NEW CORONA VIRUS DISEASE 2022: SOCIAL DISTANCING IS AN EFFECTIVE MEASURE (COVID-19) PANDEMIC

[*1]Ms. Sandhiya. R, [*2] Mrs.Bhuvaneswari R

*[*1]PG Student, Department of computer science, DKM colleges for Women (Autonomous), Vellore.*
*[*2]Assistant Professor, Department of computer science and Applications, DKM colleges for Women (Autonomous), Vellore,*

---***---

Abstract- The use of social distance as a barrier to the spread of the infectious Coronavirus Disease 2019 has been found to be successful (COVID-19). Individuals, on the other hand, are not accustomed to keeping track of the required 6-foot (2-meter) gap between themselves and their surroundings. The spread of the deadly disease could be slowed by an active surveillance system capable of identifying distances between persons and alerting them. In addition, assessing social density in a region of interest (ROI) and regulating input can reduce the likelihood of social distancing violations. Individual rights in free societies will be violated if data is collected and those who do not obey the rules are labelled. People often fail to maintain this spacing when they are in a public setting, making it difficult to implement. A Social Distancing Sensor is my solution. This system is made up of several devices that communicate with one another using both ultrasonic (US) and radio frequency (RF) signals. The devices are configured with microcontrollers such that when any two units are within six feet of each other, both units start beeping, alerting both users to take a step away from each other. My prototype exhibits successful proof-of-concept, and the devices demonstrate successful bi-directional ultrasound and radio-frequency communication, but not perfectly accurate. The technology also sends a control signal to modulate intake into the ROI if the social density exceeds a certain value. We assessed the suggested method's universality and performance using real-world datasets. The proposed approach is ready to use, and our code id open sourced

Key Words: Artificial intelligent, Infrared Sequence, Machine Learning, Covid-19, Region of interest

I. INTRODUCTION

Social isolation is a good way to combat the new Coronavirus Disease 2019 (COVID-19) epidemic. The general public, on the other hand, is not accustomed to maintaining a fictitious safety bubble around them. Individuals' perceptual abilities might be aided and enhanced by an automatic warning system. Deploying an active monitoring system like this necessitates careful ethical considerations as well as intelligent system design [5]. The first issue is one of privacy. Individuals' privacy may be infringed intentionally or unintentionally if data is recorded and retained. As a result, the system must operate in real time without the ability to save data. Second, the detector must not make any distinctions. The most secure method of accomplishing this is to constructing an artificial intelligence-based detection system It's possible that removing the human from the detecting loop isn't enough; the detector must also be design-free. Malicious designs may result from domain-specific systems with hand-crafted feature extractors. In this regard, a connectionist machine learning system, such as a deep neural network without any feature-based input space, is far more equitable, with one caveat: the training data distribution must be equitable. Another important factor is to avoid becoming invasive. The warning system should not be used to target specific individuals. To do this, a non-alarming audio-visual cue can be sent to the area surrounding the social distancing breach. It is necessary for the system to be open-sourced. This is critical for building trust between society and the active surveillance system. In light of this, we present a non-intrusive augmentative AI-based active monitoring system that sends omnidirectional visual/audio cues whenever a breach of social separation is detected[3]. To detect humans with bounding boxes in a monocular camera frame, the proposed method employs a pre-trained deep convolutional neural network (CNN). The picture detections are then converted into real-world bird's-eye view coordinates.

Create systems that allow for successful environmental interaction (rather than aiming at a particular representation) guiding an autonomous car down a road and around obstacles, recognising human motions and movements for computer control, and so on. Start with an assumption of what the system (e.g., a robot) perceives, and see if the image matches the hypothesis. Change and motion ('optical flow') are frequently more crucial than detecting shape or inferring the third dimension in dynamic vision.

Many research outcomes have been reported in recent years based on the above notions. Crowd counting has emerged as a promising research subject with numerous societal implications. By presenting alternative height homographs for head top detection and solving the occlusions problem associated with video surveillance applications, we focused on crowd detection and person count. Chen and his colleagues created an electronic advertising programme based on crowd counting. The coronavirus pandemic is a world-wide health emergency. Countries are attempting to restrict the spread of the virus by limiting travel, prohibiting big groups, and quarantining citizens in order to reduce face-to-face contact.One of the primary ways to reduce the spread of the Viruses are known to engage in social distance [1] [2]. Keeping a physical gap between yourself and others is referred to as social distancing. It's very vital to keep a distance of at least 6 feet between yourself and other individuals. In a public setting, however, it can be difficult to remember to maintain this distance, and people frequently fail to do so.

