FOCAL IMPAIRED AWARENESS SEIZURE DETECTION USING A SMARTWATCH

by
Andrew Mark Masteller

A thesis submitted to Johns Hopkins University in conformity with the requirements for
the degree of Master of Science and Engineering

Baltimore, Maryland
August 2020

Abstract

Epileptic seizures are commonly classified as either generalized (originating simultaneously in cortical neurons across the entire brain) or focal (originating in a subpopulation of cortical neurons in a focal brain region). Focal seizures are further subdivided according to whether the seizure causes impaired awareness (focal impaired), or does not (focal aware). Focal impaired awareness seizures cause more disability because patients cannot adequately react to their surroundings, increasing the risk of car crashes, burns, and other accidents. FIA seizures can also be accompanied by automatic behaviors, such as undressing or running, which can be embarrassing or dangerous. While treatments exist for epileptic patients with FIA seizures, tracking the efficacy of treatments can be difficult since patients often will not remember the episodes or recognize that one has occurred. In order to effectively track FIA seizures in epileptic patients, active detection that does not hinder daily life is needed.

FIA seizures may originate in different regions of the brain and thus may have very different clinical manifestations across individuals. However, most cause an increase in heart rate, likely through stimulation of sympathetic efferents in the hypothalamus, without a commensurate increase in physical activity. Using a common consumer-grade smartwatch, the Apple Watch, we used a refined set of engineered features derived from the photoplethysmogram (PPG) heart rate sensor date and tri-axial accelerometer data to train neural networks to detect FIA seizures associated with an increase in heart rate. This strategy was based on our judgment that FIA seizures without an increase in heart rate are not sufficiently stereotyped across patients to detect with any of the available sensors in consumer wearable devices.

A binary classification, artificial neural network was trained using leave-one-out cross-validation (LOOCV). Each of the models correctly identified the left-out FIA seizure within 30

seconds, but high false alarm rates prompted further development. Further training and the use of accumulation filtering allowed for a neural network model with 90% specificity, identifying 27 out of 30 total FIA seizures, and an overall false alarm rate of 1.65/hour with no false alarms during resting, running, or household chores.

Primary Reader and Advisor: Nathan Crone

Committee Readers: Gene Fridman, J. Webster Stayman

Acknowledgements

I would like to thank Dr. Nathan Crone for introducing me to the field of accessible seizure detection and providing guidance through the project, particularly in understanding the unmet need of patients and their healthcare providers in adequately tracking and alerting caregivers of seizure events.

My thanks to Mr. Samyak Shah for helping me to develop my understanding of seizure detection strategies and leveraging large data storage for training machine learning models.

I would also like to thank my committee readers, Dr. Gene Fridman and Dr. J. Webster Stayman, each for teaching courses that were directly valuable to this work and subsequently providing feedback and support during my thesis.

Table of Contents

List of Tables

List of Figures

Introduction

Focal Impaired Awareness Seizures

Focal onset impaired awareness seizures are a specific category of epileptic seizures. These seizures have historically been called complex partial seizures, but were renamed in 2017 to more clearly distinguish seizure categories. These seizures are often referred to as focal impaired awareness (FIA) seizures or focal unaware seizures.

The focal onset portion of the term describes where the seizure starts in the brain. FIA seizures originate from localized regions in the brain, usually in a single hemisphere (left or right). However, these seizures include those which originate at any depth or surface in the brain and can vary in terms of how much they spread across networks in the brain or whether they remain tightly localized. The major distinction between focal onset and general onset seizures is that in general onset seizures, both sides of the brain, the right and left hemispheres, experience abnormal activity at the onset of the seizure.

The impaired awareness distinction of FIA seizures describes the observed physical manifestations of the seizure. Patients become unaware of their surroundings and cannot properly react to their environment. While individual patients will often have similar responses for each FIA seizure, these responses can vary considerably from patient to patient. During a FIA seizure, patients will often perform automatisms – automatic, repetitive motions with no obvious purpose – such as smacking their lips or touching their face. Patients could also simply continue simple tasks that they were already doing, for example watching television or cleaning dishes, or wander aimlessly. FIA arrest seizures can also be present, where a patient becomes motionless regardless of their previous motor activity.

FIA seizures are notably different from classic tonic-clonic (grand mal) seizures that are most often associated with epileptic seizures in the lay public. While there are motor symptoms, the sharp stiffening tonic phase and convulsive clonic phase are not typically characteristics of focal impaired awareness seizures. Generally, seizures are classified by their onset region and awareness level (Figure 1). The seizures can be further described by their motor symptoms.

These three levels of distinction do not capture the full depth of different seizure types or how seizures can evolve from one category to another, but rather provide a high-level understanding of major differences between seizure types. FIA seizures can also end up spreading to both hemispheres of the brain, resulting in focal to bilateral tonic-clonic seizures. The events of the FIA and bilateral tonic-clonic portions of the seizure are distinct, but it can be difficult to recognize the focal impaired awareness symptoms as the tonic-clonic symptoms are more noticeable and severe.

Onset Region	Awareness Level	Motor Symptoms
• Focal • General • Unknown	• Aware • Impaired Awareness	• Motor • Non-motor

Figure 1. Basic seizure classification categories.

The varying symptoms and behaviors experienced by patients with focal impaired awareness seizures largely depend on the brain region in which they originate and the degree to which the seizure spreads to other regions. The experienced symptoms are closely tied with the normal functional properties of the affected brain region, showing the physical manifestations of the effective localized dysfunction.

Focal onset impaired awareness seizures most commonly originate in the temporal lobe, accounting for approximately 70%–90% of FIA seizures. Temporal lobe FIA seizures are often accompanied by stereotyped automatisms. The automatisms are unique to the patient and specific brain localization. Automatisms are typically oral or manual, possibly with other automatic motor behavior. Some patients experience more overt motor activities, ranging from walking or running to nondirected violent outbursts. Patients who experience temporal lobe FIA seizures may go through minutes of confusion before recovery. [1]

The second most common origin for FIA seizures is the frontal lobe. Frontal lobe FIA seizures are often brief, lasting around 30 seconds, and occurring multiple times a day in succession. Half of patients experience a loss of consciousness. The seizures are often triggered while the patient is sleeping with minimal confusion following the seizure. Hypermotor symptoms such as leg thrashing or pelvic thrusting can occur, or motor symptoms can be milder with abnormal posturing. Vocal automatisms can also manifest. [1]

Symptoms associated with parietal lobe onset focal impaired awareness seizures are not as physically pronounced. Abnormal sensorimotor experiences and vestibular hallucinations can occur during a FIA seizure in the parietal lobe. Visual or somatic illusions can be experienced as well as dizziness. Parietal lobe FIA seizures can spread to other regions of the brain, causing automatisms and loss of consciousness. [1]

Focal impaired awareness seizures in the occipital lobe often cause visual auras. Uncontrolled eye movements, rapid blinking, and temporary loss of sight can also occur. More in-depth visual hallucinations with well-formed images can arise from seizures originating in the occipital lobe. The seizure can also further spread to other regions of the brain, resulting in further symptoms. [1]

Figure 2 shows a summary of the basic symptoms for the common focal impaired awareness seizure onset localization mapped to their respective brain regions.

Frontal Lobe
- Hyper motor movements
- Posturing
- Loss of consciousness
- Vocal automatisms

Parietal Lobe
- Visual illusions/hallucinations
- Somatic illusions
- Vertigo
- Vestibular hallucinations

Temporal Lobe
- Oral automatisms
- Manual automatisms
- Varied motor activities
- Post-seizure confusion

Occipital Lobe
- Visual auras
- Eye movements
- Rapid blinking
- Loss of sight
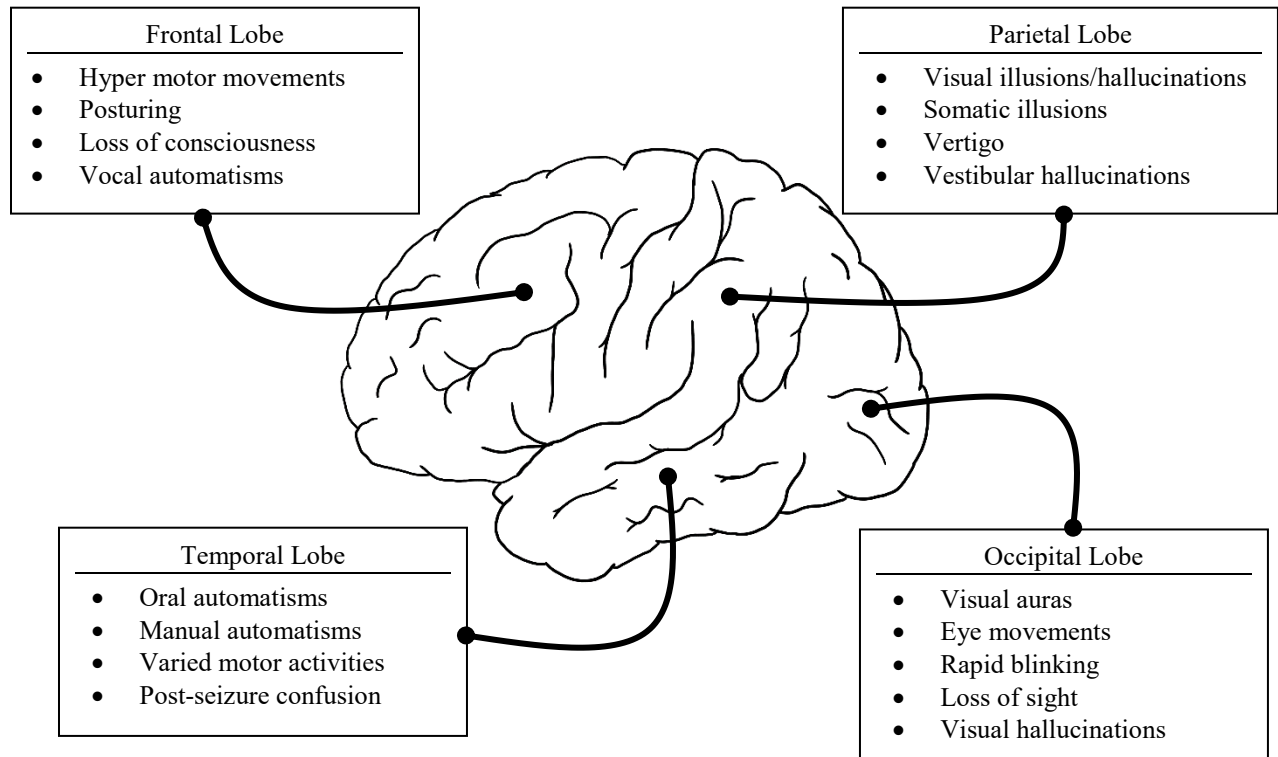- Visual hallucinations

Figure 2. Common focal impaired awareness seizure
onset regions and their characteristic symptoms.

