Project Proposal

Automatic Target Recognition and Map Generation

Target Acquired

Ryan Ylagan - Technical Co-Lead, Justin Campbell - Project Lead, Rohan Wariyar - Comm. Co-Lead, Preston Hart - Comm. Co-Lead, Nicholas Aufiero – Technical Co-Lead

March 7, 2022

COE 374

Problem Statement

Contributions: Ryan

The Austin metropolitan area has experienced flooding. The rising water levels have become so significant that the Austin Fire Department (AFD) has begun to receive distress calls from individuals stranded on the roofs of their houses trying to escape the rising waters. In order to be more efficient with their search and rescue efforts, AFD needs to be able to know exactly if they need to deploy their rescue teams and where to deploy them. To do so, the AFD wants to install an unmanned aircraft system that is equipped to provide 24/7 search, surveillance, and precision airborne delivery to aid in rescue operations.

In the water, the houses with individuals atop them create targets of interest (TOI). Some individuals may be completely uninjured and in a safe enough position to wait for the water to subside. On the other hand, some individuals may be in need of immediate medical attention. The location of these homes, an assessment of the distress, and any other information gathered from the surveillance should be transmitted to AFD.

the mission will be a joint effort with the Aerospace senior design team (ASE) and the Computational senior design team in charge of the aerial drop payload. The ASE team will be in charge of assembling an aircraft and programming its flight path for the search and surveillance stage of the mission. The aircraft will perform this stage over an approximately 300,000 ft2 mowed field with the TOI's represented by five 2 x 2 meter targets made with a large blue tarp. Atop two of these TOI's are smaller objects that mark these targets as candidate targets for potential rescue; one with a "smiley face" representing an uninjured person, and another with a "frowny face" signifying an individual in need of immediate assistance.

Problem Objectives

Contributions: Ryan

The main objective is to create a piece of automatic target recognition software that is capable of categorizing the aforementioned TOI's as a candidate target and furthermore be able to classify the type of assistance needed based on the face that lies upon it. The software also needs to return accurate GPS coordinates of each TOI that can be return to the aerial drop payload team. This information then needs to be compiled together along with images taken from the mounted camera to generate a map of the targets and their classification. The speed at which the software is able to acquire and identify these targets as well as how accurate the GPS locations are to the true location of the targets will be used to score the ASE team's mission.

Some of the objectives that are not required but may be good to include as stretch objectives are:

- An interactive map with that is paired with the respective GPS coordinates.
- Image recognition software that can identify human emotion rather than just the preset TOI's.
- A more efficient algorithm that improves on accuracy and required search time.
- An algorithm that is capable of identifying targets on more complex terrain rather than just an open field.

Introduction and Background

Contributions: Ryan, Rohan, Preston, Nicholas

Every year, the aerospace senior design teams are tasked to perform a demonstration mission that is to be conducted at the aforementioned 300,000 ft2 field. This mission includes a few important steps that include takeoff and deployment of the aircraft, which will be handled by the aerospace design team. the involvement concerns the search and surveillance steps of the mission that occur after the aircraft reaches the cruising altitude of between 200 to 300 ft. During this phase, the software should be able to automatically recognize the numerous targets laid out on the field as well as classify these targets based on their display. In previous years this phase was done manually, but due to the involvement, this is now done completely autonomously.

Overall the mission and its objectives are not necessarily a unique problem. There have been many well documented case studies of other teams creating an unmanned aerial vehicle equipped with a mounted camera software for image recognition and map generation for the purpose of search and rescue. In one particular study, a Japanese team used a small UAV to generate a map of an area recently affected by an earthquake. This map detailed the surrounding area's geological features and marked points of interest that may be considered important for a search and rescue team, such as a collapsed house. Along with the map, the aircraft would return data such as position, altitude, and time that is specific to each individual photo.