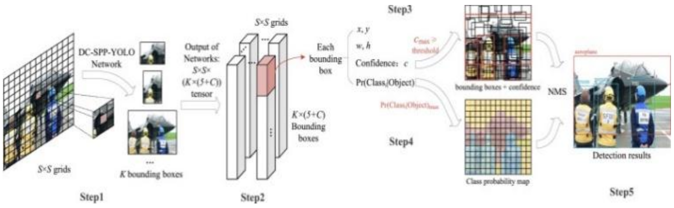

Fig1.1: Real time Data Analysis

A brief description of the proposed system. Our system operates in real time and does not save data. When an individual breach of social separation is recognised, an audio-visual cue is emitted. We also contribute something new by identifying a crucial social density value c for assessing overcrowding. This value [10] can be used to control entry into the region of interest.

II. RELATED WORK

tannguyen et.al[1] COVID-19 is spread through viral diseases, and social distance is crucial in avoiding their spread. We can lower the chances of contracting the virus and spreading it throughout the community by limiting close physical contact between people. This two-part study intends to present a complete overview of how developing technologies, such as wireless and networking, as well as artificial intelligence (AI), can facilitate, encourage, and even impose the practise of social distancing.

Marlena Robakowska et.al[2] The law on mass event security has defined the idea of a mass event. Drones can be used to monitor events for a variety of reasons. Drone surveillance is very effective when big open areas are involved. Unmanned aerial vehicles (UAVs) can detect problems in a crowd.

Sebastian Brutzer et.al[3] To address this problem, we must first identify the major obstacles to background subtraction in video surveillance. Then, based on their ability to satisfy those obstacles, we compare the performance of nine background subtraction approaches with post-processing. As a result, a new evaluation data set with accurate ground truth annotations and shadow masks is introduced. This allows us to provide a detailed analysis of the benefits and downsides of various background subtraction approaches.

Joseph Redmon et.al[4] It could use a few design tweaks to make it better. We also got this new network up and running, which is quite cool. It's a little bigger than the last one, but it's more precise. Don't worry, it'll still be quick. @ 320 320, YOLOv3 operates in 22 ms at 28.2 mAP, which is three times faster than SSD. When compared to the old.5 IOU mAP detection metric, YOLOv3 performs admirably. On a Titan X, it achieves 57:9 AP50 in 51 ms, compared to RetinaNet's 57:5 AP50 in 198 ms, for a same performance but 3.8 quicker.

[5] Zhengxia Zou et al. A number of topics are discussed. The historical milestone detectors, detection datasets, metrics, essential building elements of the detection system, speed-up strategies, and modern state-of-the-art detection methods are all covered in this study. This study also examines certain key detection applications, such as pedestrian detection, face detection, text detection, and so on, and examines their limitations as well as recent technological advancements.

guruh fajar shidik et.al [6] As application-oriented studies, video surveillance systems have attracted a lot of attention during the last decade. Recent research aims to include computer vision, image processing, and artificial intelligence into video surveillance applications. Despite the fact that there have been many advances in the acquisition of datasets, methodologies, and frameworks, there are few papers that provide a comprehensive picture of the present status of video surveillance system research.

III. PREVIOUS IMPLEMENTATIONS

Pedestrian detection can be thought of as either a subset of a larger item detection problem or as a separate task dedicated just to detecting pedestrians. In, you'll find a thorough examination of 2D object detectors, as well as datasets, metrics, and basics. Deep learning techniques for both generic item detection and pedestrian detection are the subject of another study[8]. Deep learning techniques, which are commonly classified into two types, are used by state-of-the-art object detectors. The first are two-stage detectors, which are generally based on R-CNN. These techniques' accuracy and real-time performance are sufficient for deploying pre-trained models for social distancing detection [6].