Patients with epilepsy who suffer from FIA seizures can have difficulty with living independently since episodes are often unpredictable and can result in compromising situations, leading to trauma and serious accidents. FIA seizures have a higher rate of recurrence than generalized onset seizures. Patients may suffer from anxiety due to the stress of anticipating a seizure. Driving may not be possible due to the risk of having an episode while behind the wheel. Patients are at risk for sudden unexpected death in epilepsy (SUDEP) as well. Epilepsy patients also have a mortality rate 2–3 times greater than those without epilepsy. [1]

Treatments using medication, specialized diet, or surgery can be beneficial. However, tracking the progress and efficacy of treatments can be difficult given the varied symptoms of FIA seizures, particularly the loss of consciousness and not remembering that an episode has even occurred. Relying on caregivers or family members to be with the patient at all times is taxing. Seizures can also be missed if there is no one in the same room as the patient or if the symptoms do not reliably alert those monitoring the patient at home. In order to better detect and track focal impaired awareness seizures in patients, a robust, automated detection strategy is needed.

Current Seizure Detection Strategies

Specialized algorithms and machine learning strategies have been developed and refined for seizure detection. Various modalities have been studied, balancing physiological information, impact on the patients' daily life, reliability, and feasibility. When studying a condition that affects the neural activity of the brain, measuring the electrical activity of the brain is an important method to consider. However, other indirect measurement modalities have a greater potential for adoption outside of highly specialized clinical settings.

Scalp-level electroencephalography (EEG) provides a noninvasive measure of electrical activity across the surface of the brain. Detection of seizures by EEG greatly depends on the type of seizure, whether it is focal or more generalized, and how the specific features of a given episode present over time. EEG-based detection schemes have shown to be reliable with sensitivities over 75% and up to 90%. False alarm rates are also acceptable with some as low as one every ten hours, but others are more taxing at five per hour. Still, focal activity is often more difficult to reliably detect. Using EEG for seizure detection is also limited to clinical settings. Outpatient adoption is unlikely as patients would be unwilling to wear the EEG electrodes in daily life. [2]

Electrocorticography (ECoG), a system with electrodes implanted in the brain, can also be used for seizure detection. This method is far more invasive and thus limited to special cases. The limited number of electrodes also restricts use to patients with one or two seizure foci, where the onset of the seizure reliably occurs. Detection is based on repeatably observed abnormal activity at the seizure foci for each patient. [3]

Noninvasive seizure detection methods that have minimal impact on daily life are more tractable for outpatient care. Video-based, mattress-embedded, and wearable sensors have all been used to detect seizures. Video and mattress systems rely heavily on pronounced motor symptoms in order to detect seizures. These are most useful for generalized tonic-clonic seizures, but are not reliable for focal impaired awareness seizures. [4] [5]

Wearable sensors that can monitor acceleration and physiological parameters such as heart rate are noninvasive and commonly used by the general population in everyday life. The use of low-impact, wearable monitors is more likely to be adopted by patients at home during outpatient care, especially when common consumer-grade hardware is used. Wrist-based wearable sensors have been used widely for development of seizure detection strategies for hypermotor and

generalized tonic-clonic seizures. However, detection of seizures with minimal motor activity and FIA seizures have not been reliable and are often dismissed. [5]–[9] Nonmotor seizure detection strategies have relied on heart rate features often without the active involvement of movement detection. For nonmotor seizures characterized by an increase in heart rate, the seizure can be detected. However, the most successful sensitivity rates, 70% [10] and 74% [11], also have the greatest false alarm rates, 50/day and 216/day, respectively [9].

Aim

The aim of this study was to reliably detect focal impaired awareness seizures with an increase in heart rate using sensors from a consumer-grade smartwatch. The detector should not impose excessive false alarm rates while successfully identifying as many FIA seizures as possible. The algorithm ideally should also be lightweight enough to be deployed on consumer-grade devices.

Data Collection

Patient data were collected over a 13-month period in conjunction with ongoing research studying a variety of seizure types at Johns Hopkins Hospital. Patients with epilepsy are cared for in an epilepsy monitoring unit (EMU) for the purpose of diagnosing and providing treatment to patients. Scalp EEG, general vitals, and video monitoring is used to collect data regarding observed seizures. In addition to the standard collection of video-EEG data in the EMU, a consumer-grade wearable device, an Apple Watch, was worn by patients.

Among the larger set of patients and varied seizures, 41 focal impaired awareness seizures from 17 patients were recognized by a team of epileptologists. A characteristic increase in heart rate was evident by the sensors on the Apple Watch for 30 of the FIA seizures.

The only data streams directly used for this work were the heart rate measured by the photoplethysmogram (PPG) sensor and acceleration of the tri-axial accelerometers on the Apple Watch. The PPG sensor measures changes in light absorption on the skin surface, which can in turn be used to infer blood volume changes due to pulsing blood flow, thereby effectively measuring the heart rate. The practical realization of this process merits its own study as complications can arise from movement anomalies and other confounding factors. Most solutions rely on windows of approximately seven seconds to measure the heart rate and account for factors that would otherwise affect a more naïve approach. The Apple Watch uses an onboard system to directly produce a heart rate measure from the raw PPG sensor. The resulting heart rate was used for deriving features for FIA seizure detection.

Accelerometer data from the Apple Watch are directly measured by accelerometers placed orthogonally across the three dimensions of space. The x, y, and z axes of the wristwatch were defined as being along the length of the forearm, perpendicular to the forearm along the face of the watch, and perpendicular to the watch face, respectively. General signal processing strategies were used to account for common issues with accelerometer data from wearable devices. Each of the three accelerometer readings were effectively bandpass filtered. A high-pass filter using a second-order Butterworth filter with a 0.5 Hz cutoff frequency was used to isolate the linear acceleration component of the signal from the gravitational portions while also filtering out other low-frequency artifacts [12]. High-frequency noise was filtered out with a low-pass, fourth-order Butterworth filter with a cutoff frequency at 20 Hz.

General non-seizure data were also needed for this study. While non-seizure data were collected, it was not representative of daily activities. In order to properly train and validate against ambulatory settings and common real-world activities, open data sources were searched. Given the rise in human activity recognition through applied machine learning and open datasets, there are multiple datasets available. However, the data must be carefully selected so as to reasonably represent data that could be obtained in the same manner as the seizure data, i.e. with the wrist-worn Apple Watch.

The PAMAP2 physical activity monitoring dataset contains a comprehensive group of motion and physiological sensor data during various activities including walking, running, and household chores [13]. The dataset provides over ten hours of physical activity monitoring for nine subjects. Among the different sensor data in the PAMAP2 dataset is a wrist-worn, 3D accelerometer and heart rate recordings. In order to account for different sampling rates between the seizure dataset and non-seizure dataset, data was down-sampled to the lowest sampling rate. Furthermore, the developed features used for training and seizure detection were designed to inherently account for dropped data points and to normalize across datasets.

Seizure Criteria

Focal impaired awareness seizures were identified by epileptologists based on scalp-level EEG data and video recordings in an epilepsy monitoring unit. Motor responses, dates with the time of day, and seizure duration were all recorded. Analysis was restricted to FIA seizures with a characteristic increase in heart rate during or immediately preceding the seizure. FIA seizures with no heart rate increase were not expected to be able to be reliably detected and would likely

inhibit overall seizure detection and result in greater false alarm during normal activities. Figure 3 shows examples of accepted and rejected FIA seizures.
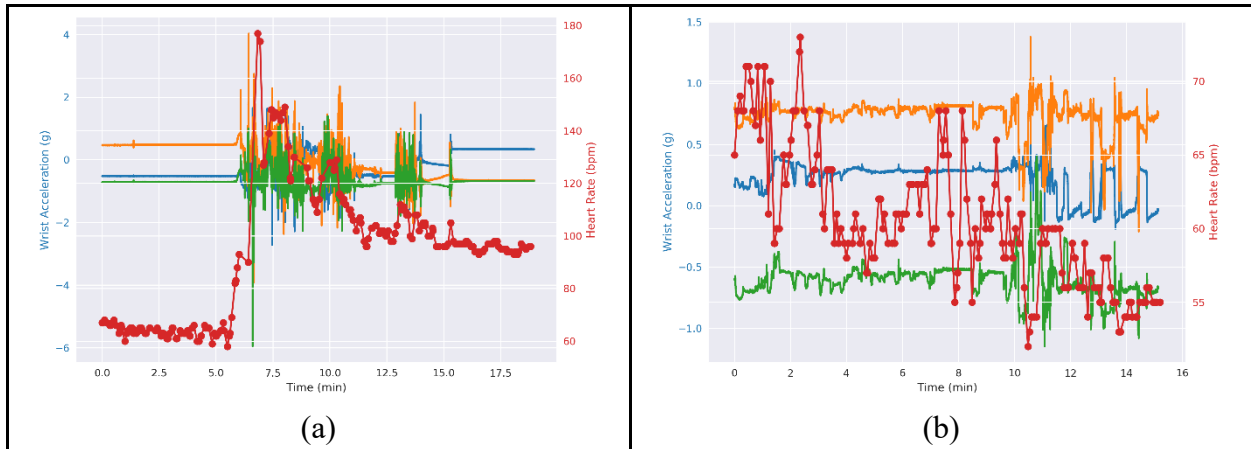

Figure 3. Example of a FIA seizure with a characteristic increase in heart rate (a) used in the study versus one with no discernable heart rate peak during or preceding the seizure (b). Orange, blue, and green traces correspond to the x, y, and z axes of the accelerometer, respectively. The red trace denotes heart rate.

