Detecting Self-harming Activities with Wearable Devices

Levi Malott*, Pratool Bharti*, Nicholas Hilbert*, Ganesh Gopalakrishna†, and Sriram Chellappan*
*Missouri University of Science and Technology
Department of Computer Science
Rolla, Missouri 65409
Email: {levi.malott, pbtn4, nsh9b3, chellaps}@mst.edu
†University of Missouri
Department of Clinical Psychiatry
Columbia, Missouri 65201
Email: gopalakrishnag@health.missouri.edu

***Abstract*—In the United States, there are more than $35,000$ reported suicides with approximately $1,800$ of them being psychiatric inpatients. Staff perform intermittent or continuous observations in order to prevent such tragedies, but a study of 98 articles over time showed that 20% to 62% of suicides happened while inpatients were on an observation schedule. Reducing the instances of suicides of inpatients is a problem of critical importance to both patients and healthcare providers. In this paper, we introduce *SHARE* - A Self-Harm Activity Recognition Engine, which attempts to infer self-harming activities from sensing accelerometer data using smart devices worn on a subject's wrist. Preliminary classification accuracy of 80% was achieved using data acquired from 4 subjects performing a series of activities (both self-harming and not). The results, application, and proposed technology platform are discussed in-depth.**

I. Introduction

Annually, psychiatric inpatient suicides account for more than 5% of the more than $35,000$ suicides in the United States [1]. Psychiatric hospitals most often implement 15-minute manual checks on inpatients as a preventive measure. In [2], the charts of 76 patients, who committed suicide while in the hospital or immediately after discharge revealed that 51% were on 15-minute checks or one-to-one observation. Additionally, in 2010, Bowers, et al. conducted a systematic study from 98 articles covering nearly $15,000$ inpatient suicides [3]. Of the reviewed cases, 20% to 62% of suicides happened on intermittent observation and 2% to 9% on constant observation [1]. Needless to say, the number of suicide attempts (some of which could be unsuccessful) will far out number actually documented cases due to stigma associated with truthfully reporting them. While manual suicide checks are effective at reducing the number of suicides, they consume significant nursing resources, and conducting checks may be sometimes neglected due to other responsibilities of staff [4]. Clearly, there is a need for an alternative or supplementary procedures to reduce the cost of suicide checks at medical facilities, while simultaneously not compromising on patient care.

Recently, smart-watches are providing new ways to view health and healthcare. Already, many smart-watches or wrist-worn devices with in-built sensors are in the market, and primarily focused on consumer health applications. Sleep monitoring and step counting are the most common functions implemented on them. Not enough of either one can lead to degradation of performance in daily activities. An accelerometer is one such sensor on modern smart-watches that can measure both sleep and physical activity (in the form of step counting). Long periods of lull movement indicate sleep (or rest) and sinusoidal patterns (from arm swinging while walking/running) indicate steps. Monitoring sleep and movement can provide unique insights on day-to-day health, while maintaining a low cost and have motivated many of the health-emphasized smart devices permeating the market.

Determining the type of activities, like sleep and movement, is achieved through complex *activity recognition* algorithms, fusing information from multiple sensors. Many research studies have been conducted to use activity recognition to produce more context-aware applications. Using a combination of smart-phones or other wearable sensors, the goal is to accurately identify user activities (cooking, cleaning, eating, etc.) from low-level sensor data. These systems have unique applications to health fields. Within the context of this paper, namely, psychiatric inpatients with suicide risk (or unknown risk), automated continuous monitoring of patient activity could be employed to alert staff of self-harming activities. If an ideal system were developed (that detects all self-harmful activities), then the need for suicide checks could be be obsolete. While this may not be feasible always, there are clear tangible benefits to at-risk patients and healthcare providers if a practically deployable self-harming activity detection system can be designed.

