Physical Activity Classification

Phanitta Chomsinsap and Sheila Ahoraian
Wireless Health Institute Summer Program 2016

***Abstract*—Personal health monitoring systems allow users to track their daily activities and exercise progress to stay motivated. This information can also benefit medical professionals who require a close monitoring of their patients' exercise regimen. Current technologies such as Fitbit and Apple Watch enable tracking of general activities such as steps taken, distance traveled, and time duration of activities performed; however, without user input, these devices are unable to classify physical exercises besides walking and running. In contrast, our device is a bodysuit that consists of three Inertial Measurement Units (IMU) placed on the wrist, shoulder, and foot. Using machine learning algorithms, we are able to accurately classify walking, running, bicep curls, elliptical cardio, and biking activities. Our goal is to make this information accessible on a mobile application and cloud platform for both personal and professional use. Based on the data collected for five different exercises, the bodysuit demonstrates a 6.7% increase in accuracy for Decision Tree classifiers in comparison to devices with only a wrist sensor.**

***Index Terms*—Activity Classification, Health monitoring, Inertial Measurement Units (IMU), machine learning algorithms**

1. INTRODUCTION

Human activity recognition using a wearable motion sensing device has many applications in health care, security, and exercise lifestyle [1]. In patient monitoring systems, activity classification also plays a vital role in fall detection for elderly patients who may need an immediate emergency treatment [2]. Other near-future applications include fitness tracker mobile apps and exercise records of patients for medical purposes.

Our main goal for this study is to achieve accurate classification of different exercises by using an Inertial Measurement Unit (IMU) chip that consists of tri-axial accelerometer and tri-axial gyroscope sensors. With the aid of Bluetooth Low Energy (BLE), the sensor data can be accessed by mobile devices and through the cloud. This accessibility of data allows not only for personal health monitoring, but also for professional health monitoring of users both in local and remote areas by health care and medical institutions.

Our work aims to maximize the accuracy with the fewest number of sensors possible. In this paper, the performance of a single sensor and a bodysuit (three-sensor device shown in Fig. 2) was compared based on the classification of five physical activities: walking, running, bicep curls, elliptical cardio, and biking. The classifiers used in the study include Decision Tree, Random Forest, and Support Vector Machine with both linear and polynomial kernels.

2. IMPLEMENTATION

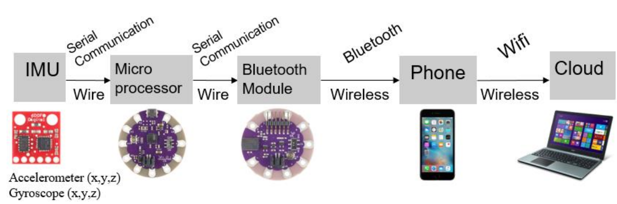

Fig. 1. System Components Overview (adapted from [3],[4],[5],[6],[7])

A. Hardware

The main objective is to be able to get physical data needed from the IMU chip, do some computation on the microprocessor board through Arduino programming, and communicate the data to the cloud through Bluetooth and wifi connections as shown in Fig. 1. Our project consists of three hardware components: IMU chip, LilyPad Arduino USB microprocessor, and Arduino LilyPad Simblee Bluetooth module. With a voltage output of 3.3V, the IMU chip has the capacity for six degrees of freedom allowing it to measure data for both rotational and translational motion through the combination of acceleration, rate of rotation, and orientation. Within the IMU, there are two sensors which are the accelerometer and the gyroscope. The accelerometer measures the acceleration in units of gravity (g) while the gyroscope measures the rate of rotation in degrees per second; both of which are measured along the x, y, and z axis. Three angles (yaw[1], pitch, roll) can be obtained by combining the readings of the accelerometer and gyroscope. The angles describe the orientation around each axis in three-dimensional space. The accelerometer and gyroscope sensors on the IMU chip play a crucial role when classifying the physical activities being performed. Since activities can be done at various accelerations, the accelerometer data was helpful for classifying between walking and running. In addition, the data collected from the gyroscope was useful for activities that required more rotational motion such as bicep curls, biking, and elliptical cardio. The accelerometer and gyroscope data such as acceleration, rate of rotation, and orientation (angles) will be used as features that allow for the classification of the exercise activities. The method for classifying data will be discussed later on under the “Software” section.

