Public Safety Alarm

ECE445 Design Document

Team 35 - Adrian Wells - Swetank Griyage - Sagar Katiyar
ECE 445 Project Proposal - Spring 2021
TA: Dean Biskup

Contents

1. Introduction

1.1 Problem and Solution Overview

In the United States and all over the world, people are victimized in silence. According to the United States Bureau of Justice Statistics, over 3.4 million violent crimes per year went unreported from 2006 until 2010 [1]. Stigma, lack of evidence, lack of witnesses, and lack of personnel remain an issue even in one of the safest countries. To add to this, only 42% of violent crime and 36% of property crime was reported to authorities in 2016, showing us that there is clearly a need for improvement in protection [2]. Although this is a very large problem, ranging across many different types of settings and spaces, it's very apparent that different situations or settings can be solved in different ways, and the best solutions will be setting dependent. Our objective is to create a simple alarm system that can be installed around schools and other public places in order to aid authorities in covering these areas without always having to be present. According to the National Center for Education Statistics and Bureau of Justice, nearly 20% of students between the ages of 12 and 18 experience bullying while only 46% of these incidents are reported to authorities [3]. More specifically, about 45% of students have experienced bullying in hallways or stairwells, and 13% of students have experienced bullying in locker rooms or bathrooms, all of which are locations that aren't constantly monitored by authorities.

In order to solve this problem, our design will feature two main components, the monitoring device, and the supervisor device. The monitoring device will be fixed at the location of importance and designed to take in both sound and infrared video of the environment. It will be constantly analyzing the sounds of the environment and making decisions on whether or not an emergency is occurring, while always protecting the identities of those on the scene. The supervisor device will then be a mobile device, designed to be carried with a supervisor at all times that is connected to the monitoring device via wifi. In the event of an emergency detected, this alarm will come to the attention of the supervisor, and show them the infrared video starting 3 seconds before the detection of the emergency. This will give them the information needed to act accordingly in order to try and aid the situation, if needed. After the event has passed, the supervisor will have an opportunity to provide feedback on whether or not the alarm was a false alarm or not. This feedback will be used to constantly improve our decision model.

1.2 Background

There are several companies that specialize in crime detection systems such as HikVision that have made cameras that claim to be able to detect potential anomalies in public areas [4]. However, there still isn't a low-cost, easily implementable and widely deployable automated surveillance system in place, and that is why we want to come up with a solution that addresses all these concerns. We are striving to automatically detect potentially dangerous situations while not being too complicated or too expensive to publicly install. While our technology is not entirely new, delivering this essential product at a fraction of the cost will be an enormous improvement given the environment it will most likely be used in.

1.3 Visual Aid

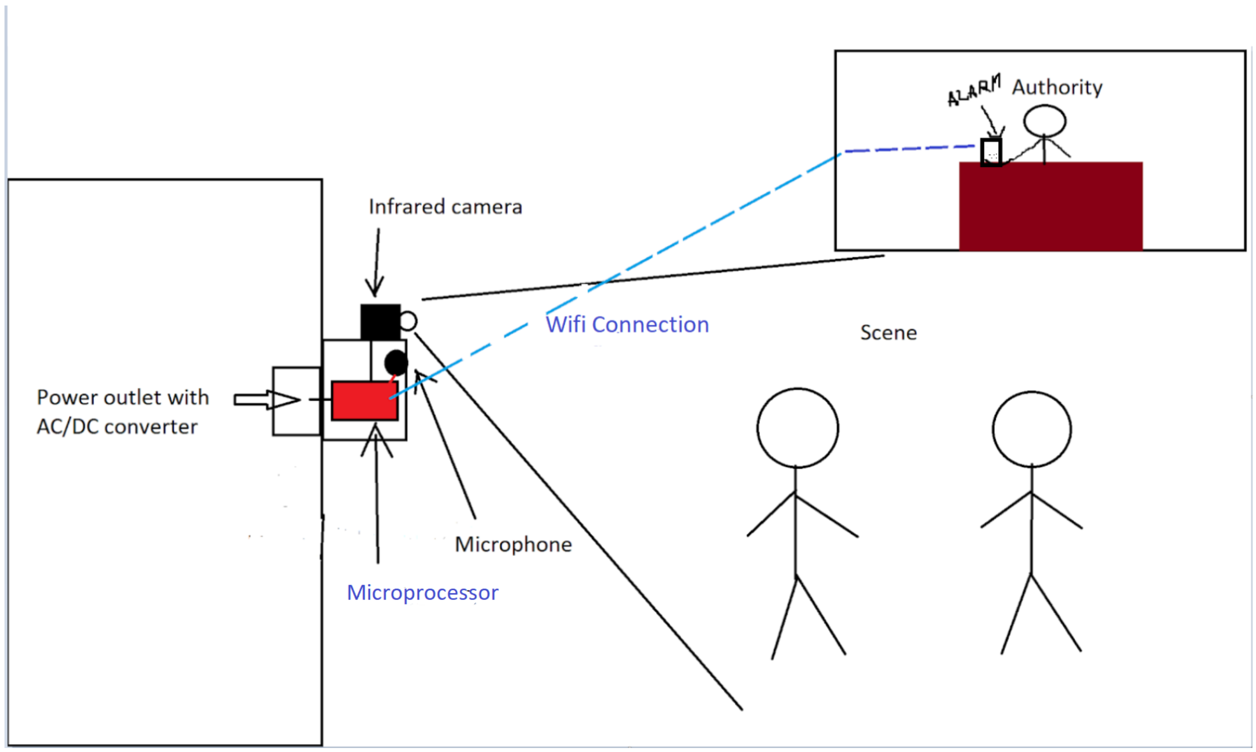

Figure 1. A visual representation of an emergency scene

1.4 High Level Requirements

For our high level requirements we will be defining each separate moment as a period of 5 seconds long. This will allow us to properly calculate the accuracy percentages.

