MID EVALUATION PRESENTATION

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

Group 03

Department of Electronic and Telecommunication Engineering

University of Moratuwa

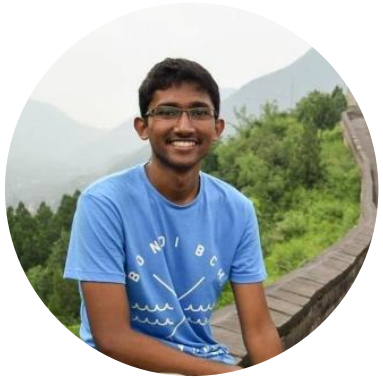
Group Members

160005C

K.G. Abeywardena

160243D

H.L.S.H. Jayasundara



160285G

G.K.S.R. Karunasena

160616B

S.K. Sumanthiran

Supervisors

Dr. Peshala Jayasekara

Dr. Ranga Rodrigo

Motivation

CNA938 | CNA Lifestyle | CNA Insider | Singapore | Asia | World | Business | Sport | Commentary | News Clips | Video on Demand | Podc | All Sections

05 Mar 2020 06:34PM

Bookmark

Asia

Sri Lanka seizes record US$33 million drugs haul at sea

ඇදිරි නීතිය නොතකා පුත්තලමේ සිට මුහුදු මාර්ගයෙන් මන්නාරමට යන්න ගිය **20**ක් අත්අඩංගුවට

Friday, 20 March 2020 - 18:27

Trending News

ඉතා අවධානම් දිස්ත්‍රික්ක 6ට අවණ්ඩව ඇදිරි නීතිය

01 April 2020

යාපනය, මරදාන සහ කුරුණෑගලින් තවත්

Sri Lanka Coast Guard
Safe, Secure & Serene Sea

HOME | ABOUT US | PROFILE | NEWS | REGULATIONS | PUBLI

OVER 3KG CRYSTAL METHAMPHETAMINE (ICE) FOUND FLOATING AT SEA OFF TRINCOMALEE

Three sealed parcels of Crystal Methamphetamine, commonly known as 'Ice' found floating in seas off Trincomalee by a returning Sri Lankan multiday Fishing Vessel handed over to Sri Lanka Coast Guard on 14th February 2020. The parcels contained 3.172 kg of Ice with a street value over Rs. 30 million handed over to Police Station, Tangalle for further investigations.

Sri Lanka Coast Guard is playing a vital role under close supervision of Director General – SLCG, Rear Admiral Samantha Wimalathunge in making a Sri Lanka a drug-free society, has been initiating a number of steps to prevent drug trafficking into the country via sea routes and peddling of drugs in the country. The recovery of drugs is a result of a series of awareness programmes conducted by Sri Lanka Coast Guard for fishery community on illegal drug trafficking by sea and the monitoring fishing movements in all major fishery harbours in Sri Lanka.

Event News >> Thirty (30) illegal Sri Lankan immigrants held by Navy in southern seas

Thirty (30) illegal Sri Lankan immigrants held by Navy in southern seas

Naval personnel attached to Fast Attack Craft (FACs) belonging to the Southern Naval Command, deployed on patrol waters, intercepted a suspicious trawler plying in southern seas this morning (07th March) and held 30 suspects onboa

Having spotted a suspicious boat movement at sea about 80 nautical miles off the Galle Lighthouse, two Fast Attack C directed to the location of the suspicious trawler. Accordingly, the naval personnel held 30 illegal Sri Lankan immigra are due to be handed over to the Galle Harbour Police after a medical examination and preliminary naval investigation

The Navy urges the general public not to involve in high risk sea-borne migration to overseas countries based on false reiterates that such attempts would finally end up behind bars.

Further the Navy reminds of its strong network of intelligence and regular patrols in place to nip such illegal migratio strict measures to curb such attempts.

Modified Deliverables (After Feasibility Study)

1

Functioning Object Detection algorithm

- For objects related to maritime environment such as boats, swimmers etc.
- Competitive mAP score on predicted bounding boxes
- Using thermal images.

2

Functioning Activity Detection algorithm

- Identifying at least 3-classes of pre-defined suspicious activities.
- With competitive mAP score.
- Using thermal image sequences.

