Fall 11-29-2021

Non-Invasive Monitoring of Human Hygiene using Vibration Sensor and Classifier

Shashank Trivedi

Maria Valero
Kennesaw State University

Hossain Shahriar
Kennesaw State University

Liang Zhao
Kennesaw State University

Non-Invasive Monitoring of Human Hygiene using Vibration Sensor and Classifier

A Thesis Presented To
The Faculty of Information Technology Department

by
Shashank Trivedi [MSIT Candidate, November 29, 2021]
Dr. Maria Valero [Committee Chair]
Dr. Hossain Shahriar [Committee Member]
Dr. Liang Zhao [Committee Member]

In Partial Fulfillment of Requirements for the Degree
Master of Science in Information Technology

College of Computing and Software Engineering, Kennesaw State University

Fall 2021

Table of Contents

LIST OF TABLES

Table References

1. World Health Organization. (2020, December 4). *Life expectancy and Healthy life expectancy Data by country*. World Health Organization. Retrieved November 26, 2021, from https://apps.who.int/gho/data/node.main.688.

LIST OF FIGURES

Figure References

2. Raspberry Shake. (n.d.). *RS1D: Turnkey*. Raspberry Shake Shop. Retrieved November 26, 2021, from https://shop.raspberryshake.org/product/turnkey-iot-home-earth-monitor-rs-1d/.

ABSTRACT

This paper presents a concept and an idea of a non-invasive monitoring system for human hygiene using vibration sensors. The approach is based on a combination of geophone sensor, a digitizer, and a cost-efficient computer board in a Raspberry Shake enclosure. People's personal hygiene habits speak volume about how they take care of their bodies and health. Maintaining good hygiene practices not only reduce your chances of contracting a disease, but it could also reduce the risk of spreading illness within your community. Given the current pandemic, daily habits such as washing hands or taking regular showers have taken major importance among people, especially for our elderly population living alone at home or in an assisted living facility. Monitoring daily hygiene routine could truly help our healthcare professionals be proactive rather than reactive in identifying and controlling the spread of potential outbreaks within our community.

Keywords: Non-invasive monitoring, human hygiene, vibration sensor, SVM

CHAPTER I.

INTRODUCTION | MOTIVATION

It may come as a surprise that the number of people over the age of 60 years will nearly double from 12% to 22% in the next 30 years [1]. This is in big part due to the increase in life expectancy from medical and technological advancements over the last few decades. The world is aging so rapidly that the number of people over the age of 60 would have outnumbered children under the age of 5 before 2021 [1]. It is great news that people are living longer, but a longer life does not necessarily mean a healthier life.

Aging is defined as “accumulation of a wide variety of molecular and cellular damage over time which leads to a gradual decrease in physical and mental capacity, a growing risk of disease, and ultimately, death [1].” Aging is often realized as a reason for concern for most people as their body starts to degrade over time and they lose confidence in their own abilities; but aging does not follow a linear or even a consistent progression for all. There are many folks over the age of 70 living a healthy life; while there are others that start to deteriorate before they even cross 60.

Aging is a phenomenon that can neither be stopped, delayed, or avoided – it is a natural progression of a human’s lifecycle in which the only option is acceptance. Old age is not an “I” problem, but rather a “we” problem as it impacts all ethnicity, genders, and social classes. As our society grows older, we need to be prepared for all the challenges that stem from aging – changes in physical appearance and mental state are only the beginning, but there are other factors such as housing, treatment, relocation, caregiving, etc that need to be taken into consideration. The moment you look at aging as a bigger picture by taking these external factors into consideration, aging sways from being a medical problem to a logistic and economical affair.

In a 2016 report by the US Census Bureau, it was declared that the US population is aging slower than the rest of the world [2] but this is all about to change as the baby boomer generation is about to age into this realm – the "number of Americans ages 65 and older is projected to nearly double from 52 million in 2018 to 95 million by 2060" and this group is more racially and ethnically diverse [3]. Immigration plays a huge role in defining the age structure in US as "there is an influx of mostly young and middle-aged adults between the ages of 25 and 45 that migrate to US on an annual basis [4]." It is a sad fact that aging is very relative – it is indirectly influenced by several socioeconomical factors such as education, wealth, employment, etc. Developed nations often have longer life expectancy than developing nations due to their reach and exposure to the global economy. Countries such as Australia, Canada, and the US often rank higher in life expectancy compared to their counterparts being some African nations, but a longer life doesn't necessarily mean a healthier life as "the U.S. spends more on health care as a share of the economy — nearly twice as much as the average OECD country — yet has the lowest life expectancy and highest suicide rates among the 11 nations [9]".

COUNTRY	YEAR	Life expectancy at birth (years)
Australia	2019	83
Canada	2019	82.2
France	2019	82.5
Germany	2019	81.7
Netherlands	2019	81.8
New Zealand	2019	82
Norway	2019	82.6
Sweden	2019	82.4
Switzerland	2019	83.4
United Kingdom of Great Britain and Northern Ireland	2019	81.4
United States of America	2019	78.5

Table 1- Life Expectancy and Healthy Life Expectancy Data by Country (WHO)

The steady rise in elderly population over the next few decades will result in an increased demand for nursing homes and assisted living facilities. – the elderly population tend to have the highest disability rate and highest need for long-term care services and is also more likely to be widowed and without someone to provide assistance with daily activities [5]. A long-term care is defined as a variety of services designed to meet a person's health or personal care needs during a short or long period of time [6]. Long term care is designed to help people (elderly, in most cases) live as independently as possible and it can be provided either at home by family members or in an assisted living facility such as nursing homes or adult daycare homes. In 2016, there were 15,600 nursing homes and 1.7M licensed beds in the United States [7].

The “sociocultural theory” states that human development is directly tied to the culture or society that people live in [8]. It is no surprise that the environment that you live in and the people that you surround yourself with can have a significant impact on your personal growth and development. This is especially true in young adults – the values cultivated during the first half of a child’s life will guide their judgement for the latter half of their life which is why early development and continued support is crucial. This is especially important when it comes to the topic of health and lifestyle. It is important for every parent or guardian to teach their young ones the benefits of practicing a healthy lifestyle since a healthy heart starts with a healthy lifestyle.

Alzheimer’s disease poses a serious risk as more than 6 million Americans are currently suffering with this disease, and it is especially deadly in the older generation as 1 in 3 seniors fall victim to Alzheimer or other forms of dementia [11]. Most people take their health for granted. As a result of this negligence in the early years, most people find themselves in situations where they must rely on others for help. Some may even need assistance with daily activities such as taking a shower or flushing toilet. As the world gets older and more people enter retirement over the next two decades, our society and our healthcare system are both faced with an enormous challenge to provide for the less fortunate. As the exact cause of Alzheimer's disease is still unknown, there's no certain way to prevent the condition, but a healthy lifestyle can help reduce your risk [10].

Over the last year and a half, we have witnessed a global catastrophe in the form of a pandemic that ravaged most countries. This pandemic put our healthcare system under tremendous amount of duress, but we also recognized our strength of working collectively together and using every tool at our disposal to its full potential. We need to continue this approach and leverage the current technology to address real-world scenarios. The idea here is to equip our healthcare workers with the right set of tools to enable them to achieve even more. By enabling them, we are building our healthcare system to be a proactive force rather than a reactive one.

This pandemic made us all realize the importance of being proactive and prepared. As a community, we were not prepared well enough for this pandemic but now is our chance to take what we learned and put that into practice. Being aware of risks such as Alzheimer that impact millions but could potentially be mitigated by practicing a healthier lifestyle – we can either sit back and point fingers at our broken system or we take this opportunity at hand and leverage the power of technology to aid healthcare.

It is a known fact that people's daily habits indicate a lot about their health and well-being – According to CDC, researchers in London estimate that if everyone routinely washed their hands, a million deaths a year could be prevented [12]. Questions such as "are you washing your hands regularly" or "are you taking enough showers" may not seem as important during the early adult years, but these questions become essential with age as they could be an indication of a developing health crises such as dementia.