1. Our system should be able to detect potentially dangerous audio situations with an accuracy of at least 95%. This means that given a situation, dangerous or not, the alarm's decision will be correct at least 95% of the time.

2. The recall of the system should aim to be at least 80%. This measure can be calculated by dividing the number of reported incidents with a clear audio indication by the total number of incidents with a clear audio indication.

3. The response delay between the occurrence of alarming audio and the supervisor alarm sounding should be 5 seconds or less.

2. Design

2.1 Block Diagram

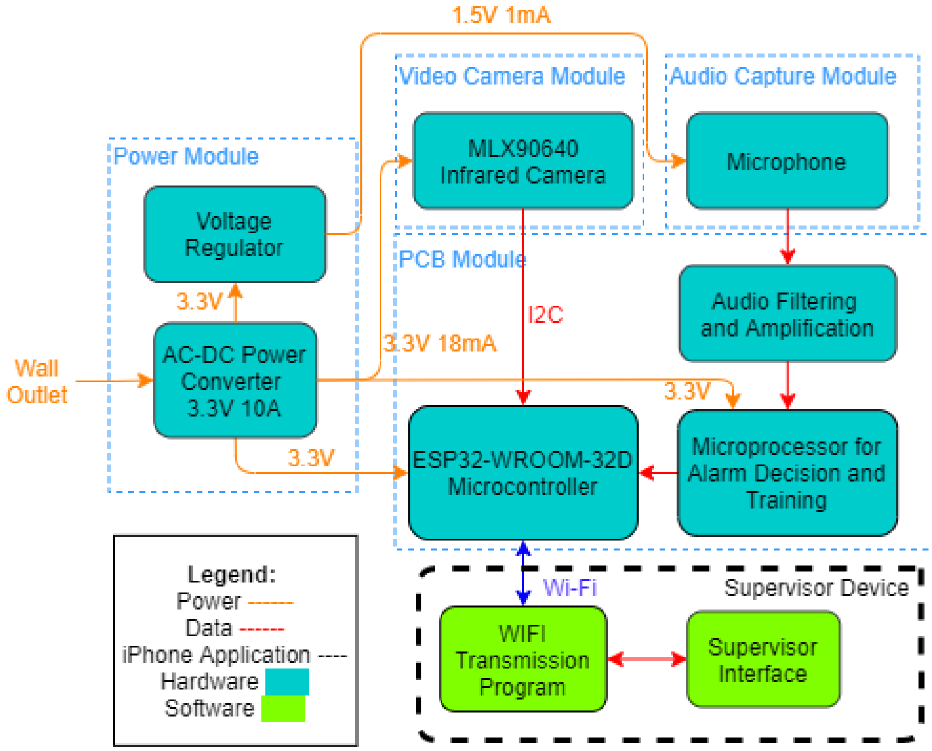

Figure 2: Public Safety Alarm Block Diagram

2.2 Physical Design

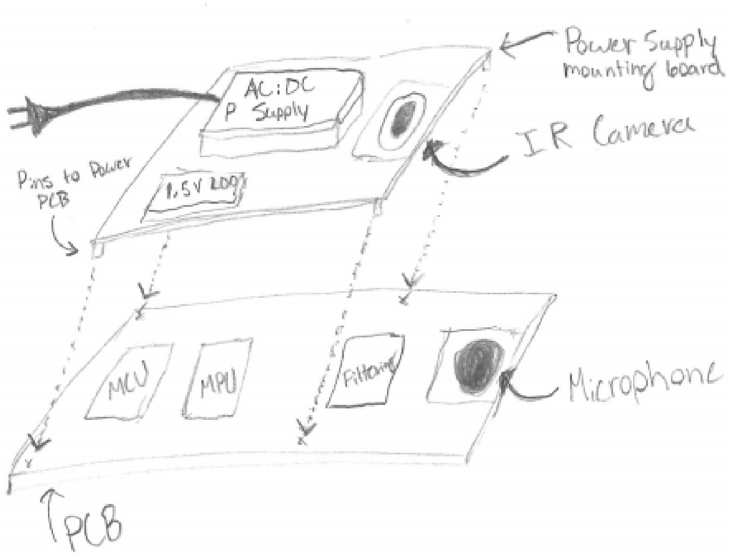

Figure 3. Physical Design plans for PCB, Camera, and Power Supply

2.3 Subsystems

2.3.1 Infrared Camera

The infrared camera block will be responsible for capturing thermal infrared video from the scene and processing the video into transmittable data. This data will be temporarily saved in the ESP-32 Microcontroller, so that it can be sent to the supervisor in the event of an emergency. The goal of this module is for the camera to be able to clearly detect human beings, and allow the supervisor to have a good understanding of what's going on when he receives an alert that there is an emergency.

We are currently looking at the Grove MLX90640 Thermal Infrared camera as our best potential option to fit our Camera needs. As you can see in Figure 4, the MLX90640 provides a clear enough picture for the supervisor to analyze, while also keeping the identity and attributes of those in range from being too obvious. It features a 110 degree field of view, allowing us to get a full view of the scene when placed in a corner and a relatively full view of a scene when placed on a wall. According to it's datasheet, the MLX90640 can be powered with 3.3V and approximately 18mA. It will easily be powered by the AC-DC Power converter we have chozen. It can process images at a frequency up to 64 Hz, which will be more than fast enough for the Supervisor to analyze the scene. The MLX90640 does all of the image processing in house and will output data in I2C format, making it easily compatible with the ESP-32 Microcontroller [5].