The LilyPad Arduino USB microprocessor, powered by a 3.7V LiPo (lithium-ion polymer) battery, was created to be sewn onto clothing. Using Arduino programming, the microprocessor is programmed to retrieve data from the IMU chip. In order for the data to be displayed, the microprocessor and the IMU chip must communicate through serial communication which sends the data one bit at a time. For these devices, we used the Inter-Integrated Circuit (I2C) protocol which requires two signal wires to exchange information: a serial clock pin (SCL) and a serial data pin (SDA). A bit of information is transferred from the IMU chip to the microprocessor through the SDA line at a clock edge. This information consists of the accelerometer, gyroscope, and angle values, resulting in nine values constantly being sent between the devices.

These values are then passed from the microprocessor to the LilyPad Simblee (Bluetooth module). The LilyPad Simblee is a BLE device that uses the Internet of Things (IoT) platform, which includes Bluetooth (TM) Smart, Mobile, and Cloud. The IoT allows one device to connect to another device as well as the Internet. However, in our case, it allowed us to connect the LilyPad Simblee to both the microprocessor and the Cloud. In order for the information on the LilyPad microprocessor to be sent to the Simblee module, a Universal Asynchronous Receiver/Transmitter (UART) serial communication was used. This protocol has a transmitter and receiver side where the UART creates a data packet on the transmit side and sends that packet on the TX line according to a specified baud rate, or the rate of data transmission in bits per second. The UART then obtains the sync bits after taking samples from the RX line according to the baud rate, and sends the data to the receiver side [8]. Physically, we had to connect the TX pin of one device to the RX pin of the other device and vice versa in order for the serial communication to function properly.

[1] Only the rate of change of the yaw angle can be obtained. The absolute yaw angle requires a magnetometer reading.

B. Software

C++ was used to program and obtain the reading from the IMU chip. Both the microprocessor and the Bluetooth use Arduino programming. A web page was written using HTML and JavaScript to retrieve data from the Simblee Cloud. Python provides machine learning algorithms necessary for data classification.

1) Getting Data from the IMU

An open source C++ and Arduino code from a bildr tutorial [9] was used to obtain 9 IMU values including the angles, acceleration, and rate of rotation. These values were read by the LilyPad microprocessor programmed in Arduino. Data from the IMU chip was sent by the I2C Protocol, which allows multiple master-slave communication through a bus as long as the clock frequency is specified by a master device.

2) Computation on Microprocessor

Using Arduino programming, we calculated the mean, variance, and standard deviation for each IMU value in 5 second intervals. At each time interval, 27 data values were obtained and later used as our features for classifiers as discussed in the section “Feature Selection.”

3) Sending Data to the Bluetooth Module

The Arduino “SoftwareSerial” library enables serial communication between the microprocessor and the Bluetooth module with speeds up to

115200 bps [10]. The SoftwareSerial object is used to

i) initialize the RX and TX pins,
ii) set the baud rate,
iii) call a write() function that sends the data from the Lilypad, and
iv) call a read() function that receives the data by the Simblee module.

4) *Sending Data to Cloud*

Sending data from the Simblee module to the cloud required two processes: 1) downloading a "Simblee for Mobile" application and 2) setting up a "SimbleeCloud" account [11].

The Simblee for Mobile application enables wireless communication between the Simblee module and a mobile phone via Bluetooth signals. This mobile app is compatible with iOS devices including iPhone, iPad, and iPod touch. Furthermore, a user interface can be written on the Simblee module using Arduino programming and displayed on a mobile device for user interaction.

A SimbleeCloud account is required for data transmission from the Simblee module to the SimbleeCloud, which is also used for data storage. Each SimbleeCloud account has a UserID or a specific address where the data is sent to (SimbleeCloud). Each Simblee module has a unique address, or Electronic Serial Number (ESN), where the data is sent from (Arduino). In addition, a web page address, which provides a destination for retrieving the data (destESN), can be assigned in the "Pools" section of the SimbleeCloud account. The SimbleeCloud library allows us to send data from Arduino to SimbleeCloud. By declaring a SimbleeCloud object, the send() function can be called to transmit a character array of data to the cloud.

5) *Retrieving Data from the Cloud on a Webpage*

HTML and JavaScript code by Simblee Corporation was used to retrieve the data from the SimbleeCloud [11], and was modified to print the calculated values on the HTML web page.