Elderly community is at the most risk for this as majority of them either live alone at home or live in an assisted living facility where the staff may be stretched too thin.

Since a proper hygiene is the first step to a healthy life, we need to be thinking of different ways that we can assist the ones that need our help the most but may not even realize it. To avoid another potential outbreak of a deadly disease, we need to take the first step of examining the current methods of monitoring seniors and research other potential methods that can be more effective in improving their health, lifestyle, and quality of life.

When you think of remote monitoring, the very first method that probably comes to mind is using cameras. CCTV cameras are one of the most widely used forms of monitoring, but there is a huge drawback with this solution: privacy. If we start to use cameras in bathrooms to monitor seniors, it is a direct attack on their privacy and could even be considered unlawful. With that in mind, the new solution needs to be nonintrusive and respectful of people’s privacy.

In this thesis, we present the concept and an idea of a non-invasive monitoring system for human hygiene using nonintrusive vibration sensor. The idea here is to place these vibration sensors around the “most used places” in a home. We have targeted specific places such as bathroom and kitchen that can be an area of concern or could help identify potential disorders. These sensors will be placed in/on/around faucet, sink, toilet, shower, etc where the sensor will capture the vibration signals from the running water and store the metadata in a log. The sensors will not only capture when the faucet was opened, but it can also tell us how long the water was running for. Once the data has been recorded, we can use signal-processing algorithms to find patterns and classify these patterns as hygiene activities. The idea here is that this data will help our healthcare professionals better understand the everyday habits of our seniors, and they will also be able to draw conclusions based on the same data about their health condition.

The experiment and related research will use a combination of both hardware and software to capture and process the data. Our goal here is to study the use and effectiveness of these sensors, and even identify other direct and indirect broader use cases. This method of monitoring is nonintrusive as it does not invade their privacy or bring any major changes to their daily routine. The target audience will continue to go about their daily life without any major interference as a result of our approach. The other advantage of this method is automation. Once we have trained our solution to be able to identify and classify patterns as hygiene activities, the sensors and the algorithm will continue to operate autonomously as designed and will continue to record and report the data without any sort of intervention.

There has been some previous research (referenced in Chapter 2 of the thesis) conducted in this area of home monitoring which will be discussed in this paper, but none have used vibration sensors and classifier in combination with a raspberry Pi. This thesis will discuss details around system architecture, implementation overview, challenges encountered, review of findings, and conclusion on acceptance and expansion. There are many factors that can affect the success rate of this research and recommendation around best practice will be shared. The data collected during the research and experiment will also be made available to the public for any future studies.

We propose to leverage the power of these sensors to help better understand the daily habits and our goal here is to raise awareness. If we are able to provide this data to our healthcare workers, they can take proactive measures in remediating a potential health crisis. When it comes to health and safety, time is always of the essence. The sooner we can detect an issue, the faster our healthcare workers can respond and help with recovery.

CHAPTER II.
RELATED WORKS | PREVIOUS WORKS

There have been some previous studies done in the area of unintrusive monitoring using vibration sensors and classifiers, but none have used the combination of geophone, digitizer, and a computer board (raspberry pi) to measure multiple activities in bathroom and kitchen, to the best of our knowledge. This related work section discusses several papers and works that are closely related and serve as an inspiration for this research. This research could be seen as a continuation and/or an extension of some of the work listed here.

The study completed by Yiyuan Zhang titled “Bathroom activity monitoring for older adults via wearable device” makes use of a wearable accelerometer to monitor six types of bathroom activities: dressing, undressing, washing hands, washing face, brushing teeth, and toilet using [13]. Two main models for validation were used: 10-fold cross validation and LOPO (Leave-One-Person-Out) model. In the 10-fold approach, all participant data was shuffled and split into 10 folds in which one-fold was selected for testing while the other 9 folds were used for training. In the LOPO model, the dataset for one participant was used for testing while the other participant’s data was used for training. The outcome of their research concluded that one wearable could be used to monitor six bathroom activities but further studies would need to be done as some activities such as washing face and undressing resulted in a lower classification score. Additional research using this technique is recommended to improve the success ratio for the activities in the lower classification category.

The study conducted by Jianfeg Chen labeled “An Automatic Acoustic Bathroom Monitoring System” focuses on the strength of acoustic signals to monitor ADL (Activities of Daily Living) in the elderly. This system is designed to recognize and classify different activities occurring within a bathroom based on the sound [14]. Their system architecture includes an infrared system which would detect entry into the bathroom, a microphone with a pre-amplifier circuit, and a laptop running the required application to record and classify the sound events in real time using HMMs (Hidden Markov Models). A personalized daily report is generated by the system which provides a daily summary of the subject’s bathroom behavior and provides recommendations based on the data recorded. This study resulted in 87% or higher accuracy in detecting activities and event classification. The next step for this study would be to improve the accuracy of the event recognition and they also recommended to enhance the system capability to monitor and identify human vocals to assess the subject’s state of mind within the bathroom.

The study done by Yingqi Hao with the subject of “Design and implementation of an Intelligent Kitchen Safety Monitor” focuses on creating a safety monitor that not only detects a safety hazard in the kitchen but will also respond to the threat and send out alerts or notifications. This system utilizes a single-chip microcomputer as the central core along with sensors to detect temperature, gas, and smoke concentration. On the software end, it uses the C-language to control the MCU along with Keil C51 for debugging purposes. The idea here is to create a threshold for normalcy and any abnormal behavior that crosses the threshold will trigger an action from the system [15]. This infrastructure combined with the hardware used for research in this paper can create a robust “monitoring and response” system in case of any potential hazard performing ADLs.

In the paper “Bathroom Monitoring with Fast-Chirp Modulation Millimeter-wave UWB Radar” by K. Jimi, the approach is using a 79GHz millimeter-wave UWB sensor with Fast-Chirp Modulation (FCM) and hidden Markov model (HMM) for bathroom monitoring. This is a non-intrusive method of monitoring as bathroom raises the concerns of privacy infringement, humidity, and fluctuations in the water levels. By installing the radar sensor on the wall, it will allow for the velocity, angle, and distance to be measured. The experiment was run in three separate scenarios which included normal bathroom without accident in scenario A, drowning in the bathtub for scenario B, and falling in the washing area for scenario C. The results showed a 95% success rate for prior learning method, and a 90% success rate on incremental learning approach [16].

The paper titled “Monitoring Hand-Washing Practices Using Structural Vibrations” by Jonathon Fagert focuses on the four activities associated with the event of hand-washing – steps, water, soap, and rinsing. The approach also utilizes recording and measuring the structural vibration signals through the means of a geophone. The vibration signal is then converted to digital using ADC (Analog-to-Digital) converter and processed using a computer. The classification of the activities is done through SVM

(Support Vector Machine) algorithm and utilizes L-1 norm for efficiency. The technique used for classification is called "one-against-one" in which six SVM models are used to label each activity – the activity is only considered as approved or classified when 3 of the 6 models classify as the same label. Overall, the results were positive with a 95.4% average accuracy which is a 1.4X improvement over the baseline of 69.1 percent [17].

The paper "Personal hygiene monitoring under the shower using Wi-Fi channel state information" by Jeroen Klein Brinke employs the use of radio-frequency based recognition to monitor the variations in several shower related activities. The trials takes into account a regular/normal shower and an instance of troublesome shower for comparison. The experimental set up included multiple custom Gigabyte Brix IoT devices with Intel Apollo Lake processor, 8GB RAM, and Intel N Ultimate network card where one acted as transmitter and the other as receiver. A Convolutional Neural Network (CNN) was used for classification of the dataset as it preserves both the structural and spatial information. Although the experiment was a success in determining whether the shower is on or off, the authors do recommend validating the results in the paper with further trials which would include more participants, different locations, and different frequencies in new settings [18].

CHAPTER III.