In this paper, we investigate the feasibility of detecting self-harming behavior using smart-phones worn on both wrists of a subject and sensing associated acceleration during movements. We conducted an experiment with 4 subjects acting as individuals with intent to inflict self harm. Each subject was instructed to perform a series of 15 activities while a smart-phone was securely attached to either wrist. Some of the activities were normal daily activities (walking, drinking, etc.) while others were intended to mimic self-harmful behavior. Using only accelerometer data, an overall classification accuracy of 80% using 1-Nearest-Neighbor with Dynamic Time Warping was achieved. Incorporating additional sensors (gyroscopes, light sensor, pressure sensor, etc.) could improve the classification

978-1-4799-8425-1/15/$31.00

rate and is a goal of future work.

The results of this paper were concluded from an experiment where the subjects were neither in a clinic, and were not known to have past self-harmful behavior. Thus, the results should be viewed through that context (non-clinical, non-suicidal). We have taken every precaution to avoid generalizing any conclusions. Instead, the contributions pertain to demonstrating the feasibility of applying activity recognition to the context of reducing inpatient suicides. The data capture, processing, and analysis are explicitly detailed as to provide similar future studies with a basis for comparison.

II. Background

A. Related Work

Human Activity Recognition (HAR) using wearable devices infer activities of users through devices attached to the body. Devices may incorporate a number of individual sensors to capture temperature, humidity, audio level, acceleration, location (GPS), vital signs, and more [5]. In smart-watches, adoption of GPS, magnetometers, accelerometers, gyroscopes, pressure sensors, and light sensors is dependent upon the product. Of these, accelerometers are the mostly widely leveraged, since they can sense motion (or lack thereof). While GPS measurements provide little information in the context of indoor location or movement, as the technology requires line-of-sight to a subset of the GPS satellite constellations, they provide rich contextual information for inferring activities outdoors. While magnetometers and gyroscopes are included in recent smart-watches, their measurements are mostly ignored, and yet to find applicability. Including those data would provide orientation information of the smart-watch, but again, the adoption rate of those sensors is currently low. By using only the accelerometer, the results of this paper could apply to any wrist-worn device capable of recording accelerometer measurements.

Recently, the use of smart-phones for activity recognition has increased in popularity. The pervasiveness of smart-phones in society have made them a desirable platform for activity detection rather than dedicated, custom wearable sensors. Previous works incorporate tri-axial accelerometer measurements and, less frequently, gyroscope measurements [6], [7], [8], [9], [10]. As the hardware platform is smart-phone based, the activity recognition in these works have pertained to ambulation activities (walking, running, going up/down stairs, etc.). Our work contains classification of some ambulatory behaviors, but also includes self-harmful behaviors. The inclusion of ambulatory behaviors provides a measure of robustness for the classifier in determining non-harmful activities.

In [11], the authors analyze activity recognition to improve hospital staff efficiency (without any wearable devices). Basically, they collected approximately 200 hours of data by shadowing hospital staff over 9 months. Paired with application logs of a mobile health application, they used Hidden Markov Models to determine if staff were performing clinical care assessment, patient care, coordination, preparation, information management, or classes and certification. Their classifier demonstrated accuracies of $> 90\%$ in four of the five activities. However, their efforts focused on the staff of the hospital and inferring high-level activities only.

Using smart-phones to aid in patient monitoring, [12] developed a system for movement activity recognition on the MATRIX telemedicine platform. Motivating their work was the desire of physicians to understand how patient mobility was progressing without direct observation. Movement activities and durations are inferred and transmitted to corresponding health professionals. Similar to the previously discussed smart-phone-based HAR systems, their classifier was only trained on walking, sitting, standing, going upstairs and downstairs.

Determining ambulatory activities from smart-phone accelerometers has been studied a number of times. The main difference of those works are the classifiers developed to infer behaviors. Our work incorporates wrist-worn tri-axial accelerometer measurement collectors to ascertain ambulatory and self-harming activities. The use of smart-watch like devices limits this study to an offline HAR system. However, many smart-watches can connect to smartphones, via Bluetooth, allowing data streaming. In such a case, our methods can work real time.