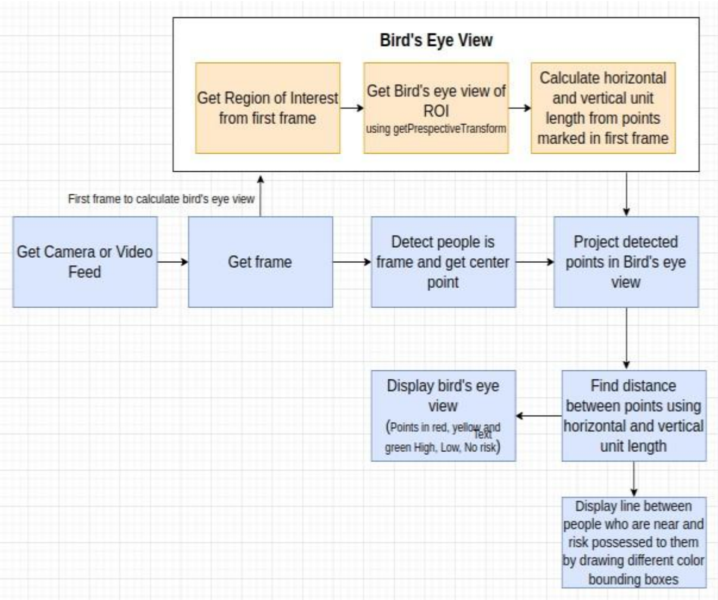

Fig1.2: Birds eye View

The first stage is to figure out how and in what format the output will be produced. There are examples of output and input. Second, input data and master files must be designed to meet the output's specifications. Program construction and testing, as well as a list of programmes required to satisfy the system's objectives and thorough documentation, are used to handle the operational phases.

3.1. Camera Perspective Transformation or Camera Calibration:

Because the input video can be from any perspective, the initial step is to convert the perspective to a bird's-eye (top-down) view. The simplest transformation approach is selecting four points in the perspective view that define ROI where we want to monitor social distancing and mapping them to the corners of a rectangle in the bird's-eye view because the input frames are monocular (collected from a single camera). In the real world, these points should create parallel lines when viewed from above (bird's eye view). This presupposes that everyone is standing on the same level ground. The points are dispersed uniformly horizontally and vertically in this top view or bird's eye view (scale for horizontal and vertical).From this mapping, we can derive a transformation that can be applied to the entire perspective image

3.2. Region of Interest (ROI)

Using the mouse click event, we draw 8 points on the first frame. The first four elements will determine where we want to track social distancing in terms of ROI. The next three locations will define a horizontal and vertical distance of 180 cm (unit length), and they should form parallel lines with ROI. Points 5 and 6 define 180 cm in real life in the horizontal direction, while points 5 and 7 define 180 cm in real life in the vertical direction [12]. Because the ROI generated by the initial four points has different lengths in the horizontal and vertical directions, the number of pixels in the 180 cm rectangle (bird's eye view) formed after transformation will be variable.**3.3. Detection**

The next stage is to identify pedestrians and draw a bounding box around each one. To reduce the risk of overfitting, we use minimum post-processing techniques including non-max suppression (NMS) and several rule-based heuristics to clean up the output bounding boxes.

3.4. Distance Calculation

Each individual in the frame now has a bounding box. We need to figure out where people are in the frame. i.e., we can use the bottom centre point of the bounding box as the location of the person in the frame. Then, by applying transformation to the bottom centre point of each person's bounding box, we estimate (x,y) location in bird's eye view, resulting in their position in the bird's eye view. The next step is to calculate the bird's eye view distance between each pair of persons and scale the distances in both horizontal and vertical directions using the scaling factor calculated via calibration[13].

3.5. YOLOV3

YOLOv3 is the latest version of the popular YOLO – You Only Look Once – object detection method. The disclosed model can recognise 80 different items in photos and videos, but it is also much faster and more accurate than Single Shot MultiBox (SSD). It begins by dividing the image into a 1313 cell grid [14]. The size of these 169 cells is determined by the input size. The cell size was 3232 for a 416 416 input size that we used in our studies. After that, each cell is in charge of estimating a certain number of boxes in the image. The network also predicts the probability of the enclosed object being a human and the confidence that the bounding box truly encloses an object for each bounding box.. For each bounding box, the network also predicts the confidence that the bounding box actually encloses an object, and the probability of the enclosed object being a particular class. Most of these bounding boxes are eliminated because their confidence is low or because they are enclosing the same object as another bounding box with very high confidence score. This technique is called non- maximum suppression.