SYSTEM ARCHITECTURE AND OVERVIEW

We present a four-stage approach to this experiment as shown in Figure 1

- Data Collection – this first stage involves configuring the hardware and software
- Data Validation – this second stage involves verification of the collected data
- Data Extraction – this third stage involves extraction of the locally stored data
- Data Analysis – this final stage involves finding features and event classification

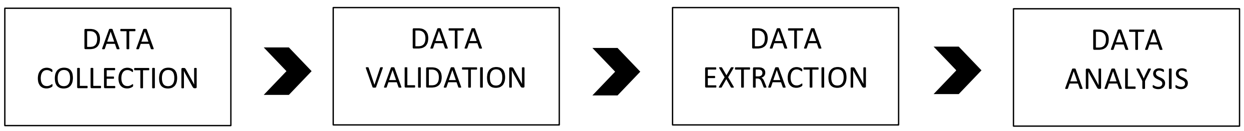

Figure 1 – Four Stages of Experiment

STAGE 1: DATA COLLECTION

In this first stage of the experiment, our focus was figuring out the type of sensor required to collect the data and the type of data to be collected including the location of the sensor placement and the data points that are most relevant to the research.

We started with a Raspberry Shake turnkey product (shown in figure 2) which included a Raspberry Shake circuit board, smart sensor composed by a geophone (vertical axis seismic sensor), a digitizer board, and a simple-board computer (Raspberry Pi 3B) all in a preassembled enclosure [19].

Figure 2 – Raspberry Shake Enclosure

Raspberry Shake I/O
1 x Ethernet Port
1 x HDMI Port
4 x USB Ports
1 x 3.5MM jack
1 x MicroSD card slot

Table 2 – Raspberry Shake I/O Ports

Before getting started with data collection, the Raspberry Pi needs to be connected to a network and a dashboard would need to be configured in Grafana for data visualization. To get the Raspberry Pi configured with network settings, connect it to an external monitor using HDMI cable and also connect a wired keyboard to be able to run the following commands –

At the Command Line, log in with –

```
U: myshake
P: shakeme
```

Once you are logged in, run the following command to access the network interface to add your own network details (SSID and passphrase)

```
sudo nano
/etc/wpa_supplicant/wpa_supplicant.conf

Network = {
ssid=”Enter your SSID”
scan_ssid=1
psk=”Enter your WiFi Passphrase”
}

CTRL + X > Y (select ‘yes’ to save) > ENTER
```

After the configuration above, the device should start recognizing your network and join automatically. Next step is to configure a local instance of Grafana to be able to visualize the data. First, will need to download and install Grafana from https://grafana.com/grafana/download for your operating system. Once Grafana is installed, open a new browser window and navigate to http://localhost:3000 which will bring you to the admin console for Grafana. You will be logging in with the default credentials (admin/admin) and be prompted to change on your initial login.

Setting up Data Source in Grafana

- Go to Settings > Data Source > Add Data Source > Influx DB
- URL is going to be the IP of the Raspberry Pi (port is 8086)

```
DB: <confidential>
U: <confidential>
P: <confidential>
```

- Save and Test > DB connection is working (confirmation)

Setting up Dashboard in Grafana

Once you have the data source configured, next step is to set up the dashboard in Grafana as shown in figure 3. Make sure that Influx DB is the selected source, and all the other configuration mirrors figure 3. Once you have the empty panel configured with your dashboard settings, click on “save” which will prompt you to provide a name for your dashboard before saving. To verify, head over to https://localhost:3000/ and confirm connectivity

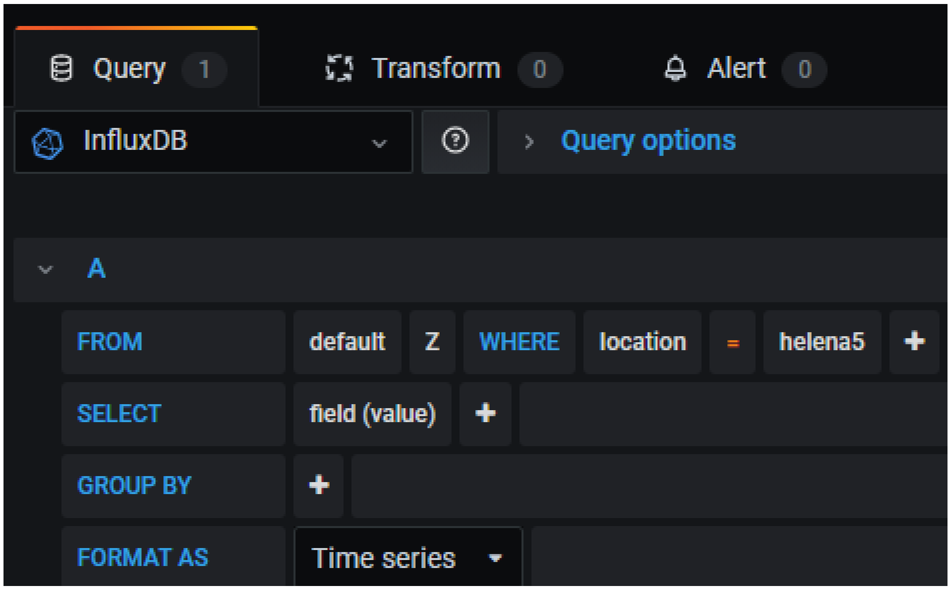

Figure 3 – Grafana Dashboard Configuration Screenshot

** If you forget admin password, navigate to the bin directory and run the command below
grafana-cli admin reset-admin-password <new password>

After setting up the Raspberry Pi and Grafana, the next part of this data collection stage is to figure out the type of data that needs to be collected. To do this, we will need to choose a location for the sensor placement and also identify some key data points that are relevant and can add value to this research. Since our main goal is to monitor the human hygiene using vibrations from water, we chose to narrow down our scope and focus on two of the six Activities of Daily Living (ADL) - personal hygiene (washing hands) and toileting (flushing toilet) [20]. There has been no previous research done that involved monitoring of vibrations through the use of toilet so that was chosen as the first location under the ADL category of toileting and the other two locations were bathroom faucet and kitchen sink under the ADL category of personal hygiene. The reasoning behind choosing these three locations is the frequency of touch through the day, and the activities associated with both as it relates to health. We wanted to select the location(s) that fell under the realm of personal hygiene but also part of the ADL scope.

LOCATION	ACTIVITY
Toilet	Toilet Flushing
Bathroom Faucet	Hand Washing
Kitchen Sink	Hand Washing

Table 3 – Categorization of Location and Associated Activities

For the scope of this study, we are assuming that the subject is washing their hands each time the bathroom faucet or kitchen sink is opened and closed. All the associated activities are related to the concept of human hygiene and includes capturing vibrations from water. While considering the sensor placement, it was also important to consider the environment the sensor is in which may include the possibility of it getting wet or challenges that come with increased humidity. It was also important to consider factors related to external interference such as people walking, sensor falling off due to off-balance, or sensor getting pulled accidentally.

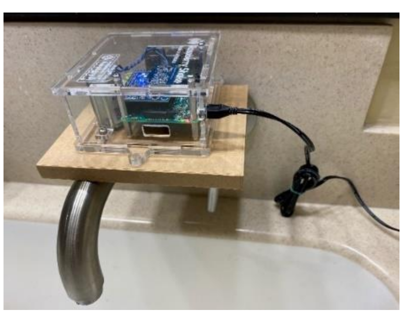

Figure 4 – Bathroom Faucet Monitoring

Figure 5 – Toilet Monitoring Setup

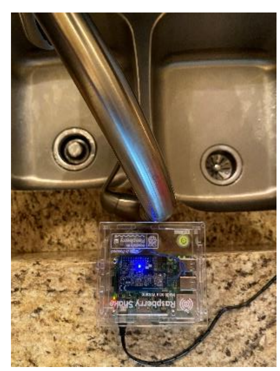

Figure 6 – Kitchen Sink Monitoring

There was some trial and error in the process of finding the ideal position to place the sensor which would allow for maximum data to be collected. Due to the size and shape of the faucet, we had to improvise and come up with a makeshift solution to hold the sensor on top of the bathroom faucet as shown in Figure 4. For kitchen sink, the shape of the faucet forced us to place the sensor on the side of the faucet, as shown in figure 6. This was not the ideal position, and it may have potentially degraded the sensitivity of the sensor, but it was the best positioning given the circumstances. Figure 5 shows the sensor placed on top of the toilet tank.