Furthermore, labeling of identified seizures should be based on the available data streams for model prediction: heart rate (from PPG sensor) and movement (from accelerometers). Depending on the timing of cardiovascular and motor symptoms with the seizure onset, it would be expected that detection could be delayed. However, with only these secondary symptomatic data sources, training based on segments that yet show no symptoms would cause class confusion between seizure and non-seizure classes when training. Therefore, labeled segments for model training were restricted to FIA seizure segments that included the initial increase in heart rate.

In order to standardize seizure labeling based on heart rate, labeled FIA seizure onset was set to the time when the heart rate increased to beyond two standard deviations from the mean over the past five minutes. For the purposes of training the machine learning models, the detection of FIA seizures with cardiovascular symptoms can be thought of as detecting whether an increase in

10

heart rate is due to a FIA seizure or normal activity. Figure 4 shows an example of a labeled FIA seizure segment with surrounding non-seizure segments.
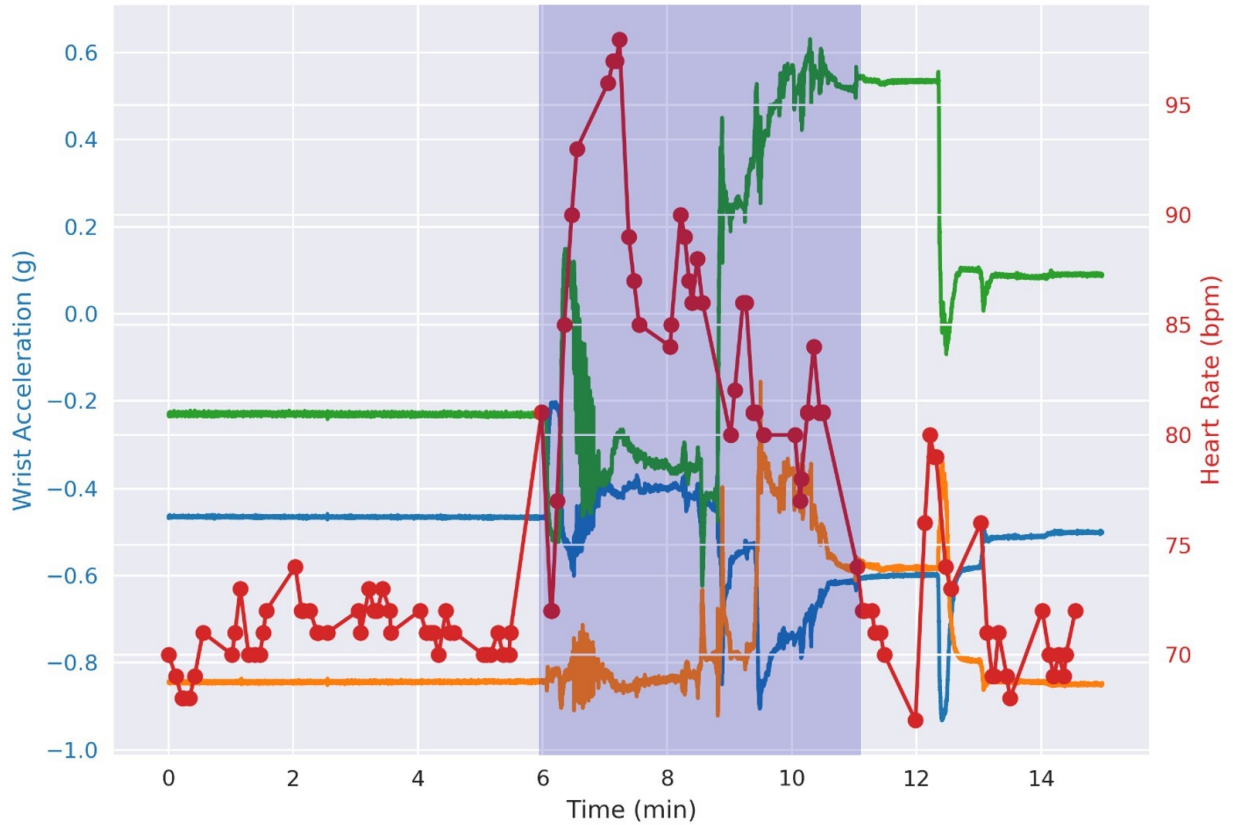
Figure 4. FIA seizure segment labeled based on heart rate increase of two standard deviations from the five-minute moving mean. Orange, blue, and green traces correspond to the x, y, and z axes of the accelerometer, respectively. The red trace denotes heart rate.

From the total of 41 identified FIA seizures, 30 seizures (73%) were characterized by an increase in heart rate of at least two standard deviations from the running five-minute mean.

Analysis Software

Python 3 was used to analyze the data for this work. The NumPy package was used for streamlined computing and data processing. Data analysis across variable moving windows data organization for machine learning training was accomplished with the Pandas package for the Python programming language. Data visualization was performed with Matplotlib package. Tensorflow 2 was used to create, train, and test machine learning models.

Feature Engineering

Heart rate and tri-axial acceleration data were obtained from the wrists of patients using the Apple Watch. From these four data streams, various features were created in order to enhance separability of focal impaired awareness seizures from a range of non-seizure activities. For this work, the goal was to detect FIA seizures with an increase in heart rate without presenting high false alarm rates during ambulatory settings and everyday life. The major premise/hypothesis that guided feature engineering was that the increase in heart rate would not be explainable by current or preceding physical activity. This distinction is made because it is assumed that different types and intensities of physical activity will have somewhat proportional heart rate responses. The benefit of using machine learning to separate FIA seizures from normal behavior is that every possible scenario does not need to have independent activation criteria. An entirely complete equation using movement and heart rate with different permutations and relationships during all normal physical activities, and then yielding a probability that a FIA seizure is occurring, does not need to be manually formulated. Instead, the analysis can be performed using a machine learning strategy.

Raw data can be used for many machine learning applications and has previously been explored in the field of seizure detection. However, raw data for seizure detection has mainly been restricted to EEG-based systems and still benefits from engineered or extracted features [14]–[16]. Data consistency can become a major confounding factor when relying on raw data. Consumer-grade wearables, even with real-time heart rate and accelerometer data, are still susceptible to measurement failures, causing missing data segments.

In order to prevent reliance on single datum instances and account for the different sampling rates between acceleration and heart rate, time-windowed, statistical features were used. Translation of the model to similarly sampled heart rate and acceleration data using different hardware also becomes available.

Statistical measures of raw data are often used for classification tasks in machine learning. For time series data, statistical measures can be applied over a moving window to enhance the predictive potential of the model. Typically, this is used for forecasting data — predicting the next value of the series before it is revealed or measured. For the purpose of seizure detection, the future physiological measures (heart rate and movement used here) are not meant to be predicted. Instead, the current and previous physiological measures are used to classify the activity and detect seizures against normal activity.

Engineered features ideally capture different aspects of the data, but do not necessarily need to be without mutual information. Rather, if two features are tightly coupled in class and become decoupled in another, this would be beneficial in creating separability during model training. Over-specific features that can be used to independently predict the class can also lead to the model training towards a local minimum, for example created by a heavily influential class, rather than training towards the optimal, more generalizable solution. This is often referred to as

the model finding a shortcut in the data. Maintaining more general features that still capture the relevant domain knowledge also provide more generalizable results with previously unseen, new data. Restricting to the fewest features that still result in optimal model performance further helps to reduce the memory and storage demands of the model, making it more available for deployment directly on wearable devices.

In order to maintain these principles for engineering the features, the only purely statistical measures performed on the heart rate and accelerometer data for feature engineering were moving means and standard deviations. These measures alone do not provide the necessary information to create a reliable model. There is no inherent history to the features nor can they be generalized across different patients with varying normal physiological levels. The history of time series data can be captured by recurrent neural networks (RNN), but the history used is typically recent predictions rather than recent features. Instead, selective time points in the recent past can be used for current features. We selected the time points to be used in our model based on initial observations of time differences between movement and heart rate changes during normal activities and FIA seizures. These were further refined by incorporating larger time differences and retaining those with the greatest positive effect on separability between FIA seizures and non-seizures while removing others.