3

Annotated maritime dataset

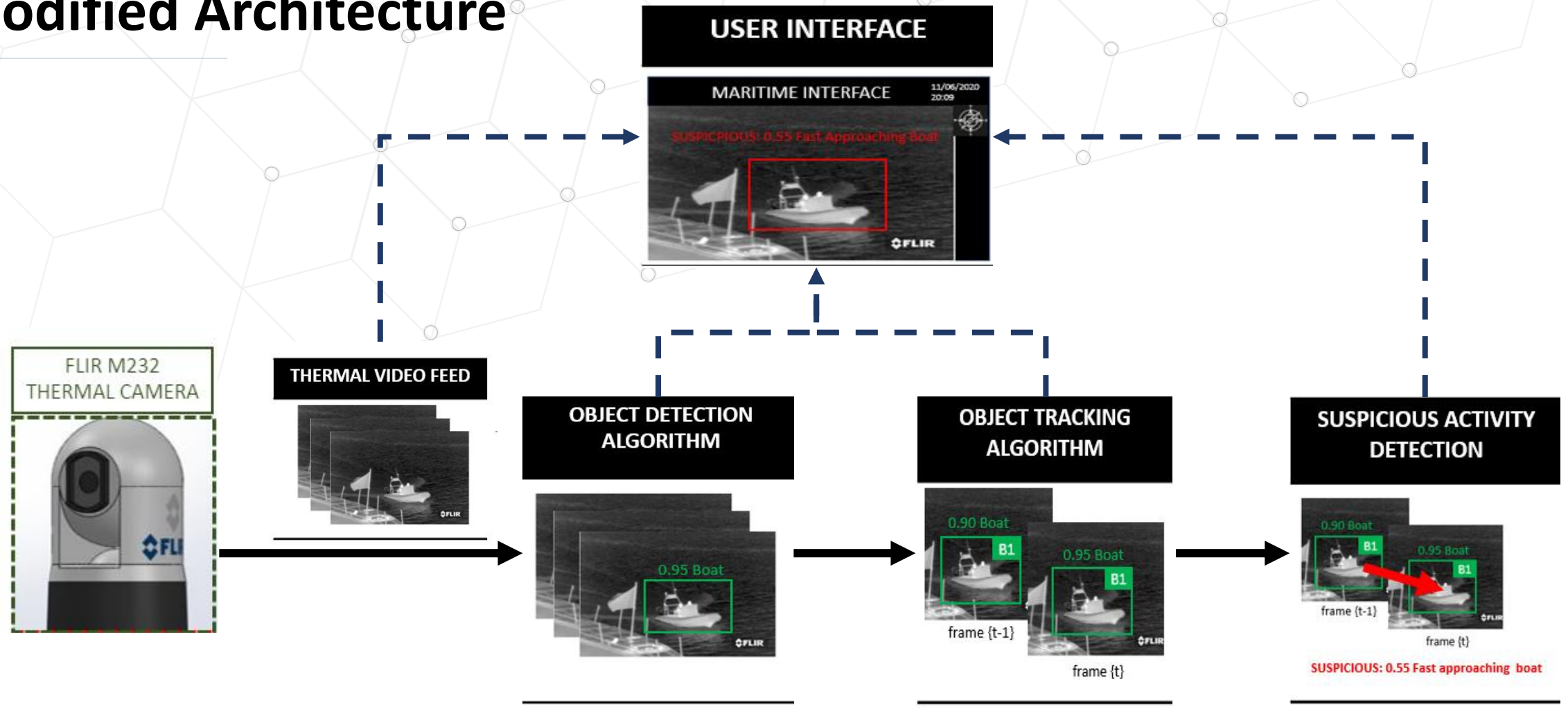
- Capable of use for object detection and activity detection in maritime environment.
- Either a locally obtained thermal dataset or already published maritime thermal dataset.
- Publication in a well-known forum.

4

Well documented software

- User-friendly

Modified Architecture



May include:

- Dumping objects
- Unauthorized fishing
- Loitering
- Evasive actions

Content

Object Tracking

Tracking of detected objects.

Paper Publication

Publication in a well-known conference.

Object Detection

RGB, Thermal, objects in Maritime, General environments.

Action Detection

Online Real Time Spatio Temporal Action Detection

User Interface

An interface to showcase the detected and tracking of objects.

Maritime Object Detection

Alternative Datasets for Maritime environments

Datasets	Link	Images/ Video type	Description
Singapore Maritime Dataset	https://sites.google.com/site/dilipprasad/home/singapore-maritime-dataset	RGB and Near IR	• RGB (Onshore and Onboard) • Near IR (Onshore) • NIR videos were captured using Canon 70D camera with hot mirror removed and Near-IR Bandpass filter (Different than actual thermal images)
SeaShips	http://www.lmars.whu.edu.cn/prof_web/shaozhenfeng/datasets/SeaShips(7000).zip	Only RGB	• Contains 31455 images (Only 7000 images publicly available) • Annotations provided
IPATCH	http://ftp.pets.rdg.ac.uk/pub/PETS2016/MidHighLevelChallengeData/IPATCH/	Both RGB and Thermal	• Contains a set of fourteen multi camera recordings (visible, thermal) collected off the coast of Brest, France • No annotations provided/ The categories of the objects

