LineWidth', 2)
    xlim([150 400])
    ylim([-70 -30])
    xlabel('Hz')
    ylabel('dB')
    title('HRV Periodogram')
```

```
    grid on;
else
    disp('VLF is: '); disp(VLF);
    disp('LF is: '); disp(LF);
    disp('HF is: '); disp(HF);
    disp('LF/HF is: '); disp(LF/HF);
    set(get(gca,'child'),'FaceColor','b');

    subplot(3,1,2);
    plot(freq, 10*log10(pxx), 'b', 'LineWidth', 2)
    xlim([40 150])
    ylim([-70 -30])
    xlabel('Hz')
    ylabel('dB')
    title('HRV Periodogram')

    subplot(3,1,3);
    plot(freq, 10*log10(pxx), 'b', 'LineWidth', 2)
    xlim([150 400])
    ylim([-70 -30])
    xlabel('Hz')
    ylabel('dB')
    title('HRV Periodogram')

    grid on;
end

%% PUTTING FEATURES INTO MATRIX
% populate dataset

features = struct();
% nominal classes
type_class = { 'stress', 'not_stressed' };

%% declare nominal specification attributes
nomspec.type_class = type_class;
features(features_index).idx = features_index;
features(features_index).avgtime = mean_rr;
features(features_index).standarddev = standard_dev_rr;
features(features_index).nn50 = nn50;
features(features_index).pnn50 = pnn50;
features(features_index).rmssd = rmssd;
features(features_index).lfhf = LFHF;

if (StressOrNormal)
    features(features_index).type_class = 'stress';
else
    features(features_index).type_class = 'not stressed';
end
% for test, change stress and not stressed to '?'
```

CREATE ARFF

```
% ACKNOWLEDGMENT: arff_read and arff_write are standard open source modules to
read and write ARFF files

%% create output filename
if GroundOrRadar
    outfile = 'Ground_ML.arff';
    relname = 'Training';
else
    outfile = 'DopplerRadar_Training.arff';
    relname = 'Training';
end

% read the file
if exist(outfile, 'file')
    [data, relname, nomspec] = arff_read(outfile);
    index = length(data);

    data(index + 1).idx = index + 1;
    data(index + 1).avgtime = features(features_index).avgtime;
    data(index + 1).standarddev = features(features_index).standarddev;
    data(index + 1).nn50 = features(features_index).nn50;
    data(index + 1).pnn50 = features(features_index).pnn50;
    data(index + 1).rmssd = features(features_index).rmssd;
    data(index + 1).lfhf = features(features_index).lfhf;
    data(index + 1).type_class = features(features_index).type_class;

    % save arff
    arff_write(outfile, data, relname, nomspec);
else
    % save arff
    % nominal classes
    type_class = { 'stress', 'not stressed' };
    nomspec.type_class = type_class;
    arff_write(outfile, features, relname, nomspec);
end
```