Computer vision is an area of technology that a dense amount of research and development efforts are currently being ptheed into. The ability for equipment to visualize its surroundings, make sense of the environment that it is in and make smart decisions from such perceptions has been a major breakthrough in autonomous capabilities. Specifically, in the case of the unmanned aerial vehicles, many military and search and rescue operations have been seeking out computer vision software to automate tasks that are time consuming and error prone for the human eye. There are numerous open sthece computer vision software available for implementation and most differ in how they through an algorithmic process learn the object(s) of interest and then proceed to evaluate and make decisions on new pieces of data. Both of these key components of image evaluation are integral parts of the run time complexity that computer vision software are judged on and ultimately a piece of the process that went into evaluating the software discussed in this project proposal. As a note, there is a difference between locating and mapping targets and areas of interest. Today, more modern computer vision algorithms are seeking to combine the process locating and mapping into one continuous process that uses minimal training data. This combined methodology is referred to as a SLAM process and is the architecture behind self-driving cars as well as autonomous ground-based robots. However other computer vision algorithms simply seek to find large localized differences in color contrasts in order to identify targets of interest. Although this category of computer vision algorithms leans toward being less computationally intensive than other methods, they are not always as practical since there is a large emphasis placed on color differences which could lead to difficulties when color is not the key factor of identification.

Because UAVs have become extremely popular over the past decade, the equipment required to run computer vision on the go has become widely available. Performing computer vision from up to 300 feet in the air requires a camera with enough resolution to accurately capture small ground targets. UAV computer vision also requires a small, light computer which can run independently but in communication with the main flight computer. This computer must be powerful enough to process the images acquired from the camera in real time. With today's processors, that is possible.

Feasibility Studies

Contributions: Justin, Rohan, Preston, Nicholas

Concerning the hardware of this project, there were decisions about what camera and what co processor to operate within the UAV. The selection process for the camera will be discussed first. The team again began the Preliminary Design Stage by assembling a table with preliminary criteria for the camera that were determined to be the most important for the camera to satisfy the project's end objectives for both the COE and ASE design teams. The ASE team emphasized the importance of weight reduction on their aircraft as weight plays a large role in their final score. As such, the weight of the camera was an important consideration. Other important considerations were the operating resolution, the frame capture rate, the output interface terminal type, the camera's physical dimensions, and the available documentation. The three cameras that were considered in the preliminary design stage were the Topotek 10x zoom, the Lumenera Lt-C, and the RunCam Zoom. After a literature review on similar ATR projects, a team successfully performed ATR at altitudes of around 200 feet using a 5 megapixel camera. 5 Mpx was then set as the minimum criteria for the operating resolution. Similarly, a frame capture rate of 20 fps was set as the minimum criteria according to documentation found in the literature process. With respect to the total weight and physical length dimensions criteria, values were set in collaboration with the ASE Design Team to operate within their weight and mass distribution constraints. This will ensure that implementation of the hardware products do not impact the vehicular performance of the UAV. With respect to the other two, the team used a more subjective approach to determine minimum criteria. For the type of interface output terminals category, the criteria was satisfied for a particular design if it contained at least one in any combination of USB, HDMI, or MIPI CSI output terminals. For the available documentation category, the criteria was satisfied if documentation was easily accessible, readable, and verified/validated.

After the preliminary criteria were assembled, each design was documented according to whether or not it satisfied the criteria. The team then set aside three critical criteria that were determined to be the most important for the camera to achieve its end objectives, namely, the operating resolution, frame capture rate, and type of interface output terminals. The cameras that satisfied a majority of the critical criteria were down-selected to the Feasibility Analysis to be considered in the final design cut. Out of the four cameras considered in the preliminary design stage, only the RunCam Zoom did not meet a majority of the criteria. In particular, it neither met the operating resolution criteria nor the frame capture rate, and was thus eliminated from consideration. After down-selecting the preliminary design ideas, the team assembled them into a weighted scoring matrix along with the criteria specified in the preliminary design phase as shown in Figure 1 below. This process marked the beginning of the Feasibility Analysis to obtain final designs from the down selection. Using a linear scoring methodology, the criteria were ranked from most to least important, and assigned numerical weight values in descending order of magnitude.