Maritime Object Detection

Alternative Frameworks for Maritime Object Detection

Dataset	Evaluation Criterion	Framework		
		SSD[1]	CornerNet-Lite (Squeeze)[2]	CenterNet[3]
SeaShips	mAP % @ IoU 0.5	28.4	59	81.8
	FPS	19	60	52
Singapore Maritime Dataset	mAP % @ IoU 0.5	27	55.3	60.7
	FPS	19	60	52

*Algorithms are trained under both SeaShips and Singapore Maritime dataset.

Video from an Onboard camera -SMD (Inferenced using CenterNet)

Video from an Onshore camera -Seaships (Inferenced using CenterNet)


Vessel/ship0.8
Vessel/ship0.3
Vessel/ship0.8


2017-01-04 星期三 09:15:51
bulk_cargo_carrier0.5
琴海北路北侧HJ005球机

Maritime Object Detection

Near IR version of Singapore Maritime Dataset

Video from an Onshore NIR camera (Without Haze) -SMD (Inferenced using CenterNet)

Video from an Onshore NIR camera (With Haze) -SMD (Inferenced using CenterNet)

Thermal Object Detection

Alternative Frameworks for Thermal Object Detection

Evaluation Criterion	Framework		
	SSD[1]	CornerNet-Lite (Squeeze)[2]	CenterNet [3]
mAP % @ IoU 0.5	28.8	81	88
FPS	19	60	52

Video from FLIR ADAS Thermal Dataset (Inferenced using CenterNet)

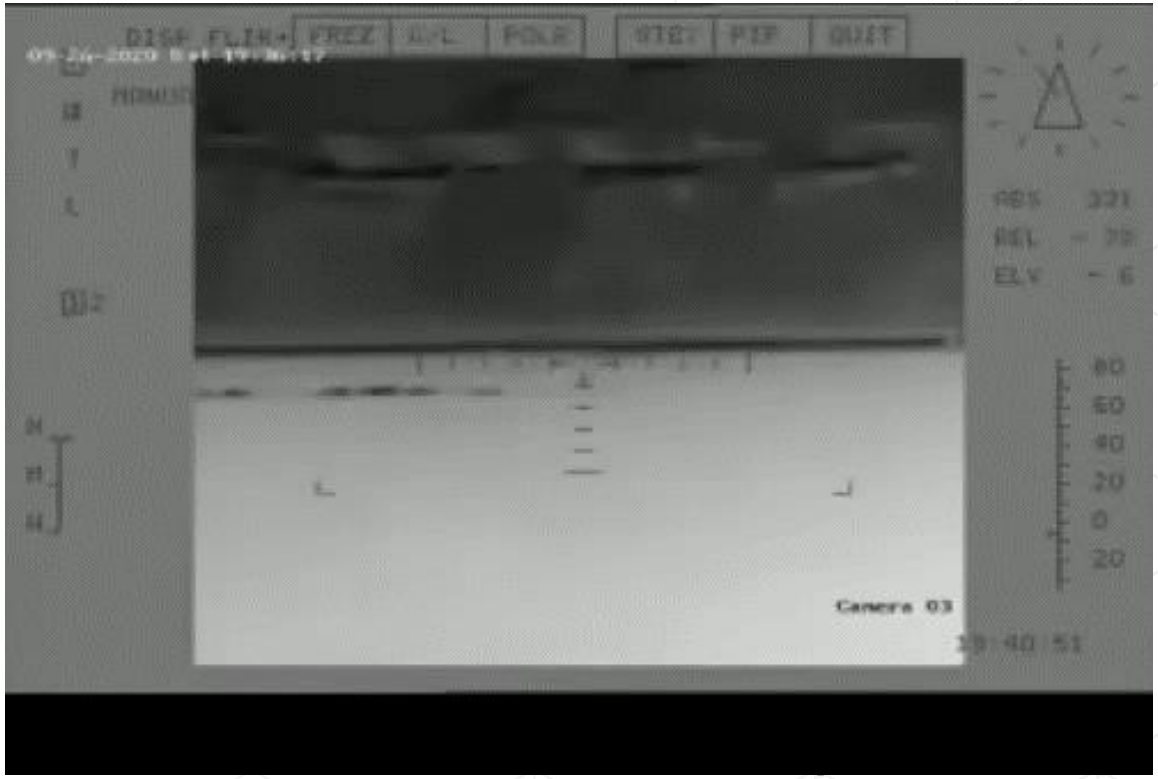
*Algorithms are trained under FLIR dataset.