3. MACHINE LEARNING

A number of research has employed machine learning methods for activity recognition. Examples of classifiers commonly used in research include Decision Tree and Artificial Neural Network (ANN) in [2] and [12], and Support Vector Machine (SVM) in [1].

A. Feature Selection

Features are attributes used to describe data. In this experiment, features are from the time domain. These include mean, standard deviation, and variance of each 9 IMU values (angles, acceleration, rate of rotation). For a single sensor, there are 27 features. The bodysuit with three sensors consists of 81 features in total.

B. Classification

The classifiers used in our study are Decision Tree, Random Forest, and SVM (linear and polynomial kernels).

1) *Decision Tree*

Decision Tree is a predictive model that classifies an unknown set of data (testing data) based on the previously trained model (training data). To classify the data, this algorithm splits the sample into two or more homogeneous sets [13]. Data is divided using the most useful decision node or feature. The population divides first at the root node, and recursively subdivides itself until all the data is fully classified into one of the categories or classes. The Decision Tree performs the binary splits on the decision node that has the highest value of Gini index given by the equation, $p(1-p)$, where p is the fraction of positive examples [14].

2) *Random Forest*

Random Forest uses ensemble learning methods for classification. Random Forest constructs many Decision Trees by splitting each node using the best among a subset of predictors randomly chosen at that node [15]. One advantage of using this algorithm is to avoid overfitting of the Decision Tree, the case of having over-complex trees that have poor predictive performance.

3) *SVM*

Support Vector Machine is a learning algorithm that optimizes data classification rather than using a greedy search. SVM finds an optimal hyperplane (for dimensions higher than 2) for linearly separable patterns. In 2 dimensions, the algorithm draws a boundary line to separate two sets of data points, where the margin between each data set and the decision boundary is maximized. Inner product kernels are used for data sets that are not linearly separable. [16]

4. DATA COLLECTION

The data was collected for five different physical activities: walking, running, bicep curls, elliptical cardio, and biking. There were a total of three female participants, who were between the ages of 19-21 years old. For each activity, 60 data sets were collected and each data set was taken on a five second interval. Out of 60 data sets, 70% was used for training (45 data sets) and 30% for testing (15 data sets).

Fig. 2. Sensors placed on wrist, shoulder, and foot

5. RESULTS

The heat maps (or confusion matrices) show the percent accuracy of each classifier including Decision Tree, Random Forest, and SVM (linear and polynomial kernels) for each physical activity. The higher the diagonal values are (represented in white) and the lower the off-diagonal values are (represented in dark blue), the higher the percent accuracy. The results of a single sensor (foot, shoulder, wrist) are compared with a bodysuit (three-sensor device) on three different classifiers.

Fig. 3 shows the performance of classifiers on individual foot and shoulder sensor data. The highest and lowest accuracy for the foot sensor are 86.67% and 72% respectively, in which the Decision Tree classifier performed best overall. The average accuracy is 81.33%. On the other hand, the highest and lowest accuracy for the shoulder sensor is 98.67% and 86.67% respectively, where the SVM (linear kernel) classifier performed best. The average accuracy is 91.67%.

Fig. 4 displays the percent accuracy of each classifier for a wrist sensor and three-sensor device. The highest and lowest accuracy for the wrist sensor is 97.33% and 90.67% respectively, in which the SVM (linear kernel) classifier performed best overall. The average accuracy is 94.67%. On the other hand, the highest and lowest accuracy for the bodysuit sensors are 98.67% and 85.33% respectively, where the Decision Tree and Random Forest classifiers performed best. The average accuracy is 93.67%.

This concludes that a single sensor worn on the wrist or shoulder can function effectively alone without the need of two additional sensors when classifying the physical activities performed in our study.

When classifying a limited number of activities, having multiple sensors may not be necessary. In [17], Wang et al. attached five sensor units to legs, wrists, and chest of eight subjects, who were to perform a total of 12 daily activities. Each sensor unit was comprised of a tri-axial accelerometer for motion sensing. The leave-one-out accuracy of each classifier was below 90%.