After identifying the locations of sensor placement, next step is to figure out the key datapoints that should be collected which can help understand how the sensor was placed and used. These datapoints needed to be logged for the purpose of data validation (as part of stage 2). The end goal of this study is to determine whether or not vibration sensor and classifiers could be used to track the usage of ADLs, so the validation piece plays a critical role in verifying the data stored on the device with the data logged manually through the life of the experiment. It was also important for the data logging to be as accurate as possible so that it aligns as closely with the actual as possible. The data validation piece is going to be extremely important in determining the success of this study.

The following datapoints were recorded for each location as part of the experiment log –

- Date – Date of the recorded event
- Start Time – When the event began

- Duration of the event (in sec) – How long the event lasted for
- Building Type – Type of building where the sensor was placed
- Location of the sensor – Where was the sensor placed within the building
- Position of the sensor – How was the sensor placed at its location
- Event Type – How did the subject interact with the sensor
- Sensor Distance (approx.) from the water source – Distance from the water source

Date	Start Time	Duration (in sec)	Bldg Type	Location	Position	Event Type	Sensor Distance (in CM)
6/14/2021	8:24AM	50	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	8:28AM	15	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	8:31AM	20	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	8:35AM	65	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	9:08AM	12	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	9:15AM	18	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	9:33AM	10	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	11:00AM	25	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	12:00PM	20	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	12:44PM	2	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	2:44PM	60	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	5:54PM	30	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	6:28PM	10	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1
6/14/2021	7:25PM	20	Residential	Bathroom Faucet	Top	Opening/Closing Faucet	1

Figure 7 – Sample of Data Log from Bathroom Faucet

There were in total of 368 total events captured from all three locations over several weeks of the experiment. It cannot be said with certainty that each and every event was captured, but enough data was captured to be able to run analysis and draw any applicable conclusions. This data was logged manually on a paper and then documented electronically in a spreadsheet as shown in Figure 7. During the data validation stage, the actual data off the sensor will be verified against this data logged to measure the accuracy of recorded data and provide validity to the experiment. In other words, it is a comparison between what the device recorded versus what actually happened. As part of this research, all the data that has been recorded will be made publicly available for others to reference for their own individual use cases.

Here are some challenges that were encountered with collection and recording of data -

- Shape and size of the device
- Shape and size of the faucet
- Positioning of the sensor (getting it as close to water source as possible)
- Open ports – water intrusion
- Wet and humid environment
- Failure to log the time
- Log sheet (paper copy) getting wet from water splashes
- Sensor going offline frequently (requiring reboot)
- Sensor getting unplugged accidentally
- Needing to adjust the sensor as opening/closing faucet makes it off-balanced
- Adjusting the intensity of the water flowing through the faucet

The device itself was portable enough to be able to be easily relocated to fit our requirements, but the shape and size of it made it somewhat difficult to be placed in an ideal location. Also, one of our requirements for an ideal scenario was to get the device as close to the "water source" as possible to be able to collect the maximum quantity of data, but this was not possible because of the shape, size, and location of the faucet in both the bathroom and kitchen. Also, there was the fear of getting the device wet because of two reasons 1) water intrusion through the open ports on the device and 2) the wet/humid environment of the bathroom. In addition to that, there were some occasions when one or more people may forget to log the time or accidentally get the paper copy of the log sheet wet from water splashes making the data harder to be identified. Since the device was placed on top of the faucet (inside bathroom), it would often get unbalanced from the frequent opening/closing of the faucet.

Research typically gets categorized into either qualitative or quantitative – the research we conducted falls into the qualitative bucket as it is more important for us get the right data than to collect abundant data. The accuracy of the collected data in this case gives this research more meaning than the quantity of the data – this is especially important during the analysis stage as the quality of the data is far more critical in training the machine learning models than the quantity. With that said, we would like to point out that enough data was collected as part of this experiment to be able to provide value for our research and recommend any further research to be done in this area.

Data was collected in a controlled environment over a period of about 33 days between June 14 and August 20, 2021. The prototype device was placed by bathroom sink, toilet tank, and kitchen sink to simulate a real-world need. All the data was collected using the same single Raspberry Shake moved between the three locations. To the best of our knowledge, the device was neither dropped, nor tampered, nor was there any damage from water intrusion.

STAGE 2: DATA VALIDATION

Data Validation is an important stage of the process as it allows for the data to be validated through the means of comparison between what the device captured vs what actually happened. Since the main purpose of this research is to assess whether this sensor in its current configuration can be used to accurately capture the events. As such, this data validation stage is crucial in determining the success ratio for the classification.

In this stage, there are two main components – view the data captured on the device and compare that data with the manual log that was maintained. This data validation stage is actually in effect through the life of the experiment. Every time the location of the device is altered or if the device is rebooted for any reason, we would need to verify if the data is still being recorded.

For data verification, open a new browser window and navigate to http://localhost:3000/ which will bring you to Grafana admin console. Select the previously configured dashboard from the "dashboard" panel to open and view it. From here, you will be able to apply filters such as data/time and absolute range from the top right corner. As shown in figure 8, the time range is in a 24-hour format so make sure to convert the data to fit the requirements.

Figure 8 – Screenshot Showing Data Captured in Grafana

As part of this process, there were several challenges encountered –

- Wireless signal strength issues causing the device to stop collecting data
    - This could be due to several factors. It is also an assumption as other devices on the network did not lose wireless connectivity. As shown in figure 9 and figure 10, the device often resulted in network time outs and the data collection would stop as a result (shown on the left).
- Device losing connectivity all together needing to be rebooted
    - The device would often lose connectivity and would need to be rebooted. Even after rebooting, it would take several attempts to join to the network again. The experience with network connectivity has been a hit or miss as the network card has limitations of its own, such as not being able to connect to a 5GHz network.

- Network configuration (IP changes) which caused the device to go offline
    - After a couple of days of the device being offline, the IP address changed which required reconfiguration of the data source through Grafana. This is not a device specific issue, but an issue with DHCP rules in general. This would be rectified by either 1) leaving the device always connected to the network or 2) reserving the IP address on the router so it does not change (even after reboots or being offline for an extended period of time).
- Lack of notification on losing connectivity
    - This was by far the biggest challenge of all as the device would not provide any sort of indication when it would go offline. This lack of notifications lead us to have a poor experience overall. If the device was able to send a notification or provide some sort of alert/indication when it went offline, it would increase the amount of data captured (since no/minimum amount of data would be missed) and it would have provided a better customer experience and quality of data.

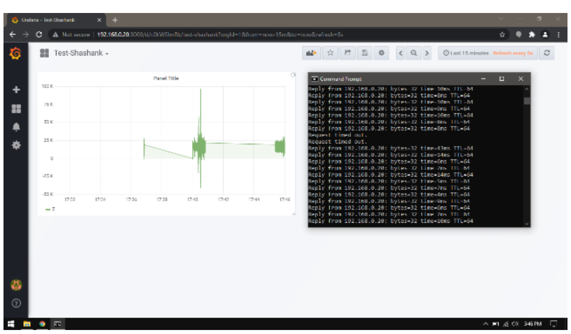

Figure 9 – Network Timeouts Screenshot

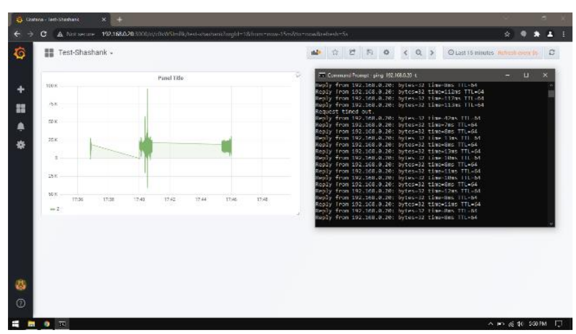

Figure 10 – Network Ping Screenshot

These network related challenges were a cause of great frustration as it would not only increase the effort involved with the data collection phase, but it would also impact the quality of the data itself as we would never know when the data capture would stop. Even though the device was verified as online and reporting at one point, it would automatically fall off the network or it would stop capturing data even while being online and connected as shown in figure 9

During the life of the experiment, we tried to maintain the manual log as close to real-time as possible for the events to match up exactly, but there was some level of “guesstimate” on event durations depending on each individual scenario. Also, the event captured by the device shows precision down to the seconds, while the log that was maintained was recorded up to the hour and minute precision. The duration of the event was also manually recorded so there may be some level of inconsistency within that data.