car
car
car
car
car

Object Tracking

Alternative Algorithms

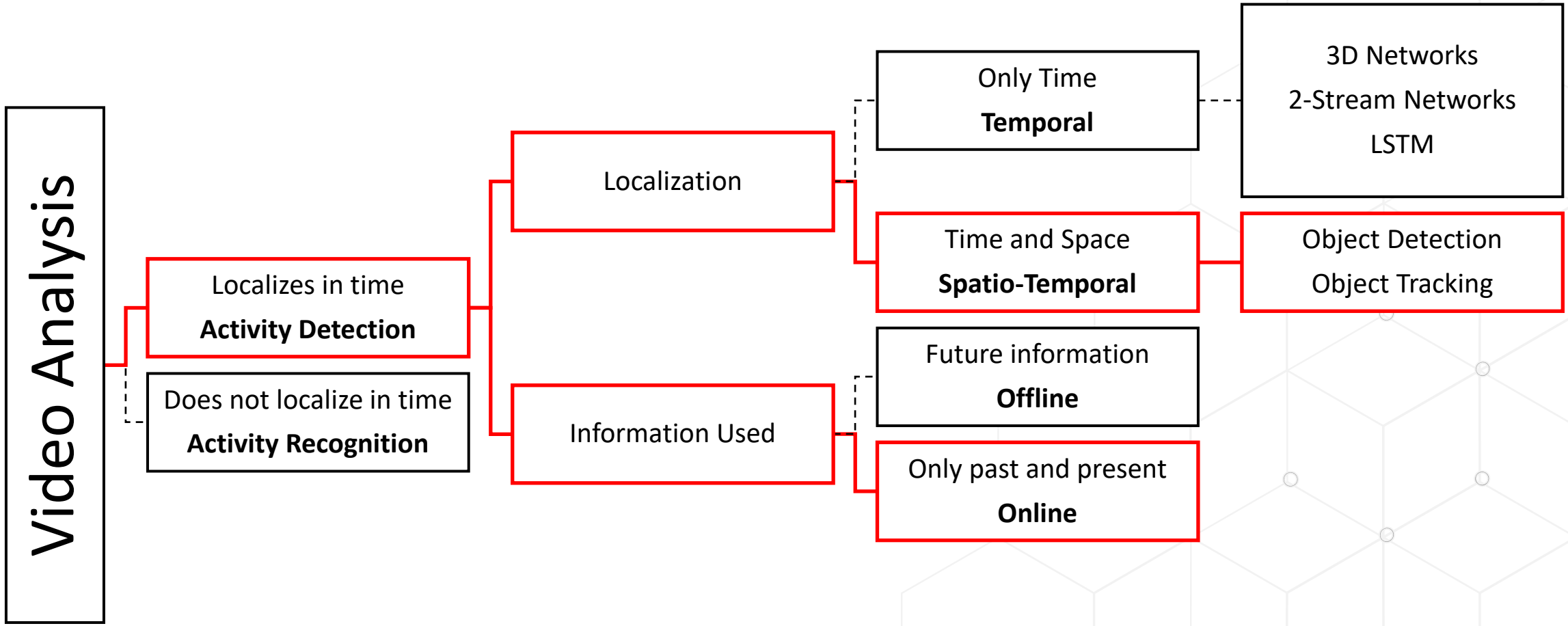
Tracker	Metric		
	Average IoU	Average CD	FPS
Boosting[4]	0.38	80	49
KCF[5]	0.22	15	102
Median Flow[6]	0.23	160	160
MIL[7]	0.35	60	20
MOSSE[8]	0.20	78	210
TLD[9]	0.24	110	27
CSRT[10]	0.50	50	25



Video from SL Navy (Tracker - CSRT)

Evaluation results taken from : *Single Object Trackers In OpenCV: A Benchmark Brdjanin et al.*

Action Detection



Action Detection

Alternative Datasets

Datasets	Link	Images/ Video type	Description
UCF 101 - 24	https://www.crcv.ucf.edu/data/UCF101/UCF101.rar	• RGB • Sports	• Extracted from a large dataset (UCF 101) - 101 action categories • 13320 videos • Annotations provided • UCF 24 – Only sports with 24 categories
HMDB 21 - 51	https://serre-lab.clps.brown.edu/wp-content/uploads/2013/10/hmdb51_org.rar	• RGB • Facial actions and Body movements	• Extracted from a large dataset (HMDB 51) - 51 action categories • 928 videos • Annotations provided