V. IMPLEMENTATION OF SOCIAL DISTANCE KIT

Emerging trends and the availability of clever technologies force us to create new models to meet the needs of the developing globe. As a result, we've created a revolutionary social distancing detector that could help with public healthcare. The model presents a real-time deep learning-based framework for automating the process of monitoring social distancing via object identification and tracking methodologies, in which each participant is identified in real-time using bounding boxes. Using the Bird's eye view approach, identifying clusters or groups of people who satisfy the proximity property. The number of violations is determined by calculating the number of groups established and the violation index term, which is calculated as the ratio of the number ofThe extensive trials were conducted with popular state-of-the-art object detection models Faster RCNN, SSD, and YOLO v3, since this approach is highly sensitive to the spatial location of the camera, the same approach can be fine-tuned to better adjust with the corresponding field of view.

4.1 Input Design

The link between the information system and the user is the input design. It entails creating data preparation specifications and procedures, as well as the steps required to convert transaction data into a usable format for processing. This can be accomplished by inspecting the computer to read data from a written or printed document, or by having people key the data into the system directly. Limiting the amount of input required, controlling errors, avoiding delays, avoiding superfluous stages, and making the process simple are all goals of input design. The input is created in such a way that it gives security and convenience while maintaining privacy. The following factors were taken into account by Input Design:

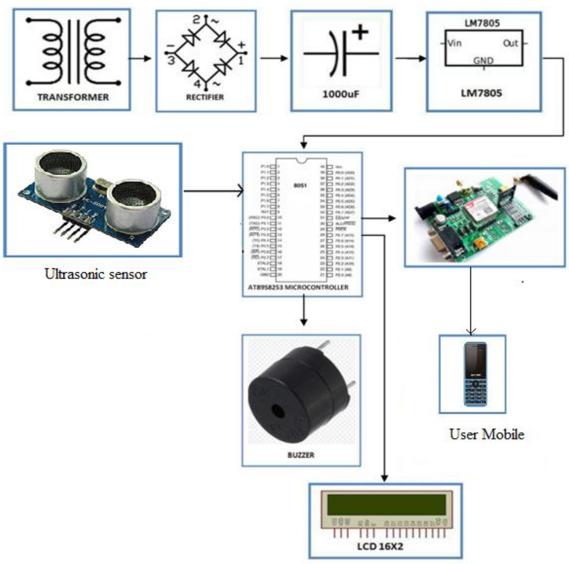

Fig 1.3 Block Diagram

4.2 System Architecture

A brief description of the proposed system. Our system operates in real time and does not save data. When an individual breach of social separation is recognised, an audio-visual cue is emitted. We also contribute something new by identifying a crucial social density value c for assessing overcrowding. This variable can be used to control the amount of time it takes to enter the region of interest. Wireless technology combines to allow people to communicate with one another [15]. In a heterogeneous wireless environment, a mobility management scheme is necessary for location management and handoff management. The ability to locate the mobile node for call delivery is enabled by location management. When a mobile node travels from one coverage area to another, a handover strategy is necessary to provide mobility management.