B. Self-harming Activities

Two highest-risk times for inpatient suicides are the week after admission (and also shortly after discharge) [13]. Hanging is the most common form of inpatient suicide [3], requiring only 4 to 5 minutes to be successful [14]. Discharged patients or patients on approved leave tend to use more violent methods. Some of these include cutting, head banging, strangulation/suffocation, insertion of foreign objects into body, reopening old wounds, burning, and self-poisoning [15]. Since inpatients in clinical (and especially in psychiatric) settings are monitored in terms of what they bring into the facilities, and after extensive discussions with a Clinical Psychiatrist, we identified a total of 7 self-harming activities in this paper for detection (see Table I and II for full activity list) that were most likely to be attempted in clinical settings.

TABLE I. Harmful Activity Set

Activity
Cutting Left Hand (CLH)
Cutting Right Hand (CRH)
Cutting Throat (CTHT)
Hanging (HNG)
Injection in Left Arm (ILA)
Injection in Right Arm (IRA)
Smothering (SMTH)

TABLE II. Non-Harmful Activity Set

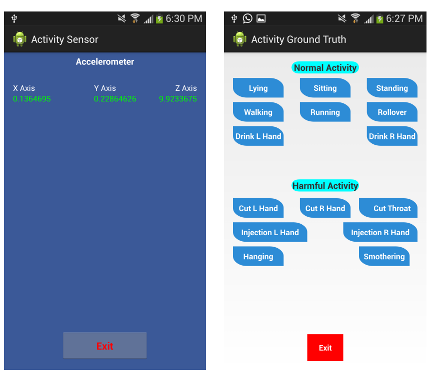
Activity
Drink with Left Hand (DLH)
Drink with Right Hand (DRH)
Lying (LYNG)
Rolling Over (RO)
Running (RUN)
Sitting (SIT)
Standing (STND)
Walking (WLK)

III. Methods

A. Subject Demographics and Data Capture

Four adults between the ages 19 and 25 were the subjects in the study. One other adult served as an instructor that collected accelerometer data from Samsung Galaxy S4 smart-phones worn on either wrist by the subjects as shown in Figure 1. The wrist-secured phones ran a simple application that recorded all accelerometer measurements with corresponding timestamps upon initialization. The experiments were administered by the instructor who maintained possession of another phone (master phone). The master phone contained a separate button for each of the 15 activities. When the instructor clicked a button, the date, time, and activity tag were stored to a file, to provide the ground truth. Also, buttons on the tagging application (master phone) would send a message to the recording applications to begin recording measurements. Sample screenshots of the recorder and tagging applications are shown in Figure 2.

Fig. 1. Acceleration acquisition smartphone secured to subject with a wrist strap.



(a) Recorder App (b) Tagging App

Fig. 2. Screenshots of the data acquisition applications developed for the Android platform.

The instructor was informed to have the subject maintain an activity for approximately three minutes before initiating a new activity. Sequences of activities did not follow any specific order as long as each one was performed by the subject. Capturing data from both wrists on subjects allows more diverse samples to train/test the classifier on, without making assumptions on the dominant hand of a subject. To illustrate the collection process, Figure 3 shows an entire trace of the tri-axial accelerometer measurements of one subject from the left-wrist phone.

