OBJECT DETECTION, TRACKING AND SUSPICIOUS ACTIVITY RECOGNITION FOR MARITIME SURVEILLANCE USING THERMAL VISION

Project Proposal and Feasibility report submitted in partial fulfillment of the requirements for the
Degree of Bachelor of Science of Engineering
in

The Department of Electronic & Telecommunication Engineering
University of Moratuwa.

Supervisors:

Dr. Peshala G. Jayasekara

Dr. Ranga Rodrigo

Group Members:

Abeywardena K.G.

Jayasundara H.L.S.H

Karunasena G.K.S.R.

Sumanthiran S.K.

July 2020

Abstract

Object Detection, Tracking and Suspicious Activity Recognition for Maritime Surveillance using Thermal Vision

Group Members: Abeywardena K.G., Jayasundara H.L.S.H., Karunasena G.K.S.R., Sumanthiran S.K.

Department of Electronic and Telecommunication Engineering

Supervisors: Dr. Peshala G. Jayasekara, Dr. Ranga Rodrigo

Keywords: Thermal Vision, Object Detection, Activity Detection, Tracking, Suspicious Activity

In a world of a globalized economy, maritime surveillance is a crucial element. Today maritime transportation is considered to carry more than 90% of the world trade when it comes to long distance transporting methods. Due to this rapid growth in sea traffic, security and safety have risen as key issues. Along with that, real time detection of maritime activities has become essential to monitor and control fishing activities, smuggling detection, human trafficking, and maritime pollution. Sri Lanka is a country where most of the coastal families depend on a daily wage incurred by fishing and the safety of these fisherman is a crucial factor not only to themselves but also to their families. Apart from that during the recent past, Sri Lanka also recorded its two largest drug busts within a period of one month within its maritime borders.

With this project we propose a system not only to detect objects within the surveillance area but also a one which is capable to early detect pre-identified suspicious activities happening within the borders. We believe this will be an ideal replacement to the current system available which is run by manual labor to detect both objects and classify activities as suspicious or not. With the detection of any such suspicious activities, the system is capable of alerting the relevant authorities real time and with that traditional method can easily be overpassed with an additional benefit of increasing the safety of the marines. One key objective of this project is to be able to detect both objects and activities happening during any time of the day. Hence, thermal imagery is used for development of the models and for real-time detection.

Many of the currently available systems are limited to object detection in marine environments yet along using RGB imagery while activity detection and object tracking as a maritime surveillance is not a very common area of research. With this project, we propose a novel deep learning solution which is capable of object detection, activity detection, tracking and early identification of suspicious activities using thermal images in maritime environments.

The final solution will run on an inference hardware where the thermal camera will be controlled by a dedicated platform which is capable of assisting the camera to align horizontally and also to track the detected objects within the target range.

Table of Contents

List of Figures

List of Tables

Acronyms and Abbreviations

ISP – Inertial Stabilized Platform

LOS – Line of Sight

PID – Proportional Integral Derivative

RANSAC – Random Sample Consensus

MEMS – Micro-Electromechanical System

SSD – Single Shot Detector

YOLO-V3 – You Only Look Once Version 3

R-FCN – Region based Fully Convolution Networks

GTAN – Gaussian Temporal Awareness Networks

TRN – Temporal Recurrent Networks

RED – Reinforced Encoder-Decoder Network

STEP – Spatio-Temporal Progressive Learning

IMU – Inertial Measurement Unit

DOF – Degrees of Freedom

DOM – Document Object Model

Chapter 1

INTRODUCTION

In the recent past we have seen an increase in illegal activities within the maritime borders of Sri Lanka. It has come to a place where we see either drug busts or illegal fishing activities or foreign immigrants trespassing in news every day. In order to monitor and prevent these illegal activities Navy personnel have to patrol Sri Lankan marine borders through day and night continuously. Their task is not only limited to detect any floating or evasive objects but also to classify them as suspicious or not. Due to the fatigue induced by the monotony of tasks involved in surveillance in maritime environments vessel crew can make errors and might lose the traceability along with ability to identify such suspicious tasks. Hence, this conventional method can be ineffective due to the heavily dependency on manual labor and can be quite expensive due to the repetitiveness of the task. Apart from efficiency and accuracy factors, continuation of such surveillance actions via manual labor can become a major security threat to the Navy personnel involved. Also, due to the monotony of the work, the surveillance routes can be identified and interested parties in such illegal activities mentioned above will try to avoid these routes as much as possible. With that in mind, our main objectives in this project can be summarized as follows.

With this project, we are introducing a novel solution to automate all the above-mentioned tasks by a deep learning solution and a stabilization platform along with a control system to control our camera and the whole system attached to it. As the first step towards automating the traditional tasks, one of our key objectives is to detect pre-defined objects using a thermal camera in seawaters. As explained earlier, we want our system to work all day long irrespective of time. In order to get informative and descriptive imagery at all times of the day, we plan to use thermal imagery for detection purposes when it comes to both objects and activities. With that, detecting suspicious activities occurring in any given area in the sea is another objective in this project. Activity detection is a very common research and used areas when it comes to deep learning. Yet classifying these identified activities into suspicious class is a novelty we are bringing along with this project. Upon detection of such activities, with the help of the fully proposed architecture which is a combination of a control system, a stabilization platform and a thermal camera, these suspicious actions will be tracked while alerting all the relevant authorities real-time. The final objective identified for this project is to create an annotated marine thermal dataset for future research purposes.

Apart from the already mentioned novelties, the proposed system would be the first of its kind to attempt to automate maritime surveillance within the borders of Sri Lanka. Also, this will be the first-time detection of suspicious activities in maritime environments will be performed using thermal images.

Looking at this project, we have identified four potential applications and areas where this can be used or further developed. The Navy and the Coast Guard are two highly potential customers that can benefit from the system as it is. With the fully user-friendly interface incorporated with the deep learning solution we provide as the end result, all of the above identified suspicious activities will be displayed to interested parties at both Navy and the Coast Guard. We also believe, the stabilization platform along with the control system we design can be used by photographers to mount their camera by replacing the thermal camera and use it for their own purposes in photography and videography. Finally, due to its early inception in the research area, we believe our system would be a guide to anyone who is ready to work with suspicious activity detection using thermal imagery in maritime environments.

This report consists of 4 chapters. In chapter 2, we will look at similar work carried out in each of the key areas in our project ranging from object and activity detection to stabilization platforms and control systems. In chapter 3, we will introduce our proposed architecture along with possible alternatives and other areas related to the methodology. Chapter 4 will be a discussion along with a conclusion where we will discuss the main findings in the literature survey along with a feasibility study on several aspects. In this chapter, we will also look at the impact that this project creates in terms of both locally and globally.

Chapter 2

LITERATURE SURVEY

In Chapter 2, we will discuss about the related works that have been carried out in, in the order of, Object Detection using Deep Learning, Activity Detection using Deep Learning, Thermal data for Deep Learning, Stabilized Platform Design, Control System Design for the Stabilized Platform and Digital Video Stabilization.

2.1 Object Detection using Deep Learning

Deep Learning based object detection has recently become highly popular owing to its high accuracy compared to contemporary image processing methods [1], [2], [3]. Deep learning-based approaches can be divided into two major areas, namely 2-stage detectors, and single-stage detectors. It is widely accepted that while two stage detectors such as [3] can achieve higher accuracy, they are slower at inference time, and therefore single stage detectors such as [1], [2] are more popular for real time applications. While region proposals [3], [4] and bounding box prior proposals [1], [2] have been widely used in the past, recently key-point detection has been utilized to achieve higher detection accuracy for the same model complexity such as in [5], [6], and [7]. Deep learning-based object detection using thermal images is an area that is not very widely explored. Therefore, our work will explore the use of both types of object detection approaches on thermal images.