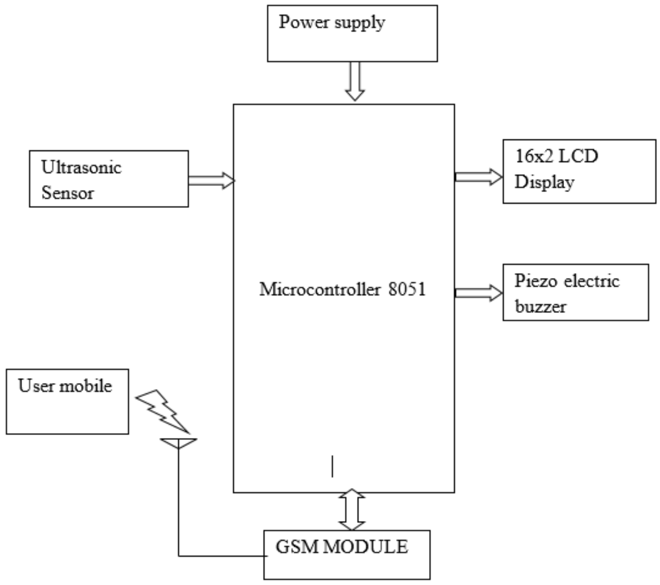

Fig1.4: System Architecture

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

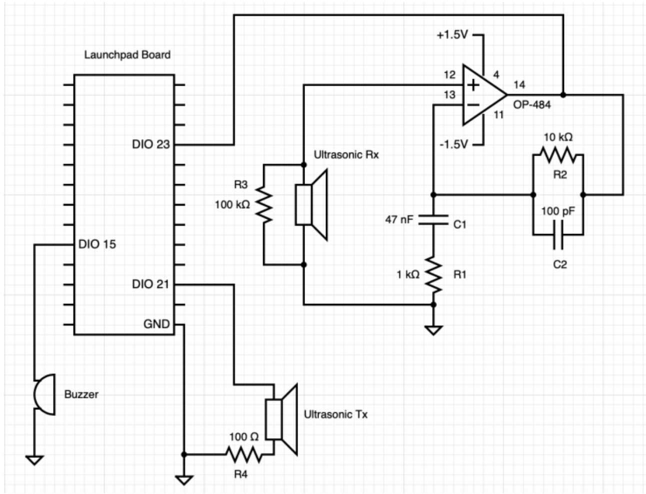

Fig1.5: circuit Diagram

The US transmitter receives a 1 ms pulse-width-modulated signal burst from the microcontroller (through DIO 21). After that, the US transmitter turns the electric signal into sound waves and broadcasts it (which are then picked up a US receiver on another device).

At DIO 23, the microcontroller additionally uses ADC sampling to read voltages from the ultrasound receiver. Sound waves are received by the ultrasound receiver, which are then converted to electrical impulses (voltages). To detect when a US signal is received, the microcontroller samples the signal at the US receiver at discrete intervals. The microcontroller, on the other hand, isn't directly connected to the output of the US receiver.

A. Micro controller

Instead of using an independent microcontroller chip, I elected to utilize a Launchpad development kit board (which incorporates a microcontroller chip) because it has numerous general input/output pins that make it easier to interface with. The device also includes built-in radio-frequency communication capabilities [3], which I needed for my project [9]. It would have been necessary to construct a separate radio antenna and receiver without this built-in capabilities. Although this would have been possible, I chose to abstract this component and use existing RF technology so that I could concentrate on the ultrasonic capabilities and microcontroller programming.

B. Ultrasonic sensor

An ultrasonic sensor for garbage bins is a device that measures the amount of rubbish in the bin. It works out distance by sending out a sound wave at a certain frequency and waiting for it to return. Sound waves travel through the air and are time and distance measured.

C. Piezo electric buzzer

An audio signalling device is a buzzer. A buzzer is a loudspeaker that generates sound via the piezoelectric effect. Applying a voltage to a piezoelectric material causes mechanical motion, which is then translated into audible sound using diaphragms and resonators.

LCD display (D)

A liquid crystal display creates a visible image by using a liquid crystal. It connects to a microcontroller. The image or message is shown using LCD technology. It shows the output of the microcontroller.

VI. EVALUATION RESULT:

One of my prototype's flaws is that the US and RF signals from device 1 do not always transmit at the same time. Due to the Launchpad's restrictions, it appears that it cannot do both RF and US capabilities at the same time. As a result, I attempted to slightly delay the RF transmission such that it would occur immediately after the PWM code that transmits the US message. However, it appears that the hardware automatically sets a 0.25 ms delay after the US signal transmission, regardless of how little a delay I programme (which is equivalent to a 1.25 ms delay from the start of the US signal transmission since the US burst is 1 ms in duration).This discrepancy in the US signal and RF signal transmission start times from device 1 leads to some complications, and this is discussed in the next section Inaccurate ADC Sampling Time Block due to Hardware Limitations.

With device 2, this isn't an issue (which is programmed with rfEchoRxFinal). This is due to the fact that in device 2, the Rx command comes first, followed by the Tx command. Because it is the end of the programme, the US signal transmission can occur immediately after the RF packet has been repeated back (the Tx command). In device 1 (which is programmed with rfEchoTxFinal), on the other hand, the Tx command comes first, followed by the Rx command.