B. Data Pre-processing

Let $\mathcal{D}$ denote a database containing all collected observations for one subject. Then $\mathcal{D}_i$ is the i-th sample containing 5-tuple $(timestamp_i, x_i, y_i, z_i, activity_i)$. The timestamps of $\mathcal{D}$ may be irregularly sampled as the Android API does not allow the enforcement of an arbitrary number as the capture frequency [16]. Resampling is necessary to conform the inter-sample time as $\frac{1}{f}$, where f is the frequency. From numerical analysis, the approximate frequency was determined to be $f = 200$Hz. The time-series is segmented into windows of size $\frac{1}{f}$ seconds. The corresponding acceleration component values (x, y, z) are average to produce a single value for the window occurring at time t where t is the start time of the window. $\mathcal{R}$ denotes the regularly sampled observations from $\mathcal{D}$ containing the observations $(x_i, y_i, z_i, activity_i)$, $\forall i \in [0, n]$. Note that the resampling process occurs separately for different activities so their measurements are disjoint. Additionally, some windows may contain no values producing Not-a-Number (NaN). Forward filling was used to replace NaN values, where the tuple of the previous measurement is replicated into the tuple containing NaN. An example annotated database $\mathcal{D}$ from the left hand phone of a subject is shown in Figure 3.

Depending on the orientation of the smart-phone, gravity influences the readings on one or more of the components. The acceleration values are transformed into linear acceleration, where linear acceleration is the acceleration measurement caused by movements of the collector without the influence of gravity. The effect of gravity is removed using a high-pass filter as defined in the Android developer reference [17]. Equations (1) and (2) show the steps to calculate the i-th linear acceleration measurement from collected samples.

$$\mathbf{g}_i = \alpha\mathbf{g}_{i-1} + (1-\alpha)\mathbf{e}_i \tag{1}$$
$$\mathbf{a}_i = \mathbf{e}_i - \mathbf{g}_i \tag{2}$$

where

$\alpha = 0.8$ (smoothing factor),
$\mathbf{g}_i =< g_{xi}, g_{yi}, g_{zi} >$ (current gravity vector),
$\mathbf{e}_i =< e_{xi}, e_{yi}, e_{zi} >$ (current acceleration),
$\mathbf{a}_i =< a_{xi}, a_{yi}, a_{zi} >$ (calculated linear acceleration).

The resulting $\mathbf{a}_i$ vectors replace the corresponding acceleration components in each observation. $\mathcal{R}$ contains the regularly spaced measurements of linear acceleration from the raw acceleration values.

The dimensionality of the database was reduced to a single variable for each time series by calculating the magnitude of the acceleration measurement vector. The magnitude was determined by obtaining the Euclidean norm of the acceleration components. Reducing the dimensionality decreased the computation requirements of classification while still allowing discrimination among classes.

Applying time series classification on streaming data requires the use of sliding window techniques. Instead of observing the entire time series then attempting classification, only a subset of the observations are used at a given time. Classification is then performed on that subset of measurements and assigned to a class. A window size of $WS = 2,000$ samples (approximately 10 seconds) with 50% overlap was used to create a new database, W, that was used as the training/testing data for the classification step. Formally, w contains the windowed acceleration magnitude time series data where W_i is the i-th sample $(mag_i, activity_i)$. The element mag_i contains the magnitude of acceleration measurements

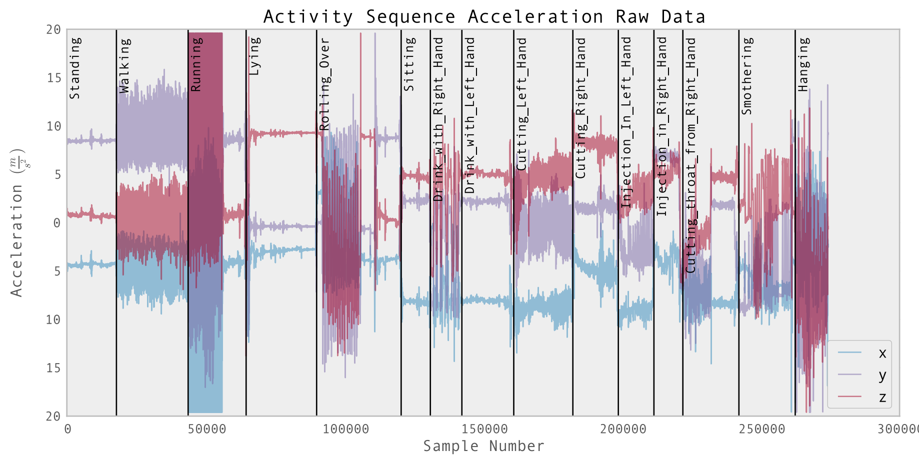


Fig. 3. Captured accelerometer components throughout an activity sequence from the phone secured to participants left hand. The blue series denotes the x component, red shows y, and purple shows z. The figure is best viewed in color.

from a window and $activity_i$ is the associated activity of the window.