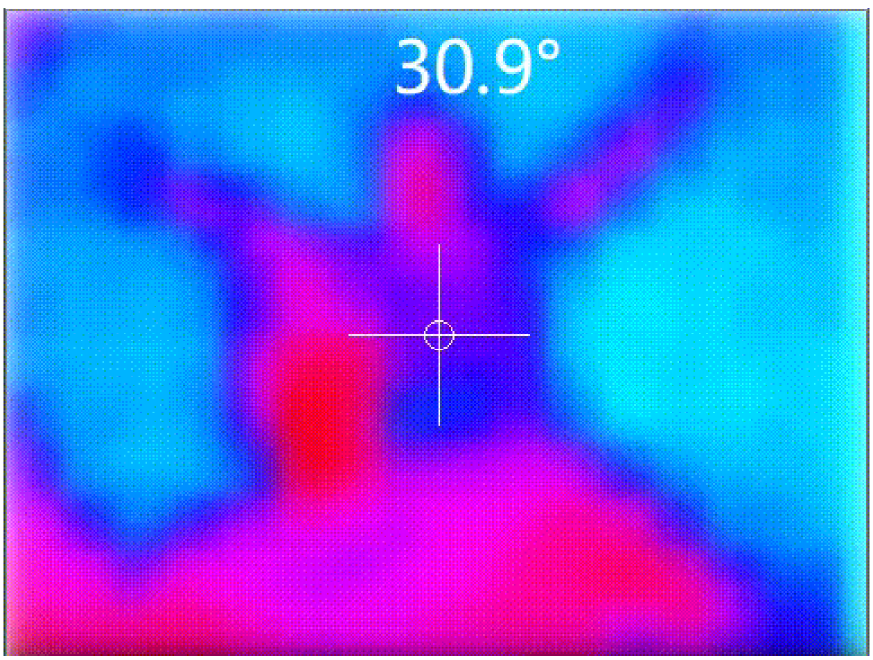

Figure 4. Example of Infrared Image when using the MLX90640 [6]

Requirements	Verification
1. IR Camera must be able to detect a human from 10 feet away. 2. Infrared (IR) Camera must capture Infrared video at a frame rate of at least 10 Hz.	1. Mount the IR Camera and mark off 10 feet from the lens. From the marking, have 2 people stand side by side moving slowly. Ensure the video output displays 2 distinct warm figures. To unit test the Infrared Camera and ESP-32, we will be using the iPhone app called Blynk which allows easy connection to ESP devices. After Unit Testing we will need to repeat the same process when the video is being displayed in the Supervisor IOS application. 2. Mount the IR camera on the wall and open the stopwatch application on a smartphone. Have someone place their hand in front of the lens for 2 full seconds. From here, you can count the number of frames sent to the supervisor and divide it by 2 to find the frequency. Ensure it is at least 10 Hz. To unit test the Infrared Camera and ESP-32, we will be using the iPhone app called Blynk which allows easy connection to ESP devices. After Unit Testing we will need to repeat the same process when the video is being displayed in the Supervisor IOS application.

Table 1. Infrared Camera Requirements and Verification

2.3.2 Microphone

The microphone block is a very important component for our project as it is responsible for listening to the scene being monitored in efforts to detect any emergency situations. We are planning on using the Projects Unlimited AOM-4542P-R because it should fit the requirements of our design well. Specifically, the AOM-4542P-R is omnidirectional, meaning it can pick up sounds in any direction which will only improve the awareness we can create in the monitored location. On top of this, the microphone is more sensitive than most with a sensitivity benchmark of -42 dB. This will allow us to pick up sounds at least 10 feet away from the microphone. Lastly, the AOM-4542P-R is designed to take in audio ranging from 20 Hz to 19K Hz, which will specifically target the frequency range humans can create [7]. Humans can create noises in the range of 125 Hz to 8kHz, with a large majority of the noises falling between 2 kHz and 4 kHz. This range includes the sounds at a raised, loud and shouted level, for males, females and children. This can be seen in more detail on the spectra in Figure 5.

Our microphone will be attached to our PCB and sending analog outputs to the audio filtering and amplification module. It will strategically be placed at the bottom of the board to ensure it's as close to the scene as possible, and also help avoid any on board interference that could occur. It will be powered by the LDO Voltage Regulator, providing it 1.5V of Voltage and 0.5mA of current [8].

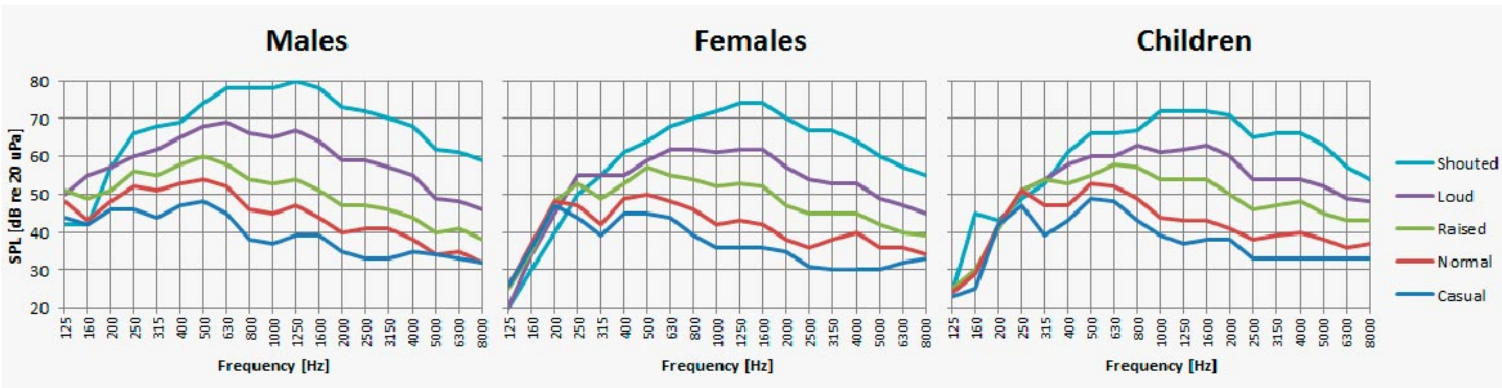

Figure 5. Voice Spectra of Males, Females, and Children [8]

Requirements	Verification
1. The microphone must be able to take sounds ranging from 125 Hz to 8 kHz. 2. The microphone must detect a moderate yell from 10 feet away.	1. First ensure that the microphone is receiving proper input voltage and current supply. Next use a phone application to sweep the frequency from 125 Hz to 8 kHz. Use an Ohmmeter to ensure that the Microphone detects inputs throughout the entire range. 2. Ensure the microphone is receiving proper input voltage and current supply. Walk 10 feet away from the microphone and use a phone application to sweep a noise at 75% from 125 Hz to 8 kHz. Use an Ohmmeter to ensure that the Microphone detects inputs throughout the entire range from 10 feet away.