Criteria Description	Topotek 10x Zoom Score	Lumenera Lt-C/M2020 Score	Linear Scoring Methodology Legend
CR - Operating Resolution Range	0	6	GREAT >= 85% criteria weight satisfied
CR - Frame Capture Rate	5	5	GOOD >= 70% criteria weight satisfied
CR - Quantity of HDMI and/or USB outputs	4	4	NEUTRAL >=50% criteria weight satisfied
CR - Available Documentation	0	3	POOR < 50% criteria weight satisfied
CR - Total Weight	2	2	
CR - Physical Length Dimensions	1	1	
Total Score	12	21	

Figure 1: Weighted Scoring Matrix for Camera Design

Total criteria scores for each down-selected design were obtained by summing the contribution from each criteria according to whether or not they were satisfied. For example, the Topotek design did not meet the criteria for resolution range, and was thus allocated a score of a 0. However, it did meet the criteria for frame capture rate, and was thus allocated the number of points corresponding to that criteria rank (since the rank is the second highest on a list of 6 criteria, this meant 5 points). Using the metric shown for evaluating the "goodness" of a design according to the percentage of criteria weight satisfied, it was determined that the Lumenera Lt-C and the Arducam IMX were the most desirable options. The Topotek was safely removed from contention because it did not meet the resolution criteria, and is poorly documented. To weigh the two remaining cameras against one another, the Lumenera Lt-C was removed because it is restricted to taking still images instead of streaming video, and, after deliberation, video became a hard requirement for the ASE team. As such, a pivot

Not Arducam. It's a Raspberry Pi Camera

was made to the Arducam camera, which meets all of the important requirements, including resolution, weight, and ease of use. As such, the team will be moving forward with the Raspberry Pi Camera in the final design.

The ATR code will run on a co processor. This co processor must fit within the middle section of the UAV, it must meet stringent weight requirements so as to not significantly shift the center of gravity of the UAV, and it must be able to read in video data from the camera as well as flight data from the Pixhawk Autopilot Ardupilot. The co processor must also require no more than 5 volts to run, and it must also be able to effectively run the ATR algorithm within reasonable time. Raspberry Pi's are a user-friendly, effective solution to edge computing. The model 4 is powerful enough to run the reasonable ATR algorithm. Most importantly, from the literature, it has been determined that many teams have been able to solve similar problems using the Raspberry Pi. the Jetson TX2 was also looked at. This co processor is a machine learning workhorse with 6 CPU's and a GPU. The Jetson might be required for highly complex ATR tasks, but it is overkill and too heavy for the project. In particular, both the Raspberry Pi model 4, and the Jetson TX2 met a majority of the critical criteria in the preliminary design phase, and were documented for Feasibility Analysis. In the Feasibility Analysis, it was determined that both the Jetson and Raspberry Pi were acceptable designs satisfying a considerable majority of the criteria weight. The differentiating factor between the two lied in the power supply voltage, and weight. In particular, the Jetson requires a much higher voltage supply and is much heavier than the Raspberry Pi. As such, the team selected the Raspberry Pi model 4 as the final design.

Amongst many different ideas that were discovered through extensive research into the problem of automatic target recognition and aerial map generation, three different algorithmic software approaches were evaluated.

RGB based recognition, another valued approach to automatic target recognition, looks to expose color contrasts within an image frame to confidently identify the object of interest. In the case of the posed project objective, this method would work well to identify the high blue contrast of the tarps that the targets of interest will rest on and are in stark contrast in color to the green grass of the ground within the search area. By filtering out each pixel's red and green values and honing in on the numeric blue value, areas of dense blue in an image can easily be identified. The RGB based recognition algorithm finds the max and mean pixel contribution values of blue in the image and if that max - mean is greater than predetermined threshold then it can be confidently assumed the high area of blue in the image is a tarp. Where the RGB based method on its own fails to meet the projects' objective is differentiating between the smile and frowny faces of the targets of interest that rest on the blue tarps. The reasoning behind this stems from the fact that the faces of the target do not differ in color, but in orientation; an RGB based method would struggle to make a distinction purely based on differing orientations.

OpenMV is another open-source computer vision software that was researched after coming across a published paper by a group of engineers that developed an unmanned aerial vehicle able to identify and track moving objects. OpenMV differs from OpenCV in the sense that it was developed as a "machine vision" solution to object detection rather than a "computer vision" solution. That is to say, where OpenCV requires a camera connection typically via USB to a co-processor, OpenMV integrates a small camera directly into a customized co-processor. This open-source software is advantageous in its low latency and minimal resolution loss when transferring images from the camera to storage on the co-processor. However, out of a desire to keep the co-processor as lightweight as possible, the camera only has a resolution of 640x480 Mp and a small field of view (FOV). Following the same procedure as above for the camera, and co-processor, OpenMV was not a solution moved into the Feasibility Studies phase primarily due to the implications this solution would have on the aerospace engineering design team that the automatic target recognition product is being built for.