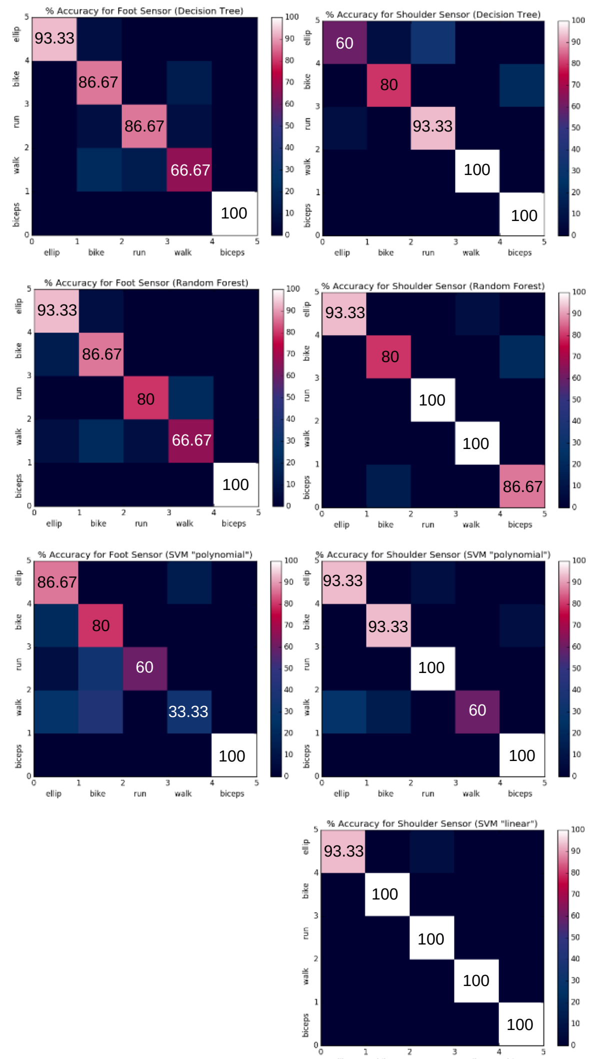

Fig. 3. Percent Accuracy Heat Maps for Foot (left) and Shoulder (right) Sensors

The highest accuracy for foot and shoulder sensors is 86.67% and 98.67% respectively. Linear kernel SVM is not applicable for foot sensor data.

The study by [18] demonstrates an accurate performance of using a single accelerometer for human activity classification. The data was collected among 50 subjects performing 10 different activities. Nevertheless, AlZubi et al. was able to achieve up to 95% in classification accuracy.

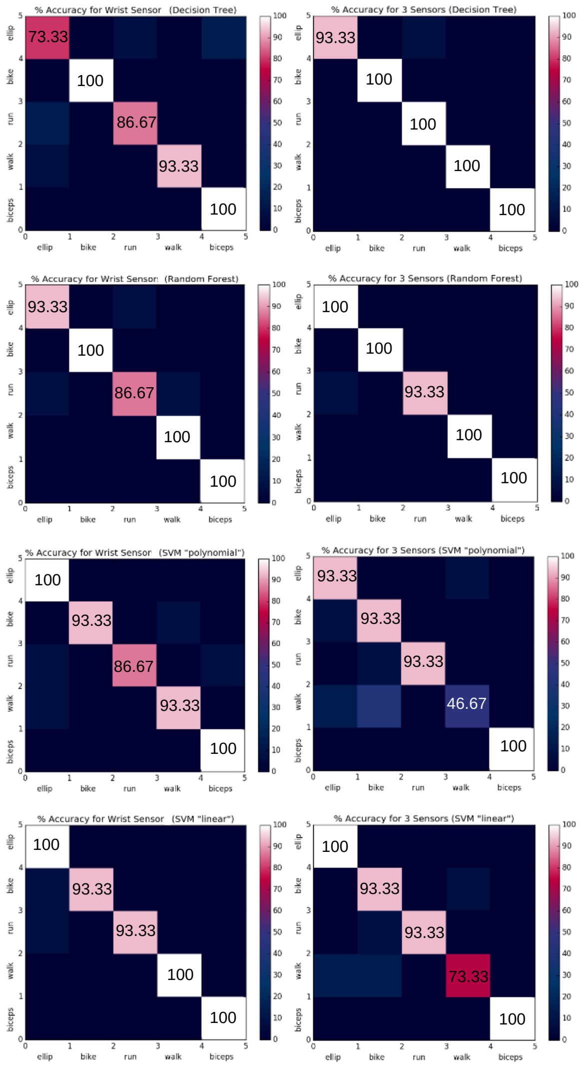

Fig. 4. Percent Accuracy Heat Maps for Wrist Sensor (left) and 3-sensor (right) Device

The highest accuracy for wrist and bodysuit sensors is 97.33% and 98.67% respectively.