The values of the statistical measures of heart rate and movement themselves are not particularly telling of whether a seizure is occurring. Rather, the progression or evolution of the statistical measures in relation to one another allows for accurate prediction. Since the watch can be worn in slightly different positions by each patient and automatisms and motor symptoms are often unique to a given patient, the L2 Euclidean norm of the three accelerometer values was used

14

to capture the meaningful component of motion and prevent confounding the model with misleading or unusable baselines.

In order to describe how these continuous, statistical features capture the onset and progression of FIA seizures, we used the following notation scheme. The data before implementing statistical measures, the heart rate and acceleration magnitude, are each referred to as x. The same methods were used on both data; therefore, a single representation is used to describe the feature engineering process for both. Subscripts denote the time in seconds from the current point, x_0. Given that the necessary data are from the recent past, ranges of time are shown by colons to avoid confusion with negative values. The past 10 seconds of data are denoted by $x_0:x_{-10}$. The $f(x)$ notation is used to represent the $mean(x)$ and $std(x)$ statistical functions for the purpose of describing the sliding window and comparison with recent past.

In the process of selecting features, the length of the sliding window, which time points to compare, and different statistical measures were tested. The final list of features was selected for brevity without sacrificing final selectivity of the resulting model. For these features, the current 10-second mean $[mean(x_0:x_{-10})]$ is compared to the 10-second mean from 10 seconds ago $[mean(x_{-10}:x_{-20})]$, 20 seconds ago $[mean(x_{-20}:x_{-30})]$, 30 seconds ago $[mean(x_{-30}:x_{-40})]$, 1 minute ago $[mean(x_{-60}:x_{-70})]$, and 2 minutes ago $[mean(x_{-120}:x_{-130})]$. This process is repeated for the current 10-second standard deviation. Both processes were used for the heart rate and acceleration magnitude. Table 1 shows a summary of the statistical measure comparisons applied. Figure 3 gives a visualization of the process applied to a sample data instance at t_0. In practice and real-time implementation, the previous raw data do not need to be stored, but only the resulting statistical measures $f(x)$.

Table 1. Engineered features of the change in windowed statistical measures over time from 10 seconds to 2 minutes for both heart rate and acceleration magnitude.

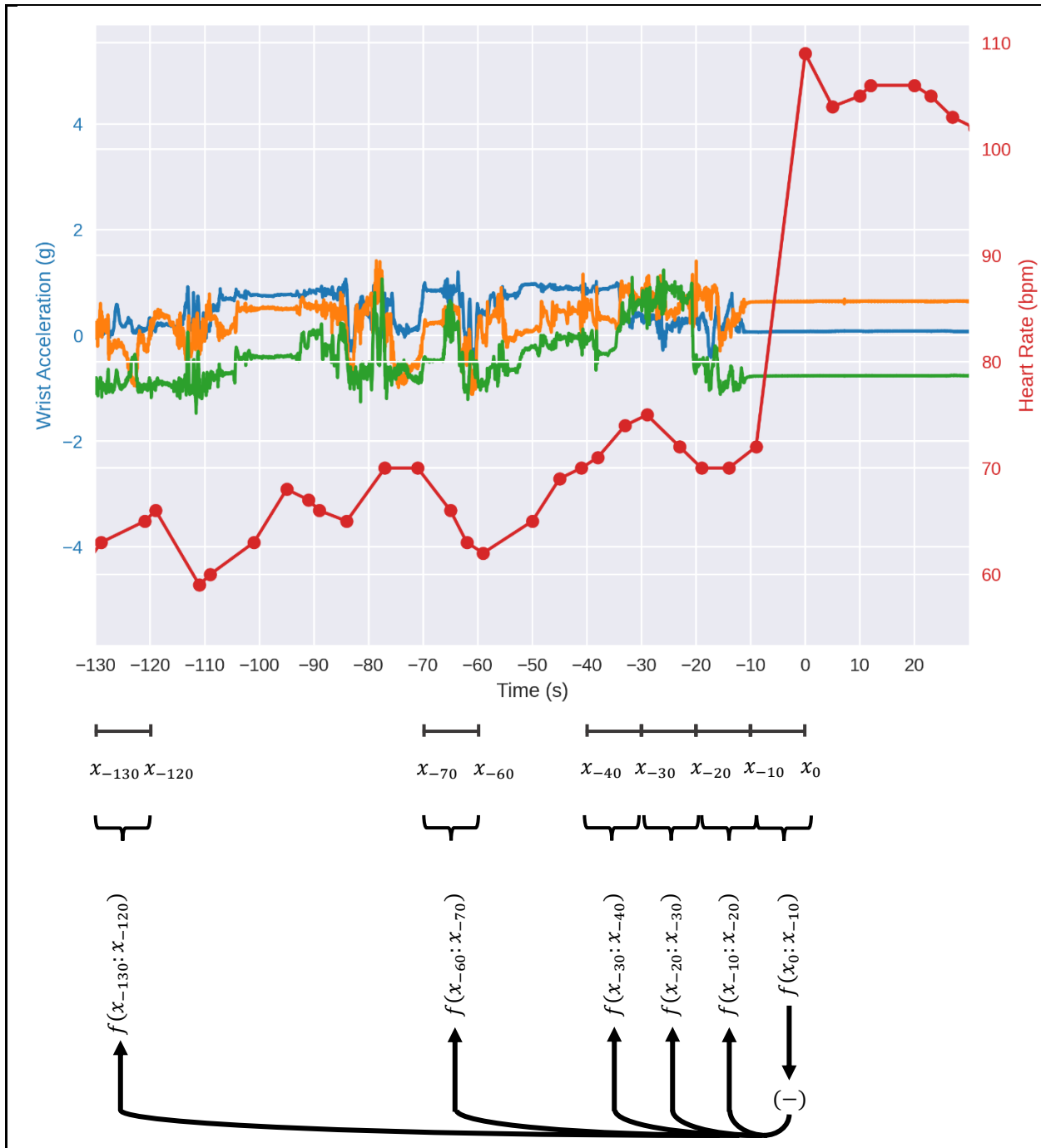| Change in the Mean	Change in the Standard Deviation
$mean(x_0:x_{-10}) - mean(x_{-10}:x_{-20})$	$std(x_0:x_{-10}) - std(x_{-10}:x_{-20})$
$mean(x_0:x_{-10}) - mean(x_{-20}:x_{-30})$	$std(x_0:x_{-10}) - std(x_{-20}:x_{-30})$
$mean(x_0:x_{-10}) - mean(x_{-30}:x_{-40})$	$std(x_0:x_{-10}) - std(x_{-30}:x_{-40})$
$mean(x_0:x_{-10}) - mean(x_{-60}:x_{-70})$	$std(x_0:x_{-10}) - std(x_{-60}:x_{-70})$
$mean(x_0:x_{-10}) - mean(x_{-120}:x_{-130})$	$std(x_0:x_{-10}) - std(x_{-120}:x_{-130})$

Figure 5. Process of creating features for data at t_0. Selected 10-second sliding windows of the data are used. The mean and standard deviation functions are represented by the $f(x)$ notation.

Machine Learning Model Design

Machine learning models have been applied to extensive fields for the purpose of analyzing data into meaningful results. Models are often highly specific to the individual task at hand and must be designed to accommodate the type of input data and yield the proper output. There is rarely, if ever, an agreed upon optimal model for a given task; many different designs will give comparable final results. There are still general principles to follow when designing and testing a new model. Tradeoffs must be made regarding complexity and computational demand versus predictive performance. We found that a binary classification artificial neural network with seven layers produced reliable FIA seizure detection without overdue complexity or computational demand.

In order to accurately detect focal impaired awareness seizures using the Apple Watch, current physiological measures of cardiovascular activity and upper body movement must be categorized as either occurring during a FIA seizure or during normal activity. The discrete measures of heart rate and wrist acceleration are being recorded continuously in time, creating time series data with multiple, concurrent channels. This type of data is often used in the context of predicting future values in the time series rather than categorizing the segments of the data into different states or classes. The data can also be analyzed in a manner similar to image classification tasks. However, these methods focus on borrowing entire model designs from other tasks rather than designing a model to specifically fit the task domain from the bottom up. Instead, engineered features that encode for the present and different time points in the recent past are continuously updated in time and used as first layer inputs in an artificial neural network.

Recurrent neural networks are a class of algorithms designed for handling time series data of variable lengths. These networks can retain states of internal neurons and results from previous

time points in order to aid in prediction for future time points [17]–[19]. While the true transition from pre-seizure to seizure and then back to normal activity may be continuous, the data have only been labeled as one of the discrete selections, seizure or non-seizure. Furthermore, the input segments are consistent in length and history has already been embedded in the engineered features.