2.2 Activity Detection using Deep Learning

The area of activity detection has experienced greatly improved performance in the recent years owing to the introduction of deep learning-based approaches. Various methods have been proposed to carry out temporal video processing using neural networks including, but not limited to: 3D inflations of 2D-Convolutional Neural Network (CNN) algorithms [8], using information from optical flow [9], [10] and Long-Short-Term-Memory (LSTM) cell-based architectures [8]. Similar to image classification vs. object detection, there are two types of temporal video processing: activity recognition and activity detection. Activity recognition seeks to identify the action taking place in an entire video, such as in [11] whereas activity detection seeks to localize that activity in time, such as in [9] and [12]. Activity detection may work on a frame level [8], or an object level [10], which requires not only temporal localization of the activity, but also spatial localization. Additionally, while most contemporary work like [12] focuses on post processing in order to identify activities, there are some more recent works which looks at online, real time activity detection [10]. For our application, we require online real time activity detection. We will explore both temporal-only localization [8], and also spatio-temporal localization [10], as each has its advantages and disadvantages for our application.

2.3 Thermal Data for deep learning

All of the algorithms that we will be using will be run on thermal data. While there has not been a lot of work done on thermal data using deep learning, recent literature demonstrates that using thermal data for object detection and activity recognition is possible and produces results which are better than those of existing statistical methods.

Due to the simplicity of thermal features, traditionally statistical methods have proven sufficient to carry out object detection on thermal images. Popular methods such as C-means clustering for object detection have been used on thermal images in [13] and have produced good results. Pixel intensity histogram-based segmentation has also been utilized to great success [14]. However, in the recent past, using small neural networks for object classification [15] has been successful, and since then, the use of extreme machine learning for object detection has been developed in [16], and therefore it has been shown that the use of deep learning-based computer vision algorithms on thermal data outperforms that of existing statistical methods. It is worthwhile noting that [16] deals specifically with maritime objects, namely, ships.

For the more challenging task of activity recognition, [17] has shown that Convolutional Neural Networks that are used for activity recognition on RGB images can easily be trained and used for activity recognition on thermal images as well. These works show the effectiveness of deep-learning based computer vision algorithms on thermal data, which our project will utilize extensively, and develop upon.

2.4 Stabilized Platform Design

To compensate for the undesired movements due to the motion of vessel against the tides and micro vibrations of the vessel itself, self-stabilized platform is required to mount the camera on. Further the platform needs to be externally controlled based on the tracking information provided by the deep learning algorithms. Inertial Stabilized Platforms (ISP) stabilize and hold or control the line-of-sight (LOS) of one object relative to another object or inertial space. Usually ISPs are integrated into tracking or pointing system [18], [19], [20] where in our case it is rather a tracking system than a pointing system such as a laser targeting. In the domain of ISP design for optical imaging systems [19], two fundamental objectives should be fulfilled: obtaining good quality images of the targets or target regions and determining the location of the target with respect to a prescribed frame of reference. To achieve the two fundamentals, an electromechanical structure (Gimbal) needs to be designed based one of the two main approaches in [19]: Platform Stabilization and Steering stabilization.

During platform stabilization, the image sensor that needs to be stabilized is mounted on the Gimbal and controlled using direct forces (mass stabilization) as opposed to steering stabilization where the imaging sensor is mounted on the moving vehicle while controlling a mirror placed in front of it to stabilize and track the objects (mirror stabilization [21]). Apart from these stabilization methods, momentum wheel stabilization [22], feedforward techniques [23] are also used where no feedback gyros are required in stabilization. These [23] and [22] have been used prior to the availability of instrument gyros for stabilization hence the current commercial utilization of them is limited. Since the objective of having a platform for our project is to mount the camera and stabilize LOS and/or track objects using instrument gyros, mass stabilization is the most suited option.

Based on the design specification in [18], [19] for multi-axis gimbal design, the ability to control large loads (~3kg) requires powerful motors at each axis to be controlled and a complex electromechanical design that can handle the large torques created due to the weight displacement of the camera mounted on it when the base of the platform is rotated with respect to the reference inertial frame as the structures that controls the imaging sensor around each axis acts as independent gimbals.

When designing ISP for imaging sensors, parallel manipulators pay a vital role as well. Stewart platform [24], [25] is the first and most famous parallel manipulator on which many researchers have worked on during the last few decades. This is a 6-dof mechanism that consists of two rigid bodies (referred as base and platform) connected through six extensible legs, each with spherical joints at both ends or with spherical joint at one end and with universal joint at the other [24]. The Stewart platforms are usually designed with hydraulic prismatic actuators which are large in size and is capable of handling larger loads [26]. With the improvement of technology, the hydraulic prismatic actuators have been replaced with the fixed electrical rotary actuators [27], [28].

Compared to the Stewart Platforms with prismatic actuators, the Stewart platform with rotary actuators can be made small in size, with higher bandwidth and handle more challenging motions and control problems. Moreover, the platform can be designed with lower cost compared to the platform with prismatic actuators [28]. Further compared to the serial manipulators, solving inverse-kinematics for the Stewart Platform [29], [30] is easier compared to the solving of direct-kinematics that facilitates adjusting the orientation of the platform for a desired orientation based on the tracking and stabilization loops easy.

Since our solution requires 3-axis stabilization and to control a load of 3kg with a height of 22cm, serial manipulator-based ISP designs are not feasible. Even though the Stewart platform with rotary actuators addresses these issues, the bandwidth of the motions compensated is less especially with yaw which is required to have a 360-degree continuous movement. Hence, we will explore certain modifications to the original rotary Stewart platform design to increase the bandwidth of motion specially to allow for 360 degrees of yaw when achieving our final design.

2.5 Control System Design for the Stabilized Platform

Stabilization of the platform is not achieved solely from the electromecanical design described in Section 2.4. Instead, the control system desiging plays a vital part in stabilizing the designed electrmechanical platform against unwanted motions as well as to achieve an orientation with minimum response time and minimum overshoot.

[31] describes a two axis gimbal mounted on an Unmanned Aerial Vehical (UAV) controlled using image processing algorithm, where the algorithm continuously provides the centroids of the target identified to create the error signal to fed into the simple Proportional-Integral-Derivative (PID) control loop. This system is ideal to keep a target idenitified at the center of the LOS of the camera but does not provide a mechanism to stabilize the LOS with respect to inertial reference frame when a target is not identified by it. [18], [19] describes the control system design for stabilization based on feedback gyro which uses both orientation of the platform (angle) and the rate of change of angle of the gimbal. Further [18] describes about how a multi-axis control system can be designed for the multi-axis ISP design which is having independent

control loops for each gimbal facilitating the stabilzation around specific axis of rotation.

[32] describes the general overview for control system design. As described in [32], [18], [26], when designing and simulating the performance of the control system, the dyanmics of the gimbal, motors and gyros should be considered to model them mathematically.