There was an assumption made at the beginning of the experiment that the noise surrounding the event would be at a minimum – or – if not minimum then the actual start/end of the event would “overpower” the noise surrounding it which would allow us to easily distinguish between the actual event and the noise. The sensitivity of the sensor on the device was strong enough to be able to pick up even the slightest motion which allowed us to come to this judgement on assumption.

STAGE 3: DATA EXTRACTION

Once there is enough data collected, next step is to extract the data off the device for analysis purposes. This step could be performed in parallel with stage 1 as new data could continue to be collected while analysis is conducted on the existing data that has already been collected. All the data is locally stored on a MicroSD card on the device itself. Extracting this data is a simple yet tedious process as it requires further validation on each individual event.

Before extracting the data, it is important to understand the data itself and what is being extracted. It is essentially a drill-through version of the graph that is being plotted during the process of the event occurring – the data is shown as X and Y values with X being the date and time stamps, while the Y being the actual “value” that represents the intensity of the vibration. The data is not very complex, but the sheer volume of data that is being outputted for each event is immense. The idea here is that the exported data should theoretically be able to recreate the graph by plotting the X and Y values.

Once the data has been exported, stage 4 is about studying that data and running analysis to find features relevant to the data and event. This extraction process is not complicated, but it is time consuming as it involves creating individual excel files for each event. While extracting the data, it is important to label each file with appropriate date, time, and location details to be able to later identify and validate it.

Steps to extract data from the device

The data extraction process has some similarities with the data validation phase when it comes to accessing the data itself. Instead of accessing the local instance of Grafana, you will need to go to the Grafana on the device from https://DeviceIP:3000/ and logging in with the admin credentials. The default credentials should be admin/admin and the interface should look fairly similar to the Grafana installed locally. Once you get logged in successfully, navigate to your dashboard via "Dashboards" and select the date and time range (Note: this will be in 24-hour format) from the top right corner. When exporting data, it is advisable to capture some additional data before and after the event. To do this – click and grab your cursor on the graph to zoom in on a particular event. Once you feel like you have enough data for a particular event, click on the drop-down menu for Panel Title > select More > and choose the option to Export CSV. The "Export CSV" dialogue box is shown in figure 11 for your reference on the settings.

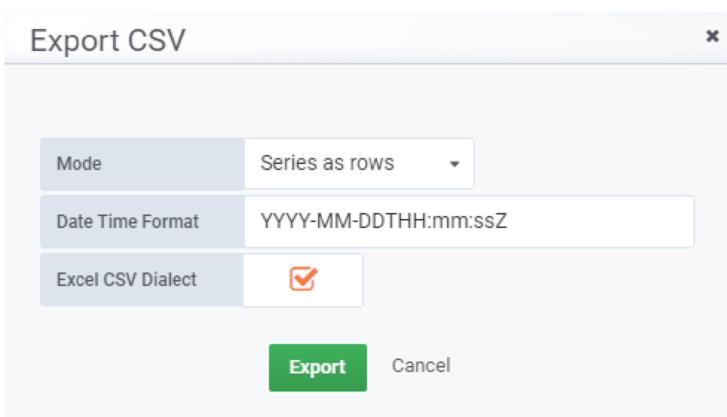

Figure 11 – Screenshot of Data Extraction from Grafana

This data extraction only works when access Grafana on the device itself (no export option available on the local install of Grafana). Once you have the data exported and saved at a location of your preference, make sure to open the file for verification. These steps will need to be performed for each event that you would like to export for future analysis. When exporting an event, it is important to zoom in on the event itself to eliminate the surrounding noise as much as possible. This will improve the quality of your sample. It is also advisable to leave some activity before and after the actual event to analyze the behavior that leads up to the event and after the event. This exported data will need to be imported into MatLab in stage 4 for feature extraction.

STAGE 4: DATA ANALYSIS

The last stage of the process is the study of the data. All the data that was extracted in the previous stage will need to be imported into MatLab. Before we can import the data, we will need to download and install the MatLab application and also format the data in a way that can be read into MatLab.

Open a browser and head over to https://www.mathworks.com/ to download the latest stable release of MatLab for your operating system. Please note that MatLab does require a valid license to be installed, but you may be able to obtain this through your educational institution. Once the executable has been downloaded, please run the file and install the application on your local computer. During the installation process, make sure you select the following plug-ins that will be needed for our type of data: Signal Processing Toolbox, Wavelet Toolbox, and Simulink. These three plug-ins are essential to interpret the digital signal data. If you forget to choose them during the installation, you can always add them in after the application has been successfully installed. It may prompt you to reboot your computer to finish the install process.

Once the application has been installed, it is now time to create a new project and start to import the CSV files exported as part of the Data Extraction phase. Before you import, you will need to reformat the files by removing column A which has the value of “Z” and the first row which contains the headers for each column. This information is redundant and unnecessary for the analysis. Once you have removed and reformatted each file, we will read the file into our new MatLab project. There is quite a bit of data collected as part of this experiment so we will only choose a fraction of events from each of the locations for this initial analysis. The idea here is to read the data first and then extract some features from that data that could potentially help us train the machine learning algorithm in detecting the events and classifying them as one of the ADLs. It is important to note that not every feature will be applicable in the classification of the event. There will be some trial and error in the process of figuring out which feature is able to be used for the classification purposes.

Once you have each of the CSV files imported into your project, first step is to map the data. The reason you would want to do this is for verification and confirmation purposes. Once you plot the data in MatLab, the corresponding graph should be an exact replica of what you saw in Grafana. If the two graphs are not identical then there is some issue with the data itself. Figure 12 shows comparison between the data plotted in Grafana (top) and MatLab (bottom) and both are an exact copy of each other which signifies the data has been exported and imported correctly. By using MatLab to plot the data, you also have the ability to remove the noise around the event to smoothen out and enhance the actual signal for the event itself.

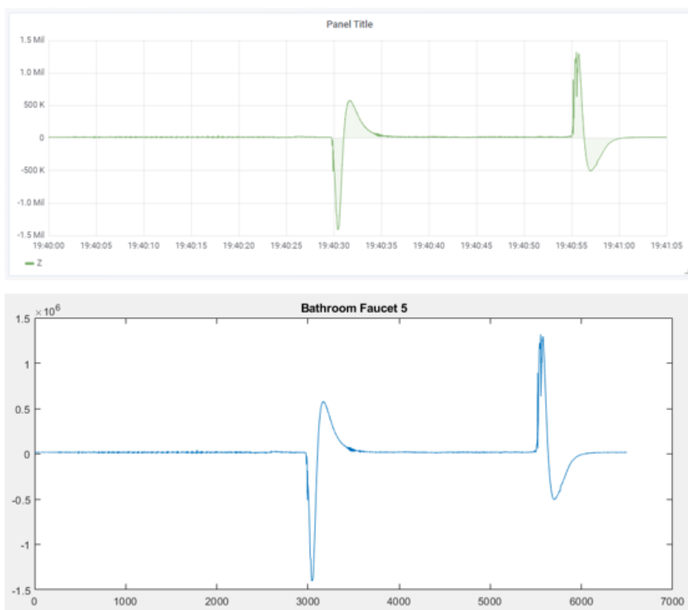

Figure 12 – Screenshot Comparing Data Plotting in Grafana (top) and MatLab (bottom)

After the plotting and verification of data, next step in the process is to do feature extraction. This piece is very important as it allows us to manipulate the data in a way that can highlight the similarities and differences between the datasets. Feature extraction is a critical step not only for us to be able to interpret and comprehend the data, but it is also necessary for the SVM as the results from this feature extraction will be fed into the machine learning algorithm for the automated classification of the hygiene event. Without feature extraction, it is very difficult to make much sense out of the raw data as that is only values tied to a point in time.