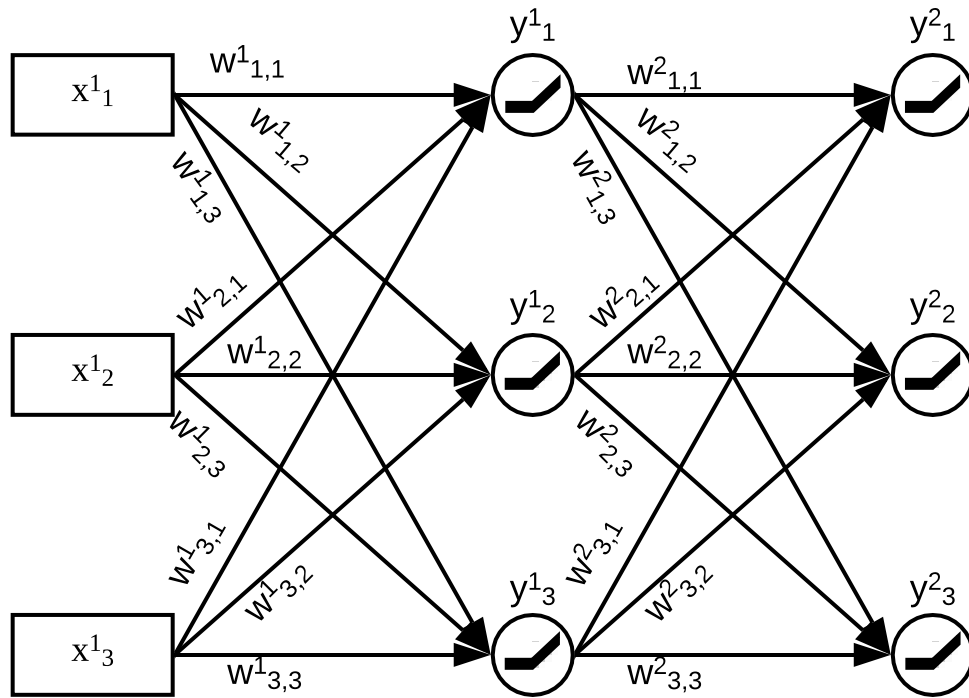Figure 6. Diagram of a simplified neural network segment.

The basic units of an artificial neural network are the individual neurons. At each layer, the neurons take input from the neurons or engineered features from the previous layer. A weighted sum of all inputs is performed and then mapped to an activation function for the given neuron. The weights are systematically updated during training in order to optimize the

performance of the overall network. Figure 6 shows a segment of simplified neural network. Notation of the weights w shows the layer in the superscript and the input neuron followed by the output neuron in the subscript ($w_{input\ neuron,\ output\ neuron}^{layer}$).

The activation function a is applied to the weighted sum input for each neuron. As with the variable weights, the use of an activation helps to accommodate for interactions between input neurons. Activation functions also support the model in addressing non-linearities needed to increase separability for the overall predictive performance. For all of the neurons in the hidden layers, the rectified linear unit (ReLU) activation function was used. The ReLU function allows for faster training during gradient descent with limited slopes of zero or one. Variable bias terms for each neuron (i.e. where the linear unit portion reaches the x-axis) allow for non-linearities to be addressed.
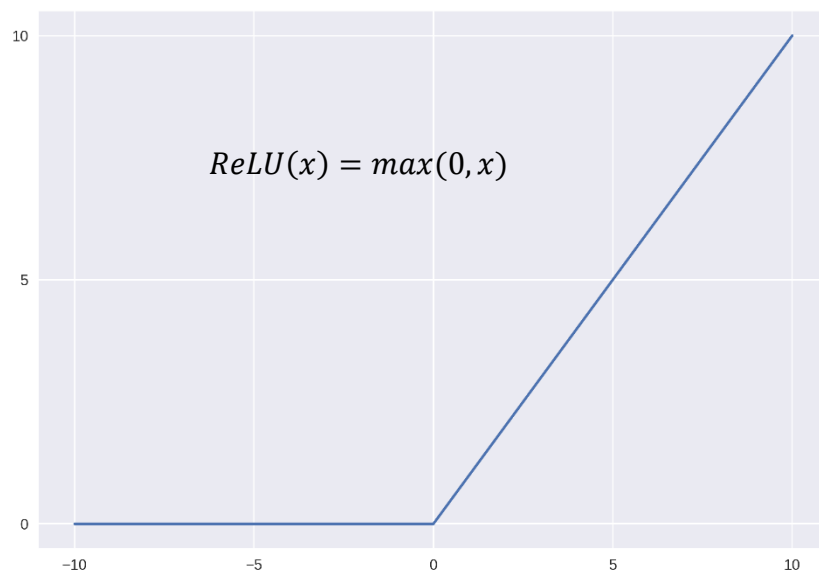

Figure 7. Rectified linear unit (ReLU) activation function.

The depth of the artificial neural network should be deep enough to allow for sufficient interactions between inputs and effective model training without being so long as to create unnecessary computational demand with minimal to no improvement on performance. More layers create further abstraction from the input data. Since the input layer already uses domain-specific engineered features, extensive abstraction is not needed. Following the input layer of engineered features is a 10-neuron dense layer, three consecutive 30-neuron layers, then another 10-neuron layer followed by a final single neuron layer.
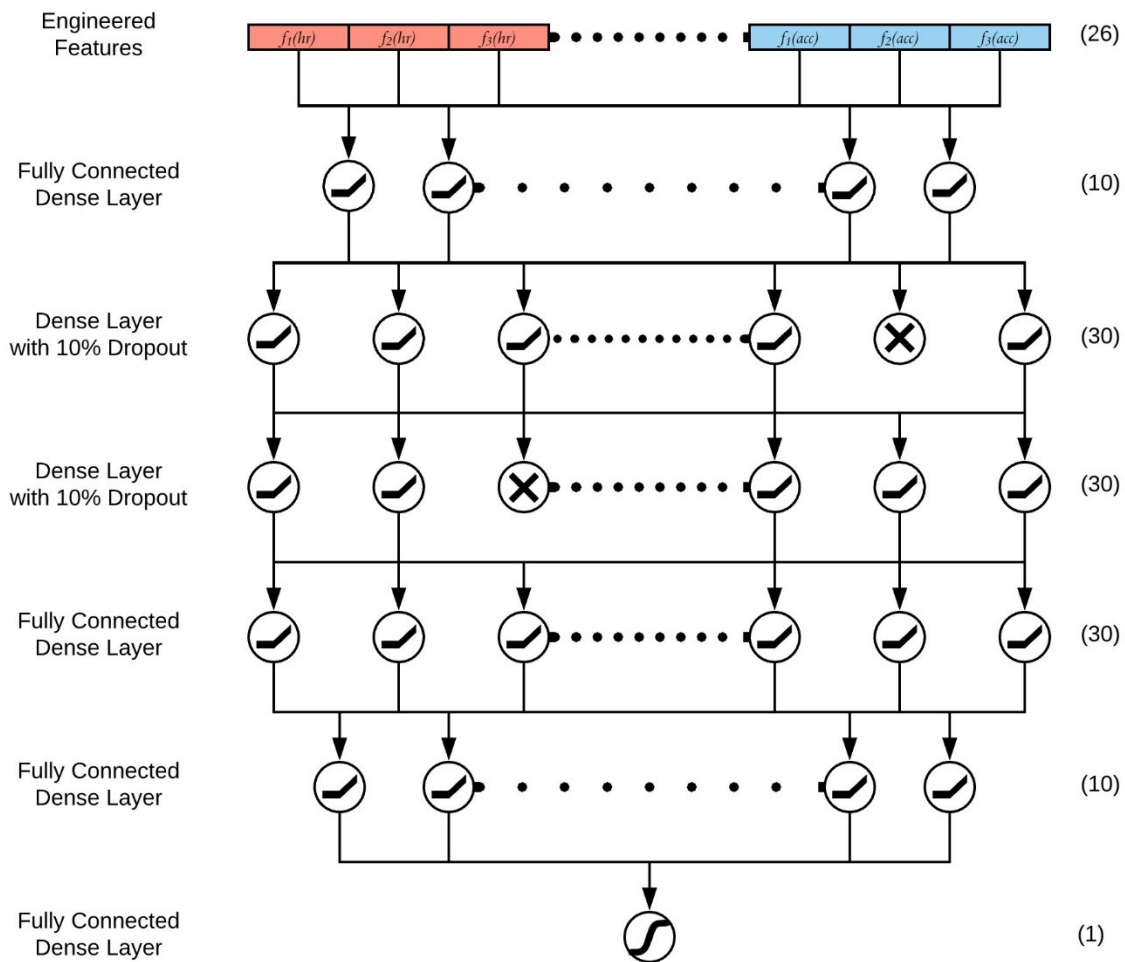

Figure 8. Diagram of the binary classification neural network
with activation functions and dropout shown.

Dropout between layers during training prevents the model from overfitting and heavily relying on individual neurons or connections. Connections can also be codependent without dropout, resulting in a lower learning potential. During a given training phase, individual neurons and their associated connections are dropped or zeroed out. Dropout was used at a rate of 10% on the first two 30-unit layers, effectively creating layers of 27 neurons that interchange with each propagation of training.

The final layer of the neural network uses the more computational sigmoid function in order to separate the two classes, FIA seizure versus non-seizure, with a single neuron. Combined with the binary cross-entropy loss function, the final layer allows for effective binary classification.

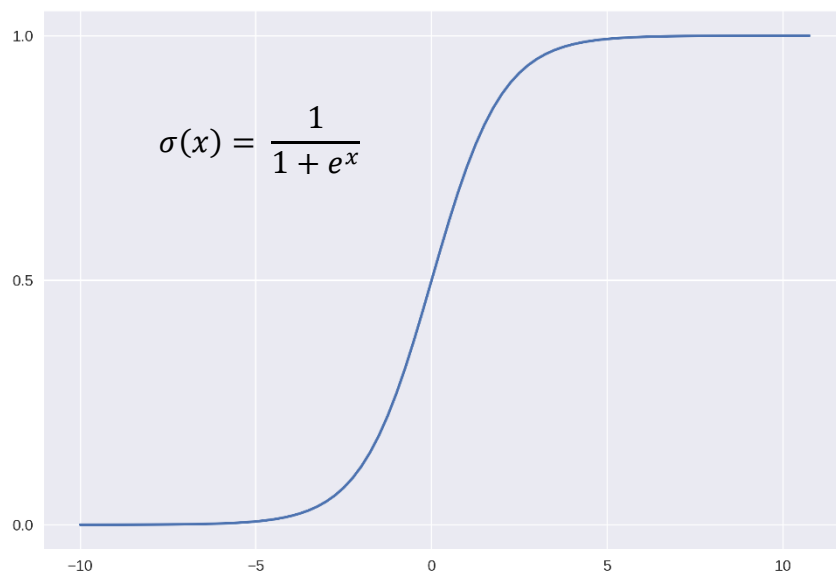$$\sigma(x) = \frac{1}{1 + e^x}$$

Figure 9. Sigmoid activation function.