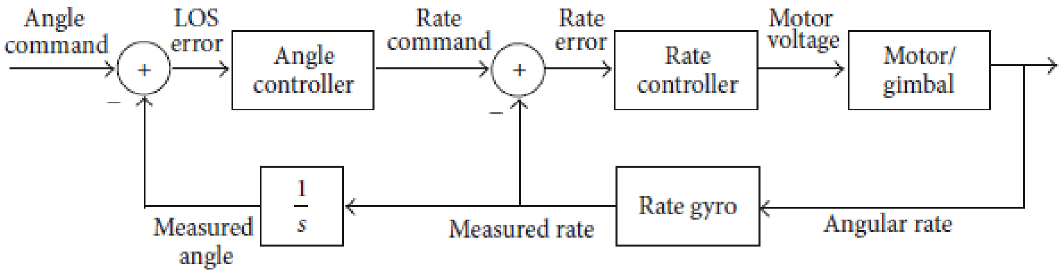

Fig. 2.1. Control System for Stabilization

[32] describes several critical factors to look for when selecting an instrument gyro as the mass stabilization solely depends on the feedback given on the platform orientation to the control loop. Based on the basic control system in Fig. 2.1 as proposed by [32], the Angle Controller consists a PI^2 controller while Rate controller consists a PID controller to increase the stability of the tracking loop and the compensation loop. The desired orientation of the gimbal, given as an angle command, and the measured orientation from the rate gyro creates the error signal that is to be compensated and controlled by the PI^2 controller and generate a Rate command. Along with the measured rate, the rate error is calculated and fed into PID controller which outputs a voltage signal the motor to control its rotation speed to ultimately achieve the desired angle of the gimble.

As described in section 2.4, this corresponds to the control of individual gimbals (inner and outer gimbals) that compensates the unwanted motion which is not feasible for our application.

Authors of the paper [32] use a single motor to rotate the gimble, but in our task we need to use several servo motors and as inputs, an angle command is required. If we are going to use this control system to control our stabilized platform, we have to come up with a methodology of controlling several servo motors and arms using Inverse Kinematics.

[33], [34] describes the control mechanisms researched with the parallel manipulators specifically with the Stewart Platform. Both these describes the control loops designed for the prismatic actuators where no publications related to the general control loop design is published for Stewart platform with rotary actuators. Instead of having a feedback from the rate gyro about the platform orientation, rate gyros are used at each prismatic actuator to measure the individual actuations. Once the desired position is given as a command to the control loop, it is converted into joint parameters through Inverse-Kinematics and the error signal for six actuators are created and fed into PID controller. Rather than using the Joint-space controlling mechanism in [33], [34], using the Cartesian space control [35] to measure the orientation of the platform limits the number of rate gyros needed in the design. Inspired from the designs described in [34], [33], [26], an adjusted control loop design is required for the controlling and stabilizing of the platform design for our project.

2.6 Digital Video Stabilization

Apart from the mechanical stabilization, digital stabilization can provide certain level of stabilization on a software level. Since some micro-vibrations may not be stabilized mechanically, having digital video stabilization in place can improve the video quality we receive.

As a solution for the presence of warping when capturing a video from a mobile device, we can capture the motion of the device while capturing a video and stabilize the video based on sensor data. [36] has carried out this method in three main stages: camera motion estimation, motion smoothing and image warping. It has used RANSAC algorithm, which is computationally light, to detect point feature matches. After comparing the results with famous image stabilization techniques, [36] demonstrates the work using a MEMS Gyroscope comes with iPhone 4. This method is tested on regular RGB images as well as for a handheld device. But in our case, we deal with thermal images in real time and with high frequency vibrations of the vessel. The main problem with this method is, this performs poorly in high vibrations.

Many video stabilization techniques we found were offline techniques as well as they require some additional sensors such as gyroscopes. On the other hand Mesh Flow [37] stabilization, stabilize the video with only one frame latency by path smoothing carried out on the *vertex profiles* (Motion vectors obtained at the same location in the Mesh

Flow). The smoothing is carried out using *Predicted Adaptive Path Smoothing* (PAPS), which is a novel technique uses only past frames. Based on these reasons we decided to use the Mesh Flow method as our digital stabilization technique.

Chapter 3

METHODOLOGY

In Chapter 3, we start our discussion by introducing the Proposed System Architecture. Along with that, an analysis of Possible Alternatives is described on every aspect mentioned in the System Architecture. Then a Risk Analysis is performed on the project covering every aspect ranging from Technical, Feasibility, Sustainability to Health factors. The chapter concludes by describing the Budget, Task Allocation, Timeline and Initial Work.

3.1 Proposed System Architecture

3.1.1 Components

The system architecture that we propose consists of the following main components:

- A thermal camera
- An object detection algorithm (to detect object in maritime environment)
- An object tracking algorithm (to track the identified objects in successive frames)
- An activity detection algorithm (to detect the suspicious activities in maritime environment)
- A user interface software
- A platform (to mount the camera and perform active tracking)
- A control system (for stabilization and tracking using the platform)

3.1.2 Pipeline Design

The components are connected as shown in Fig. 3.1 into one pipeline that is to run in real time. The stages are as follows:

- The thermal camera captures and feeds a sequence of frames into the system
- On each of the frames fed by the thermal camera, the object detection algorithm runs and detects objects and sends the information of object locations (bounding box coordinates) to the object tracking algorithm
- The object tracking algorithm keeps a track of all objects in the scene and sends signals via the control system to the platform in order to actively track the objects using the camera.

- The tracked objects and the temporal information are sent to the activity detection algorithm to detect whether any actions are taking place and if so whether they are suspicious or not.
- The outputs of each of these algorithms is sent to the user interface to be recorded and monitored

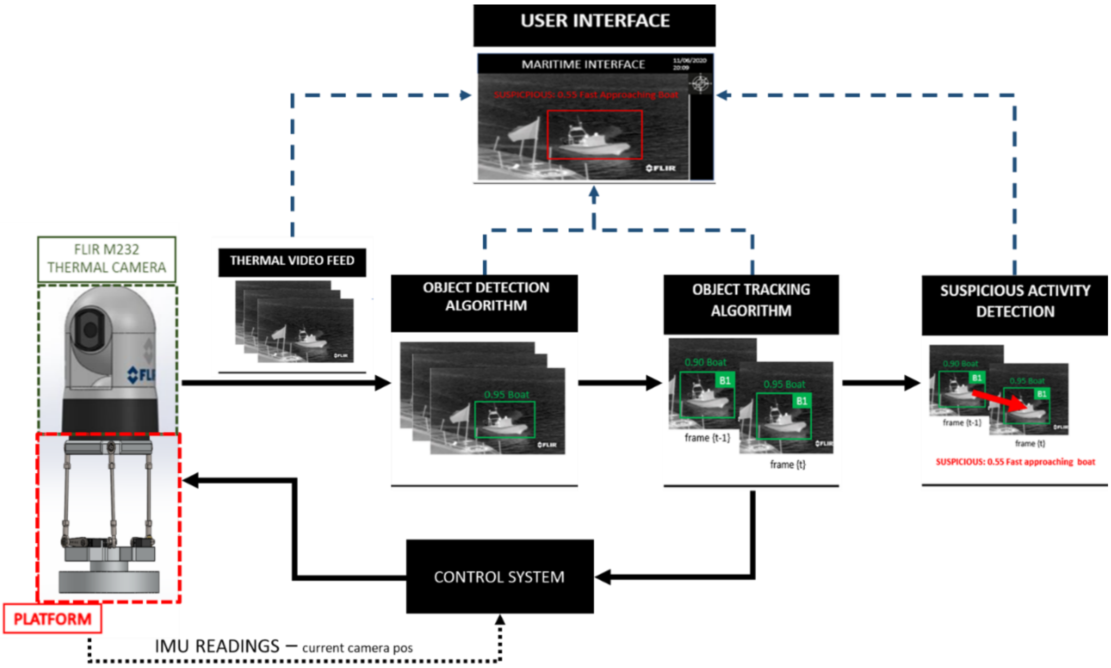

Fig. 3.1. Proposed System Architecture

3.1.3 Detection Targets

There are 2 sets of detection targets that have to be defined in the project. The first is Objects of Interest (OI), and the second is Activities of Interest (AI).

In general, for any given project, these targets have to be defined in discussion with all related parties, as they are applications specific. For our application, they must be defined in discussion with the Navy. However, we have been able to create a preliminary list for testing purposes which can be refined and confirmed as we move forward.