The logic behind feature extraction is identifying the features that can expose the qualities of the data that can be used in training of the SVM. These features need to be relevant to the data, activity, and our end goal for the research. It is completely acceptable to find the features that are relevant to the data and research but does not add much value in determining the classification – not every feature will be used for classification. The main goal here is to understand the data and make sense of how the data is tied to the hygiene event.

Here is a list of features that were extracted

Kurtosis	Location of the Highest Peak
Standard Deviation	Location of the Lowest Peak
Entropy	Peak Difference
Highest Peak	Mean
Lowest Peak	Period

Table 4 - List of Extracted Features

CHAPTER IV.

RESEARCH FINDINGS

During the previous analysis phase, we quickly realized that the quality of data is more important than the quantity of data; for example, you can have 1000 events captured, but if these 1000 events do not correspond accurately to the log that you maintained then you would have extracted incorrect data which would then lead to incorrect feature identification resulting in inconsistent findings. All the issues encountered during the initial capture resulted in a lack of confidence with the first batch of data so additional data was collected (while verifying each event in real time) to increase the accuracy and credibility of the data. First batch of data was discarded.

The new batch of data included 30 new events which were further broken down into a new category of "Behavior" that would indicate the level or amount of faucet opened which represents the "intensity" of the water through the faucet. The very first step on the data is to plot out the events to visualize the similarities and differences between the various samples for the same event at the same location. This visualization, as shown in figure 13**,** of the data would help in better understanding how factors such as angle, intensity, and even duration could influence the dataset. It can also help find anomalies such as Kitchen Sink 7 event in figure 13.

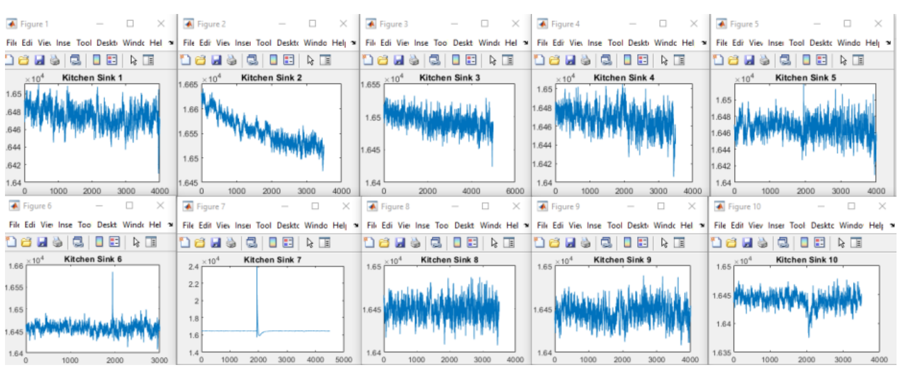

Figure 13 – Screenshot Comparing Samples of the Same Event

As we compiled the results from feature extractions on each of the hygiene events, it was pretty apparent to a human eye which features would be ideal for autonomous identification and classification, but the SVM would not be able to differentiate at a glance like how a human can. In order for the SVM to be able to distinguish between the hygiene events, each of the values from the feature extraction would need to be fed into the machine learning model for training and cross validation.

Here are some initial observations from the analysis of the features –

- Most of the values seem to have fallen within a range, except a few outliers – this is to be expected as the results would need to be somewhat consistent for event classification
- The standard deviation for the bathroom faucet was very high – this was also expected as the sensor was sitting on top of the faucet and the opening/closing of the faucet would have created increased movements
- The highest and lowest peak numbers were also distinct for bathroom faucet as it involved the movement of the sensor with each opening/closing event which would have resulted high/low frequencies

Support Vector Machine (SVM) is a machine-learning model used for classification and regression problems using a hyperplane in an N-dimensional space (N is the number of features) that is ideal in relation with the support vectors [21]. This is why feature extraction is a necessary step in classification via the SVM as it emphasizes the data points that adds the most value and could be used in determination of the support vectors. In our case, we will be using a binary classification schema which is an OR model that will help us differentiate between two classes which are the different hygiene events and/or locations.

Before diving into the specifics of the configuration of the classification model, the dataset will need to be reformatted in a way that is readable by the learning model. The data will need to be split in two ways: 1) Labels – numeric values (1, 2, 3) corresponding to the three locations (kitchen, bathroom, toilet) of the events and 2) Values – the results of the feature extractions from each of the events. Both files need to be saved as Comma Separated Values (CSV) format so they can be imported into MatLab. When creating these files, consistency is of utmost importance as the values in the second file need to be in exact order to accurately correspond to the labels in the first file. Also, the features would need to be lined up in order as well so SVM does not read jumbled data as this would throw off the classification model when the values are mixed. Since this is a binary classification model, we had three separate files containing 60 events each which would compare data from kitchen sink with bathroom faucet, bathroom faucet with toilet flushing, and toilet flushing with kitchen sink to measure accuracy of the classification from each batch.

Next would be to divide up the dataset into training and test batches – the training set would be the larger set of the two as you need more data to train the model than to test. The split of 80:20 was chosen which divided 24 events from each location for training and 6 events were left for testing the trained set so each file had 48 events for training and 12 for test. To further increase the accuracy of results, the data was split in a random fashion which means the algorithm was responsible for choosing which data to test at random. To do this, the function of random number generator was used which divided up the events in a random order for both the training set and the test set. Furthermore, the cross-validation set was prepared with a K-5 fold to train using the entire dataset to improve accuracy by minimizing errors.

Next part of the process was to allow the SVM to choose the three best features out of the ten features used for input – more features does not necessarily mean more accuracy. You may have features extracted that are relevant to the kind of data that you are working with, but the features itself do not add much value to the classification. For each of the three datasets used for comparison, the SVM chose separate features as prominent for classification. No predefined features were selected, so the SVM was allowed to choose the features that it selected as the most ideal for classification.

LOCATION	SVM SELECTED FEATURES
Kitchen Sink – Bathroom Faucet	[1,2,10] Kurtosis, Standard Deviation, Period
Bathroom Faucet – Toilet Flushing	[1,9,10] Kurtosis, Mean, Period
Toilet Flushing – Kitchen Sink	[1,2,10] Kurtosis, Standard Deviation, Period

Table 5 – List of SVM Selected Features

It was interesting to see the SVM select these features as two of the three features that it selected were the ones that were correctly predicted by us. The algorithm tries to find the features that are the most ideal to be able to create the hyperplane which is essentially the decision boundary with maximum margins possible to be able to predict successfully and classify with highest theoretical accuracy.