Remove this

The final, and most feasible, option when approaching target recognition software is OpenCV, or open-source Computer Vision Library. After researching the available algorithms and looking at codes from OpenCV, it was decided to utilize the Haar Cascades algorithm for target recognition. Haar Cascade classifiers are machine learning object detection programs that can potentially identify objects in photos and videos. The algorithm begins by calculating Haar features in ? the image or video resolution. Haar features are calculations that are performed at specified locations in a detection (position) window. The calculations involve comparing the differences between the sums of pixel intensities. Examples of Haar features include edge features, line features, and four-rectangle features. Internal images are created with the inclusion of Haar features. However, since the majority of the internal images created will be irrelevant to the users project, something like Adaboost can be used to determine the "best" features. Once these features are chosen, they are used to train the classifiers. Weak classifiers are often combined together to create stronger classifiers. Finally, the classifiers are implemented and they're able to identify objects. One important item to note, these algorithms will require training by using positive and negative images.

Unfortunately, Haar Cascades has a high false positive rate. This is due to the age and nature of the algorithm; more recent ones offer a smaller margin of error. However, with enough tuning, the Haar Cascades code can be optimized against false positives. Furthermore, despite the potential pitfalls of larger margins of error and high false positive rates, this algorithm continues to remain computationally inexpensive while operating at high speeds. Especially since the team will be using a Raspberry Pi, it is important that the co processor is not overworked. As such, OpenCV is currently one of the only options available that meets all of the software and hardware requirements. Additionally, these algorithms are well documented and developed. Python has an OpenCV library that allows for the potential inputs of images, videos, and live feed to the algorithm.

Consequently, the team proceeded to down-select both the OpenCV and RGB Recognition designs into the Feasibility Studies. It was determined that the OpenCV software is the optimal standalone design while the RGB Recognition software satisfied a moderate amount of the criteria. However, given that the RGB Recognition system can be integrated with the OpenCV software for image identification/classification, the team will be proceeding with a combined Haar Cascades and RGB Recognition algorithm as the final choice on how to address the issue of object recognition in photos and videos.

Proposed Plan and Timeline

Contributions: Justin, Preston

In collaboration with the ASE Design Team, and Aerial Drop Payload (ADP) Team, the ATR team has established a high-level Gantt Chart that provides a foundation for major project tasks and their deadlines throughout the course of the project. In referencing the figure below, the bars coded in blue are specific to the ASE deadlines, while those shaded in red, pertain to the ADP and ATR deadlines. To expand, the tasks set aside in the chart are structured around major project deadlines such as presentations, project proposals, design reviews, and flight tests. Moving forward, the team will consider documenting accomplished tasks in a separate organizational framework, and will use an additional Gantt Chart to overlay internal team deadlines that are specific to particular team members.

This chart will likely be implemented on a week-by-week basis, and will be a fluid/dynamic project management tool whose utility will depend on overall team workload, conflicts, and dependencies with the other two teams during a particular week.There are two categories of conflict that the team will consider to minimize during the project life cycle. The first category is concerned with personal conflicts that arise from individual team members due to a combination of the following: disinterest in work, sub-standard effort, poor punctuality,