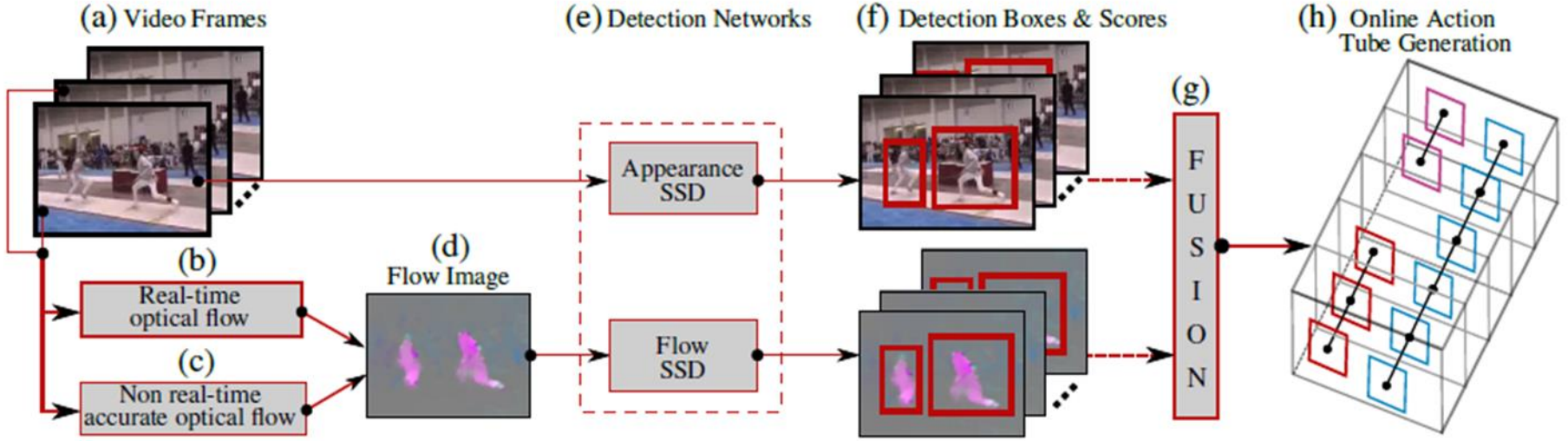
Action Detection

Evaluation of alternatives – Action Detection Algorithms

	Temporal Recurrent Networks (TRN)[11]	**Information Discrimination Unit (IDU)** [12]	**Spatio-Temporal and Motion Encoding (STM)** [13]	**A structured Model for Action Detection**[14]	**Spatio-Temporal Progressive Learning(STEP)** [15]	**Online Real-time Multiple Spatio-temporal Action Localization and Prediction (ROAD)** [16]
Temporal/Spatio-temporal	Temporal	Temporal	Temporal	Spatio-Temporal	Spatio-Temporal	Spatio-Temporal
Backbone	VGG-16 / ResNet-200	VGG-16 / ResNet-200	ResNet-50	ResNet-50/Mask-RCNN	VGG-16	VGG-16
Online/Offline	Online	Online	Offline	Offline	Offline	Online
FPS	24	24	-	12	21	28
Dataset	THUMOS'14	THUMOS'14	UCF101	UCF101	UCF101	UCF101
mAP score	47.2	60.3	96.0	77.9	75.0	43.0

Action Detection

ROAD Architecture



Issues	Solutions Worked-On
No implementation of linking algorithm in python	Implementing the linking algorithm in python (90% of the conversion is done)
Lack of end-to-end pipeline from action localizations to action linking	Building the end-to-end pipeline in python (100% implementation with SSD detector)
Duplication of results for Fast Optical Flow implementation is not provided	

Paper Submission on Action Detection Domain

IEEE International Conference of Image Processing (ICIP) - 2021

Paper Title: Online Real Time Spatio-temporal Action Detection Using Key Points

Our Contributions:

- A key-point based Spatio-Temporal (ST) Action Localization.
- A two-frame cascaded input to compute required temporal features.
- An improved tube-linking algorithm.