Since our approach is binary classification, we used “fitcsvm” to build our model with the best hyperparameters. This allows us to use kernel functions such as Radial Basis Function (RBF) kernel [22] as the learning mechanism, calculated as: $G(x_j, x_k) = \exp(-\|x_j - x_k\|^2)$. Support Vector Machines utilize kernel functions to find the support classifiers within a higher dimension. In our case, the RBF relies on observations that are closest to the one that you are trying to classify for prediction purposes. The goal here is to minimize the loss during cross-validation of the data. Figure 14 shows how the model evaluates 30 iterations to try and get as close to no errors as possible

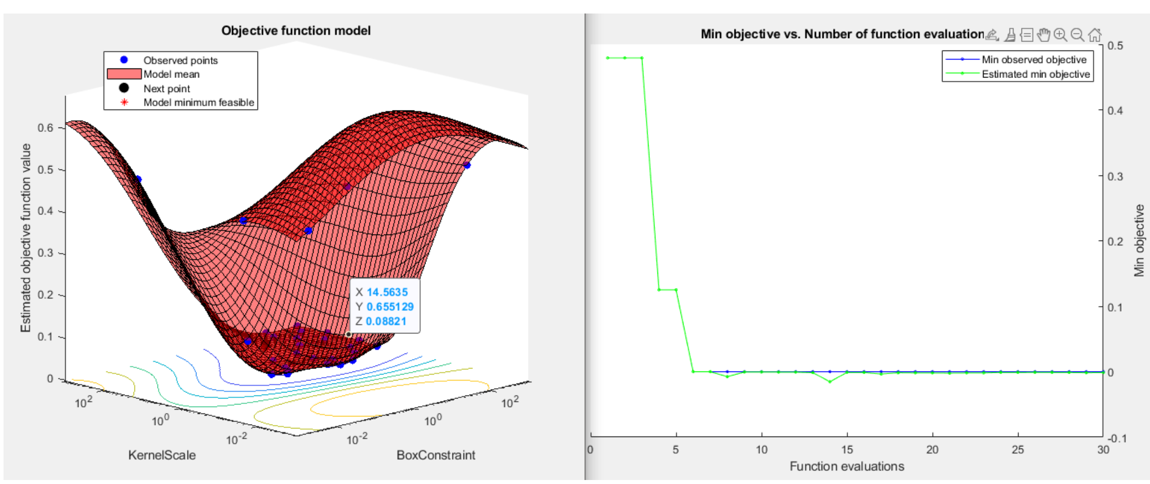

Figure 14 – Screenshot of Objective Function Model

The model was cycled 10 times, so we had a total of 120 events for each of the three classification scenarios. Overall, the training model gave us a very consistent and high accuracy each time with an overall accuracy of above 90% as shown in Figure 15.

Kitchen Sink and Bathroom Faucet - 95%

Attempt	Features	Accuracy	Input	Misclassification
1	1, 2, 10	100%	4KS, 8BF	0
2	1, 2, 10	83.33%	8KS, 4BF	2
3	1, 2, 10	100%	6KS, 6BF	0
4	1, 2, 10	91.67%	4KS, 8BF	1
5	1, 2, 10	100%	8KS, 4BF	0
6	1, 2, 10	83.33%	6KS, 6BF	2
7	1, 2, 10	100%	6KS, 6BF	0
8	1, 2, 10	91.67%	6KS, 6BF	1
9	1, 2, 10	100%	7KS, 5BF	0
10	1, 2, 10	100%	4KS, 8BF	0

Bathroom Faucet and Toilet Flushing - 93.33%

Attempt	Features	Accuracy	Input	Misclassification
1	1, 9, 10	91.67%	6BF, 6TF	1
2	1, 9, 10	91.67%	7BF, 5TF	1
3	1, 9, 10	83.33%	4BF, 8TF	2
4	1, 9, 10	100%	6BF, 6TF	0
5	2, 9, 10	100%	8BF, 4TF	0
6	1, 9, 10	100%	7BF, 5TF	0
7	1, 9, 10	100%	6BF, 6TF	0
8	1, 9, 10	83.33%	5BF, 7TF	2
9	1, 9, 10	100%	8BF, 4TF	0
10	1, 9, 10	83.33%	6BF, 6TF	2

Kitchen Sink and Toilet Flushing - 99.17%

Attempt	Features	Accuracy	Input	Misclassification
1	1, 2, 10	100%	8KS, 4TF	0
2	1, 2, 10	100%	4KS, 8TF	0
3	1, 2, 10	100%	6KS, 6TF	0
4	1, 2, 10	100%	3KS, 9TF	0
5	1, 2, 10	100%	8KS, 4TF	0
6	1, 2, 10	100%	8KS, 4TF	0
7	1, 2, 10	100%	6KS, 6TF	0
8	2, 9, 10	100%	7KS, 5TF	0
9	1, 2, 10	91.67%	6KS, 6TF	1
10	1, 2, 10	100%	6KS, 6TF	0

Figure 15 - Screenshot Comparing the Success Ratio from 10-Run Trial

These results help demonstrate the true potential and the value this approach can add to our end goal of automating the detection of hygiene events to make our healthcare professionals be proactive. Having said that, we do recommend this experiment be conducted under a different set of circumstances to further validate the results above.

CHAPTER V.

CONCLUSION | RECOMMENDATIONS

The results obtained through this experiment validate this research and the concept of using vibration sensors to monitor the hygiene activities and automate the classification process. With the accuracy that we received in our controlled environment, we can confidently conclude that the model does work which would justify further research to advance this. To the best of our knowledge, this is the first work that distinguishes the various activities based on vibrations only.

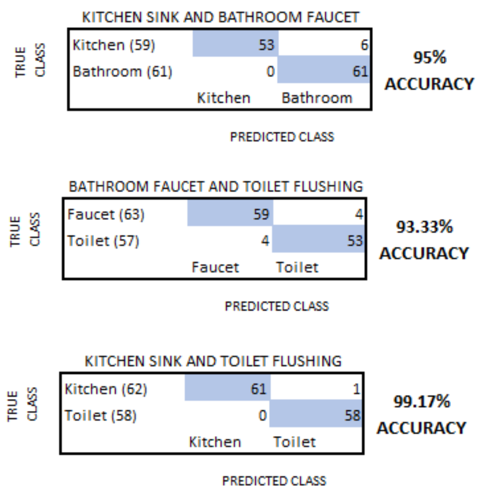

Figure 16 – Confusion Matrix Screenshot

Figure 16 shows the success ratio through confusion matrix for each of the three comparison sets. Between kitchen and bathroom, six of the kitchen events were misclassified as bathroom giving us a 95% success ratio for classification. For bathroom faucet and toilet flushing, four of the faucet events were miscategorized as toilet, while 4 of the toilet events were miscategorized as faucet resulting in a 93.33% success. The highest accuracy was achieved when comparing kitchen sink and toilet flushing where only one kitchen sink event was misclassified as toilet flushing resulting in a 99.17% success.

As the foundation has been set through this research, we hope to see this advance further and potentially lead to a real-world implementation. It is understood and acknowledged that the quantity of data was a potentially limiting factor which could have influenced our results. Also, the fact that the second batch of data was observed and collected in a very controlled and monitored environment which may have helped with the accuracy was another limiting factor which would need further testing. As such, we recommend this experiment and research to be conducted again to replicate as close to a real-world scenario as possible – this would not only provide a larger set of data for analysis, but it would also help understand the behavior around the event itself.

As far as challenges with network connectivity go, our recommendation would be to try wired connection using Ethernet instead of solely relying on wireless as that would improve connection and consistency. We also recommend verifying the collected data more often to identify any gaps whilst in the process of data collection so there wouldn't be a need to discard any data or recollect data. As the need to collect "real-world data" is critical for the success of this research, we need to factor in external influences such as people walking in/out of the designated areas, objects being placed or moved around the sensor, etc. All of these external factors would create frequencies which would add some level of noise to our data, so our recommendation is to run the dataset through a frequency analysis filter to remove any high/low frequencies that are not relevant to the hygiene activity itself. Additionally, there are different shapes and sizes of faucets and toilets, including different build and material, which would need to be tested for to further validate the results.

As we were able to consistently generate accuracy higher than 80% for each individual attempt and higher than 90% overall for each of the three scenarios, we truly believe this research has the potential to add value in both healthcare and technology sectors. With the help of continued research and through a real-world application of this solution, our healthcare professionals will be able to keep track of their patient's daily hygiene behavior and detect any abnormalities that may be a cause for concern. Keeping our end goal of "helping our healthcare professionals be proactive rather than reactive" in mind – Our approach here of "using vibration sensors as a non-intrusive method of monitoring" would equip our healthcare professionals with tools to address any growing health concerns before they become life threatening for the patient and our community.