Table: Kit Approximate part and cost

Part	Approximate Total Cost (# of units x unit cost)
100 ⌧ resistor*	x 0.10 = 0.20
1 k⌧ resistor*	x 0.10 = 0.20
10 k⌧ resistor*	x 0.10 = 0.20
100 k⌧ resistor*	x0.10 = 0.20
100 pF capacitor*	x 0.20 = 0.40
47 nF capacitor*	x 0.20 = 0.40
OP484FPZ op-amp*	x 12.47 = 24.94
6" male-to-female*, male- to-male, & female-to-female jumper wires	packs x 1.95 = 5.85
22 AWG Solid Hook Up Wire	spools x 2.50 = 5.00
Solderless Breadboard*	x 2.93 = 5.86
Piezo Buzzer	x 2.13 = 4.26
AA Alkaline 1.5 V Batteries	x 0.325 = 1.30
AA Battery Holder w/ 6" Leads	x 1.18 = 2.36
Ultrasonic Transmitter	x 5.13 = 10.26
Ultrasonic Receiver	x 7.42 = 14.84

- • Diligent Analog Parts Kit (P/N: 240-000): I used an analogue parts kit that comprised many fundamental electronic building blocks such as resistors, capacitors, op-amps, and jumper wires because I was working remotely and didn't have access to an electronic lab.
- • CUI Devices Piezo Buzzers (P/N: CPI-2212-85PM): These buzzers are powered by a DC voltage and are used to alert users when something is wrong.
- • Prowave Ultrasonic Transmitters (P/N: 400ST100) and Receivers (P/N: 400SR120): These transducers alter the medium through which a signal is sent (between electrical and acoustic). These were chosen based on specifications such as beam angle, transmitting sound pressure level (for the transmitter), and receiving sensitivity (for the receiver). The central frequency of both the transmitter and receiver is 40 +/- 1 kHz.
- • The waveguide is an S-band WR-284 with dimensions of a = 2.84 and b = 1.37. The design is for a frequency of 3.4 GHz.
- One broadwall has 10 elliptical slots with a slot length of 0.98mm. One end of the waveguide is shorted, while the other is fed.

Main Process

- λ = 1.088, λ guide= 2.393, λ freespace = 2.84, λ cut-off = 0.514 (INTO dB= 6 dB)
- The centre of the first slot, Slot 1, is placed at a distance of quarter guide wavelength ($\lambda g/4$), or $2\lambda g/4$, from the waveguide feed.
- The center of the last slot, Slot 10, is placed at $3\lambda g/4$, from the waveguide SLR.
- The distance between the centers of two consecutive slots is $\lambda g/\mathbf{2}$

5.2 Tools and Materials

• AstroAI Digital Millimeter: I primarily utilised this to determine resistance values.

• WaveForms Software + Digilent Analog Discovery 2 (AD2): I used the AD2 as an oscilloscope, function generator, and power supply. Wave Forms, its companion programme, enabling me to create and visualise signals on my laptop.

• USB Isolators: When powering the microcontrollers and AD2 from my laptop, I used these to safeguard my laptop from surge attacks and short circuits.

• Digilent Power Supply: I needed this to power the AD2 and guarantee that it received adequate power, as the USB isolator only allowed the AD2 to get a limited amount of power.

5.3 Prototype Setup

This prototype comprises of two devices that communicate via radio frequency and ultrasonic signals. Each device consists of a Launchpad board (which houses a microcontroller), an ultrasound transmitter that sends sound waves to other devices, an ultrasound receiver that receives sound waves from other devices, a band pass filter that amplifies and filters the output of the US receiver, and a buzzer that beeps when triggered (if the detected sound is loud enough).