Our defined OI's are as follows:

- Boats (Small, Medium, and Large)
- Humans
- Unusual floating objects

Our defined AI's are as follows:

- Unauthorized Fishing (whether it is authorized or unauthorized can be determined using other information such as location, time, method etc.)
- Human trafficking (This can be determined as a function of the number of detected humans on a vessel and the vessel size)
- Dumping objects in the sea
- Loitering
- Unusual maneuvers (such as evasive maneuvers or high-speed approaches)

3.2 Possible Alternatives

The following describes the possible alternatives for each component in the pipeline design shown in section 3.1.2.

3.2.1 Object Detection Algorithms

With a vast variety of algorithms to choose from, there are a few that can be shortlisted as having provided state-of-the art results over the past years. These alternatives are compared in Table 3.1.

Table 3.1 Object Detection Algorithms

	SSD512 [1]	**YOLO-V3** [2]	**R-FCN** [4]	**CornerNet** [5]	**CenterNet** [7]	**CornerNet-Lite** [6]
Backbone	VGG-16	DarkNet-19	ResNet-101	Hourglass-104	Hourglass-104	Squeezed Hourglass
Single/ Double stage	Single	Single	Double	Single	Single	Single
FPS	19	45	6	1	1	60
mAP%	28.8	21.6	29.9	37.8	41.6	34.0

Our major requirements for the object detection algorithm are high accuracy and real-time inference. Many of the 2-stage algorithms have a high inference time and therefore are not suitable. Of the single stage algorithms, YOLO-V3 [2] and SSD [1] use bounding box prior proposals which need to be manually defined in order to provide good performance. Additionally, the algorithm needs to evaluate many thousands of these boxes to have high accuracy, which increases latency. A newer technique is to use key-point detection to construct bounding boxes. Here, there is no need for a manual definition of possible object sizes and shapes, and less key points need to be evaluated each forward pass, providing high accuracy with reduced latency. Therefore, a good selection for our project is the CornerNet-Lite [6], which is an optimized version of the CornerNet [5] and provides high accuracy predictions in real time.

3.2.2 Activity Detection Algorithms

Deep learning algorithms for activity detection are relatively new, owing to the higher complexity of the task. However, the algorithms that have been developed have proven to produce very good results. Table 3.2 shows a comparison between the state-of-the-art algorithms for activity recognition and activity detection.

Our application requires that the algorithm we select is able to work in real-time and use only past information. It also must be able to isolate an activity in time (temporal localization) and ideally in space as well (spatial localization). We selected [10], which satisfies all of these requirements.

Table 3.2 Action Detection Algorithms

	GTAN [12]	**TRN** [8]	**RED** [9]	**STEP** [38]	**Online Real-time Multiple Spatio-temporal Action Localization and Prediction** [10]
Temporal /Spatio-temporal	Temporal	Temporal	Temporal	Spatio-Temporal	Spatio-Temporal
Backbone	Pseudo-3D	VGG-16/ ResNet-200	VGG-16/ ResNet-200	VGG-16	VGG-16
Online/ Offline	Offline	Online	Online	Offline	Online
FPS	8	24	24	21	28
Dataset	THUMOS’14	THUMOS’14	THUMOS’14	UCF101	UCF101
mAP%	38.8	47.2	45.3	75.0	43.0

3.2.3 Inference Hardware

A significant challenge is to run the entire pipeline, which has multiple components, in real time. In order to do this, we need sufficiently powerful hardware: specifically, a Graphics Processing Unit (GPU). We evaluated several different GPU’s. Our main criteria are that the hardware should have enough RAM to load and simultaneously run all the models that we are using in the pipeline. The other considerations were price and power consumption. We selected the GeForce GTX 1660Ti because despite its low cost, it satisfies all other requirements. All the alternatives are compared in Table 3.3.

Table 3.3 Inference hardware

	GeForce GTX 1660 Ti	GeForce RTX 2060	GeForce GTX 1650 Ti	Radeon RX 590
Cuda Support	Yes	Yes	Yes	No
RAM (GB)	6	6	4	8
Memory Bandwidth (GB/s)	192	336	128	256
Power (W)	120	160	75	175
Cost ($)	280	325	150	180

3.2.4 Platform Design

To facilitate the mounting of the camera on to the vessel and actively track the objects using the information provided by the object tracking algorithms, we explored several alternative platform designs. When selecting the best suited design for our project out of the alternatives we looked at the structural stability, mass of load which the platform can handle, positioning accuracy, scalability of the platform to fit into our interests, cost of building, speed of action and bandwidth that each platform design is capable of compensating. The comparison in terms of these features were summarized in Table 3.4 and are based on the recent blog posts of implementation of platform designs [39], [40], git-hub repos [41] and literature survey described in section 2.4.

Table 3.4 Platform Designs

	Serial-Link based Gimbal	Stewart Platform (Prismatic actuators)	Stewart Platform (Rotary actuators)	Delta Robot
DoF	3	6	6	6
Stability	Moderate	High	Moderate	Moderate
positioning accuracy	Low	High	High	High
Size	Small – Medium	Large	Medium	Small-Large
Cost	Low	High	Moderate	Moderate
Mass of load	Low	High	Moderate	Low
Speed of action	Moderate	Low	High	High
Bandwidth of compensation	High	Low	Moderate	Moderate

Since our application requires handling of a load with moderate mass (~3kg) and high positioning accuracy, Serial-Link based Gimbal and the Delta Robot designs were found not suited. Since the Stewart Platform designs with Prismatic actuators are usually designed to handle very large loads, usually the size of them are quite big and they use high powered hydraulic or electrical actuators which in turns increases the cost of building it. Hence the Stewart Platform with rotary actuators design was selected with possible improvements to be incorporated to achieve the final platform design for our project.

Fig. 3.2. Selected platform design - Rotary Stewart Platform

3.2.5 Hardware – Servo Motors

As discussed in section 3.2.4, we chose rotary Stewart Platform as the basis for our platform design. Classical rotary Stewart platform consists of 6 servo motors to change the orientation of the platform by adjusting the roll, pitch and yaw angles. It is critical to evaluate alternative servo motors in terms of maximum torque, speed, voltage and the cost when selecting the most suited servo motor for our project. The evaluation among the most popular industrial servo motors in terms of the features mentioned was carried out and the results of the top three servos are shown in Table 3.5.

Table 3.5 Industrial servo motors

	RKI - 1211	Dynamixel – AX12A	DS3218
Gear Material	Metal	Metal	Metal
Stall Torque	14 kgf.cm (4.8V) 16 kgf.cm (6V)	15.3 kg.cm (12V) 212 oz.in	18 kgf.cm (6V) 20 kgf.cm (7.2V)
Operating Speed	0.18 s/60^0 (4.8V) 0.14 s/60^0 (6V)	59 RPM 0.169 s/60° (12V)	0.14 s/60^0 (6V) 0.16 s/60^0 (7.2V)
Operating Voltage	4.8V – 6.0V	9V - 12V	4.8V-7.2V
Dead Bandwidth	5μs	2μs	3μs
Weight (g)	48	55	60
Dimensions (mm)	41 x 20 x 36	50 x 26 x 37.5	40 x 20 x 40.5
Cost ($)	22	47	12

Since the load that is required to be handled is 3kg, sufficiently high torque is required by the motors in order to change the orientation of the platform with the camera mounted on it. Further, by considering the selected platform design, minimum of 6 servos are required to build it which makes cost an important feature as well. Further the operating speed is also critical as platform will need to correct its orientation due to the unwanted motions (due to waves) to keep its LOS focused on the target it tracks. Thus, Dynamixel AX12A and DS3218 motors are close candidates for our project. Even though the Dynamixel AX12A motor is more professional and more popular among robotics field, DS3218 was selected for our project factoring the cost and the complexity on communication with the motor in.