Table 2. Microphone Requirements and Verification

2.3.3 Audio Filtering and Amplification

The Audio Filter and Preamplifier are attached to the microphone and its main objective is to remove some of the noise from the background by the use of a band pass filter that only allows frequencies that can be produced by the human voice to go through and clean the rest as well as increase the volume of the signal picked up by the microphone. The noise filter looks like the schematic in figure 6, with the frequency range determined by the resistance values.

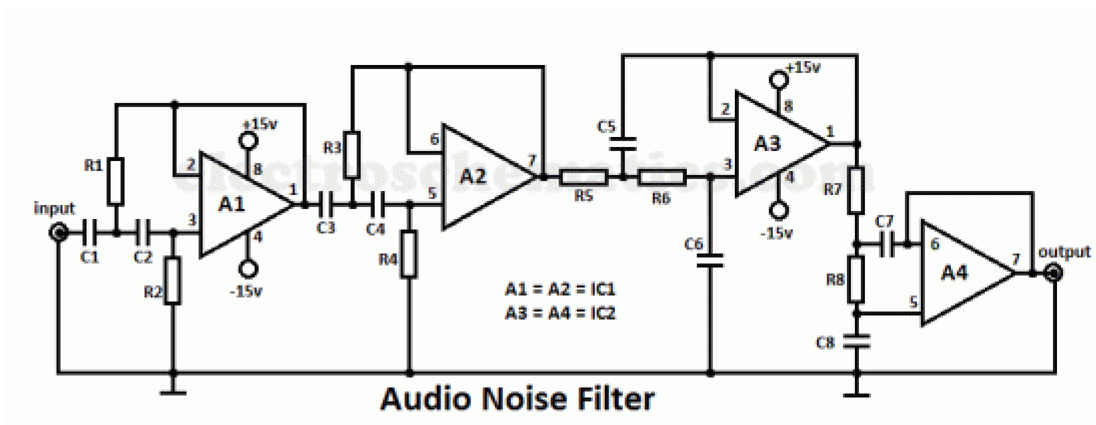

Figure 6: Schematic for an Audio Noise Band-pass filter [9]

Requirements	Verification
1. Audio Filter must only allow frequencies between 100Hz and 3Khz to be present in the input sound. 2. The pre-amp must be able to raise the input volume to the point that sounds from 10m away must be clearly audible.	1. Make some sounds and let them be sampled. Then analyse the frequencies of the sample and see if the noise is being filtered out (consider Audacity software to analyse frequencies). 2. Sounds must be fed into the microphone from the maximum range and checked to see if they are easily audible and distinguishable.

Table 3. Audio Filtering and Amplification Requirements and Verification

2.3.4 Alarm Decision and Decision Training

Possibly the most crucial part of our entire system is the alarm decision. This will be the logic that decides when a supervisor may be needed at the scene. The alarm decision will be based on a training model, which will take in recorded .wav files of 5-second intervals since sound is the primary sense to decide whether or not a situation is dangerous. The sound analysis has two major components - one being the scream detection, second being speech recognition. The reason both parts of this sound analysis are important is because one could easily abuse the system if simply words like ‘help’ or ‘stop’ are recognized. In addition, such keywords are used in everyday speech, which is why relying on simply detecting words would lead to many false positives. Instead, scream detection would take into account the emotion or the context in which such words are being used. For example, using ‘help’ in a sentence would not be dangerous, whereas screaming or shouting for help would be. While there are many reasons why one may hear a scream, it is a better indicator of a dangerous situation.