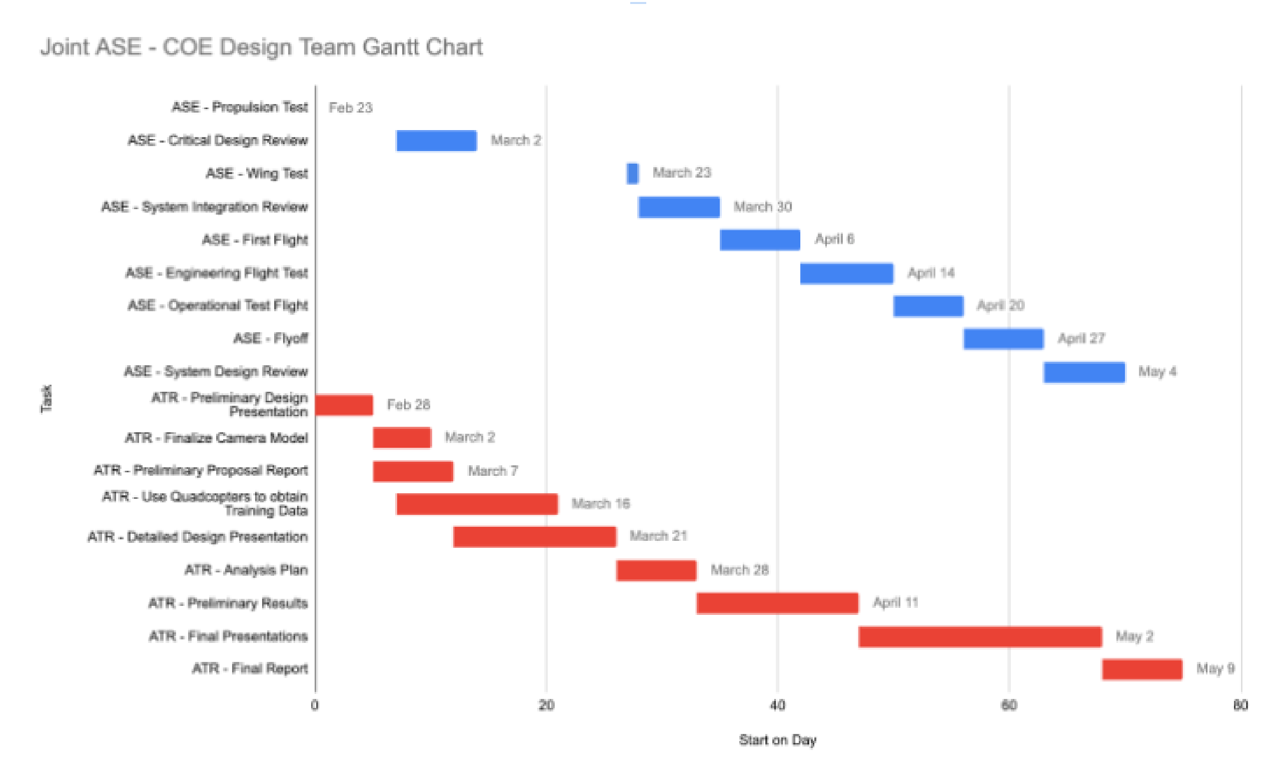

Figure 2: Proposed Gantt Chart

and other related concerns. Considering that the ATR team is responsible for working closely with both the ASE and ADP teams throughout the product development cycle, there will inevitably be instances where tasks will need to be modified in scope, timeline, and utility because of constraints posed by other teams. It is likely that similar conflicts arising from sub par execution and/or planning of tasks will become realised across the three teams.

A primary strategy that the team will invoke as a preventative measure is conducting bi-weekly meetings where the team will have open discussions about each team members' progress, and concerns, and where a document of action items and team member roles will be clearly delineated. Additionally, through instituting a second Gantt Chart that will break down the workload for each team member according to fine details such as duration of completion, and team member percentage contributions, the team will have a project management tool that minimizes ambiguity in team member roles and responsibilities for tasks that range from high-level to narrow in scope.

References

Hoed T., Brandt D., Soerge U., and Wiggenhagen M. (2012). "REAL-TIME ORIENTATION OF A PTZ-CAMERA BASED ON PEDESTRIAN DETECTION IN VIDEO DATA OF WIDE AND COMPLEX SCENES". *Intercommission Working Group III/V.*

Lt-C/M2420 (n.d.). https://www.lumenera.com/products/industrial-scientific/ltx2420.html. Accessed: 2022-03-01.

Machine Vision with Python (n.d.). https://openmv.io/. Accessed: 2022-03-01.

Sun J., Li B., Jiang Y., and Wen C. (2016). "A Camera-Based Target Detection and Positioning UAV System for Search and Rescue (SAR) Purposes". *Sensors (Basil)* 16.11, p. 1778.