C. Activity Detection

Shape-based time series classification uses a set of templates to determine the label of some test sample. Test samples are compared to the templates using a distance metric, or dissimilarity measure. The label from the most similar template is assigned to the test sample. In reality there are many advanced techniques to extend this concept, but the general ideology remains the same. One of the most widely used time series classifier is the one-nearest-neighbor with Dynamic Time Warping (1NN-DTW) [18]. Since the origination of 1NN-DTW, it has been regarded as one of the most accurate time series classifiers and is widely used [19].

Dynamic Time Warping (DTW) is a technique that measures the dissimilarity of two sequences while handling local temporal distortions [18]. Euclidean distance between two time series sequences can provide misleading results if the time series are very similar but one is temporally shifted. Such cases can occur naturally where observations are collected and compared using sliding windows. DTW identifies misalignments of sequences by finding the minimum cost warping path between them. The resulting distance, or dissimiliarity, of the sequences is determined by the distance of corresponding locations in the warping path.

Given two time series Q and C, where

$$\begin{aligned} Q &= \{q_1, q_2, \ldots, q_i, \ldots, q_p\} \\ C &= \{c_1, c_2, \ldots, c_j, \ldots, c_m\} \end{aligned} \tag{3}$$

a p-by-m matrix is constructed. Each (i^{th}, j^{th}) element contains the distance $d(q_i, c_i)$, or the alignment between q_i and c_j. A warping path is a contiguous set of matrix elements that defines a mapping between Q and C.

The optimal warping path between Q and C minimizes the warping cost as:

$$DTW(Q, C) = \min \left\{ \sqrt{\sum_{k=1}^{K} w_k} \right\} \tag{4}$$

Constructing an optimal path can be found with the recurrence

$$\gamma(i,j) = d(q_i, c_j) + \min\{\gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1)\} \tag{5}$$

where

$$\begin{aligned} \gamma(i,j) &: \text{cumulative distance of the path to } (i,j), \\ d(q_i, c_j) &: \text{distance between points } q_i \text{ and } c_j. \end{aligned}$$

Evaluation of the recurrence is solved using dynamic programming on the p-by-m distance matrix [19].

k-Nearest Neighbor classifiers compare test samples to all training samples to determine dissimilarity. The k least dissimilar training samples are used to "vote" the class label of the test sample. Whichever class, of those k candidates, holds the majority determines the group of the test sample. In 1-Nearest Neighbor with Dynamic Time Warping (1NN-DTW), the least dissimilar sample, as determined using DTW, is used to obtain class labels for testing samples.

The benefit of this scheme is that classifying samples is not determined by computing statistics or determining some common structure among sequences. Since we captured data from two phones simultaneously, two traces for a single activity may look entirely different. Say, a subject takes a drink with their left hand. The acceleration trace from the left-wrist phone will most likely look different than the right-wrist phone. Attempting to determine correlations or determine common structure would be extremely difficult as the traces are very different (see Drinking with Right Hand and Drinking with Left Hand in Figure 3). Feature-based classifiers operate by requiring specific attributes of a sample window, such as the mean, standard deviation, entropy, frequency components, etc. Ascertaining class labels is performed through comparision of test sample features to the training samples or some representation of the span of classes on the instance space. Using the *Drinking* activities in Figure 3, it is clear that the standard deviation of DRH is much greater than DLH. But the samples from the opposite phone (left-wrist worn) would show the opposite. Building feature-based classifiers to generalize to activity samples independent of the location of the wrist-worn smart device would be very difficult. Because of this problem, we chose to use the 1NN-DTW classifier which is featureless and robust to the location of the wrist-worn device.