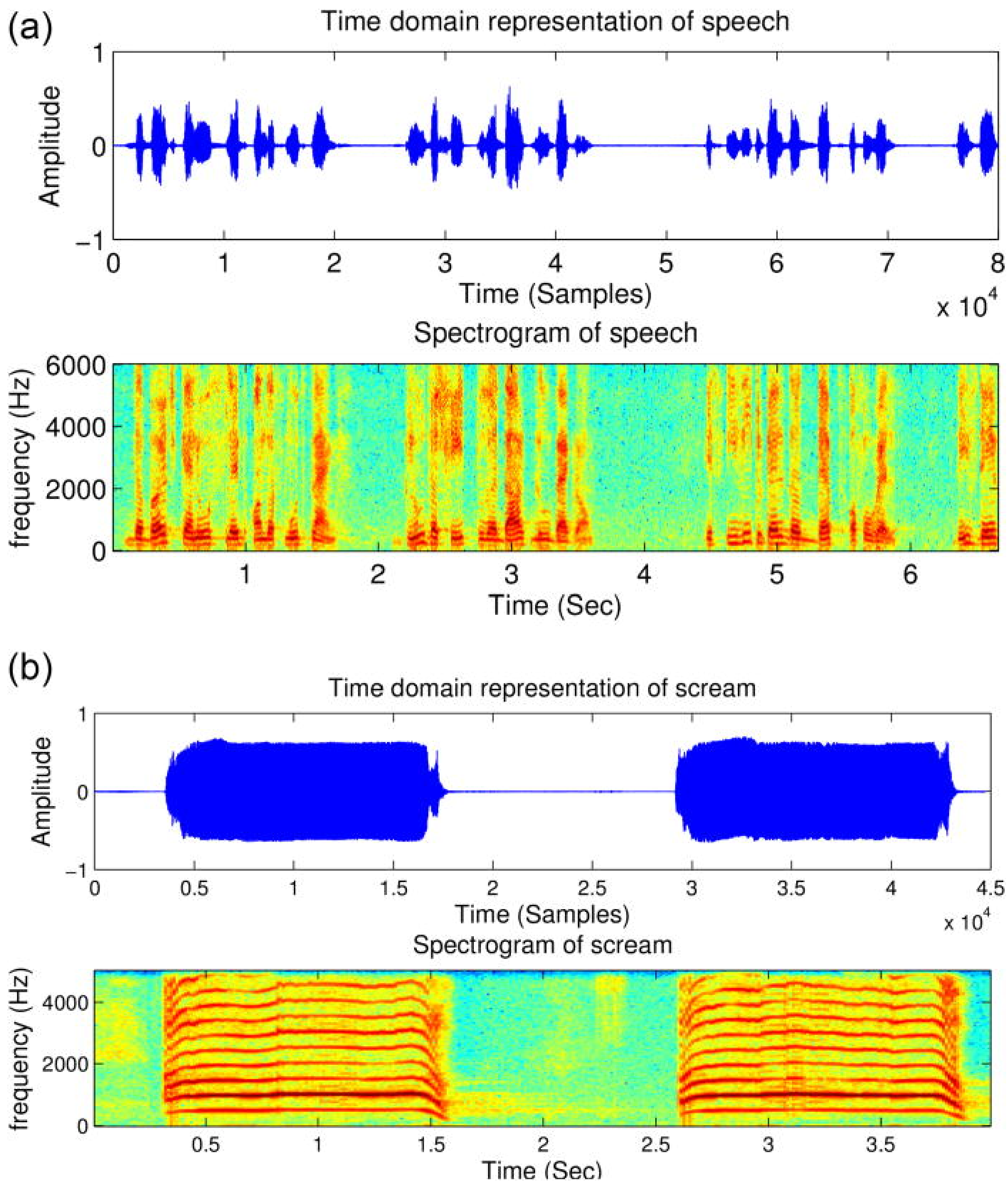

Figure 7: Time and Frequency domain representations for speech and screams. [10]

A scream is defined by Oxford Dictionary to be a long, loud, piercing cry expressing extreme emotion or pain, which is reflected in the time-domain representation shown above. This very clearly contrasts with normal speech, which is varied and irregular. This is also different from sobbing or laughter, which displays a similar irregular pattern [10].

The entire decision making logic will be hosted on the STM32F405RGT6TR microprocessor. The MPU will receive the sound input from the audio filtering and amplification module and will send the trigger signal to the ESP-32 when the model deems the sound input as a scream. The reason behind choosing this microprocessor is the low cost, data storage capacity and processing power making it a viable choice to host our decision making logic [11].

Requirements	Verification
1. The decision making logic must be able to identify the code words like help. 2. The decision making logic must be able to identify scream waveforms correctly.	1. Test the entire system by using the code words such as "help", "stop", "NO!" in a very loud and concerned manner. Check that the alarm at the supervisor interface is triggered. 2. Play sample scream noises in front of the microphone along with some non malicious audio samples and check if the voice samples are correctly categorized by the decision making logic.

Table 4. Alarm Decision and Decision Training Requirements and Verification

2.3.5 ESP-32 Microcontroller

The Microcontroller is essentially the point of communication between the on-site devices and the supervisor, it receives the video feed from the infrared camera as well as the trigger signal from the decision making logic and sends it via Wi-Fi to the supervisor. It will also receive the feedback from the supervisor and send that to the microprocessor to update the decision making logic accordingly. In order to accomplish these tasks, we have selected the Espressif ESP-32 Microcontroller. This MCU fits our needs exceptionally well because it has Wi-Fi communication built in. It also has multiple I2C ports allowing it to communicate with the Infrared Camera without any data manipulation needed. On top of this, it has built in memory, which will be used to save the most recent video footage so that a full context can be communicated to the supervisor. The ESP-32 will be powered by our power supply at 3.3V and an average of 80mA [12].

Requirements	Verification
1. The ESP-32 Microcontroller must be able to send the video data and alarm trigger signal to the IOS device. 2. It must be able to receive the feedback signal from the IOS device.	1. Without the decision making logic, just program the microcontroller to send the video feed from the infrared camera directly to the IOS device and see if it is being correctly received. 2. Send a false alarm and true positive feedback messages from the IOS device and see if the decision making model is updated via the terminal.

Table 5. ESP-32 Microcontroller Requirements and Verification

2.3.6 Wi-Fi Transmission Program

The Wi-Fi transmission program will be happening on the back end of the supervisor interface at all times. It is the block of our design that communicates with the ESP-32 Microcontroller in order to receive emergency signals, video data and also to send supervisor feedback. In order to accomplish all of this, we will be using the public ESPProvision provisionary library which is open source on github. The library provides a simple mechanism for communicating with the ESP-32 via either bluetooth or Wi-Fi. It also provides end to end encryption in order to protect the data being transmitted back and forth, which is very important for our project [13]. The transmission program will be integrated seamlessly with our Supervisor Interface in order to allow for all of this communication to happen behind the scenes in order to not disturb the Supervisor's normal workflow.

Requirements	Verification
1. IOS Application is able to receive a trigger signal over WiFi. 2. IOS Application is able to send feedback to the microcontroller	1. Program the ESP-32 Microcontroller to send a test signal over to the IOS Device. Have the IOS Interface stripped down to merely display the alert was received upon arrival. 2. Program the ESP-32 Microcontroller to send a high voltage to an unused output pin prior to dropping back down to ground when a signal is received. Then, strip down the IOS device to merely press a button to send a signal. Use a voltmeter to ensure that the MCU Output pin is raising to high voltage.

