# Unmanned Aerial Vehicle for Search, Evacuation, and Rescue to Detect Human and Animal Using Deep CNN Detectors

By:

Alden Evangelista

Jewel Kate D. Lagman

ECE/3

A Research Proposal Submitted to the School of ECE in Partial Fulfilment of the Requirements for the Degree

Bachelor of Science in Electronics Engineering

Mapua University

June 2021

**TABLE OF CONTENTS**

## LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1

## INTRODUCTION

The Philippines is one of the world's most vulnerable countries to climate impacts and natural disasters like typhoons, repeated earthquakes, volcanic eruptions as well as sea-level rise, storm surges, landslides, floods, and drought are prone due to the country's location. High population, illegal housing and environmental degradation are some factors why the NDRRMC find it difficult to respond quickly. A professional operation provided by a group of specially trained personnel who rescue and care for wounded in risky conditions is known as search and rescue. Search and rescue operations are carried out in close collaboration with the community and as a team effort. When a community is affected by natural disasters the rescue operation takes place to save people and animals that may be trapped by falling debris. Assistance from outside the community is also required when local rescues lack the required efficiency and equipment. Unmanned Aerial Vehicles (UAVs) can help with search and rescue operations due to being fast and capable of autonomous behavior, allowing them to carry out actions that would be difficult for humans to carry out at low cost and can protect rescuers' lives out of extreme danger. It can also explore the area and collect evidence about the whereabouts of a missing person or animal, specifically a dog. One of the most powerful storms to hit the Philippines is Typhoon Yolanda, wherein the Government and the locals are not prepared for this disaster which causes a huge destruction and casualties. It is hard for the rescuers to search due to no electricity, light and lack of manpower; they can only check limited areas to rescue people but with the use of UAV that can detect humans and animals, as well as a thermal camera that is more accurate than a regular camera, this UAV can scan a larger area in a specific places, has high flight altitude, that increased the chances of finding a person or animal and shortened the search time even in dark situations.

The study uses object detectors to identify people in videos taken by a drone in non-urban areas during search and rescue operations. The SARD dataset was used for transfer learning and fine-tuning for person detection in search and rescue scenes using the following state-of-the-art person detectors: Faster R-CNN, YOLOv4, Retina Net, and Cascade R-CNN. The drone flew at various altitudes (ranging from 5 to 50 meters) and camera angles (ranging from 45 to 90 degrees). In search and rescue situations, the SARD database was created to identify casualties and people in drone photos and videos. The main object in SAR operations is the human, it is recorded from a bird's eye view, and such records are not included in the large data sets on which these state-of-the-art person detectors are trained [1]. Another study identifies a person in a low-resolution image captured by webcam or in various frames of CCTV footage by using deep learning convolutional neural networks (CNN). The CNN model is designed and trained on these images from the database. The trained model is tested for its performance using separate test images from the database as well as test images extracted from low resolution CCTV cameras. The presented work can be used for surveillance and criminal identification in CCTV footage. However, the recognition of a person in low resolution cameras like CCTV is still a challenging problem [2].

From the previous research they are struggling when it comes to a detailed examination of a large amount of recorded material using a DJI Phantom 4A drone with one camera that can only detect a person [1]. The other research shows that the use of a CCTV camera is still an issue; the reason for this is that it is not in high resolution and has a limited viewing area [2]. To improve these problems, in our research, unmanned aerial vehicles (UAVs) or drones will be used, and two cameras will be employed: one for the human and animal detection specifically dogs and the other one is the thermal camera since they have the ability to identify the missing human or animal as well as other people in the immediate vicinity and improve detection accuracy. When the drone detects a person or an animal, it will show up on the FPV screen and identify the object with the use of YOLOv5. It will also capture any human or animal that is recognized. People counters will be added wherein the number of persons detected for the live person count (LPC) and the overall person count (OPC) will be recorded, and the data gathered will reveal the total number of individuals detected on that day. The LPC stands for the current number of people in the frame, while the OPC stands for the total number of people who have been detected so far. The videos taken by the drones are automatically transferred and stored to electronic devices such as phones, tablets, and laptops.

The main objective of this study is to develop a system that automatically detects a human and animal for search, rescue, and evacuation using deep CNN detectors. Specifically, this study aims (1) to create a prototype that automatically detects a person and animal specifically dogs using a deep CNN detector, (2) to develop a GUI/software for search, evacuation, and rescue, (3) to test the performance of the device in automatic detection of person and animals.

This study will benefit rescuers as it helps them to react and respond much better in emergency situations such as natural disasters, fire rescue, and search and rescue because the rescuers will know what your situation is and the drone can count total people who are stranded in that certain area. In addition, it will also help them to find people and animals who are stranded because the drone is equipped with a thermal camera. It will also benefit Filipino people and animals because they will be much easier to be rescued with the use of this technology.

This study will focus on creating a prototype that automatically detects a person and animal by a UAV and to confirm the detected object with the use of a thermal camera. This automatic human and animal detection will be tested during good weather conditions. Also, it will only be tested in Barangay Tibig, Lipa City Batangas. It will not be tested in any other areas. In terms of design implementation, only dogs will be detected by the prototype. Other animals such as cats, birds, rabbits etc. will not be detected by the prototype. For testing the performance, it will be simulated through a bird's eye view in evaluating the accuracy of the prototype.

# CHAPTER 2

# REVIEW OF RELATED LITERATURE

A variety of materials were used in the process of this study. This chapter describes the main contributions and properties of these resources. The supporting literature for the study is discussed in this chapter.

## *2.1 Person Detection*

Automatic person detection methods track the person in the video in real-time, location and movement direction, this could be extremely helpful to operators. The paper describes a method for detecting people using their faces extracted from a video frame. The authors present a method for detecting criminals using the Viola Jones algorithm in order to perform face matching operations on video surveillance frames [3]. The use of convolutional neural network-based models for automatic human detection in thermal images, which were originally designed for RGB image detection, is studied. With a limited training set and the standard YOLOv3 model, the study shows great results in terms of average precision for all examined situations, as well as a significant improvement in performance for human detection in thermal imaging [4].

### *2.1.1 Overhead View*

Detecting the region of interest and localizing the individual in the input image and video sequence are the steps in locating a person in top view photos and videos. In overhead, the individual in the image might be identified by their head, head-shoulder, or the entire overhead body, as shown in Figure 2.1. From an overhead view, it is easy to see how the person's shape, size, and body orientation vary. It's important to understand how people are discovered in overhead photos and videos in computer vision. In the paper, the use of blob-based algorithms for detecting people in overhead views is applied. Background subtraction, foreground extraction, segmentation, and pre-processing procedures are among the approaches used. A foreground image is usually obtained through background subtraction in blob-based techniques [5].

*Figure 2.1 Overhead View [5]*

### *2.2 Animal Detection*

Manual observation has typically been used to detect animals, which requires the use of skilled professionals and a significant amount of time. The use of deep convolutional neural networks is essential to modern research in the field of detecting animals on images and videos as shown in Figure 2.2. Animal detection and classification can assist in the prevention of animal-vehicle accidents, the tracking of animals, and the prevention of theft [6]. Animal detection on photos obtained by unmanned aerial vehicles is the focus of some research wherein they present a Convolutional Neural Network-based pipeline for animal (object) detection (CNNs) and enable repeated acquisitions at sub-decimeter resolution over large areas. The samples are sorted according to their likelihood of being animals using the CNN scores in the source data set, and this ranking is transferred to the target data set [7].