Fig. 3.3. Selected servo motor - DS3218

3.2.6 Hardware – Inertial Measurement Unit (IMU)

Since the design of the control system of our platform depends on feedback from IMU, selection of a proper IMU for the project plays a critical role. There are different types of IMU sensors available on market which are differentiated by key features such as

Angle Random Walk (ARM), Sensitivity and Non-Linearity. Table 3.6 shows the comparison we did to choose the most suitable IMU for our task.

Table 3.6 Inertial Measurement Units

	MPU6050	**MPU9250**	**Mini IMU 9**
Sensitivity $(\frac{LSB}{deg}/s)$	131, 65.5,32.8, 16.4	131, 65.5,32.8, 16.4	114.29, 57.14, 14.29
Non-Linearity (%)	0.005	0.01	0.03
ARW $(\frac{deg}{s}/\sqrt{Hz})$	0.2	0.1	0.2
Cost ($)	3	7	15.95

Based on the results, Mini IMU 9 was selected for our project as it has reasonable Non-Linearity, ARW values. Even though it is pricier compared to the other two options, Mini IMU 9 consists a magnetometer other than the usual gyroscope and accelerometer and is vended by “Pololu” which is a renowned vendor for electronic equipment.

Fig. 3.4. Selected IMU - Mini IMU 9

3.2.7 Hardware – Microcontroller

Microcontroller is like the brain for the stabilized platform where the running of control loop, communication with the mainframe and if needed running of digital video stabilization algorithms take place. For the selection of suitable microcontroller for our application is first evaluated based on cost, computations power and the ability to communicate over a network. Table 3.7 shows the comparison of possible alternatives for a microcontroller.

Table 3.7 Microcontrollers

	Arduino Mega	STM32	Raspberry Pi 3B+
Cost (relatively)	Low	Low	Moderate
Computational Power	Low	Moderate	High
Most Optimized Task	Logic Programming	Logic Programming	General Single Board Computer
WiFi and Bluetooth	No	No	Yes
Power Consumption	~0.5W	~0.5W	~5W

When selecting the best out of the three options, we looked for a low cost, high computational power, and the ability to communicate with camera and mainframe regularly without overwhelming as the key features. Based on the above comparison we choose Raspberry Pi 3B+ which is the most powerful microcontroller with lots of capabilities.

Fig. 3.5. Selected microcontroller - Raspberry Pi 3B+

3.2.8 Software and User Interface

Building a software can be done in 2 main ways: Web based application and Standalone application. Here we discuss our possible alternative frameworks that can be used for development of backend and frontend of both web-based application and standalone application.

Backend is the part where we implement all our Deep Learning (DL) algorithms for doing real time prediction. To make the operation happen in real time we need a framework which is simple and fast. We implement our entire DL algorithm uses python which is the most popular language in the field now a days. Since it is important that the backend supports our algorithms and we choose a python based simple framework for the backend among alternatives such as Django [42], ExpressJS [43], Phoenix [44] etc. We choose the framework Flask [45] because it is simple, easy to build prototypes, flexible and more importantly it is lightweight.

As the frontend framework we selected React [46] which is the most popular and easy to learn framework. More importantly the virtual Document Object Model (DOM) stores the page and do only necessary changes when the backend changes without changing everything entirely. This makes the React faster than any other frontend framework. More importantly, React is ideal for applications involving high data traffic. Good documentation, large community and high performance can be stated as the reasons for why we are choosing this framework.

As a method of building a standalone application we decided to use PyQt [47] which is the python-based Graphical User Interface (GUI) building library. Basically, we choose this, because our entire codebase is written in python and it will be easy to integrate if we use a python-based method to build our final application.

3.3 Risks Associated with the Project and Risk Management Plan

Table 3.8 Risks and mitigation methods

Type of risk	**Possible Risk**	**Method to mitigate the Risk**
Technical	May not be able to use pan and tilt abilities of the camera	Controlling the orientation of camera using the designed platform
	Stabilization of platform will not be possible entirely from Gyro-PID based mechanical stabilization	Using digital video stabilization alongside with mechanical stabilizer
	Workspace volume and bandwidth of compensation of platform design not sufficient	Designing and simulating multiple platforms designs with different dimensions and different restrictions (DoF)
	Unable to run the currently proposed pipeline in real time	Use smaller, faster algorithms and use other techniques to improve accuracy of the model
	Documented motor and servo parameters may differ from their actual performance parameters	Perform physical testing of the components and adjust parameters accordingly
Financial	Reordering cost for failing components.	Order with a buffer margin at first

	Cost limit for the project exceeding the university stipend.	Self-funding
	Delivery cost of different vendors	Selecting a vendor based on the price of the component and customer feedback.
Sustainability	Dataset collection on Sri Lankan sea with sufficient level of suspicious activities may take longer time	Use publicly available maritime datasets on foreign seas to develop the algorithms and system Use datasets already collected by SL Navy using their own thermal cameras
	Components may not arrive on time and depreciated components	Order components early and check whether they are in good quality
Health (Due to COVID 19)	Received packages from overseas may not have adhered to the safety measures from the vendors	Sanitization of the packages and selecting the vendors more carefully
	Getting firsthand experience with SL Navy about the motion and vibrations that needs to be compensated is not possible due to the current context.	Contact personnel from SL Navy through a Zoom meeting and get their input when the situation is better

3.4 Budget

Table 3.9 Proposed Budget

	Item	**Unit Price($)**	**No of Units**	**Total ($)**	**Total (LKR)** ($ 1= LKR185)	**Sponsored**
01	FLIR M232 Thermal Camera	3499	1	3499	647315	Yes
02	DS3218 Servo	11.20	8	89.6	16576	No
03	Connecting Rods (4 per set)	9	4	36	6660	No
04	2s1p 5200mAh LiPo Battery	40	1	40	7400	No
05	Mini IMU 9 - Pololu	16	2	32	5920	No
06	Raspberry Pi 3	40	1	40	7400	No
07	Base and top platforms Laser Cut (approx.)			28	5,185	No
08	GPU	300	1	300	55,500	Yes
	Total			3799	702815	Yes
				265.6	49136	No

Total Budget = LKR 751,951 (Non-Sponsored = LKR 49136)

3.5 Task Delegation Among the Group Members

Table 3.10 Task delegation among members

Task Group	Responsible Person	Tasks	Delegation			
			Kalana	Sakuna	Sachira	Shechem
Object Detection and Activity Detection	Shechem	Data Preprocessing		✓	✓	✓
		Data Annotation	✓	✓	✓	✓
		Algorithm Development and Training	✓	✓	✓	✓
		Object Tracking	✓		✓	✓
Platform and Control System Design	Sachira	Platform Design	✓		✓	
		Simulation	✓		✓	
		Prototyping	✓	✓	✓	✓
		Control System and Programming	✓	✓		✓
Integration	Kalana	Integration of Hardware and Software	✓	✓		✓
Software Development	Sakuna	Backend Development		✓		✓
		Frontend Development		✓	✓	