IV. Results

The preprocessing resulted in a total of 287 windowed samples for training and testing. The samples were split using 70% as training and 30% testing. Cross-validation was not performed due to the computational requirements of DTW calculations. Each activity classification performance and number of test samples in the group (support) are detailed in Table III. Overall, the classification accuracy was 80% denoted by the average F_1-score of the 15 activities.

TABLE III. 1NN-DTW Classification Performance Results

Activity	Precision	Recall	F_1-score	Support
CLH	0.82	0.56	0.67	16
CRH	1.00	0.78	0.88	9
CTHT	0.25	0.50	0.33	2
DLH	0.86	1.00	0.92	6
DRH	0.75	1.00	0.86	6
HNG	1.00	1.00	1.00	6
ILA	0.75	1.00	0.86	3
IRA	0.50	1.00	0.67	2
LYNG	1.00	0.25	0.40	4
RO	1.00	1.00	1.00	6
RUN	1.00	1.00	1.00	5
SIT	0.50	1.00	0.67	3
SMTH	0.80	1.00	0.89	4
STND	0.83	0.83	0.83	6
WLK	0.83	0.62	0.71	8
Avg/Total	0.85	0.80	0.80	86

The *cutting throat* activity had the worst performance with an F_1-score of 0.33. The confusion matrix (not shown) revealed the classifier has trouble discriminating CTHT from the other two Cutting activities. Combining all Cutting activities into a super group may be more beneficial than determining cutting of individual locations. In a practical setting, location of the cutting has little importance compared to detecting cutting and generating a notification. Similarly, most of the degraded performance of some activities is due to confusion among related activities (*i.e.,* standing with lying or sitting). Forming super groups among similar activities may improve the overall performance. Finally, the support for most of the activities is relatively low. More samples would benefit in determining the efficacy of classification for self-harming behaviors.

Along with combining similar activities, using additional sensors could improve the accuracy of classifying human activities. For instance, obtaining magnetometer and gyroscope data provides orientation information. Certain accelerometer characteristics may be significant in specific orientations while insignificant while in others. Currently, accelerometers are highly adopted in smartwatches/wrist-worn health devices while other sensors are sparse. Restricting the classification to accelerometers only improves the practical applicability of the results.

V. Discussion

A. Application

Implementation of a self-harming activity detector within a hospital setting would require wrist-worn devices capable of wireless transmission. Streaming the acceleration data to a server would allow online detection and notification. The server reads samples for comparison against the classification model to infer the presence of self-harming behaviors. Upon detection, the server could notify the staff responsible for the patient from which the notification was generated. We call the system Self-Harming Activity Recognition Engine (SHARE). As stated in [4], the use of 15-minute checks generates a level of obtrusiveness for new patients leading to a difficult adjustment for some. SHARE would allow patients to obtain higher levels of privacy and comfort from repeated human interventions as the frequency of staff observation is reduced.

To ensure patients do not remove the devices, a special locking wrist strap or device with biological sensors are necessary. Using locking wrist straps may cause another source of discomfort for patients. Biological sensors, such as heart rate monitors or skin temperature sensors (see Basis watch [20]), provide the same functionality of ensuring the device is not removed. Biological sensors require contact of the skin to record measurements. Extended periods of measurement inactivity would alert staff when a patient has removed the device. Biosensors should be preferred over locking mechanisms as patients may desire to relocate the wrist-worn device to the opposite wrist. Additionally, the patient may wish to re-adjust the strap in order to alleviate any discomfort from extended use. Momentary absence of biological measurements must be accounted, through configurable parameters, within the SHARE system to reduce the number of false negatives.