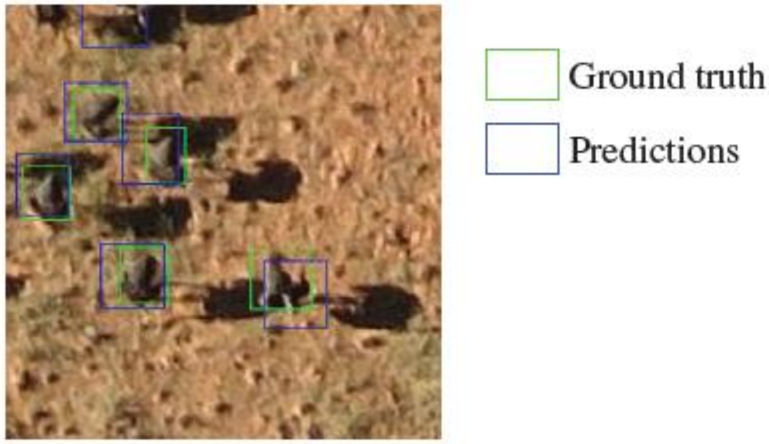


*Figure 2.2 Example of Result on Animal Detection [6]*

This study intends to create an algorithm to recognize animals in the wild. It can be challenging to manually identify all the different species because there are so many. This system classifies animals based on their pictures so that we can better monitor them. In this paper they have discussed the feature-based template matching technique using NCC. A method for detecting small parts of an image that should match the template image is known as template matching [8]. Domestic dogs have been living with humans for at least 15,000 years and it is widely distributed in continents and on most islands. In the Philippines, there are approximately 11.6 million dogs including the owned and stray dogs [9].

### *2.3 Search, Evacuation and Rescue Operation*

The use of autonomous UAVs to survey the area and collect evidence about the whereabouts of a missing person can substantially assist search and rescue activities. The primary goal of SAR operations in the wilderness is to locate a missing person. The main challenge is to use professional knowledge to determine the search area where the person might be found [10]. It is essential to properly use information from drone sensors in order to maximize its potential. This requires the development of a system capable of detecting people in the gathered images and videos. Person detection in aerial images captured by drones during search and rescue missions is one of the most difficult tasks in object detection since it involves a number of issues [11].

### *2.3.1 Deep CNN Detectors (Deep Convolutional Neural Networks Detectors)*

A neural network that is made up of multiple layers, including an input layer, at least one hidden layer, and an output layer, as shown in Figure 2.3. These are best for detecting patterns like edges (vertical/horizontal), forms, colours, and textures in object detection [12]. The backbone of most object detectors is a CNN network that was trained to extract features, and the head predicts the class and boundary box of the detected objects. In terms of localization and classification accuracy, two-stage detectors are typically more accurate [1]. Object detection locates a specific object from a given image; several network models with high detection performance have become the mainstream. This is a significant improvement in performance compared with conventional image recognition (utilizing engineer-designed feature quantities and identification logic). These new models are created for the purpose of general object detection, including “human” detection. Following the detection of areas or sub-images in which an individual is present, these areas or sub-images are submitted to a CNN classifier [13][14].

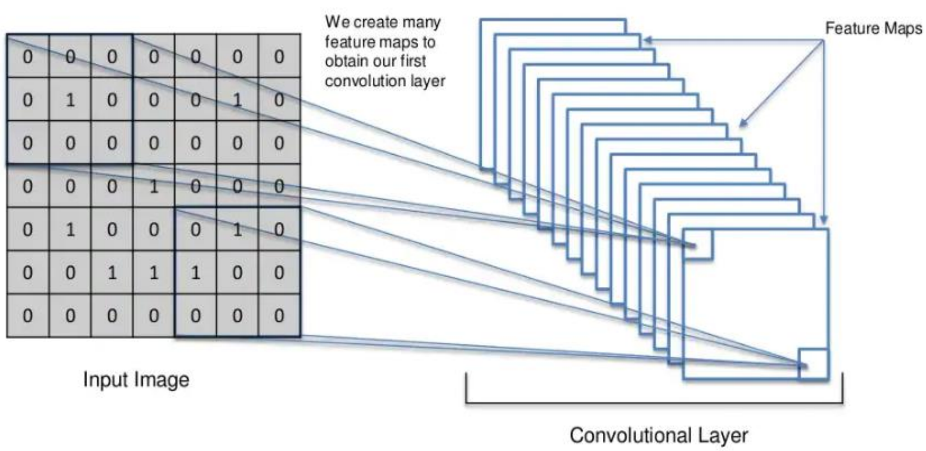


*Figure 2.3 Convolutional Layer*

*2.3.2 YOLOv5*

Yolov5 is a region-based object detection network with a one-stage detector. The Yolo classifies object detection as a regression problem with a high processing speed [15]. The backbone, head, and detection are the three fundamental components of YoloV5, as shown in Figure 2.4. A CNN serves as the backbone, gathering and shaping image features at various levels of specificity. To develop image features, the YoloV5 uses the CSP (Center and Scale Prediction) Bottleneck. The head is made up of layers that integrate image features before being sent into a prediction algorithm. The PA-NET is also used by YoloV5 for feature aggregation. The detection is a process that takes box and class prediction steps and uses features from the head [16].

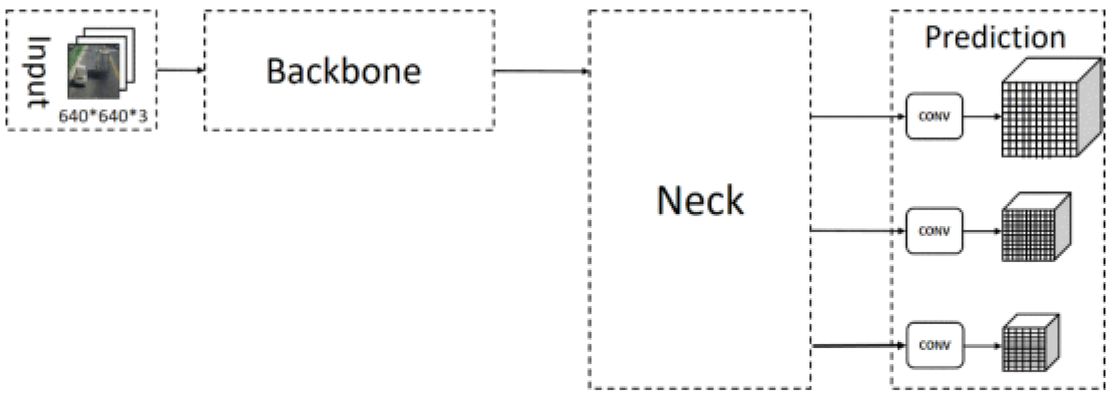


*Figure 2.4 YOLOv5*

The YOLOv5 Network Model is shown below in Figure 2.5, Ultralytics released YOLOv5 in May of this year, and it outperforms the previous YOLOv3 and YOLOv5 algorithms on the COCO dataset in terms of detection speed and precision. The weight data for the YOLOv5 version is 27M, which is 1/9 of the weight data for the YOLOv4 version, allowing for a faster detection speed. It can process a picture in 0.007 seconds, allowing real-time detection. YOLOv5 is more effective at identifying large targets, but due to the detection scale's limits in terms of compact organization and several overlapping small targets, it's easier to miss detection [17].