Table 6. Wi-Fi Transmission Program Requirements and Verification Table

2.3.7 Supervisor Interface

The Supervisor interface will consist of three main components: the alarm itself ringing the users phone through push notifications, the visual display of the infrared video, and finally an opportunity for supervisor feedback after the emergency has transpired and been dealt with. The interface will be implemented on an ios application that is triggered by a message from our monitoring device. The interface will be designed to alarm similar to a phone call, and then swiftly move into a video display once the supervisor has given the application their attention. The video display will begin 2 seconds prior to the audio that triggered the decision making, allowing for the supervisor to gather context of the situation. It will then play through the event, allowing the supervisor to then respond as needed. Whenever the supervisor comes back to their phone and the situation has concluded, the supervisor feedback will be set up as a yes or no button that allows the supervisor to provide feedback to the decision logic on whether each event is in fact an emergency. This loop will allow for the model to be constantly improving. A visualization of this process can be seen in Figure 8.

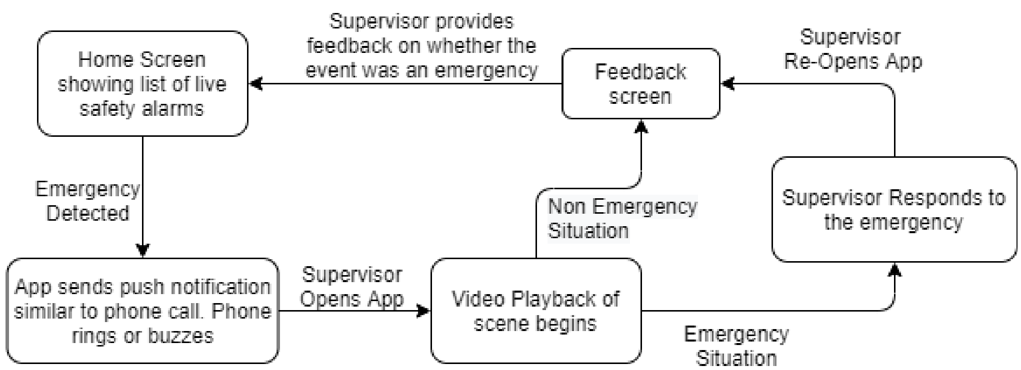

Figure 8. Supervisor Interface Usage Flow

Requirements	Verification
1. IOS Application is able to send an audible push notification to alert the user. 2. IOS Application is able to display emergency video upon receiving it. 3. IOS Application provides two options for feedback after the video is done playing.	1. In order to verify all 3 requirements, we will need all of our system working besides the decision logic and microphone . We can program the ESP-32 Microcontroller to send an alert notification and then follow that notification with video data. We can then turn the notification volume on high and set the phone down. Upon receiving the signal the phone should ring. When we open the app, we should then check for infrared video. After 10 seconds of video, we can then expect the feedback questionnaire.

Table 7. Supervisor Interface Requirements and Verification

2.3.8 AC-DC Power Supply

The AC-DC Power Supply Module is the power supply for the entire monitoring system. It's role is to convert alternating current from a wall outlet into stable direct current that can be used to power each of the components within. In order to accomplish this we are using the Mean Well LRS-50-3.3 which supplies 3.3V and 10 Amps of direct current. 3.3 Volts is the required operating voltage for the infrared camera, the ESP-32 MCU, and the MPU we've selected, so it will be an ideal fit for these components. We also made this selection because 10 Amps will be more than enough to power all of our current components, and will provide excess if we decide our system needs a stronger processor or any other more demanding components [14].

Requirements	Verification
1. Power Supply must output 3.3V +- 0.1V. 2. Power Supply must output 5A of current.	1. Plug in the Power Supply and allow a few moments for it to reach a stable setting. Probe both terminals using an oscilloscope to ensure the voltage remains between 3.2V and 3.4V. 2. Attach Power Supply to constant current testing circuit. Adjust the variable resistor to ensure that 5A of current. This can be measured using a multimeter. In the process, also ensure that the Voltage supplied remains between 3.2V and 3.4V.

Table 8. Power Supply Requirements and Verification

2.3.9 Voltage Regulator

The voltage regulator is an extension of the power supply that aids in powering the microphone that requires less potential energy. The voltage regulator converts the 3.3V of DC from the power supply to 1.5V DC, the operating voltage of the Microphone. In order to accomplish this, we are using the Texas Instruments TPS7A2015PDQNR, because it is very simple and very cost effective [15].

Requirements	Verification
1. Must Supply 1.5 V +- 0.1V to the Microphone Module. 2. Must Supply 1mA of current to the Microphone Module	1. Plug in the Power Supply and connect the Regulator. Allow a few moments for it to reach a stable setting. Probe both terminals using an oscilloscope to ensure the voltage remains between 1.4V and 1.6V. 2. Attach Voltage Regulator to constant current testing circuit. Adjust the variable resistor to ensure that 1mA of current. This can be measured using a multimeter. In the process, also ensure that the Voltage supplied remains between 1.4V and 1.6V

Table 9. Voltage Regulator Requirements and Verification

2.4 Tolerance Analysis
A very important tolerance that we must maintain in our system is microprocessor memory. Our model must have adequate space to store a 15-second audio clip and contain enough data about it to perform analysis. Typically, a machine learning model for such a system is expensive, but for our system, scream detection depends mostly on having enough samples, hence, while it is important that we have enough buses for all that data, but also the appropriate program memory space. Uncompressed audio is in the .wav format and for typical CD's and system sounds, it's frame rate is 44.1 kHz. A frame rate lower than this decreases the quality of the sound, but conserves memory space.