Due to the relatively low amount of FIA seizure data available, leave-one-out cross-validation (LOOCV) was used to train the network. In the LOOCV training paradigm, the model is trained on all but one of the data samples and then tested on the remaining one. In order to simulate a new FIA seizure event and avoid training the model on part of a seizure that it would be tested on, an entire seizure was left out of training for validation, not just a segment of the seizure. If part of the same seizure were to be used for training and validation, it would result in misleading results favoring improved predictive performance. In total, 31 models were trained on the same architecture (Figure 8) for the 30 FIA seizures. One model was trained on all of the available seizures to measure false alarm rates in a larger representative dataset of everyday activities including household chores and exercise [13].

Results

The metrics used to train the model were limited to the optimization of accuracy through the binary cross-entropy loss function. When evaluating the overall performance of the trained models, further analysis is needed. The aim for seizure detection is to identify as many of the seizures as possible while minimizing the number of false alarms. True positive (TP), false positive (FP), true negative (TN), and false negative (FN) metrics are used in terms of the positive detection of a FIA seizure, as labeled by 10-second binned increments. These values were used to calculate the sensitivity, specificity, precision, and accuracy of each model with training and testing data. The number of independent FIA seizures recognized is needed to understand if there are any missed seizures for a given model. The false alarm rate, defined here as the number of false alarms during normal activity over a given period of time, is also used as it is more meaningful in practice than the false discovery rate, which does not account for time.

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$False\ Alarm\ Rate = \frac{FP}{Time}$$

The performance of the binary classification artificial neural network was assessed in regards to how well focal impaired awareness seizures were identified against normal activities. The same architecture was trained on 30 combinations, each with 29/30 of the FIA seizures using leave-one-out cross-validation. When local minima resulting in below 70% training accuracy prevented further training, the given model was retrained with a new initialization. Metrics of the training sets show how well the given neural network converged towards optimal separation of the two classes, seizure and non-seizure. The performance of each model on the remaining FIA seizure during validation tests showed how well each model can generalize to new data. This was used to estimate how well the models would perform when applied to new patients.

The artificial neural networks were trained over 60 epochs in an effort to capture when the performance of the model reached its maximum. Figure 10 shows the average accuracy and loss of the models over each training epoch with one standard deviation highlighted. Most models reached near their optimal accuracy at 30-40 epochs. However, the trend towards improved accuracy continued into 60 epochs of training, albeit with severely diminished returns per epoch. A similar trend is apparent in the binary cross-entropy loss function as it is minimized over the epochs. At 60 epochs an average training accuracy of 84% is reached and the average binary cross-entropy loss goes to 0.553.
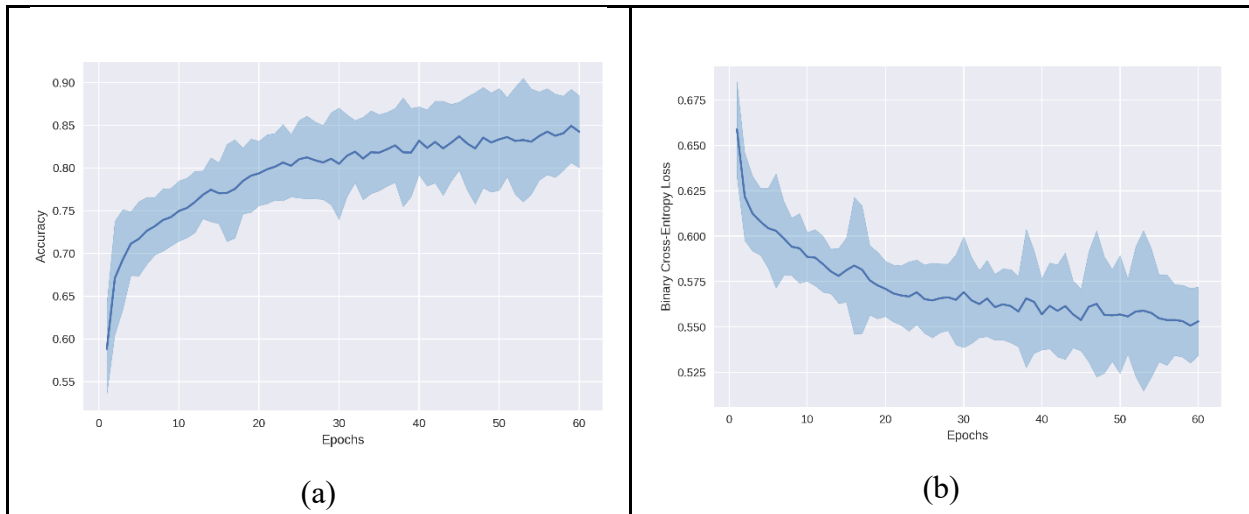
Figure 10. Average training accuracy and loss over 60 epochs
with one standard deviation errors highlighted for the 30 LOOCV neural networks.
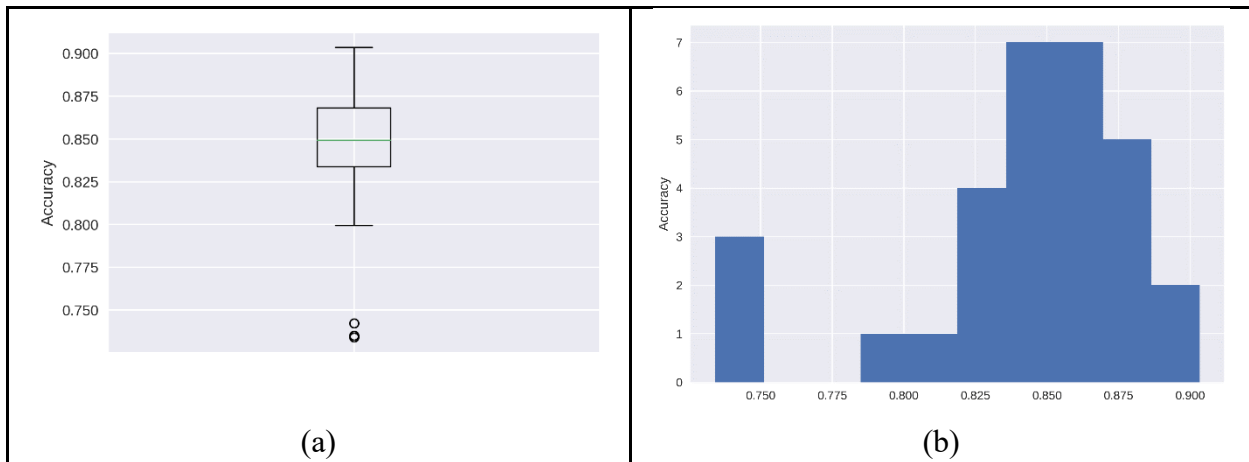

Figure 11. Boxplot (a) and histogram (b) of the final training accuracies
after 60 epochs for the 30 LOOCV neural networks.

The average training accuracy and standard deviation does not clearly show the disparity in final training accuracies across the models. A boxplot and histogram of the final training accuracies is presented in Figure 10. The majority of training accuracies fall between 82% and 88% with the maximum reaching as high as 90%. There are three outliers among the neural networks with accuracies of 73%–74%.

In an effort to obtain the greatest separability between FIA seizures and normal activity, the binary classification neural network was also trained on all available seizure data without withholding a separate validation or test set. This would be expected to yield the lowest false alarm rate and greatest FIA seizure detection accuracy, but is highly susceptible to overfitting. Without a validation or test set of the artificial neural network, the presence of overfitting cannot be ruled out. As shown later, the best-performing networks trained using LOOCV were able to perform similarly to the best-performing iterations of the network trained on all available data.

Figure 12 shows the sensitivity, specificity, precision, and accuracy for 10 iterations of the binary classification neural network. Local minima in the loss function were present, including one resulting in an accuracy of 55% that the model was not able to escape when reached. Due to this and other expected local minima, models trained to accuracies below 70% were retrained.