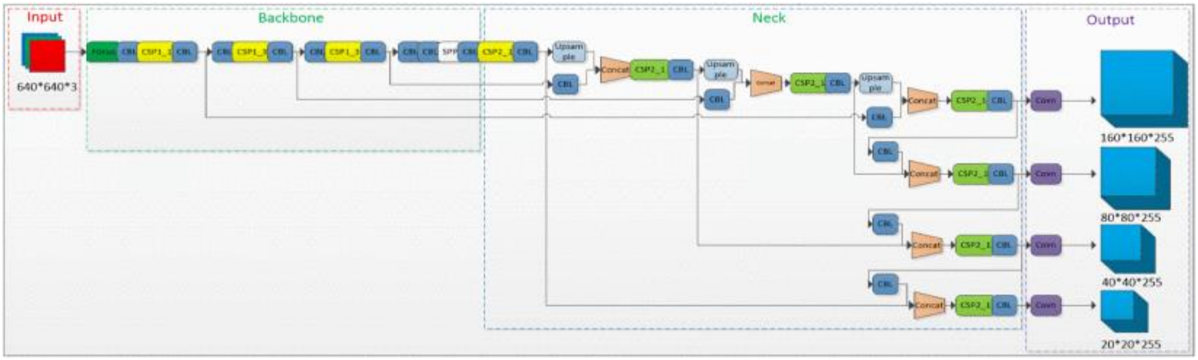


*Figure 2.5 YOLOv5 Network Model*

### *2.3.3 People Counting*

Many applications involve identifying people and the accurate counting of people in real time or near real time. A reliable system for people detection is required for people counting. When a person enters a supervised area, the trajectories are generated, and people are counted directionally based on their trajectories. The position and orientation of the camera is a key feature of all counting systems. These characteristics are examined, and it is discovered that the camera's position allows for a better view of the person's head and body [18].

### *2.3.4 Image Processing*

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms a core research area within engineering and computer science disciplines too [19].

There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital techniques are pre-processing, enhancement, and display, information extraction [20].

### *2.4 Drones/UAV (Unmanned Aerial Vehicle)*

Unmanned aerial vehicle (UAV) or also known as drones is a type of plane which could land and take off deprived of company of the onboard pilot. It is composed of aircraft components, sensors, payloads, and a human that facilitates to control of the Unmanned aerial vehicle. Furthermore, UAVs require dedicated control systems to control UAV with limited range or communication capabilities [21]. In addition, the majority of UAVs now are integrated with a camera which can record videos and capture images which can transfer the data in real-time with the use of Bluetooth to other electronic devices such as Laptops, Personal Computers, and mobile phones [22].

Due to its ease of operation, low maintenance costs, great mobility, and capability to move around, UAVs can be applied in a variety of civil applications. These UAVs are usually being used for real-time traffic monitoring, wireless coverage, remote sensing, search and rescue operations, products delivery, security and surveillance, precision agriculture, and civil infrastructure inspection, among other things [23]. Drones can carry a variety of technology, including live-feed video cameras, infrared cameras, heat sensors, and radar, and are capable of extremely complex surveillance. Some military versions can stay in the air for hours or even days at a time, using high-tech cameras that can scan entire towns or zoom in from 60,000 feet [24].

Search and rescue operations can benefit greatly from the use of unmanned aerial vehicles (UAVs). UAVs are agile, quick, and capable of autonomous behavior, allowing them to carry out actions that would be difficult for humans to carry out at minimal costs. In a typical situation, UAVs will be deployed in a target region, execute sensory operations to collect evidence of a victim's presence, then communicate their findings to a ground station or rescue team. UAVs have already proven their worth in search and rescue missions, assisting responders in focusing their efforts while avoiding risks [25]. Figure 2.6 shows the different angles taken by drones in a search and rescue operation.

*Figure 2.6 Different Shots taken by drone in Search and Rescue Operation*

### *2.4.1 Thermal Camera*

A thermal camera is a device that uses infrared (IR) radiation to make an image. The main feature of thermal cameras is that it can detect motion as well as the heat that radiates in an object [26]. The thermal camera does not work like the usual cameras that capture visible light which the human eyes can see. It captures images from heat which is also called infrared/thermal energy. It composes a lens, thermal sensor, processing electronics, and a mechanical housing. The infrared energy is focused onto the sensor by the lens [27].

In order to create an image, thermal cameras capture electromagnetic radiation. The IR wave band which is included in the elements of the electromagnetic spectrum is classified into subregions which are Short-wave infrared (1.4-3µm), Mid-wave infrared (3-8µm), Long-wave infrared (8-15µm), and Far-wave infrared (15-1000 µm). The Mid-wave infrared (MWIR) is commonly referred to as thermal. However, it is virtually unusable because of the high spectral absorption of the atmosphere in this range. So, long-wave infrared (LWIR) is the wave band that a thermal camera will work in the spectrum of light. Images are created by operating the thermal IR spectrum. Because the surrounding light intensity is unnecessary, this performs well even in complete darkness. It makes thermal imaging possible because all of the objects which are organic, or inorganic emit a certain amount of IR radiation as a function of temperature. This applies to all objects having a temperature greater than absolute zero, or 0 K (273.3°C or 459°F) which means that even very cold objects like ice still emit thermal radiation. The sensitivity of a camera can be defined as its capacity to detect temperature differences. The thermal images will be clearer if there is a great temperature difference in the situation. However, the emissivity of the objects in a thermal image affects the contrasts. Figure 2.7 shows the images in both the spectrum i.e., visible and IR. The purpose of introducing other sensors in the vision system is to overcome some of the restraints mentioned above and acquire scene's images and corresponding information [28].

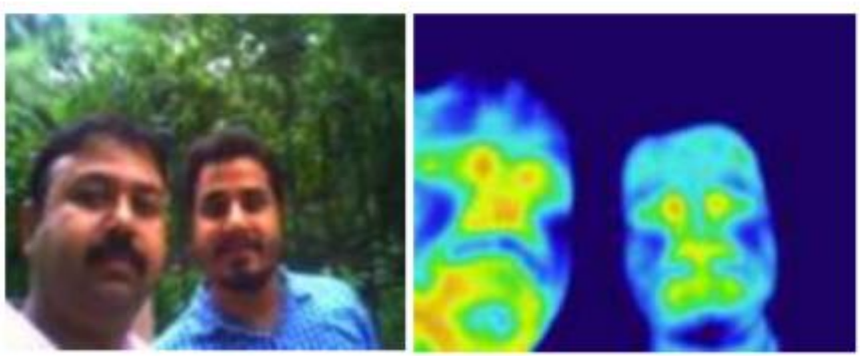

*Figure 2.7 Thermal Imaging*

There are now many applications that a thermal camera can be use such as it can be used for healthcare, which is use for detection of cancer, it can also be used for mechanical inspection and preventative maintenance to see overheated machines before it causes fire, and it can detect energy, leakage and insulation issues etc. Moreover, thermal cameras now are usually used as surveillance cameras for safety as it is used by many police because it is effective to track or to find people [29].

There are three basic parameters of thermal cameras: sensor resolution, measurement precision, and temperature sensitivity. Moreover, several secondary parameters should also be taken into consideration. These secondary parameters are the spectral range which is typically 7-14 µm, field of view (through lens selection), temperature range, the presence of a RGB camera, etc. But measurement uncertainty, and sensor resolution are the two most significant in body temperature measurement. Table 2.1 shows the basic parameters of the thermal camera which is designed to measure the human body's temperature [30].

| | |
|---|---|
| Thermal camera resolution | 640 x 512 pixels |
| Temperature range | 0°C to +50°C, the thermal camera is specifically calibrated to measure body temperature |
| Temperature sensitivity | 0.03°C (30 mK) |
| Uncertainty of measuring device | ± 0.2 ° C (reduction from standard value of ± 2 ° C due to the use of a black body for continuous recalibration) |
| Detector and spectral range | 7.5 – 13.5 µm / uncooled VOx microbolometer |
| Black body for recalibration | Continuous recalibration of the thermal camera using a black body every 5 seconds |
| Lens | 45° |

*Table 2.1 Parameters of Thermal Camera*

A thermal camera is used to detect the distribution of heat in a given environment. In basic terms, a thermography tool detects heat rather than light, displaying this information as a heat-map image with relevant temperature data. A visual marker is allocated to each pixel in a digital image, whereas image and temperature data are assigned to each pixel in a thermal image, allowing the heat map to be constructed. However, there are things that thermal cameras cannot see such as thermal cameras cannot see through walls and concrete because walls and concrete are thick enough and insulated enough to block any infrared radiation from the other side which can also seen in Figure 2.8. Also, thermal cameras cannot detect objects through the trunk of the trees. However, it can help to recognize people and animals because of its heat signature which is why it is usually used for search and rescue operations. A sample figure is shown in Figure 2.9. In addition, thermal cameras can see objects through smoke which is shown in Figure 2.10 because smoke particles effectively block visible light while allowing infrared radiation to get through, allowing firefighters and other first responders to navigate smoke-filled areas [31].