Calculating the size of such a .wav file is not too difficult, it's given by the equation below.
$file\ size\ =\ bit\ depth\ *\ sample\ rate\ *\ duration\ of\ audio\ *\ number\ of\ channels$

Explanation of each
file size- value in bits of how large the file is. This is the most important point to consider because program memory size is limited

bit depth - This is the number of bits of information in each sample. It is meaningful when we are looking at PCM digital signals.

sample rate - This is self-explanatory. The standard rate for CD's and audio on computers is 44.1 kHz. However, this can be decreased if audio quality is not as important of a factor.

duration of audio - Longer clips are bound to take more space in the memory

number of channels - Some audio files designed in such a way that certain sounds play more loudly or softly in one speaker as compared to another. Specifically, monophonic has one channel while stereo sound has 2. For this project, there is no sound being played back. Therefore, monophonic sound will be used.

Apart from calculating the memory that would have to be available, we will also consider Signal Processing methods to determine whether an audio clip indeed contains a scream. One such method is to use the frequency sampling and binning method which involves a simple average. In the below equation,

$$v_k(t) = \frac{1}{F} \sum_{f=1}^{F} h_k(t, f),$$

This method involves the use of spectroscopy. It was employed to obtain the signals shown in Figure 6.

In this equation, F is the number of frequency bins taken of the audio signal while hk is the segmentation mask. The average over each frequency bin gives us the frame-wise prediction score, which, beyond a certain value, we mark it as a scream.

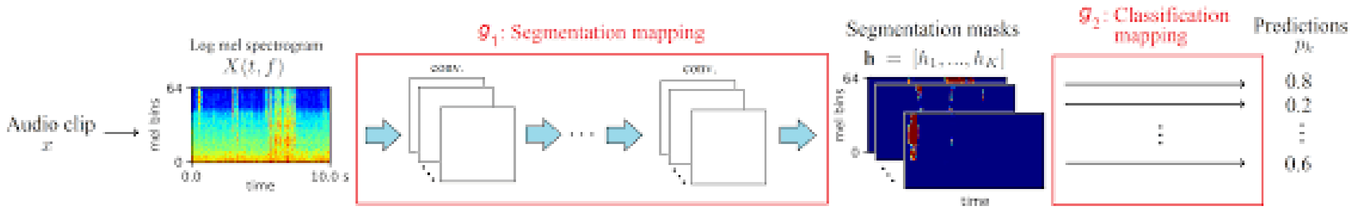

Figure 9 Visual Stages of Decision Model [16]

Here, we see the process used to obtain the segmentation masks. This is known as segmentation mapping wherein the spectrogram is trained on CNN (Convolutional Neural Networks) method to "tag" the audio [16]. That way, the sound samples can be converted into labelled data that can be used for analysis.

The reason why this method is important is because, apart from the audio, our processors should contain both the processing power and the memory to perform such computations so that our system can give timely and accurate alerts.

3. Cost and Schedule

3.1 Cost Analysis

3.1.1 Labor Costs

Team Member	Wage	Hours per Week	Weeks	Multiplier	Total
Adrian Wells	$50	15	12	2.5	$22,500
Swetank Griyage	$50	15	12	2.5	$22,500
Sagar Katiyar	$50	15	12	2.5	$22,500
Total	**$50**	**45**	**12**	2.5	**$67,500**

Table 10. Breakdown of Labor Costs

3.1.2 Part Costs

Part Number	Description	Manufacturer	Quantity	Unit Cost	Total Cost
MLX90640	Infrared Camera	Grove	1	$129.00	$129.00
ESP-32-WROOM	Wifi MCU	Espressif Systems	1	$4.20	$4.20
STM32F405RGT6TR	Microprocessor	STMicroelectronics	1	$11.21	$11.21
AOM-4542P-R	Microphone	Projects Unlimited	1	$1.44	$1.44
RS-50-3.3	Power Supply	Mean Well	1	$12.40	$12.40
TPS7A2015PDQNR	LDO Voltage Regulator	Texas Instruments	1	$0.46	$0.46
Total			**6**		**$158.71**

Table 11. Breakdown of Part Costs

3.1.3 Total Cost

We are projecting to spend $67,500 on labor and $158.71 on parts. Therefore, our total project cost will be **$67,658.71**.

3.2 Schedule

Week	Adrian	Swetank	Sagar
March 8th	Design Review Finalize PCB Schematic	Design Review Finalize PCB Schematic	Design Review Finalize PCB Schematic
March 15th	Finish PCB Layout Order PCB Begin Supervisor IOS UI development	Finish PCB Layout Order PCB Order IR Camera Order Microphone	Finish PCB Layout Order PCB Order IR Camera Order Microphone
March 22nd	Complete Supervisor Application and begin unit testing with ESP-32	Begin building audio decision model	Begin building audio decision model
March 29th	Combine IR Camera, ESP-32 and Supervisor interface for testing	Unit test decision model Combine microphone and, on board filtering and decision model for testing	Unit test decision model Combine microphone and, on board filtering and decision model for testing
April 5th	Finish Unit Testing Full System Debugging	Finish Unit Testing Full System Debugging	Finish Unit Testing Full System Debugging
April 12th	Full System Debugging	Full System Debugging	Full System Debugging
April 19th	Full System Debugging Mock Demo	Full System Debugging Mock Demo	Full System Debugging Mock Demo
April 26th	Demonstration Final Presentation Preparation	Demonstration Final Presentation Preparation	Demonstration Final Presentation Preparation
May 3rd	Final Presentation Finish Final Paper	Final Presentation Finish Final Paper	Final Presentation Finish Final Paper

Table 12. Week by Week Schedule Breakdown

4. Ethics and Safety

The main ethical concern when discussing our project is protecting the privacy of the people in the monitored location at all costs. We are entirely responsible for the data being captured and stored in our system. This valuable data is an implementation of the IEEE Code of Ethics #5: “To improve the understanding of technology; it’s appropriate application, and potential consequences” [17]. Our overall goal to use this data to build a safety alarm that can help bring a safer environment to the masses and help anyone feel comfortable at all times in protected public spaces, whether there is a supervisor present or not.