The results of the fully-trained model show variable sensitivity, suggesting that some seizures could have been entirely missed in certain models. Still, average sensitivity for the networks was 80%, with the highest reaching 91%. Specificity, precision, and accuracy, all showed good performance with average values of 93%, 94%, and 86%, respectively. The top-performing model trained with all of the available data was able to identify every seizure, as expected, and had a false alarm rate of 1/hour for normal activity including walking, running, and household chores. However, it is expected that this model could be overfit and not perform well

if presented with new data. In order to address this concern, LOOCV was used to train further neural networks.
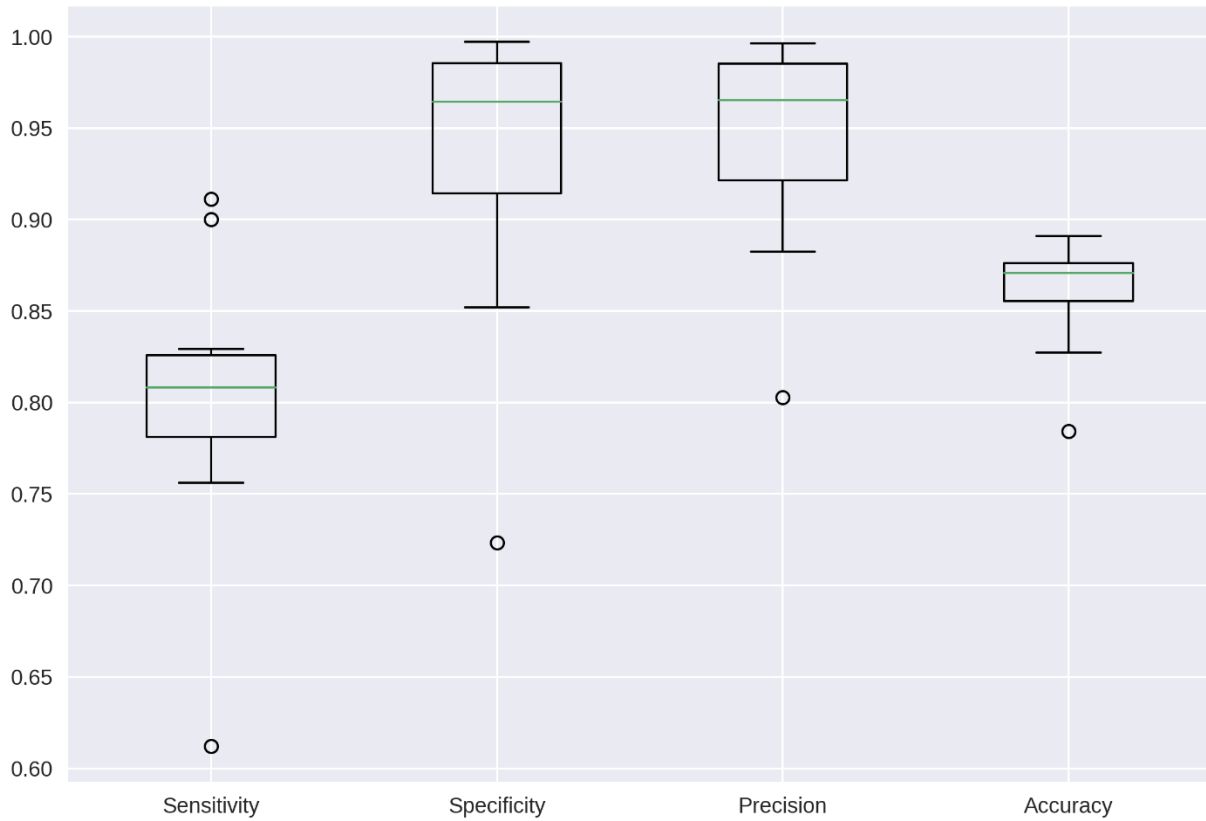
Figure 12. Metrics from 10 iterations of the binary classification neural network trained on all available with no separate validation data.

The training dataset results are able to show how each of the neural networks converge to a final combination of weights between neurons in an attempt to optimize the final accuracy through the binary cross-entropy loss function. These do not show how well the models would perform on unseen data. For this, each neural network was tested on the FIA seizure that it was

not trained on. Note that these are not randomized segments of seizures, but completely independent seizures, simulating an entirely new FIA seizure event for each of the models.

Figure 13 shows the summary of the validation set results for the 30 LOOCV neural networks. The final results are disparate in terms of the final metrics. This can be due to both how the model converged on a final set of weights and the individual seizures they were each validated against. Still, each model was able to identify the respective previously unseen FIA seizure. However, the individual seizure length labels vary from 6 minutes to 40 seconds, allowing for greater FN rates in longer seizures given that all seizures were still identified. Furthermore, 25 FIA seizures were detected directly at the onset of the labeled seizure based on heart rate increase, which would follow the 10-second delay due to feature windows. Four of the seizures were detected at the second binned window, 20 seconds following the initial increase in heart rate from the FIA seizure, and the remaining seizure was identified after 30 seconds. This strategy was useful in rejecting non-seizure activity at rest, with false alarm rates as low as 1/hour. However, when tested against the larger validation set including a range of normal activity, the false alarm rate was shown to be rather high at 83/hour.

The false alarm rate can be greatly reduced using accumulation filtering. Accumulation filtering is applied after the artificial neural network has made its predictions regarding where the time segments will be identified as FIA seizure events or non-seizure events. The final identification of a FIA seizure with accumulation filtering requires that continuous 10-second segments are identified as seizures. For a 30-second accumulation filter, the neural network would need three positive predictions in a row for the segment to be identified as a seizure.
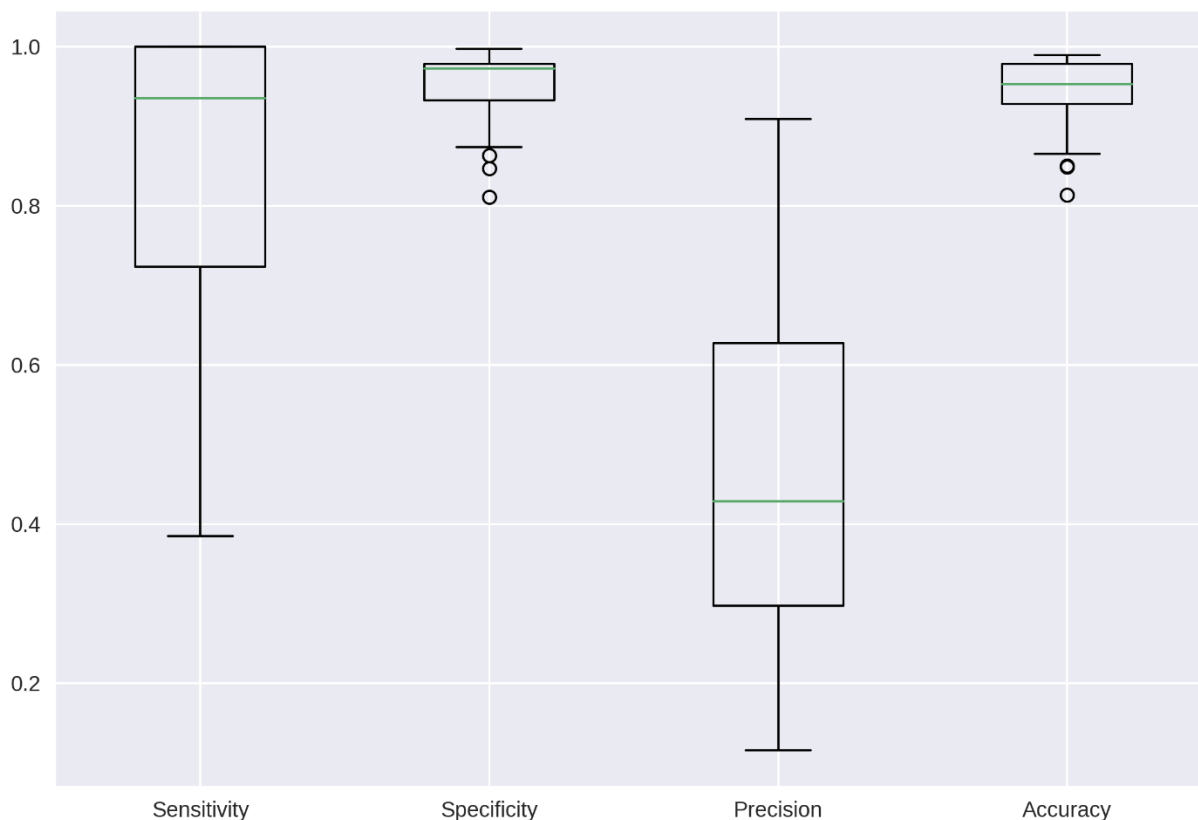
Figure 13. Validation set results for the 30 LOOCV neural networks
tested on complete FIA seizures that were not trained on.

Using a 30-second accumulation filter, all but one of the FIA seizures are still able to be

identified for the given model. The single missed seizure was the one with the shortest labeled

segment, lasting only 30 seconds itself. In order for this seizure to be identified, it would need

each 10-second interval to be properly identified. The false positive rate is reduced from 83/hour

to 11.7/hour when using the accumulation filtering. This would still be approximately one false

alarm every five minutes. Furthermore, false alarms were set off during some of the least physical

activities including lying down, sitting, and standing. A false alarm every five minutes would

likely be cumbersome for the patient and could easily result in user rejection. The false alarm

occurrence for activities with low physical demand should be preferably lower than when the user is active.

In order to reduce the false alarm rate to a manageable value that would not cause undue strain on the user, an acceptable balance between identifying every FIA seizure as quickly as possible and maintaining a low false alarm rate is needed. Ideally, every seizure would be identified and there would be no false alarms. However, this is not reasonably expected and would rather suggest strong overfitting of the data to the training and test sets.

The final model with a balanced FIA seizure detection sensitivity and false alarm rate had per-seizure sensitivity of 90%, identifying 27 out of 30 FIA seizures with a cardiovascular component. The false alarm rate for the model, based on validation time segments unseen during training, was 1.65/hour, or one every 36.45 minutes. The false alarm rate was even more promising when analyzing results from the activity-labeled PAMAP2 dataset for further possible false alarms.

No false alarms were present while subjects were lying down, sitting, or standing. However, there were false alarms during periods when subjects got up from after lying down at an event-based rate of 25% for the limited number of eight recorded events. There was a limited number of the transient events since the dataset was intended for the recording of discrete activities rather than the transition between activities. Lying down and sitting up were back-to-back events, each with considerably more data (Table 2), but recordings of the transition allowed for the capture of how the models react during transitions.