*Figure 2.8 Studs inside the wall (vertical lines) are colder than the insulation, causing a temperature difference on the surface of the wall.*

*Figure 2.9 Thermal imaging cannot see through trees (or wood), but it can be helpful for spotting people in forested areas where their heat signatures stand out much more than a visible image might.*


*Figure 2.10 The person in the doorway is concealed by smoke in the visible light spectrum, but easily detected by thermal imaging.*

# CHAPTER 3

# METHODOLOGY

This chapter provides an overview of the capabilities of human and animal detection from the list of its components to the various processing techniques to keep track of the people for search, evacuation and rescue operation. The methodology gives an in-depth view of the conceptual framework, system processing flow, list of materials, prototype design, the graphical user interface, testing setup, calibration of sensor, and statistical analysis of data.

### *3.1 Conceptual Framework*

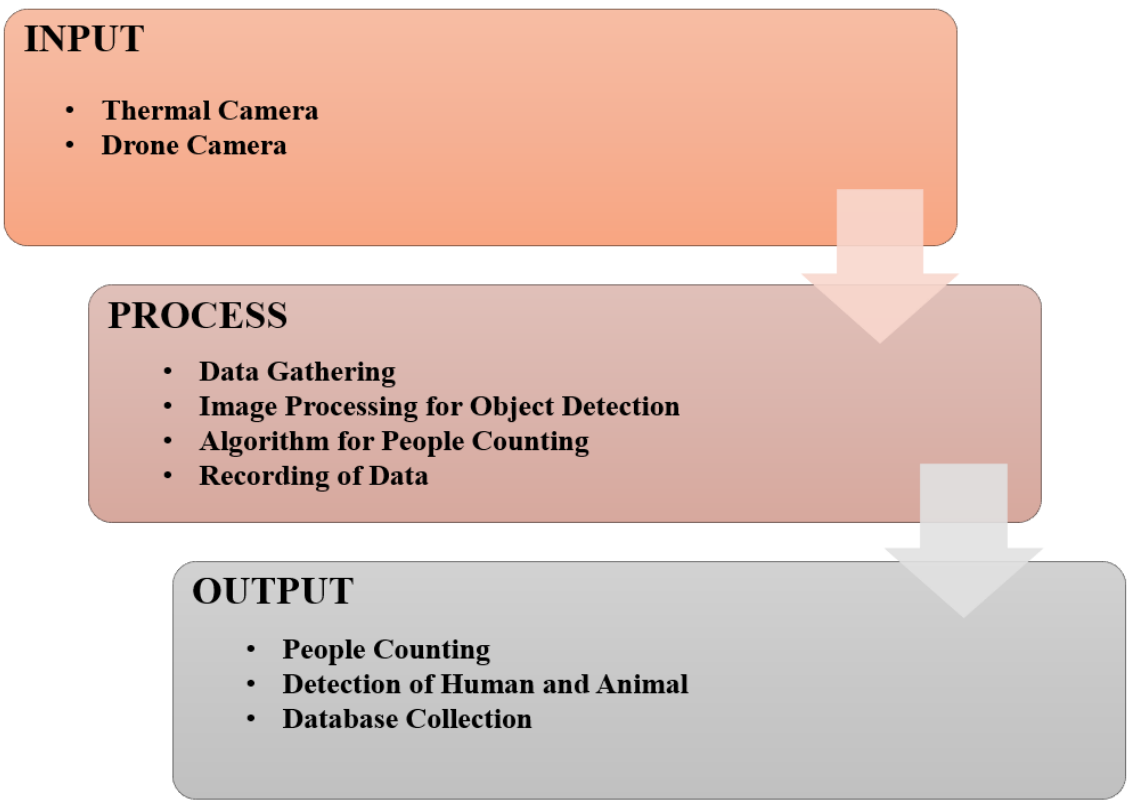


*Figure 3.1 Conceptual Framework*

Figure 3.1 shows the conceptual framework of the study. The input for the person and animal detection will be the two (2) cameras which are the thermal camera and drone camera. For the process, it starts by data gathering which is the videos that the two (2) cameras will capture.

While the data is being gathered, the system can already record the data using the mobile application. The data gathered will undergo an image processing to identify if the object is human or animal. If the system detects a person, it will undergo an application of a connected algorithm in the drone camera which is for people counting. The output of the system includes detection of humans and animals, specifically dogs. Also, it counts all the humans that were detected by the system. Lastly, all of the captured images and videos of the thermal camera and drone camera will be saved to the laptop and mobile device respectively.

### *3.2 System Process Flow*

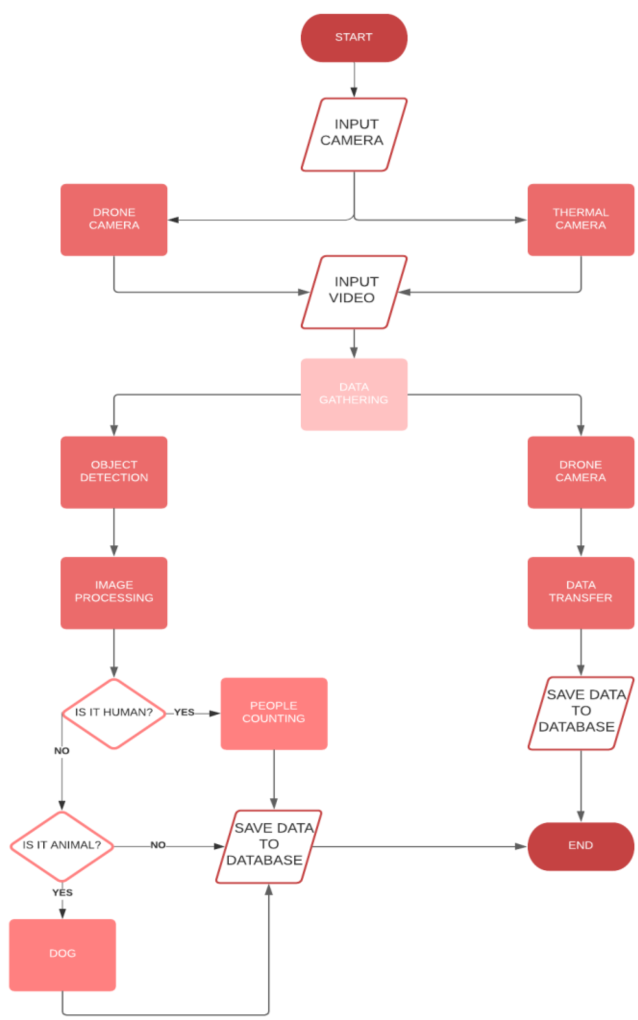


*Figure 3.2 System Process Flow*

In Figure 3.2 the starting phase of the UAV detecting system is to attach the thermal camera to the upper part of the drone. The moment the camera displays the video capture using the GUI of each camera the data gathering will start immediately. The drone will be designed to recognize objects, while Deep CNN will be used to detect humans and animals and the YOLOv5 algorithm for recognizing whether the object is a human or an animal, specifically a dog, and people counting will be added to the drone's features. For the secondary camera the ESP32 will be programmed to operate the thermal camera and connect it to the second device using Wi-Fi for data transfer. All the data collected will be saved in a database.