Paper Status

Paper Number: 1412

Paper Title: KEY-POINT DETECTION BASED ONLINE REAL-TIME SPATIO-TEMPORAL ACTION LOCALIZATION

Initial Submission Time Stamp: 1/11/2021 6:51:07 AM

Revised Submission Time Stamp: 1/25/2021 10:14:47 AM

Current Server Time: 3/17/2021 12:21:03 AM

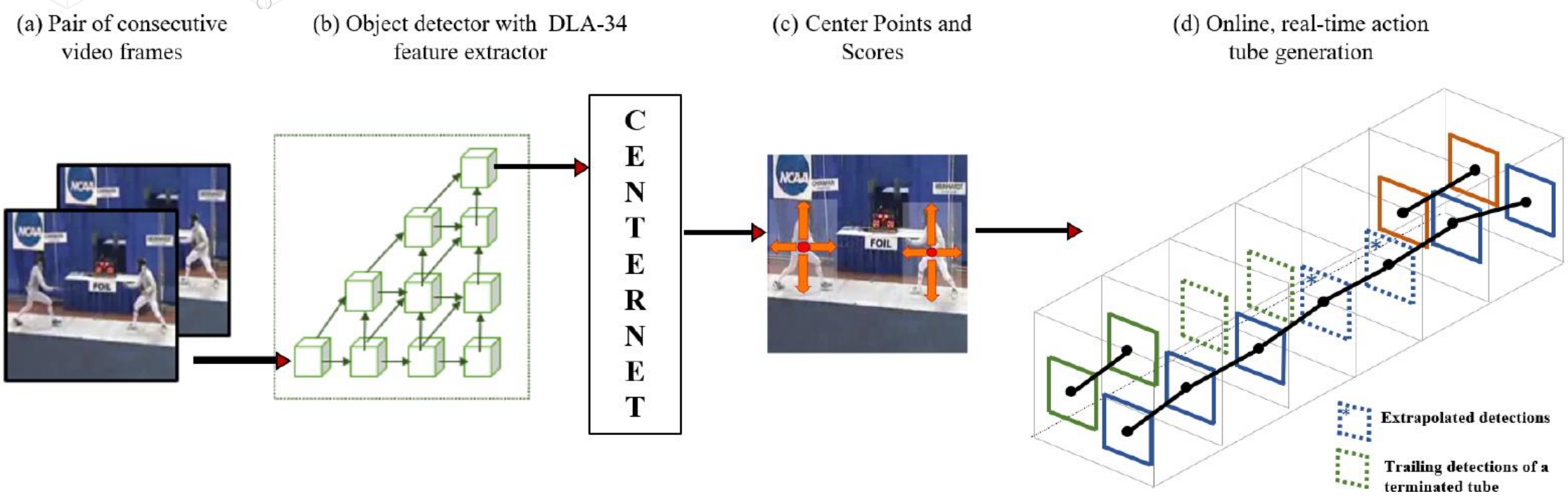
Submission Topic

Code	Category	Topic
10.2	Image and Video Analysis, Synthesis, and Retrieval	Image & Video Interpretation and Understanding::
4.7	Applications of Machine Learning	Machine learning for image & video analysis, synthesis, and retrieval::

Status	Item	Result
COMPLETE	Copyright Form	Copyright form has been received.
COMPLETE	Paper Upload	Document has been uploaded.
COMPLETE	Paper Inspection	Document passed inspection.
IN PROGRESS	Technical Committee Review	Document is under review. No action is required by you at this time.
PENDING	Final Paper Upload	Document inspection review pending results.
PENDING	Final Paper Inspection	Document inspection pending final paper upload.