3.6 Timeline

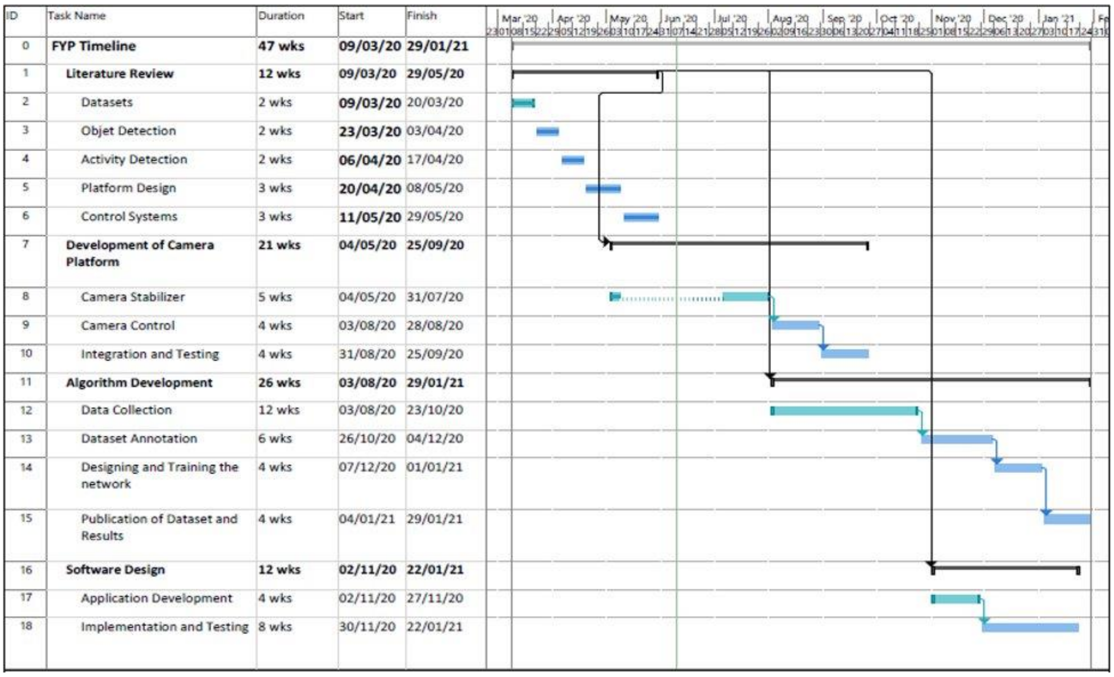

Fig. 3.6. Proposed Timeline

3.7 Initial Results

We were able to obtain initial results for the following components of the proposed pipeline in section 3.1.1.

3.7.1 Platform design and simulation

Based on the design selected in section 3.2.4, we started off with the hand-sketches of the base and platform of the rotary Stewart platform with desired parameters based on the dimensions of the thermal camera which we are supposed to use. Using MatLab 2016a [48], the initial calculation of bandwidth of compensation for each axis of rotation was done to identify the possible ranges that can be compensated by our design.

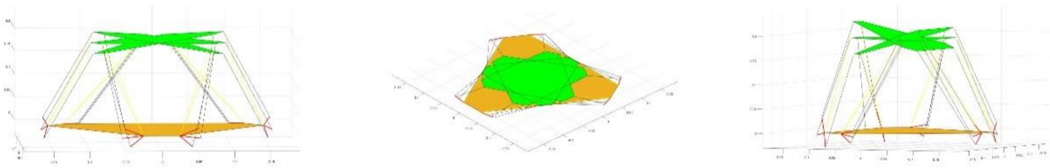

Fig. 3.7. Initial results from MatLab on Roll, Pitch and Yaw

The hand-sketches were then designed using the computer-aided design tool SolidWorks 2017 [49]. We started off with designing a 6 DOF Stewart Platform that is suited for our thermal camera dimensions. For this initial stage, the designs were made without considering industry standards and only to be used for the purpose of simulation. Then from the initial stage, the design was modified using off the shelf equipment when it comes to all servo motors, servo horn, ball joints, servo brackets and threaded rods.

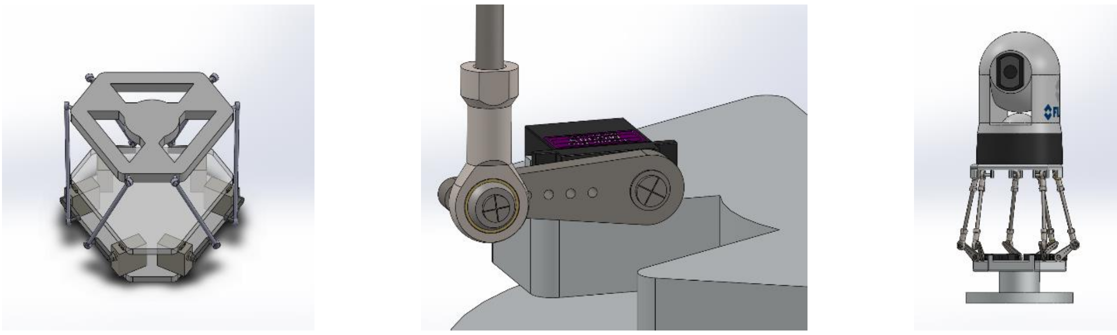

Fig. 3.8. Development of the 6-DOF Stewart Platform Design

Due to the bandwidth limitations of compensation with 6-DOF design and also due to its design complexity in terms of platform dynamics, we opted to explore a 3-DOF version of the Stewart Platform. Just like the development of the 6-DOF, we started from scratch and designed a 3-DOF version as well.

3.7.2 Control System Design

Inspired by the work on the control system design for prismatic Stewart platform, we designed control system for the rotary Stewart platform. One deviation that was incorporated into our design is, instead using joint-space control loop design, we used cartesian space control loop design as it is much convenient for our application. The Fig. 3.9 shows the theoretical design of the control loop which needs to be tested to verify its functionality.

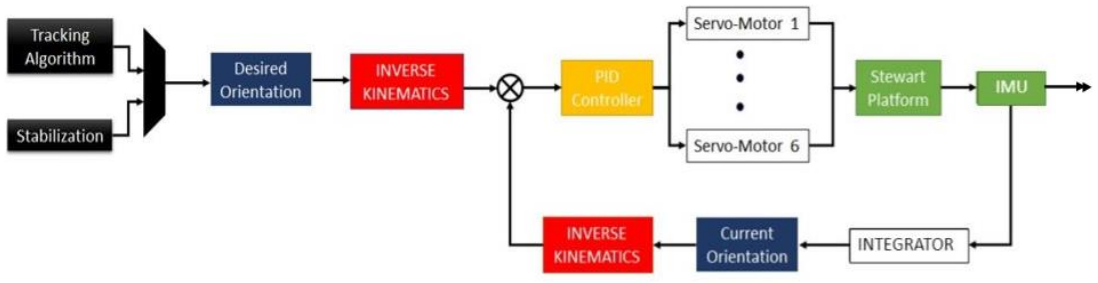

Fig. 3.9. Theoretical design of control system for rotary Stewart platform

3.7.3 Object detection algorithms tested on thermal images

As the published object detectors are based on the RGB images instead of thermal images, it is likely that the pretrained models provided may not perform well on thermal images as the two distributions are different from each other. Since there are no publicly available annotated thermal dataset for maritime environment, we use FLIR thermal dataset [50] related to Autonomous Vehicles to explore the performance of some selected object detectors.

As expected, the results were poor when the inferencing was done without performing transfer learning as the weights learnt by the networks corresponds to a totally different distribution of data compared to the thermal images. Since the dataset is not large enough to train an algorithm from the scratch, we performed transfer learning by training the pre-trained object detectors on a subset of the FLIR thermal dataset. The inference results based on the remaining images of the FLIR dataset is shown in Table 3.11.

Table 3.11 Inference results with FLIR thermal dataset

	Faster-RCNN	**Yolo-V3**	**R-FCN**	**CenterNet**	**CornerNet-Lite**
FPS	5	45	6	1	60
mAP %	65	85	78	88	81

Chapter 4

DISCUSSION AND CONCLUSIONS

In Chapter 4, we analyze the related works described in Chapter 2 and lists down the key-finding relating to each factor mentioned in System Architecture. Along with that, a feasibility study is carried out in terms of Financially, Technically and Socially. After a brief discussion about Global and Local Impact we move onto conclusions based on results from Chapter 3.