### *3.3 List of Materials*

1. ***Unmanned Aerial Vehicle***


*Figure 3.3 Tello Drone*

Tello is a mini drone with an HD camera that gives you a bird's eye perspective of the world. It is user-friendly even a beginner can easily manage it. Also, it has two antennas for additional stability in video transmission and a high-capacity battery for impressively long flight periods. Tello is easy to program and develop. The drone's software will be programmed to detect humans and animals, as well as to count people throughout the search, evacuation, and rescue area. It will also be used to record videos and photos that can be directly saved to your smartphone using its application. The secondary camera or the thermal camera will be attached to this drone. It has a maximum flight distance of 100 meters, a maximum speed of 8 meters per second, a maximum flight time of 13 minutes, and a maximum flight height of 30 meters. The camera has 5MP for photos and HD720P30 for video, as well as a detachable battery that can be instantly changed and charged by an USB charging port. Figure 3.3 shows the actual Tello Drone that will be used.

***2. Smartphone***


*Figure 3.4 Smartphone*

Figure 3.4 is a smartphone that is a touchscreen phone with internet connectivity and an operating system capable of running downloadable programs that performs many of the tasks of a computer. It will be used to control the UAV via Bluetooth or Wi-Fi connected via application. The smartphone will serve as the FPV (First Person View) display. The FPV screen shows the pilot what the drone sees while it is in the air via a live transmission. You cannot fly your drone too far away from you without risking it getting lost without a first person display to provide you accurate information about where it is. The smartphone will also be used as the storage of the recorded videos and images from the UAV.

### *3. Thermal Camera*

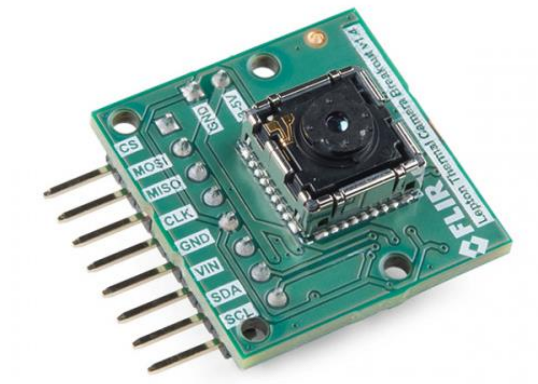


*Figure 3.5 FLIP Lepton 2.0*

Figure 3.5 shows the FLIR Lepton 2.0, the Lepton is available in a variety of field-of-view angles and with an inbuilt mechanical shutter as an option. The FLIR Lepton 2.0 camera core is a thermal imaging long-wave infrared (LWIR) breakthrough in terms of size, power consumption, and affordability. It now fits an 80x60 pixel resolution into an 8.5 x 11.7 x 5.6 mm device that draws only 150 mW. Every pixel of each image is recorded with accurate, calibrated, and noncontact temperature data. It will be used as the secondary camera that will be attached above the drone to give more accurate results.

### *4. ESP32*

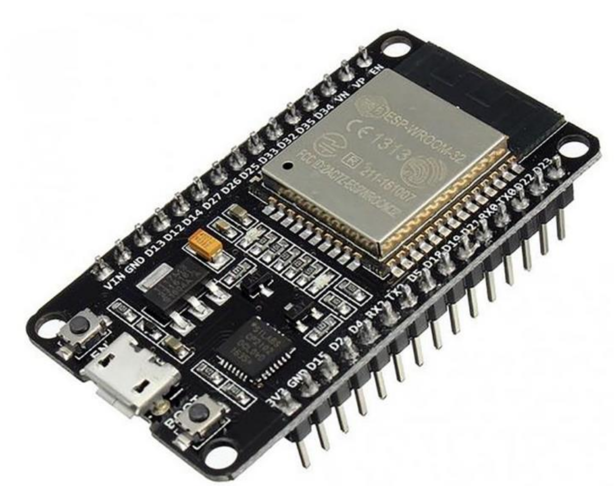


*Figure 3.6 ESP32*

ESP32 is a single 2.4 GHz Wi-Fi and Bluetooth combo chip, that is low-cost and has a low-power system in chip series as shown in Figure 3.6. It is highly suited to a variety of IoT applications due to its value for money, small size, and low power consumption. It is a microcontroller that can be connected to a laptop with the use of an USB connector or Bluetooth module. It will be used with the thermal camera so that the data that was captured by the thermal camera will be saved to ESP32 and the data will be transferred to the laptop.

5. ***Laptop***

*Figure 3.7 Laptop*

Figure 3.7 is a laptop that are small portable computers that have the same capabilities as desktop computers. The laptop will be used to connect the thermal camera and see the video of the thermal camera in real time. Also, the device will be used to collect the data that was captured by the thermal camera and drone camera.

6. ***Battery***

*Figure 3.8 LiFePO4 Battery*

Figure 3.8 is a LiFePO4 battery that has a very flat discharge curve, which means the voltage lowers extremely slowly during discharging. The nominal voltage of a lithium iron phosphate battery (LiFePO4 battery) is 3.2 volts, and the maximum voltage is 3.65 volts. It will be used as the battery ESP32 Thermal Camera. The LiFePO4 battery is perfect for the ESP32, especially for running the circuit for a maximum time and it is rechargeable. Two batteries will be needed, the other one as a back up to quickly change the drain battery and charge.

7. ***Battery Shield***


*Figure 3.9 Battery Shield*

Figure 3.9 is a battery shield that is used to charge and power the ESP32 to run the thermal camera that is connected to the ESP32. It will be connected via Type A USB output or can be directly using the 6 output soldering pins that can be seen on each side of the board (5v or 3v). There is a charging circuit on the shield but it is not protected against polarity inversions.

### *3.4 Prototype Design*

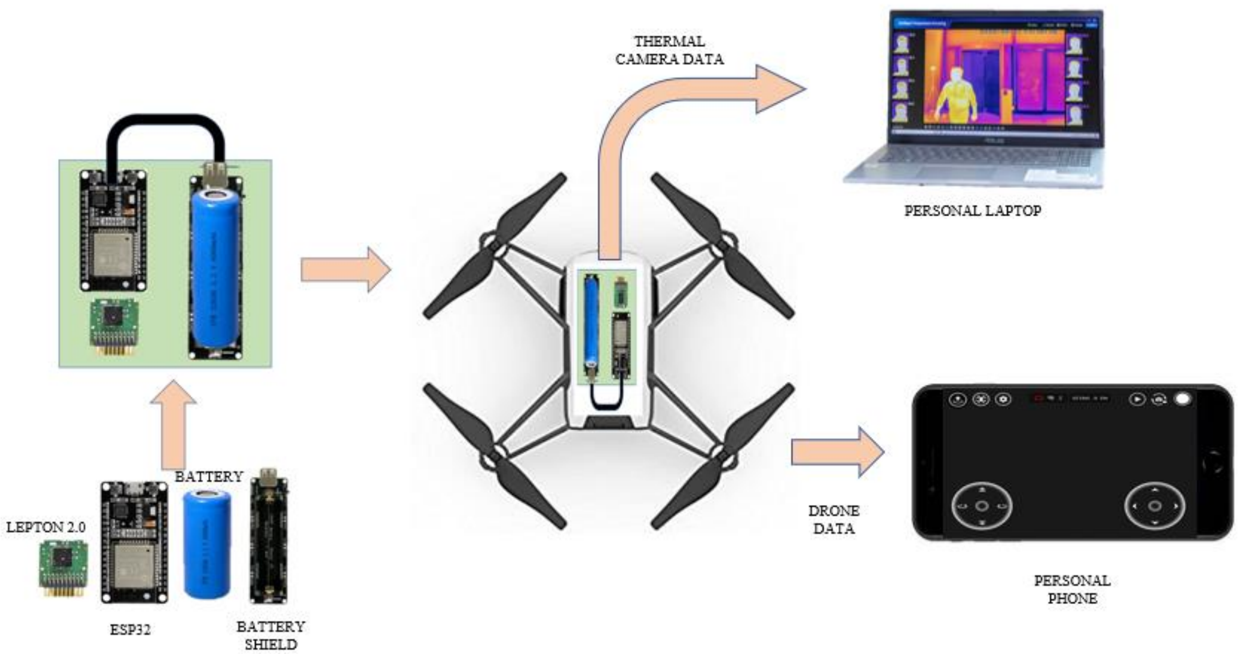


*Figure 3.10 Proposed Prototype*

Figure 3.10 is the sample prototype of human and animal detection. The thermal camera is composed of ESP32 which has a built-in Wi-Fi and Bluetooth for connection of thermal camera to the laptop and a rechargeable battery. The thermal camera will be added to the top of the tello drone. The Drone displays its camera view, and it can be controlled through the smartphone, The User's Manual for the proposed prototype is shown below:

**User’s Manual**

1. Download the TELLO app in the mobile phone
2. Turn on the TELLO drone
3. Connect the TELLO drone to the mobile phone via Wi-Fi
4. Place the TELLO drone into a flat surface
5. Open the TELLO app
6. Press the video recording button on the mobile app
7. Press the take-off button on the mobile app to make the TELLO drone fly
8. Hold the down button in the mobile app to make the TELLO drone to land

### *3.5 Application Software*

Python and Arduino Integrated Development Environment (IDE) were used to create the software and the graphical user interface (GUI). Using Python which is a programming language that is used in web development, machine learning, operating systems, and mobile app development, it will be used to add additional features to show the detected human and animal on the screen and also it will display the live person count as well as the total number of people detected by the drone’s camera. Next is the Arduino IDE which is a software program that allows users to program in the Arduino programming language or boards, it will be used to program the ESP32 to operate the thermal camera.

### *3.6 Graphical User Interface*

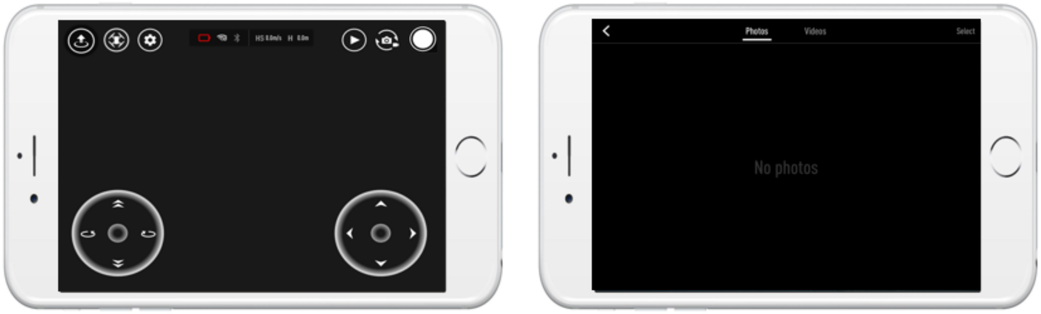


*Figure 3.11 Camera View*

The GUI is done by using the Drone Application, to control the camera and the aircraft functions. It displays a real-time view and shows if there is a detection of human or animal and the total human count. It has many different functions that can be seen on the screen in Figure 3.11. Below are the different functions of the Drone Camera view:

1. Auto Takeoff/Landing - to Land the aircraft automatically.
2. Settings - to adjust flight speed, and Wi-Fi settings.
3. Battery Level - displays the current battery level.
4. Bluetooth Status - displays the Bluetooth connection status.
5. Wi-Fi Status - display the Wi-Fi connection status.
6. Flight Speed - displays the aircraft’s horizontal speed.
7. Flight Altitude - displays the altitude above the surface below the aircraft.
8. Playback - to preview photos and videos as soon as they are captured.
9. Photo/Video Toggle - to switch between photo and video recording modes.
10. Shoot/Record Button - to start shooting photos or recording video.
11. Virtual Joysticks - to control the aircraft

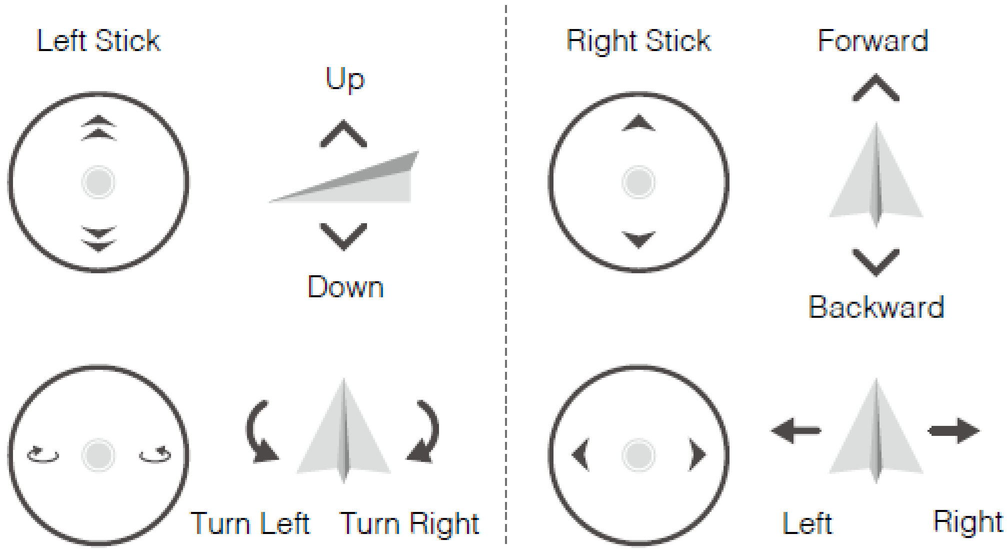


*Figure 3.12 Virtual Joysticks*

The Virtual Joysticks that will be used to control the aircraft are shown in Figure 3.12. The altitude of the aircraft can be changed by moving the left stick up or down. The orientation is controlled by moving the left stick to the left or right. The pitch of the plane can be changed by moving the right stick up and down. The roll of the aircraft is changed by moving the right stick to the left or right.

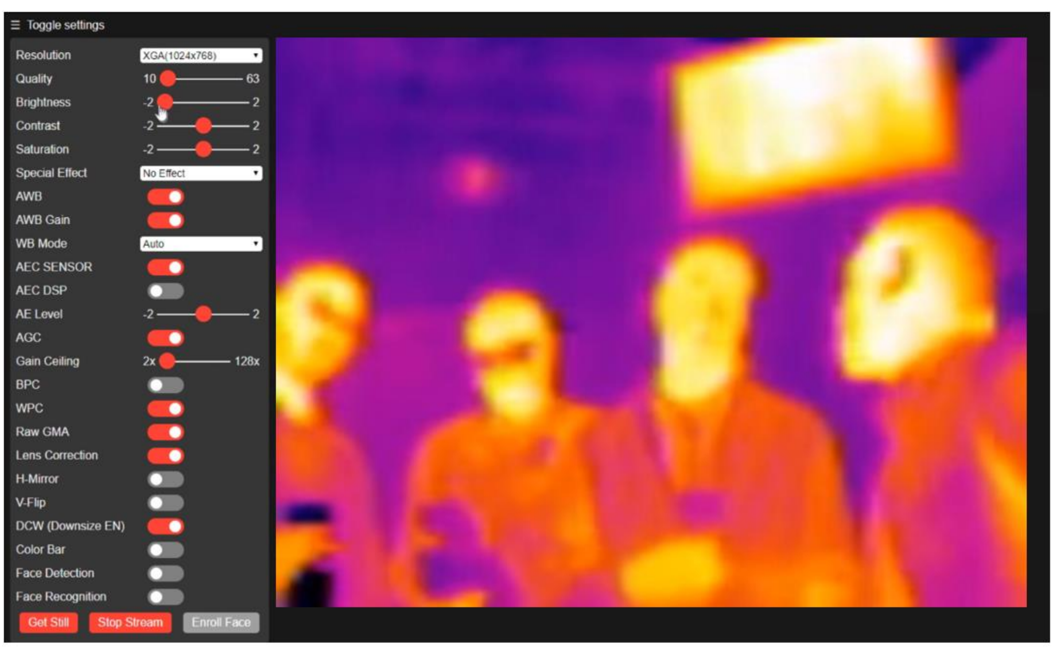


*Figure 3.13 Thermal Camera GUI*

The next GUI for the Thermal camera, or secondary camera, that is designed using the Arduino IDE is shown in Figure 3.13. It also displays a real-time view and shows the thermal image output that is recorded using temperature data that is accurate, calibrated, and noncontact. This will help the rescuers to confirm if the detected object from the drone camera is a human or animal.