REFERENCES

1. World Health Organization. (2021, October 4). *Ageing and health*. World Health Organization. Retrieved November 20, 2021, from https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.
2. Public Information Office. (2016, March 28). *U.S. population aging slower than other countries, Census Bureau reports*. United States Census Bureau. Retrieved October 21, 2021, from https://www.census.gov/newsroom/press-releases/2016/cb16-54.html.
3. Mather, M., Kilduff, L., & Scommegna, P. (2019, July 15). *Fact sheet: Aging in the United States*. PRB. Retrieved November 20, 2021, from https://www.prb.org/resources/fact-sheet-aging-in-the-united-states/.
4. National Center for Biotechnology Information. (n.d.). *Size and demographics of aging populations*. Providing Healthy and Safe Foods As We Age: Workshop Summary. Retrieved November 20, 2021, from https://www.ncbi.nlm.nih.gov/books/NBK51841/.
5. Harris-Kojetin L, Sengupta M, Lendon JP, Rome V, Valverde R, Caffrey C. Centers for Disease Control and Prevention. (2019, February). *Long-Term Care Providers and services users in the United States, 2015-2016.* Centers for Disease Control and Prevention. Retrieved November 20, 2021, from https://stacks.cdc.gov/view/cdc/76253.
6. U.S. Department of Health and Human Services. (2017, May 1). *What is long-term care?* National Institute on Aging. Retrieved November 20, 2021, from https://www.nia.nih.gov/health/what-long-term-care.
7. Centers for Disease Control and Prevention. (2021, March 1). *Nursing Home Care*. National Center of Health Statistics. Retrieved November 20, 2021, from https://www.cdc.gov/nchs/fastats/nursing-home-care.htm.
8. Maryville University. (n.d.). *Stages of human development: What it is & why it's important*. Maryville Online. Retrieved November 20, 2021, from https://online.maryville.edu/online-bachelors-degrees/human-development-and-family-studies/stages-of-human-development/.
9. Tikkanen, R., & Abrams, M. K. (2020, January 30). *U.S. health care from a global perspective, 2019: Higher spending, worse outcomes?* Improving Health Care Quality. Retrieved November 20, 2021, from https://www.commonwealthfund.org/publications/issue-briefs/2020/jan/us-health-care-global-perspective-2019.
10. National Health Services. (2021, July 5). *Alzheimer's disease - Prevention*. NHS choices. Retrieved November 20, 2021, from https://www.nhs.uk/conditions/alzheimers-disease/prevention/.
11. Alzheimer's Association. (n.d.). *Alzheimer's Disease Facts and Figures*. Alzheimer's Disease and Dementia. Retrieved November 20, 2021, from https://www.alz.org/alzheimers-dementia/facts-figures.

12. Centers for Disease Control and Prevention. (2016, July 26). *Hygiene Fast Facts: Information on Water-related Hygiene*. Water, Sanitation & Environmentally-related Hygiene. Retrieved November 20, 2021, from https://www.cdc.gov/healthywater/hygiene/fast_facts.html.
13. Zhang, Y., Wullems, J., D'Haeselee, I., Abeele, V., & Vanrumstee, B. (n.d.). *Bathroom activity monitoring for older adults viawearable device* (thesis).
14. Chen, J., Zhang, J., Kam, A. H., & Shue, L. (n.d.). *An Automatic Acoustic Bathroom Monitoring System* (thesis).
15. Hao, Y., Zhang, G., Jiang, J., Shangguan, F., & Zhang, Z. (n.d.). *Design and Implementation o f An Intelligent KitchenSafety Monitor* (thesis).
16. Jimi, K., Seto, H., & Kajiwara, A. (n.d.). *Bathroom Monitoring with Fast-Chirp Modulation Millimeter-wave Uwb Radar* (thesis).
17. Fagert, J., Mirshekari, M., Pan, S., Zhang, P., & Noh, H. (2017). *Monitoring Hand-Washing Practices Using Structural Vibrations* (thesis). Destech Publications, Lancester.
18. Brinke, J. K., Chiumento, A., & Havinga, P. (n.d.). *Personal hygiene monitoring under the shower using Wi-Fichannel state information* (thesis).
19. Raspberry Shake. (n.d.). *RS1D: Turnkey*. Raspberry Shake Shop. Retrieved November 26, 2021, from https://shop.raspberryshake.org/product/turnkey-iot-home-earth-monitor-rs-1d/.
20. Edemekong, P. F., Bomgaars, D. L., Sukumaran, S., & Levy, S. (2021, September 26). *Activities of Daily Living*. National Center for Biotechnology Information. Retrieved November 26, 2021, from https://www.ncbi.nlm.nih.gov/books/NBK470404/.
21. Gandhi, R. (2018, June 7). *Support Vector Machine - Introduction to Machine Learning Algorithms*. Medium. Retrieved November 26, 2021, from https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47.
22. Mathworks. (n.d.). *fitcsvm*. Mathworks Help Center. Retrieved November 26, 2021, from https://www.mathworks.com/help/stats/fitcsvm.html.

APPENDIX A.

Manual (electronic) Log of the Events

Date	Start Time	Duration (in sec)	Bldg Type	Location	Position	Event Type	Behavior	Sensor Distance (in CM)
10/2/2021	5:03:30PM	20	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/2/2021	5:07:20PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/2/2021	5:08:10PM	30	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/2/2021	5:09:30PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/2/2021	5:10:15PM	20	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/2/2021	5:11:10PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/2/2021	5:12:20PM	25	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/2/2021	5:13:10PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/2/2021	5:14:15PM	20	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/2/2021	5:15:20PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/2/2021	7:32:15PM	30	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/2/2021	7:34:20PM	20	Residential	Bathroom Faucet	Top	Opening/Closing	Hard	2
10/2/2021	7:36:25PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/2/2021	7:38:15PM	35	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/2/2021	7:40:20PM	25	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/2/2021	7:43:15PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/2/2021	7:45:30PM	20	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/2/2021	7:50:20PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/2/2021	7:52:25PM	30	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/2/2021	7:53:40PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/10/2021	3:43:15PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	3:53:25PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	3:55:35PM	20	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/10/2021	3:57:15PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/10/2021	3:58:20PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	3:59:30PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/10/2021	4:01:10PM	25	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/10/2021	4:02:20PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/10/2021	4:03:15PM	30	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/10/2021	4:04:30PM	20	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	4:07:40PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/10/2021	4:08:20PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	4:09:30PM	20	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/10/2021	4:10:35PM	5	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/10/2021	4:12:10PM	20	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	4:13:30PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Soft	1
10/10/2021	4:14:20PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/10/2021	4:15:25PM	25	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	4:16:30PM	15	Residential	Kitchen Sink	Side	Opening/Closing	Normal	1
10/10/2021	4:19:40PM	10	Residential	Kitchen Sink	Side	Opening/Closing	Hard	1
10/10/2021	4:48:10PM	25	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/10/2021	4:50:35PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Hard	2
10/11/2021	3:33:25PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:12:10PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:1415PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:16:20PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:19:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:20:20PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1

10/11/2021	4:23:20PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:28:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:29:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:30:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:31:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:42:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:43:20PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:44:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	4:59:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:01:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:02:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:03:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:04:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:08:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:09:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:10:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:12:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:13:15PM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/11/2021	5:21:20PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	5:22:30PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	5:23:25PM	20	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/11/2021	5:24:20PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Hard	2
10/11/2021	5:25:15PM	30	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	5:51:30PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/11/2021	5:52:40PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/11/2021	5:53:35PM	20	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	5:55:20PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Hard	2
10/11/2021	5:57:10PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	6:01:10PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/11/2021	6:02:20PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	6:03:15PM	30	Residential	Bathroom Faucet	Top	Opening/Closing	Hard	2
10/11/2021	6:05:10PM	20	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	6:06:15PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/11/2021	6:08:10PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/11/2021	6:09:25PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Hard	2
10/11/2021	6:11:30PM	15	Residential	Bathroom Faucet	Top	Opening/Closing	Soft	2
10/11/2021	6:12:40PM	10	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/11/2021	6:14:20PM	30	Residential	Bathroom Faucet	Top	Opening/Closing	Normal	2
10/31/2021	11:21:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:23:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:25:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:27:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:29:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:31:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:34:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:36:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1
10/31/2021	11:39:15AM	30	Residential	Toilet Flushing	Top	Flushing	Normal	1

Note: The first batch of recorded events have not been included in this dataset