Paper Submission on Action Detection Domain

Proposed ST Action Localization Architecture

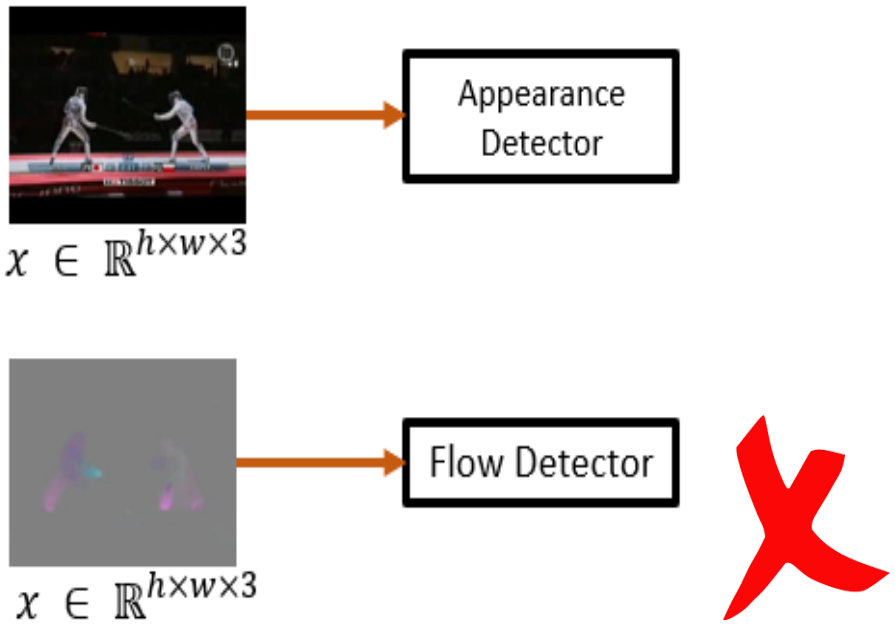


Paper Submission on Action Detection Domain

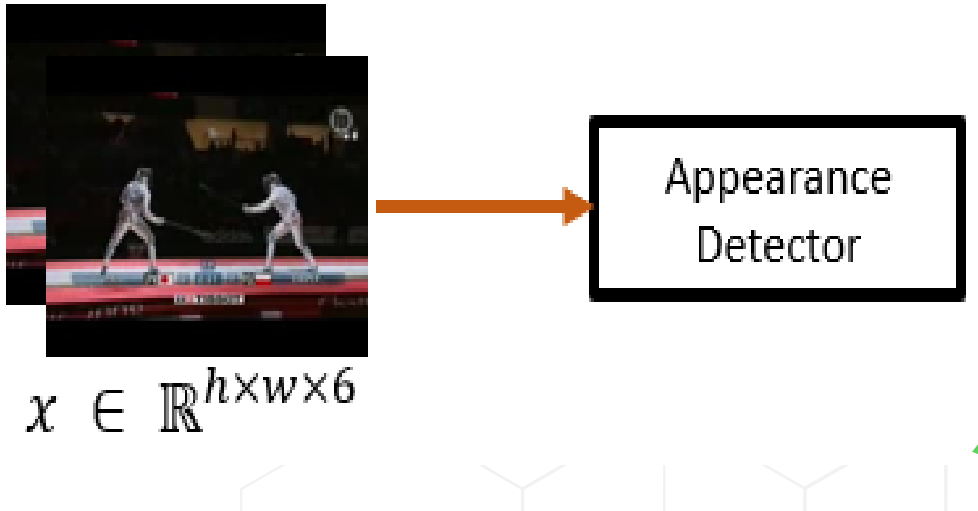
Novelty in our work

1. Cascaded two frame input for feature extractor to learn motion features.

Objective – To remove and replace the need of computationally expensive optical flow calculation



Two Stream RGB and Optical Flow Input



Proposed Cascaded Input

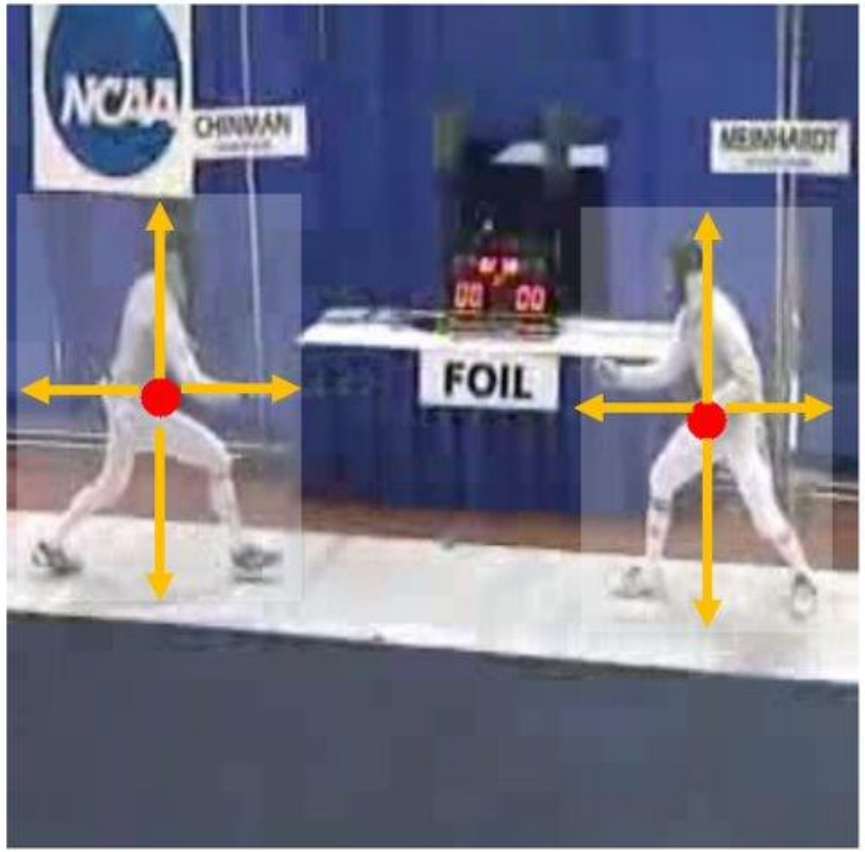
Paper Submission on Action Detection Domain

Novelty in our work

2. Key points for the Action Detection

Key-point based Action Localization has not being exploited in any of the past work.