SHARE could improve the comfortability of patients and reduce the consumption of resources involved with suicide checks. Though, implementing SHARE would require an initial cost to purchase the wrist-worn devices and hardware for the underlying infrastructure. Many hospitals already contain WiFi for medical equipment and patient use. Streaming SHARE devices over WiFi could cause increased latency or even denial-of-service to critical medical equipment or services over WiFi. Instead, rooms could be fitted with special processing units that locally process those streaming data from patients. Detection of self-harming activity then generates an alert over WiFi, or similar system, to a server. The server is then responsible for determining response actions, such as notifications. As with any health-care technology, patient privacy is of utmost concern. SHARE is in the conceptual stage and privacy is not discussed in this paper. Privacy is a major concern of the development of SHARE and will be included at all stages of subsequent refinement.

B. Limitations

Ideally, self-harming activity recognition is required in real-time for an actual implementation of a SHARE system. The experiment we performed was analyzed using post-processing. Our focus was to implement a proof-of-concept for SHARE without occurring a lengthy research and development cycle. Additionally, the 1NN-DTW classifier is computationally expensive as DTW reaches a worst-case asymptotic complexity of O(n^2), if the two time-series samples are of length n. In most cases, the optimal warping path lies within some bound of the diagonal of the p-by-m matrix of all possible warping paths. Bounding the number of steps to move from the diagonal (called the warping window) significantly reduces the number of necessary computations. It was shown in [19] that relatively small warping windows ($< 10\%$ of n) provide the optimal solution. During our analysis, the warping window was set to 10% of the largest time-series input sample being compared.

Recent advances in DTW calculations have greatly increased the practicality of its use in time series analysis. Rakthanmanon, et al., combined 4 techniques to drastically improve the speed of DTW calculations while still providing the optimal path within a given warping window. One of their test data sets included $8,518,554,188$ datapoints collected from one year of electrocardiogram (ECG) data sampled at 256Hz. Incredibly, their revised algorithm demonstrated processing rates $29,219$ faster than real-time (256Hz in this case) [21]. The result of revising DTW allows motif discovery, classification, and clustering of time series in real-time. The result of their work is promising into improving the performance of SHARE by enabling online classification in the future.

As it would be very difficult and unethical to attempt acquiring data from subjects explicitly, intentionally causing self-harm, the activities subjects performed may not accurately depict real-world scenarios. Obtaining real-world datasets would impose additional challenges, but is something we are exploring. Additionally, label activities in such large data sets could prove difficult if individual activities should be identified. It may prove more fruitful to label only self-harming activities and everything else as "other". The data collected for this experiment is minute to the data necessary for real systems. First, the sample size for participants was small and may not reflect real self-harming activities. Second, the selection of sample activities are only a subset of those that a person could perform to hurt themselves. Third, as the activities were essentially simulated self-harming events, real events may have different trace characteristics. The benefit of NN-DTW classifiers is that training samples can be iteratively added without re-training the classifier. As more samples are collected, they can immediately be added to the training database. Obtaining more samples, in quantity and diversity, are important for the success of a SHARE system. Future endeavors will focus on collecting data from health-care professionals and improving SHARE.

VI. Conclusion

In this paper, we have presented a technology to supplement/augment traditional suicide checks through self-harmful activity recognition using smart wrist-worn devices. The Self-Harm Activity Recognition Engine (SHARE) aims to reduce inpatient suicides while simultaneously reducing human resources expended in hospitals. There have been many critics of the 15-minute suicide checks or one-to-one observation in hospitals with possibly suicidal patients. Automating the process through a Human Activity Recognition (HAR) system could reduce the number of monitored patients and, ideally, reduce the number of successful suicide attempts of inpatients. Our classifier demonstrated a high degree of accuracy only accelerometer readings from wrist-worm smart-phones. Future work will look into fusing data from multi-modal sensors, improving classification set, and classification accuracy.