### *3.7 Testing Setup*

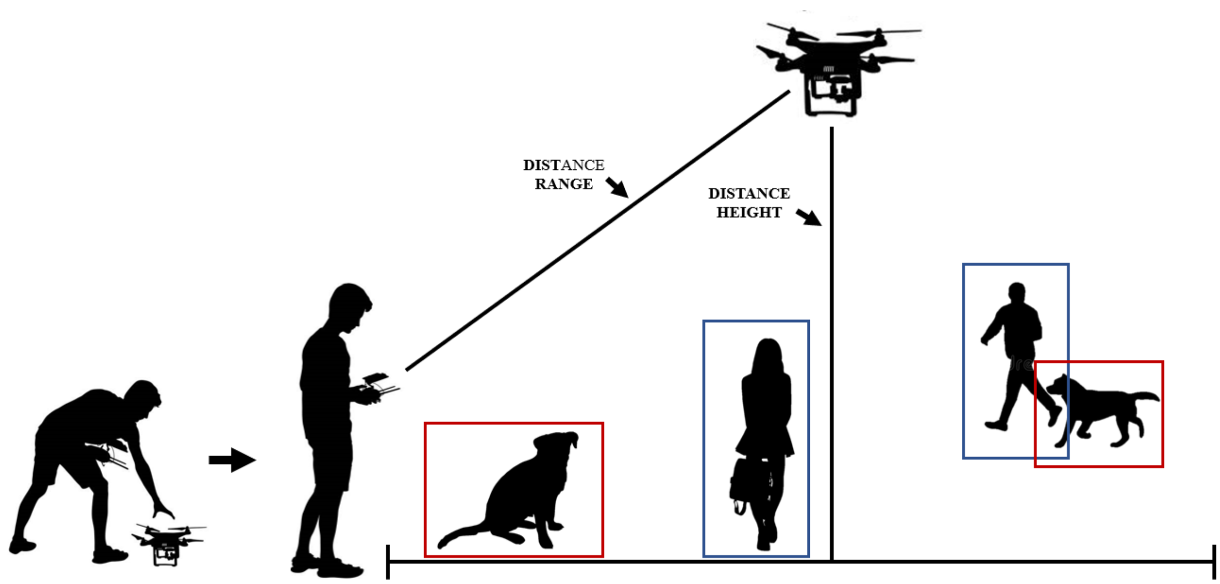


*Figure 3.14 Prototype Testing Setup*

The testing will take place in an open space location, as indicated in Figure 3.14. To take off, the drone will be placed on a flat surface.  The drone's maximum flight range is 100 meters, but the researchers will not test it to that extent to prevent losing connectivity. The distance height will be examined to see if the drone can detect humans and animals at various distances and to test the accuracy of the drone.

### *3.8 Calibration of Sensor*

1. *Thermal Camera*

For this study, FLIR Lepton 2.0 Thermal Camera will be used for the secondary camera of the prototype. The FLIR Lepton 2.0 Thermal camera will be subjected to a calibration test to determine its accuracy. To test the camera, it will be programmed to identify the temperature of humans and assign different colors that depend on the temperature range.

| Color assigned on Temperature range | Temperature Range | Actual Temperature using Thermal Scanner | Color output of FLIR Lepton 2.0 | Does it match the color assigned in the temperature range? (Yes/No) |
|---|---|---|---|---|
| Black | 0°C Below | | | |
| Purple | 1°C -10°C | | | |
| Blue | 11°C -20 ° C | | | |
| Green | 21°C -30°C | | | |
| Yellow | 31°C -36.4°C | | | |
| Orange | 36.5°C -37.5°C | | | |
| Red | 37.6°C -40°C | | | |
| White | 40°C Above | | | |

*Table 3.1 Calibration of FLIR Lepton 2.0 Thermal Camera*

Table 3.1 shows the results for calibration of FLIR Lepton 2.0 Thermal Camera. The accuracy of FLIR Lepton 2.0 Thermal Camera will be tested by comparing the result of the FLIR Lepton 2.0 Thermal camera and a commercially available thermal scanner. Moreover, it will be checked by the results if the assigned color matches to the temperature.

*2. Drone Camera*

| Distance Height | Number of actual counts of Human | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | |
| | Number of Detected Human | %Error | Number of Detected Human | %Error | Number of Detected Human | %Error | Number of Detected Human | %Error | Number of detected Human | %Error |
| 5m | | | | | | | | | | |
| 10m | | | | | | | | | | |
| 15m | | | | | | | | | | |
| 20m | | | | | | | | | | |
| 25m | | | | | | | | | | |
| 30m | | | | | | | | | | |

*Table 3.2 Calibration of Camera Drone for Human*

| Distance Height | Number of actual counts of Animal | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | |
| | Number of Detected Animal | %Error | Number of Detected Animal | %Error | Number of Detected Animal | %Error | Number of Detected Animal | %Error | Number of detected Animal | %Error |
| 5m | | | | | | | | | | |
| 10m | | | | | | | | | | |
| 15m | | | | | | | | | | |
| 20m | | | | | | | | | | |
| 25m | | | | | | | | | | |
| 30m | | | | | | | | | | |

*Table 3.3 Calibration of Camera drone for Animal*

Table 3.2 shows how the Camera Drone will be evaluated during calibration. It will be tested by counting the number of humans that were detected by the system. The sample size for humans that will be used for the calibration will be one (1) to five (5) and each trial the distance height of the drone will be adjusted 5m to 30m with an increment of 5m. The %Error will be calculated as shown in equation 3.1. The measured value will be the number of humans that were detected by the drone and the real value will be the number of actual counts of humans. Table 3.3 shows the calibration of Camera Drone for Animals. It will have the same parameters for humans.

$$\text{Percent Error} = \frac{|\text{measured} - \text{real}|}{\text{real}} \times 100\% \qquad (1)$$

Figure 3.14 *Percent error formula*

| Distance Height | Number of actual counts of Human | | | | | | | | | | Number of actual counts of Animal | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | 1 | | 2 | | 3 | | 4 | | 5 | |
| | Number of Detected Human | %Error | Number of Detected Human | %Error | Number of Detected Human | %Error | Number of Detected Human | %Error | Number of detected Human | %Error | Number of Detected | %Error | Number of Detected | %Error | Number of Detected | %Error | Number of Detected | %Error | Number of detected Animal | %Error |
| 5m | | | | | | | | | | | | | | | | | | | | |
| 10m | | | | | | | | | | | | | | | | | | | | |
| 15m | | | | | | | | | | | | | | | | | | | | |
| 20m | | | | | | | | | | | | | | | | | | | | |
| 25m | | | | | | | | | | | | | | | | | | | | |
| 30m | | | | | | | | | | | | | | | | | | | | |

*Table 3.4 Calibration for Human and Animal*

### *3.9 Statistical Analysis*

The t-test is a statistical approach for determining if two populations are statistically different. Three key data values are required to calculate a t-test. The mean difference (the difference between the mean values in each data set), the standard deviation of each group, and the number of data values in each group.