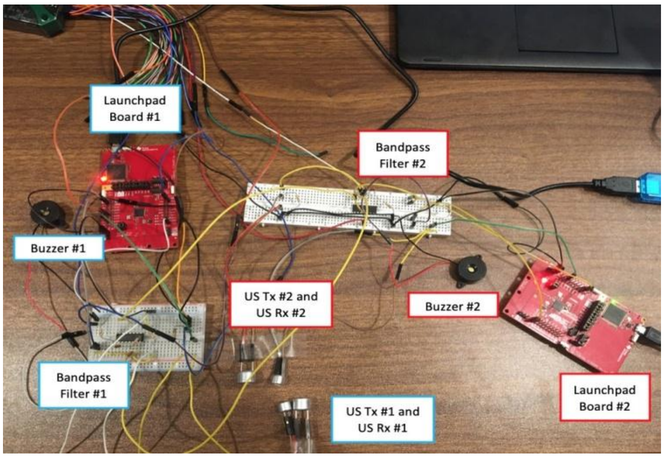

Fig1.6: Prototype Model

5.4 High-Level Block-Diagram

Below is a high-level block diagram that shows the path of communication between two units. A lot of details have been abstracted out in this simplified model, but this diagram gives a top-level overview of how two units communicate to verify the distance between them.

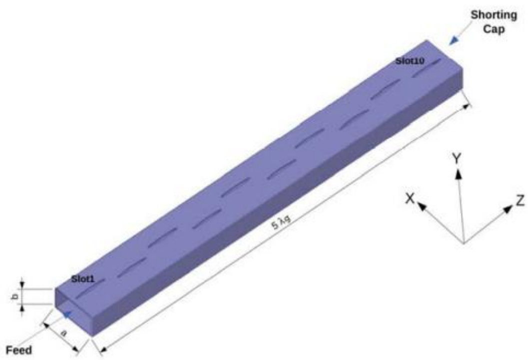

Fig1.7: cut-off wavelength

- • Device 1 enters transmission mode and simultaneously transmits a radio-frequency signal and an ultrasonic signal. To transmit or receive RF signals, the LaunchPad boards include the ability to switch between transmission and reception mode built-in.
- • Device 2, which starts in reception mode, receives the RF packet and the US signal (at different times because the RF signal travels at the speed of light and the US signal travels at the speed of sound), reads the voltages at its US receiver, verifies the distance between the two devices, and beeps its buzzer if the calculated distance is within 6 feet.
- • Device 2 enters transmission mode, echoes the RF packet, and simultaneously transmits a new US signal. necessary is because we want both devices to beep if they are in close proximity. If we stop at Step 2: Only the buzzer on device 2 will be triggered, and device 1 will have no means of knowing how far apart the two units are.
- Device 1 receives the RF packet and US signal, reads the voltages at its US receiver, validates the distance between the two devices, and beeps its buzzer if the determined distance is within 6 feet.

5.5Validation Result

A 10*10 SWA KIT has been created. When operational, the classic rectangular holes utilised in the traditional SWA design may exacerbate electrical breakdown problems. As a result, minimising sharp corners in slots provides better high-power operating and easier manufacturing than rectangular slots used in this study. To get the values of the displacement and length for each slot, the Chebyshev and Tylor methods are solved simultaneously. To offer a simplified approach for determining the slots non-uniform displacements using closed form equations for a chosen SLL. The results of designing, simulating, and measuring a prototype SWA with 10*10 running at a frequency of 3.4 GHz are presented.Slotted waveguide antenna arrays are used in radar, communication and remote sensing systems for high frequencies. They have linear polarization with low cross-polarization and low losses but can also be designed for dual polarizations and phase steered beams.

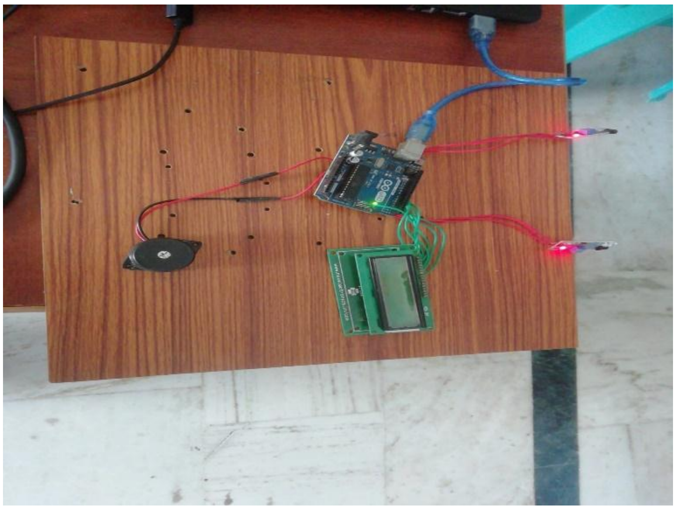

Fig 1.8: Main Board

A slot antenna consists of a metal surface, usually a flat plate, with one or more holes or slots cut out. When the plate is driven as an antenna by an applied radio frequency current, the slot radiates electromagnetic waves in a way similar to a dipole antenna.