6. Future work

Our device is currently able to classify five different physical activities which are walking, running, biking, bicep curls, and elliptical cardio. We hope to expand this number in the future in order to cover a wide range of activities including muscular strength and endurance exercises such as sit ups and push-ups as well as flexibility exercises like stretching. We also plan to even further classify specific activities within certain sports in order to identify, for example, yoga positions or tumbling moves in gymnastics. This will be helpful for both athletes and fitness enthusiasts interested in challenging themselves and achieving total body fitness through a broad range of movements. We plan on making our device real-time so that the data collected will automatically be sent to a text file since, as of now, we had the time consuming task of manually copying data from the cloud and pasting it onto a text file in order to run it on Python. Depending on the percent accuracy using our current classifiers, we may add features from the frequency domain to increase the percent accuracy of our classifiers. Since we gathered data from only a few test subjects, our main goal is to increase this number as well as the diversity of participants. In other words, we hope to gather a mass amount of data from people of different ages, sizes, genders, and ethnicities to create a more accurate device that tailors to the needs of all individuals.

The most significant improvement that we will be making to our health monitoring system is programming our device to display the activity classified and time duration of the activity onto a mobile device. We plan to create an application that will be available to both the Android and Apple stores. At the palm of their hands, the user will be able to keep an accurate record of their physical activities when the internet is unavailable. Using both HTML and JavaScript programming languages, the same information will also be available to the cloud. Furthermore, having access to this information on the cloud will provide the users with a platform that displays graphs or charts of their daily progress, so they can plan their exercise regimen accordingly in order to tailor to their specific body fitness needs. Information on the cloud also allows medical professionals easy access to continuously see an accurate record of physical activities done throughout the day of patients suffering from chronic illnesses, who may need close monitoring of their daily activity or inactivity.

A review published by the National Center for Biotechnology Information [19] examined the uses of smart wearable body sensors and came to the conclusion that this innovation will change relationships between physician and patient, develop a greater involvement of patients in their healthcare, and allow for change in healthcare monitoring and spending through the use of remote monitoring techniques. In the Journal of Neuroengineering and Rehabilitation [20], there is a review discussing wearable sensors in the application of rehabilitation. The article found that health monitoring systems can cover a broad range of applications including health and wellness monitoring, safety monitoring, home rehabilitation, assessment of treatment efficacy, and early detection of disorders. Monitoring a patient's physical activity on a daily basis as well as collecting physiological data for a long time period provides medical professionals with the tools to diagnose and treat various diseases. The use of wearable accelerometers and gyroscopes can be used for fall detection for both the elderly as well as those with impairments who are constantly losing their balance and need medical attention promptly when a fall occurs. For those who have limited mobility in certain parts of their body, wearable sensors can be used to make the rehabilitation process a more enjoyable experience. There are systems that combine body sensors with virtual reality where the user can interact with therapeutic games to increase their mobility. These devices can save a substantial amount of lives by providing early detection for diseases when there is a known correlating set of symptoms. All in all, there are numerous functions for wearable sensors that can be useful for both doctors in their professional lives and patients in their day to day living experiences.

7. Conclusion

Health monitoring systems are very important advancements in our society. There are numerous features that allow users to stay conscious of their overall health as well as allowing medical professionals to stay up to date with their patient's exercise regimen when necessary. This is why we created a wireless activity recognition system consisting of three sensors each of which were placed on different parts of the body: the wrist, shoulder, and foot. This bodysuit can classify five different physical activities which are walking, running, bicep curls, elliptical cardio, and biking. This classification was made possible by connecting a 6DOF IMU chip to an Arduino LilyPad USB microprocessor through the use of I2C serial communication. Using BLE allowed us to send the information from the sensor to the mobile device and cloud. The features we used to classify the data are mean, variance, and standard deviation. In order to classify the data, we used 3 machine learning algorithms: Decision Tree, Random Forest, and SVM. Taking efficiency of the device and comfortability of the users into account, we aim to maximize the classification accuracy with the fewest number of sensors possible. In the results, we found the SVM classifier to be the most accurate for a single sensor attached to the wrist and shoulder with an accuracy above 97%.