4.1 Summary of Literature Survey

Based on the Literature Survey done prior to the commencement of the project, the following summarizes the key findings.

- Even though there are several publications available on object detection using deep learning, no work has been published on object detection using thermal images. Since the thermal images are different in format to that of RGB images, using available state-of-the-art algorithms on thermal images can be challenging.
- Even though Activity Detection has been a recent topic in the deep learning domain, they have achieved promising results compared to the statistical approaches such as [12]. Majority of the published works corresponds to offline activity detection which uses both past and future information to predict any activity on the current frame. Online activity detection on the other hand is very challenging task which predicts the activities only using the past information. The results on online activity detection in [9], [10] shows that accuracy on detecting activities is lower compared to the accuracies of offline activity detection in [11], [38].
- Stabilization of the camera against the unwanted motion of vessels and tides is critical in surveillance as it is required to keep the LOS of the camera steady with respect to the inertial reference frame. Design of platform to achieve that is not merely designing an electromechanical structure but an integrated electromechanical structure with tracking loop and compensation loop. During the design of the compensation loop, fundamentals of control system designing need to be considered for different type of platform designs that achieves the stability and/or the desired orientation based on commands from tracking loop.

- Also the stabilization of the video fed by the camera is not necessarily be mechanical but can be digital as well. Since this is a real time application, digital video stabilization should be online and accurate. Digital stabilization can be used to suppress the micro vibrations of the boat as well as the oscillations due to tides and should be able to stabilize high frequency oscillations.

4.2 Technical Feasibility

- The conducted literature survey shows proof of concept for all individual components of our project. The question of technical feasibility is then dependent on the ability to integrate these systems and the specifics of the application that we are going to use the system for.
- Feasibility of integration of the individual components can be confirmed by the information gathered on similar projects which do both activity detection and object detection. Given sufficient hardware capabilities, we have already shown that such a system can be run in real time with good results. Therefore, from the viewpoint of system integration, the project is technically feasible.
- A large part of the novelty of our project deals with the type of data that we are working with, namely, maritime thermal video feeds. Preliminary work that we have carried out, and other work on maritime and thermal data in the literature [14], [16] show that using deep learning-based computer vision algorithms for these scenarios is feasible, and produces good results. Therefore, we can conclude that this component of the project is technically feasible as well.
- In conclusion, preliminary work carried out by us, in addition to current literature and existing applications in the market provide confirmation that the project is technically feasible.

4.3 Financial Feasibility

- Even though the estimated budget is higher than the allocated amount for a project, still the project feasible with self-funding. Aforementioned budget is a rough calculation and we will be able to borrow some of the expensive components for the project. Hence the budget can be reduced.
- Most expensive items (Camera and GPU) in the budget will be covered by a capital grant.

4.4 Social Feasibility

- Our project targets some specific social groups: SL Navy and Coast guards. Since these social groups are knowledgeable and have experience in working with surveillance systems and that our system is fully automated, they will require little or no training in adjusting to the new system from our project.
- Further specific components of our project such as the platform design and deep learning pipelines separately can be influential to social groups like photographers and researches where the profession photographers can use self-stabilized platform from our project when photographing in more challenging environmental conditions.
- Further our project does not go against any government rule, regulation or any kind of social belief and does not displease any social group. Hence the project is socially feasible.

4.5 Local and Global Impact of the Project

Our project holds greater importance in local context by assisting Sri Lanka Navy and Coast Guard to maintain a safe marine border around the country by detecting the suspicious and illegal activities that take place in it. Even though the large vessels of Sri Lanka Navy are equipped with thermal vision, our system flags suspicious activities which the Navy personnel will need to attend to. Further our project can be extended to facilitate the design of small autonomous surveillance boats which can be deployed into the sea and inform the nearest Navy vessels the coordinates of a detected suspicious activity from our system.

Further, the maritime surveillance has been a hot topic in many countries and organizations specially in France, Italy, European Union(EU) that many of these countries single-handedly or collectively deploy vessels to patrol the sea borders to prevent any illegal activities happening. Our project poses a great global importance as well as our method explores the way of automating the detection of suspicious activities and alerting the necessary authorities while performing the object detection and tracking which are the generic features given in more expensive maritime surveillance systems from vendors like FLIR. Having an end to end pipeline which can perform both object detection/tracking and action detection will inspire many researches in exploring the idea further.

4.6 Conclusion

Based on the initial results that we have obtained (section 3.7) the following conclusions were drawn.

- The Stewart platform with the parameters we used gives a modest range of compensation against unwanted oscillations for roll and pitch but fails to give 360-degree rotation. Since we do not have firsthand experience or an expert's view from a Navy personnel on the range of compensation required this may not be adequate in stabilizing the platform against unwanted oscillations. Hence, we concluded that certain changes are required to improve the range of compensation and should be incorporated into our final design of the platform.
- Using the FLIR thermal dataset, we were able to achieve promising results for the object detection using the state-of-the art deep learning algorithms. Hence, we concluded that the published deep learning algorithms work well with the thermal images and can be generalized well for the real-world datasets as well. However, since the resolution and the pixels per object in the images we use can make an impact on the detection of objects we decided to explore techniques on detecting small objects in thermal images.

In conclusion the project shall deliver an overall system containing an object detection and tracking algorithm specifically trained for maritime environment, action detection algorithm that can flag suspicious activities defined, a user interface and a self-stabilized platform along with a publication.

Chapter 5

REFERENCES

[1] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg, "SSD: Single shot multibox detector," *European conference on computer vision,* pp. 21-37, 2016.

[2] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv: Computer Vision and Pattern Recognition,* 2018.

[3] S. Ren, K. He, R. Girshick and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems,* pp. 91-99, 2015.

[4] J. Dai, Y. Li, K. He and J. Sun, "R-FCN: Object detection via region-based fully convolutional networks," *Advances in neural information processing systems,* pp. 379-387, 2016.

[5] H. Law and J. Deng, "CornerNet: Detection objects as paired keypoints," *Proceedings of the European Conference on Computer Vision,* pp. 734 - 750, 2018.

[6] H. Law, Y. Teng, O. Russakovsky and J. Deng, "CornerNet-Lite: Efficient Keypoint Based Object Detection.," *arXiv: Computer Vision and Pattern Recognition,* 2019.

[7] K. . Duan, S. . Bai, L. . Xie, H. . Qi, Q. . Huang and Q. . Tian, "CenterNet: Keypoint Triplets for Object Detection," *arXiv: Computer Vision and Pattern Recognition,* vol. , no. , p. , 2019.

[8] M. Xu, M. Gao, Y. Chen, L. Davis and D. Crandall, "Temporal Recurrent Networks for Online Action Detection," in *IEEE/CVF International Conference on Computer Vision (ICCV),* 2019.

[9] J. Gao, Z. Yang and R. Nevatia, "RED: Reinforced Encoder-Decoder Networks for Action Anticipation," 2017. [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr1707.html. [Accessed 30 6 2020].

[10] G. Singh, S. Saha, M. Sapienza, P. H. S. Torr and F. Cuzzolin, "Online Real-time Multiple Spatiotemporal Action Localisation and Prediction," *arXiv: Computer Vision and Pattern Recognition,* 2016.

[11] Y. Zhang, P. Tokmakov, M. Herbert and C. Schmid, "A Structured Model for Action Detection," in *Conference on Computer Vision and Pattern Recognition (CVPR),* 2019.

[12] F. . Long, T. . Yao, Z. . Qiu, X. . Tian, J. . Luo and T. . Mei, "Gaussian Temporal Awareness Networks for Action Localization," *arXiv: Computer Vision and Pattern Recognition,* vol. , no. , p. , 2019.