**T-Test**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (2)$$

*Figure 3.15 T-test Formula*

Where:

$X_1$ = Mean of first sample

$X_2$ = Mean of second sample

$N_1$ = Sample size of first sample

$N_2$ = Sample size second sample

$S_1$ = Standard deviation of first sample

$S_2$ = Standard deviation of second sample

$S_p$ = Pooled standard deviation

REFERENCES:

[1] S. Sambolek and M. Ivasic-Kos, "Automatic Person Detection in Search and Rescue Operations Using Deep CNN Detectors," in IEEE Access, vol. 9, pp. 37905-37922, 2021, doi: 10.1109/ACCESS.2021.3063681.

[2] S. Powale, A. Dhanawade, S. Bagwe, S. Kawale, N. L. Chutke and S. Chavan, "Person identification in low resolution CCTV footage using deep learning," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2020, pp. 236-240, doi: 10.1109/ICACCCN51052.2020.9362764.

[3] Singh M., Sahran R., “A New Approach for Suspect Detection in Video Surveillance”, Information and Communication Technology for Sustainable Development. Lecture Notes in Networks and Systems, vol 10. Springer, Singapore, 2018.

[4] Ivasic-Kos, M., Kristo, M., & Pobar, M. (2020). Person detection in thermal videos using YOLO. In *Advances in Intelligent Systems and Computing* (Vol. 1038, pp. 254–267). Springer Verlag. https://doi.org/10.1007/978-3-030-29513-4_18

[5] Ahmad, M., Ahmed, I., Ullah, K., Khan, I., Khattak, A., & Adnan, A. (2019). Person detection from overhead view: A survey. *International Journal of Advanced Computer Science and Applications*, *10*(4), 567–577. https://doi.org/10.14569/ijacsa.2019.0100470

[6] B. Kellenberger, M. Volpi and D. Tuia, "Fast animal detection in UAV images using convolutional neural networks," *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 866-869, doi: 10.1109/IGARSS.2017.8127090.

[7] B. Kellenberger, D. Marcos, S. Lobry and D. Tuia, "Half a Percent of Labels is Enough: Efficient Animal Detection in UAV Imagery Using Deep CNNs and Active Learning," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 12, pp. 9524-9533, Dec. 2019, doi: 10.1109/TGRS.2019.2927393.

[8] Banupriya, N., Saranya, S., Jayakumar, R., Swaminathan, R., Harikumar, S., & Palanisamy, S. (2020). Animal detection using deep learning algorithm. *Journal of Critical Reviews*. Innovare Academics Sciences Pvt. Ltd. https://doi.org/10.31838/jcr.07.01.85

[9]*Statistics: Philippine Canine Club, Inc. (PHILIPPINES)*. (2020). http://www.fci.be/en/statistics/ByNco.aspx?iso=PH

[10] Koester, R.J. A Search and Rescue Guide on where to Look for Land, Air, and Water; dbS Productions: Charlottesville, VA, USA, 2008.

[11] Vasić, M. K., & Papić, V. (2020). Multimodel deep learning for person detection in aerial images. *Electronics (Switzerland)*, *9*(9), 1–15. https://doi.org/10.3390/electronics9091459

[12] Bhavsar, P. (2019). Object Detection with Convolutional Neural Networks. Retrieved from Data Driven Investor: https://medium.datadriveninvestor.com/object-detection-with-convolutional-neural-networks-dde190eb7180

[13] Wel, L. et al., SSD: Single Shot MultiBox Detector, ECCV. (2016)

[14] (Hao Wu, 2018) Redmon, J. et al., You only look once: Unified, real-time object detection, CVPR. (2016)

[15] Ye Li, Kangning Yin, Jie Liang, Chunyu Wang and Guangqiang Yin, "A Multi-task Joint Framework for Real-time Person Search", 2020.

[16] J. Ieamsaard, S. N. Charoensook and S. Yammen, "Deep Learning-based Face Mask Detection Using YoloV5," 2021 9th International Electrical Engineering Congress (iEECON), 2021, pp. 428-431, doi: 10.1109/iEECON51072.2021.9440346.

[17] S. Tan, G. Lu, Z. Jiang and L. Huang, "Improved YOLOv5 Network Model and Application in Safety Helmet Detection," 2021 IEEE International Conference on Intelligence and Safety for Robotics (ISR), 2021, pp. 330-333, doi: 10.1109/ISR50024.2021.9419561.

[18] S. D. Pore and B. F. Momin, "Bidirectional people counting system in video surveillance," 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2016, pp. 724-727, doi: 10.1109/RTEICT.2016.7807919.

[19] J. M. R. S. Tavares, “Image processing and analysis: Applications and trends,” *AES-ATEMA Int. Conf. Ser. - Adv. Trends Eng. Mater. their Appl.*, no. May, pp. 27–41, 2010.

[20] V. Hlavac, “Fundamentals of Image Processing,” *Opt. Digit. Image Process. Fundam. Appl.*, no. July, pp. 71–96, 2011, doi: 10.1002/9783527635245.ch4.

[21] R. G. L. Narayanan and O. C. Ibe, “Joint Network for Disaster Relief and Search and Rescue Network Operations,” *Wirel. Public Saf. Networks 1 Overv. Challenges*, pp. 163–193, Jan. 2015, doi: 10.1016/B978-1-78548-022-5.50006-6.

[22] J. Andersson and D. Persson, “Real-time image processing on handheld devices and UAV,” 2016, [Online]. Available: http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A947096&dswid=-2146.

[23] H. Shakhatreh *et al.*, “Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges,” *IEEE Access*, vol. 7, pp. 48572–48634, 2019, doi: 10.1109/ACCESS.2019.2909530.

[24] “Surveillance Drones | Electronic Frontier Foundation.” https://www.eff.org/issues/surveillance-drones (accessed Jul. 19, 2021).

[25] S. Waharte and N. Trigoni, “Supporting search and rescue operations with UAVs,” *Proc. - EST 2010 - 2010 Int. Conf. Emerg. Secur. Technol. ROBOSEC 2010 - Robot. Secur. LAB-RS 2010 - Learn. Adapt. Behav. Robot. Syst.*, pp. 142–147, 2010, doi: 10.1109/EST.2010.31.

[26] “Infrared Radiation Theory.” [Online]. Available: https://www.azosensors.com/article.aspx?ArticleID=339.

[27] FLIR System, “Thermal Imaging Guidebook for building and renewable energy applications,” *Guid. B.*, p. 68, 2011, [Online]. Available: www.flir.com%5Cnwww.flirmedia.com/MMC/THG/Brochures/.../T820264_APAC.pdf.

[28] M. Rai, T. Maity, and R. K. Yadav, “Thermal imaging system and its real time applications : a survey Thermal imaging system and its real time applications : a survey,” no. June, 2018.

[29] J. Barela, M. Kastek, K. Firmanty, and P. Trzaskawka, “Accuracy of measurements of basic parameters of observation thermal cameras,” no. February 2018, 2016, doi: 10.21611/qirt.2016.004.

[30] “Everything You Need To Know About Thermal Imaging Cameras | RS Components.” https://uk.rs-online.com/web/generalDisplay.html?id=ideas-and-advice/thermal-imaging-cameras-guide (accessed Jul. 19, 2021).

[31] “Can Thermal Imaging See Through Walls? And Other Common Questions | Teledyne FLIR.”https://www.flir.com/discover/cores-components/can-thermal-imaging-see-through-walls/ (accessed Jul. 20, 2021).