References

[1] J. Knowles. (2012) Inpatient suicide: identifying vulnerability in the hospital setting. [Online]. Available: http://www.psychiatrictimes.com/suicide/inpatient-suicide-identifying-vulnerability-hospital-setting

[2] K. A. Busch, J. Fawcett, and D. G. Jacobs, "Clinical correlates of inpatient suicide." *Journal of Clinical Psychiatry*, 2003.

[3] L. Bowers, T. Banda, and H. Nijman, "Suicide inside: a systematic review of inpatient suicides," *The Journal of Nervous and Mental Disease*, vol. 198, no. 5, pp. 315–328, 2010.

[4] G. Jayaram, H. Sporney, and P. Perticone, "The utility and effectiveness of 15-minute checks in inpatient settings," *Psychiatry (Edgmont)*, vol. 7, no. 8, p. 46, 2010.

[5] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 3, pp. 1192–1209, 2013.

[6] Y.-S. Lee and S.-B. Cho, "Activity recognition using hierarchical hidden markov models on a smartphone with 3d accelerometer," in *Hybrid Artificial Intelligent Systems*. Springer, 2011, pp. 460–467.

[7] Z. Yan, V. Subbaraju, D. Chakraborty, A. Misra, and K. Aberer, "Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach," in *Wearable Computers (ISWC), 2012 16th International Symposium on*. Ieee, 2012, pp. 17–24.

[8] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[9] N. Győrbíró, Á. Fábián, and G. Hományi, "An activity recognition system for mobile phones," *Mobile Networks and Applications*, vol. 14, no. 1, pp. 82–91, 2009.

[10] A. M. Khan, Y.-K. Lee, S. Lee, and T.-S. Kim, "Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis," in *Future Information Technology (FutureTech), 2010 5th International Conference on*. IEEE, 2010, pp. 1–6.

[11] D. Sanchez, M. Tentori, and J. Favela, "Activity recognition for the smart hospital," *Intelligent Systems, IEEE*, vol. 23, no. 2, pp. 50–57, March 2008.

[12] S. L. Lau, I. Konig, K. David, B. Parandian, C. Carius-Dussel, and M. Schultz, "Supporting patient monitoring using activity recognition with a smartphone," in *Wireless Communication Systems (ISWCS), 2010 7th International Symposium on*. IEEE, 2010, pp. 810–814.

[13] Q. P and N. M, "Suicide risk in relation to psychiatric hospitalization: Evidence based on longitudinal registers," *Archives of General Psychiatry*, vol. 62, no. 4, pp. 427–432, 2005.

[14] P. Burgess, J. Pirkis, J. Morton, and E. Croke, "Lessons from a comprehensive clinical audit of users of psychiatric services who committed suicide," *Psychiatric Services*, vol. 51, no. 12, pp. 1555–1560, 2000.

[15] K. James, D. Stewart, and L. Bowers, "Self-harm and attempted suicide within inpatient psychiatric services: A review of the literature," *International journal of mental health nursing*, vol. 21, no. 4, pp. 301–309, 2012.

[16] Google, Inc. Sensors Overview, Android. [Online]. Available: http://developer.android.com/guide/topics/sensors/sensors_overview.html

[17] Google, Inc. SensorEvent, Android. [Online]. Available: http://developer.android.com/reference/android/hardware/SensorEvent.html#values

[18] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series." in *KDD workshop*, vol. 10, no. 16. Seattle, WA, 1994, pp. 359–370.

[19] X. Xi, E. Keogh, C. Shelton, L. Wei, and C. A. Ratanamahatana, "Fast time series classification using numerosity reduction," in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 1033–1040.

[20] Basis. Basis peak, the ultimate sleep and fitness tracker. [Online]. Available: http://www.mybasis.com/

[21] T. Rakthanmanon, B. Campana, A. Mueen, G. Batista, B. Westover, Q. Zhu, J. Zakaria, and E. Keogh, "Searching and mining trillions of time series subsequences under dynamic time warping," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 262–270.