Keeping data from for example, a locker room, would be a highly sensitive matter, and not one we take lightly. We must pay very close attention to #7, #8 and #9 of the IEEE Code of Ethics in order to protect the people our system is designed to protect [17]. To address these concerns, any data we save will not be time stamped or labeled in any way in order to maintain entire anonymity. Our data will strictly contain the infrared video, audio and a binary assertion of whether or not the specific event was an emergency. On top of not labeling data, only a basic infrared camera will be used, which will not take the details of the people’s faces, clothes or any other accessories and hardly provide any detail on the physical appearance of the people on screen.

Expanding on the sensitivity of video recordings, one very likely application of our device is inside locker rooms and bathrooms in public schools. According to the laws varying by state, we have found that infrared cameras are acceptable in multiple states, including Illinois. They are not legal within restrooms in every state in the United States though, which is something we will need to document should we reach the mass production and distribution phase of our project. [18]

One safety concern that’s always important to address when using a wall socket is protecting the user from possible electrocution on use or installation. Our Monitor is designed to be plugged into a wall socket which can be very dangerous if mismanaged. To help prevent this, we decided to use a premade, fully enclosed, AC:DC power supply that will allow for all high voltages and larger currents to be untouchable by the user. This will prevent any issues or mishaps that could potentially happen during installation or maintenance of the monitor if these parts were not enclosed. On top of that, our power supply is designed to protect from short circuit, overload, and over voltage. [14]

5. References

[1] B. of J. Statistics, “Victimizations Not Reported to the Police, 2006-2010,” *Bureau of Justice Statistics (BJS)*. [Online]. Available: https://www.bjs.gov/content/pub/press/vnrp0610pr.cfm. [Accessed: 02-Mar-2021].

[2] Dispatch, “Why do so many crimes go by unreported in the states?,” *The NYU Dispatch*, 31-Aug-2018. [Online]. Available: https://wp.nyu.edu/dispatch/2018/08/31/why-do-so-many-crimes-go-by-unreported-in-the-states/. [Accessed: 02-Mar-2021].

[3] Wang, Ke, et al. “Indicators of School Crime and Safety.” *National Center for Education Statistics*, nces.ed.gov/pubs2020/2020063.pdf.

[4] D. Faggella, “AI for Crime Prevention and Detection - 5 Current Applications,” *Emerj*, 02-Feb-2019. [Online]. Available: https://emerj.com/ai-sector-overviews/ai-crime-prevention-5-current-applications/. [Accessed: 02-Mar-2021].

[5] Inspired Engineering, Melexis. “MLX90640 32x24 IR Array Datasheet .” Melexis, 2021.

[6] “Grove - Thermal Imaging Camera / IR Array MLX90640 110 Degree.” *Seeed Studio*, www.seeedstudio.com/Grove-Thermal-Imaging-Camera-IR-Array-MLX90640-110-degree-p-4334.html.

[7] Unlimited, Projects. “AOM-4542P-R Datasheet.” *Farnell*, 2006, www.farnell.com/datasheets/40068.pdf.

[8] DPA Microphones. “Facts about Speech Intelligibility: Human Voice Frequency Range.” *DPA*, DPA Microphones, 3 Mar. 2021, www.dpamicrophones.com/mic-university/facts-about-speech-intelligibility.

[9] “Audio Noise Filter Circuit.” *ElectroSchematics.com*, 8 Jan. 2013, www.electroschematics.com/audio-noise-filter/.

[10] Hansen, John H. L., et al. ``Analysis of Human Scream and Its Impact on Text-Independent Speaker Verification).” *Scitation*, Acoustical Society of America ASA, 1 Jan. 1970, asa.scitation.org/doi/10.1121/1.4979337.

[11] “STM32F405RGT6TR STMicroelectronics: Mouser.” *Mouser Electronics*, www.mouser.com/ProductDetail/STMicroelectronics/STM32F405RGT6TR?qs=MavH1QghrkHshpR6GkvPIg%3D%3D.

[12] Systems, Espressif. "ESP-WROOM-32 Datasheet." *Espressif Systems*, 2021, www.espressif.com/sites/default/files/documentation/esp32-wroom-32_datasheet_en.pdf.

[13] "Espressif/Esp-Idf-Provisioning-Ios." *GitHub*, Apache License, github.com/espressif/esp-idf-provisioning-ios.

[14] "LRS-50-3.3 - MEAN WELL - TRC Electronics." *TRC Electronics, Inc.*, www.trcelectronics.com/View/Mean-Well/LRS-50-3.3.shtml?gclid=CjwKCAiAkJKCBhAyEiwAKQBCkuaD7wG2k6sbGxKLME0dvwVt_WAtj05QbjtOM1IMLAnWOFLSIMvRsRoC_OwQAvD_BwE.

[15] "TPS7A2015PDQNR Texas Instruments: Mouser." *Mouser Electronics*, www.mouser.com/ProductDetail/Texas-Instruments/TPS7A2015PDQNR?qs=DPoM0jnROULcheg7K1pTg%3D%3D.

[16] Kong, et al. "Sound Event Detection and Time Frequency Segmentation from Weakly Labelled Data." *ArXiv.org*, Cornell University, arxiv.org/pdf/1804.04715.pdf.

[17] "IEEE Code of Ethics," *IEEE*. [Online]. Available: https://www.ieee.org/about/corporate/governance/p7-8.html. [Accessed: 02-Mar-2021].

[18] Schirm, Benjy. "Can Retail Stores Videotape Their Dressing Rooms? - Super Lawyers Illinois." *Super Lawyers*, Super Lawyers, 7 Feb. 2021,