The majority of false alarms occurred while subjects were ascending stairs with a false alarm rate of 0.41/minute, or one every two minutes and 26 seconds. This is reasonable occurrence based on the motivation of the engineered features. There would likely be an increase in heart rate that is higher for the given wrist movements when compared to other physical activities. This

would also be an event that the patients would be more alert to the alert of a predicted seizure with few distractions. The remaining false alarms occurred during cycling at a rate of 2.19/hour. Table 2 shows a summary of the false alarm rate for the final model with accumulation filtering broken down by each activity as labeled in the PAMAP2 dataset.

Table 2. False alarm rates of the final neural network with accumulation filtering broken down by activity for the available data labeled by activity.

Activity	Time (s)	Events	FAR
Lying down	1925	8	0
Sitting	1851	8	0
Standing	1899	8	0
Walking	2387	8	0
Running	981	6	0
Cycling	1645	7	0.04/min
Nordic walking	1881	7	0
Computer work	3099	4	0
Ascending Stairs	1172	8	0.41/min
Descending Stairs	1049	8	0
Vacuuming	1753	8	0
Ironing	2386	8	0
Folding laundry	998	4	0
Cleaning	1871	5	0
Jumping Rope	493	6	0
Getting up from lying down	430	8	25%
Remaining transient/unlabeled	6610	-	0

FAR – false alarm rate

Conclusions

A binary classification, artificial neural network architecture was used to create a focal impaired awareness seizure detector using sensors from a smartwatch, the Apple Watch. Focal impaired awareness seizures were limited to those with a characteristic increase in heart rate. Custom features were engineered from heart rate and wrist acceleration in order to capture the difference between heart rate spikes due to FIA seizures versus normal activity including walking, running, and household chores. Leave-one-out cross-validation was used to ensure reliable testing of the model and prevent overfitting. Results of the network depend on how training progresses using the binary cross-entropy loss function. Each of the models trained with LOOCV were able to identify the left-out FIA seizure within 30 seconds, 12 seconds on average. However, high false alarm rates are persistent across models optimized for identifying every FIA seizure with a cardiovascular component. A balance between identifying as many FIA seizures as possible while maintaining acceptable false alarm rates was made with a neural network and accumulation filtering that resulted in 90% precision on a per-seizure basis and an overall false alarm rate of 1.65/hour with the majority of false alarms occurring during ascending stairs. No false alarms were present during resting, running, or household chores. Despite validation on separate data from training, overfitting is still a concern, as with all neural networks. The premise for FIA seizure detection with a cardiovascular component using a consumer-grade smartwatch was shown using a transferrable model for real-time deployment.

This work is a novel step in accessible seizure detection, expanding from current detectors which focus on generalized tonic-clonic seizures with dramatic motor symptoms, allowing for detection of focal impaired awareness seizures with heart rate symptoms and often minimal motor symptoms. Future work will be needed in transferring the model to wearable devices, such as the

Apple Watch. Using a pretrained model on mobile devices is currently accessible. Still, real-time testing is needed for validation. The model is designed to easily use real-time data with minimal memory restraints. Performance metrics are needed for both the performance of the model and the resources it consumes when used on consumer-grade wearable devices.

References

[1] A. Kumar and S. Sharma, "Complex Partial Seizure," in *StatPearls*, Treasure Island, FL: StatPearls Publishing, 2020.

[2] C. Baumgartner and J. P. Koren, "Seizure detection using scalp-EEG," *Epilepsia*, vol. 59, pp. 14–22, Jun. 2018, doi: 10.1111/epi.14052.

[3] B. B. Ma and V. R. Rao, "Responsive neurostimulation: Candidates and considerations," *Epilepsy & Behavior*, vol. 88, pp. 388–395, Nov. 2018, doi: 10.1016/j.yebeh.2018.09.032.

[4] A. Ulate-Campos, F. Coughlin, M. Gaínza-Lein, I. S. Fernández, P. L. Pearl, and T. Loddenkemper, "Automated seizure detection systems and their effectiveness for each type of seizure," *Seizure*, vol. 40, pp. 88–101, Aug. 2016, doi: 10.1016/j.seizure.2016.06.008.

[5] J. B. A. M. Arends, "Movement-based seizure detection," *Epilepsia*, vol. 59, pp. 30–35, Jun. 2018, doi: 10.1111/epi.14053.

[6] F. S. S. Leijten and the Dutch TeleEpilepsy Consortium, "Multimodal seizure detection: A review," *Epilepsia*, vol. 59, pp. 42–47, Jun. 2018, doi: 10.1111/epi.14047.

[7] X. Zhao and S. D. Lhatoo, "Seizure detection: do current devices work? And when can they be useful?," *Curr Neurol Neurosci Rep*, vol. 18, no. 7, p. 40, Jul. 2018, doi: 10.1007/s11910-018-0849-z.

[8] G. Regalia, F. Onorati, M. Lai, C. Caborni, and R. W. Picard, "Multimodal wrist-worn devices for seizure detection and advancing research: Focus on the Empatica wristbands," *Epilepsy Research*, vol. 153, pp. 79–82, Jul. 2019, doi: 10.1016/j.eplepsyres.2019.02.007.

[9] S. Beniczky and J. Jeppesen, "Non-electroencephalography-based seizure detection:," *Current Opinion in Neurology*, vol. 32, no. 2, pp. 198–204, Apr. 2019, doi: 10.1097/WCO.0000000000000658.

[10] K. Vandecasteele *et al.*, "Automated Epileptic Seizure Detection Based on Wearable ECG and PPG in a Hospital Environment," *Sensors*, vol. 17, no. 10, p. 2338, Oct. 2017, doi: 10.3390/s17102338.

[11] R. S. Fisher *et al.*, "Automatic Vagus Nerve Stimulation Triggered by Ictal Tachycardia: Clinical Outcomes and Device Performance-The U.S. E-37 Trial: VNS IN RESPONSE TO ICTAL TACHYCARDIA," *Neuromodulation: Technology at the Neural Interface*, vol. 19, no. 2, pp. 188–195, Feb. 2016, doi: 10.1111/ner.12376.

[12] Y. Oshima *et al.*, "Classifying household and locomotive activities using a triaxial accelerometer," *Gait & Posture*, vol. 31, no. 3, pp. 370–374, Mar. 2010, doi: 10.1016/j.gaitpost.2010.01.005.

[13] A. Reiss and D. Stricker, "Introducing a New Benchmarked Dataset for Activity Monitoring," in *2012 16th International Symposium on Wearable Computers*, Newcastle, United Kingdom, Jun. 2012, pp. 108–109, doi: 10.1109/ISWC.2012.13.

[14] A. Page, J. Turner, T. Mohsenin, and T. Oates, "Comparing Raw Data and Feature Extraction for Seizure Detection with Deep Learning Methods," p. 4.

[15] J. T. Turner, A. Page, T. Mohsenin, and T. Oates, "Deep Belief Networks used on High Resolution Multichannel Electroencephalography Data for Seizure Detection," p. 7.

[16] K. J. Cios, L. A. Kurgan, and M. Reformat, "Machine learning in the life sciences," *IEEE Eng. Med. Biol. Mag.*, vol. 26, no. 2, pp. 14–16, Mar. 2007, doi: 10.1109/MEMB.2007.335579.

[17] F. A. Gers, "Learning to Forget: Continual Prediction with LSTM," *Learning to Forget*, p. 21.

[18] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *arXiv:1406.1078 [cs, stat]*, Sep. 2014, Accessed: Apr. 24, 2020. [Online]. Available: http://arxiv.org/abs/1406.1078.

[19] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning," *arXiv:1506.00019 [cs]*, Oct. 2015, Accessed: Apr. 24, 2020. [Online]. Available: http://arxiv.org/abs/1506.00019.

Curriculum Vitae

Andrew Masteller

Education

Master of Science and Engineering in Biomedical Engineering 2020

Neuroengineering Focus

Johns Hopkins University

     Thesis: "Focal Impaired Awareness Seizure Detection Using a Smartwatch"

     Thesis Committee: Nathan Crone, MD; Gene Fridman, PhD; J. Webster Stayman, PhD

Bachelor of Science in Biomedical Engineering 2016

Medical Imaging and Devices Focus

The Pennsylvania State University

Schreyer Honors College

     Thesis: "Real-Time Optical Neuroimaging"

     Thesis Committee: Patrick Drew, PhD; William Hancock, PhD; Nanyin Zhang, PhD

Research Experience

Graduate Student Researcher 2019 – 2020

Johns Hopkins Cognitive Neurophysiology and Brain-Machine Interface Lab — Nathan Crone, MD

     Processed wearable sensor data streams from the Apple Watch using Python and further developed machine learning strategies for seizure detection with the TensorFlow framework

Graduate Research Assistant 2019

Johns Hopkins Cardiac Bioelectric System Lab — Leslie Tung, PhD

     Constructed and tested an automated beat detection system for cardiomyocytes using applied circuit design theory and MATLAB software programming to be used in studying cardiomyocyte beat rate variability due to drug application for extended durations

Undergraduate Research 2014 – 2016

Penn State Center for Neural Engineering, Drew Lab — Patrick Drew, PhD

     Designed and built an optical neuroimaging system to map functional regions of the mouse somatosensory cortex during cranial surgeries using automated stimulations and real-time analysis of Hb and HbO_2 relative concentrations via optical absorption

Honors and Awards

Biomedical Engineering Thesis Award 2016

     Biomedical Engineering Department, The Pennsylvania State University

College of Engineering Research Initiative (CERI) Grant 2014

     College of Engineering, The Pennsylvania State University