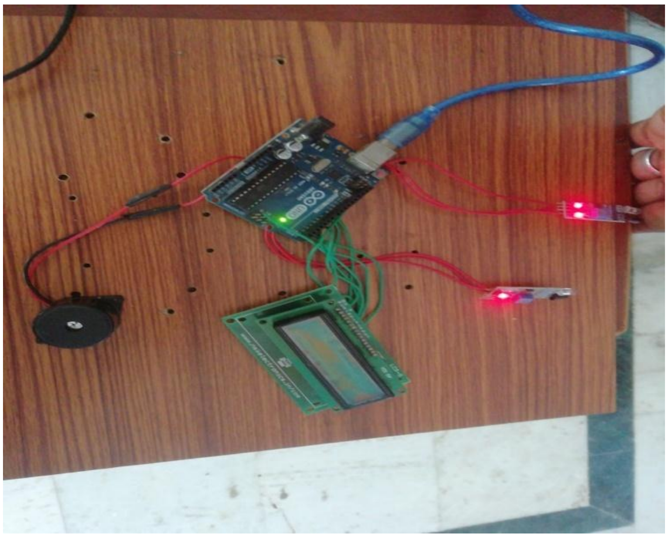

Fig1.9 : Sensor Activated

The distance between two equal phase planes along the waveguide is defined as the guide wavelength. It is derived using the equation below as a function of the operating wavelength (or frequency) and the lower cut-off wavelength.

CONCLUSION

The most difficult aspect of this project for me was programming the microcontrollers. I spent a lot of time completing example projects and reading documentation on the Texas Instruments website to learn about the Launchpads because these were new devices to me. Furthermore, despite my prior knowledge with software programming, I have discovered that hardware programming is rather different. Individual code components would frequently work, but when they were combined, there would be a functional issue since one portion of code might have been blocking another. Instead of blindly creating code, a grasp of the hardware (for example, comprehending the RAM memory while troubleshooting the ADC component) was required. Working remotely was also a difficulty for me. Because I was working from home, I didn't have access to an electronics lab's resources. For instance, I utilised the AD2 as an oscilloscope, function generator, and power supply, and it frequently did all three at the same time. Although everything worked flawlessly, it became clumsy and inconvenient at times. In addition, I had to be careful when ordering parts and tools to ensure that I had everything I needed to finish the project while also not going over budget.

REFERENCES:

[1] S. Bell, C. L. Zitnick, K. Bala, and R. Girshick. Insideoutside net: Detecting objects in context with skip pooling and recurrent neural networks. arXiv preprint arXiv:1512.04143, 2015. 6

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. FeiFei. Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 248–255. IEEE, 2009. 1

[3] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. International journal of computer vision, 88(2):303– 338, 2010. 1

[4] P. F. Felzenszwalb, R. B. Girshick, and D. McAllester. Discriminatively trained deformable part models, release 4. http://people.cs.uchicago.edu/ pff/latent-release4/. 8

[5] R. B. Girshick. Fast R-CNN. CoRR, abs/1504.08083, 2015. 5, 6

[6] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learningforimagerecognition. arXivpreprintarXiv:1512.03385, 2015. 2, 5, 6

[7] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015. 2, 5

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012. 2

[9] M. Lin, Q. Chen, and S. Yan. Network in network. arXiv preprint arXiv:1312.4400, 2013. 4

[10] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll´ar, and C. L. Zitnick. Microsoft coco: Common objects in context. In European Conference on Computer Vision, pages 740–755. Springer, 2014. 1, 6

[11] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. E. Reed. SSD: single shot multibox detector. CoRR, abs/1512.02325, 2015. 5, 6

[12] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller. Introductiontowordnet: Anon-linelexicaldatabase. International journal of lexicography, 3(4):235–244, 1990. 6

[13] J. Redmon. Darknet: Open source neural networks in c. http://pjreddie.com/darknet/, 2013–2016. 5

[14] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640, 2015. 5, 6

[15] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. arXiv preprint arXiv:1506.01497, 2015. 2, 3, 5, 6