8. Acknowledgements

We would like to acknowledge the Wireless Health Institute and the Networked and Embedded Systems Lab at the University of California, Los Angeles for contributing to our research endeavors. We would like to thank our mentors Paul Martin and Bo-Jhang Ho as well as Professor Mani Srivastava for providing us with guidance and support throughout this experience.

9. References

[1] S. Chernbumroong, A. S. Atkins, and H. Yu, "Activity classification using a single wrist-worn accelerometer," *Software, Knowledge Information, Industrial Management and Applications (SKIMA), 2011 5th International Conference on*, 2011.

[2] C. Doukas and I. Maglogiannis, "Advanced patient or elder fall detection based on movement and sound data," *Pervasive Computing Technologies for Healthcare, 2008 Second International Conference on*, 2008.

[3] SparkFun. *6 Degrees of Freedom IMU Digital Combo Board - ITG3200/ADXL345* [Online]. Available: https://www.sparkfun.com/products/10121. [Accessed: 11- Aug- 2016]

[4] Arduino. *LilyPad Arduino USB* [Online]. Available: https://www.arduino.cc/en/Main/ArduinoBoardLilyPadUSB. [Accessed: 11- Aug- 2016]

[5] SparkFun. *LilyPad Simblee BLE Board - RFD77101* [Online]. Available: https://www.sparkfun.com/products/13633. [Accessed: 11- Aug- 2016]

[6] Apple Inc. *iPhone 6s* [Online]. Available: http://www.apple.com/iphone-6s/specs/. [Accessed: 11- Aug- 2016]

[7] Acer Inc. *Acer-Aspire-E1-570-6803-15.6-Inch-Laptop* [Online]. Available: https://www.amazon.com/Acer-Aspire-E1-570-6803-15-6-Inch-Laptop/dp/B00H3R8JD8/ref=cm_cr_arp_d_product_top?ie=UTF8. [Accessed: 11- Aug- 2016]

[8] SparkFun. *Serial Communication* [Online]. Available:

https://learn.sparkfun.com/tutorials/serial-communication

[9] Bildr community. (2012, Mar. 14). *Stable orientation -- digital IMU 6DOF + Arduino* [Online]. Available: http://bildr.org/2012/03/stable-orientation-digital-imu-6dof-arduino/

[10] Arduino. *SoftwareSerial Library* [Online]. Available: https://www.arduino.cc/en/Reference/SoftwareSerial

[11] Simblee Corporation. (2016, Jan. 11). *Simblee Cloud Reference v1.0* [PDF]. Available: https://www.simblee.com/Simblee%20Cloud%20Reference%20v1.0.pdf

[12] Parkka et al., "Activity classification using realistic data from wearable sensors," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 10, no. 1, pp. 661-669, Jan. 2006.

[13] M. Saraswat. (2016, Apr. 12). *A complete tutorial on tree based modeling from scratch (in R & Python)* [Online]. Available: https://www.analyticsvidhya.com/blog/2016/04/complete-tutorial-tree-based-modeling-scratch-in-python/#one

[14] R. Schapire. *Machine learning algorithms for classification* [PDF]. Available: http://www.cs.princeton.edu/~schapire/talks/picasso-minicourse.pdf

[15] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News*, vol. 2, pp. 18-22, 2012.

[16] R. Berwick. (2003). *An idiot's guide to support vector machines (SVMs)* [PDF]. Available: http://www.svms.org/tutorials/Berwick2003.pdf

[17] Wang et al., "An incremental learning method based on probabilistic neural networks and adjustable fuzzy clustering for human activity recognition by using wearable sensor," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 16, no. 4, pp. 661-669, 2012.

[18] AlZubi et al. "Human activity classification using a single accelerometer," *Computational Intelligence (UKCI)*, *2014 14th UK Workshop on,* 2014.

[19] Appelboom et al. (2014, Aug. 22). *Smart wearable body sensors for patient self-assessment and monitoring* [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4166023/

[20] Patel et al., "A review of wearable sensors and systems with application in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 21, no. 9, April 2012.