[13] V. John, S. Mita, Z. Liu and B. Qi, "Pedestrian detection in thermal images using adaptive fuzzy C-means clustering and convolutional neural networks," in *14th IAPR International Conference on Machine Vision Applications (MVA)*, Tokyo, 2015.

[14] J. Baek, S. Hong, J. Kim and E. Kim, "Efficient Pedestrian Detection at Nighttime Using a Thermal Camera," in *Sensors*, Basel, 2017.

[15] C. D. Rodin, L. N. de Lima, F. A. d. A. Andrade, D. B. Haddad, T. A. Johansen and R. Storvold, "Object Classification in Thermal Images using Convolutional Neural Networks for Search and Rescue Missions with Unmanned Aerial Systems," in *International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, 2018.

[16] A. Khellal, H. Ma and Q. Fei, "Convolutional Neural Network Based on Extreme Learning Machine for Maritime Ships Recognition in Infrared Images," in *Sensors*, Basel, 2018.

[17] M. Z. Uddin and T. J. , "A Deep Learning-Based Human Activity Recognition in Darkness," in *Colour and Visual Computing Symposium*, 2018.

[18] J. Hilkert, "Inertial stabilized platform technology concepts and principles," *IEEE Control Systems magazine,* vol. 28, no. 1, pp. 26-46, 2008.

[19] M. K. Masten, "Inertially stabilized platforms for optical imaging systems," *IEEE Control Systems Magazine,* vol. 28, no. 1, pp. 47-64, 2008.

[20] Z. Hurak and M. Rezac, "Image-based pointing and tracking for inertially stabilized airborne camera platform," *IEEE Transactions on Control Systems Technology,* vol. 20, no. 5, pp. 1146-1159, 2011.

[21] J. DeBruin, "Derivation of line-of-sight stabilization equations for gimbaled-mirror optical systems," in *Active and Adaptive Optical Components*, vol. 1543, International Society for Optics and Photonics, 1992, pp. 236-247.

[22] C. J. Heiberg, D. Bailey and B. Wie, "Precision spacecraft pointing using single-gimbal control moment gyroscopes with disturbance," *Guidance, Control, and Dynamics,* vol. I, no. 23, pp. 77-85, 2000.

[23] J. C. DeBruin, J. M. Royalty, M. Wand and E. Allen, "Feedforward stabilization test bed," in *Acquisition, Tracking, and Pointing X*, vol. 2739, International Society for Optics and Photonics, 1996, pp. 204-214.

[24] T. S. Mruthyunjaya and B. Dasgupta, "The Stewart platform manipulator: a review," *Mechanism and machine theory,* vol. 35, no. 1, pp. 15-40, 2000.

[25] D. Stewar, "A platform with six degrees of freedom," *Proceedings of the institution of mechanical engineers*, vol. 180, no. 1, pp. 371-386, 1965.

[26] Y. Huang, D. M. Pool, O. Stroosma, Q. P. Chu and M. Mulder, "A review of control schemes for hydraulic stewart platform flight simulator motion systems," in *AIAA Modeling and Simulation Technologies*, 2016.

[27] J. Johnson, R. Pillai, S. Vignesh and A. L. Srinivas, "Design and kinematic analysis of a 6-DOF parallel manipulator for pharmaceutical application," *J. Chem. Pharm. Sci*, vol. 9, pp. 3283-3288, 2016.

[28] F. Szufnarowski, "Stewart platform with fixed rotary actuators: a low cost design study," *Advances in medical Robotics*, no. 4, 2013.

[29] K. Liu, J. M. Fitzgerald and F. L. Lewis, "Kinematic analysis of a Stewart platform manipulator," *IEEE Transactions on industrial electronics*, vol. 40, no. 2, pp. 282-293, 1993.

[30] M. Mamoon, "Inverse kinematics and path planning of stewart platform using crank arm actuation system," in *International Bhurban Conference on Applied Sciences & Technology (IBCAST)*, Islamabad, 2014.

[31] B. Indhu and V. Naidu, "Gimballed Camera Control for On-Point Target Tracking," *American Research Journal of Electronics and Communication*, no. 1, 2012.

[32] R. Jia, V. K. Nandikolla, G. Haggart, C. Volk and D. Tazartes, "System performance of an inertially stabilized gimbal platform with friction, resonance, and vibration effects," *Journal of Nonlinear Dynamics*, vol. 2017, pp. 1-20, 2017.

[33] C. Huang, C. Chang, M. Yu and L. Fu, "Sliding-mode tracking control of the Stewart platform," in *5th Asian Control Conference (IEEE Cat. No. 04EX904)*, 2004.

[34] A. Şumnu, H. G. İbrahim and V. Ç. Mehmet, "Simulation and PID control of a Stewart platform with linear motor," *Journal of mechanical science and technology*, vol. 31, no. 1, pp. 345-356, 2017.

[35] Y. Wen, "Task Space," in *PID Control with Intelligent Compensation for Exoskeleton Robots*, 2018, pp. 35-53.

[36] K. Alexandre , D. Jacobs and J. Baek, "Digital video stabilization and rolling shutter correction using gyroscopes," *CSTR*, vol. 1, p. 2, 2011.

[37] S. Liu, P. Tan, L. Yuan, J. Sun and B. Zeng, "Meshflow: Minimum latency online video stabilization," in *European Conference on Computer Vision*, Springer, 2016, pp. 800-815.

[38] X. Yang, M. Liu, F. Xiao, L. S. Davis and J. Kautz, "STEP: Spatio-Temporal Progressive Learning for Video Action Detection," in *Conference on Computer Vision and Pattern Recognition (CVPR),* 2019.

[39] Dejan, "DIY Arduino Gimbal | Self-Stabilizing Platform," 2019. [Online]. Available: https://howtomechatronics.com/projects/diy-arduino-gimbal-self-stabilizing-platform/.

[40] T. KNR, "Arduino Controlled Rotary Stewart Platform," 2017. [Online]. Available: https://www.instructables.com/id/Arduino-controlled-Rotary-Stewart-Platform/.

[41] NicHub, "NicHub/stewart-platform-esp32," 01 09 2019. [Online]. Available: https://github.com/NicHub/stewart-platform-esp32.

[42] "django:The web framework for perfectionists with deadlines.," Django Software Foundation, [Online]. Available: https://www.djangoproject.com/start/overview/. [Accessed 3 July 2020].

[43] "Fast, unopinionated, minimalist web framework for Node.js," OpenJS Foundation, [Online]. Available: https://expressjs.com/ . [Accessed 3 July 2020].

[44] "Peace-of-mind from prototype to production," phoenixframework.org, [Online]. Available: https://phoenixframework.org/ . [Accessed 3 July 2020].

[45] "Web development one drop at a time," Flask, [Online]. Available: https://flask.palletsprojects.com/en/1.1.x/. [Accessed 3 July 2020].

[46] "A JavaScript library for building user interfaces," Facebook, [Online]. Available: https://reactjs.org/. [Accessed 3 July 2020].

[47] "One framework. One codebase. Any platform.," Qt, [Online]. Available: https://wiki.python.org/moin/PyQt. [Accessed 3 July 2020].

[48] "R2016a - Updates to the MATLAB and Simulink product families," Mathworks.com, [Online]. Available: https://www.mathworks.com/products/new_products/release2016a.html.

[49] "SolidWorks," [Online]. Available: https://www.solidworks.com/.

[50] "FREE - FLIR Thermal Dataset for Algorithm Training | FLIR Systems," Flir.com, 19 06 2018. [Online]. Available: https://www.flir.com/oem/adas/adas-dataset-form/.