Objective – To reduce the complexity and to improve inference time with key point detection using CenterNet with DLA-34.



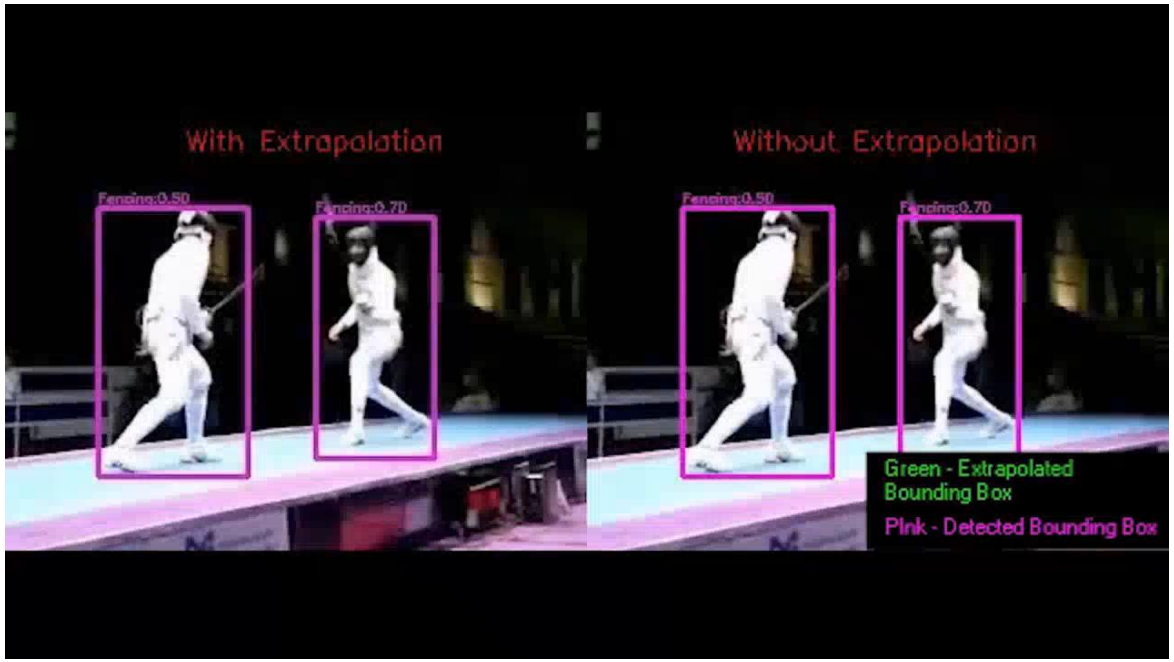
Key Point Detection of Action Instances

Paper Submission on Action Detection Domain

Novelty in our work

3. Improved Tube Linking Algorithm

Objective – To link the actions localized in the video frames and reduce the miss rate through extrapolation based on movement of the bounding boxes of past detections and their class scores



Demonstration of the Proposed Linking Algorithm

Algorithm 1: Online tube generation

Input: $\mathcal{T}^{t-1}, \mathcal{D}^t, c, \lambda, k$
Output: $\mathcal{T}^t$
for $T_j^{t-1} \in \mathcal{T}^{t-1}$ **do**
  $s \leftarrow 0$;
  $m \leftarrow 0$;
  for $D_i^t \in \mathcal{D}^t$ **do**
    if $IoU(b_{D_i}^t, b_{T_j}^{t-1}) \geq \lambda$ **and** $s < s_{D_i}^t(c)$ **then**
      $b_{T_j}^t \leftarrow b_{D_i}^t$;
      $s \leftarrow s_{D_i}^t$;
      $m \leftarrow i$;
      $\tau \leftarrow 0$;
    end
  end
  if $m = 0$ **and** $\tau < k$ **then**
    $b_{T_j}^t \leftarrow \texttt{predict\_bbox}(b_{T_j}^{t-1}, b_{T_j}^{t-2})$;
    $\tau \leftarrow \tau + 1$;
  end
  $s_{T_j}^t, c_{T_j} \leftarrow \texttt{update\_label}(s_{T_j}^{t-1}, s_{D_m}^t)$
end

Proposed Linking Algorithm

Paper Submission on Action Detection Domain

Results Obtained

Experiments were done on two datasets:

1. UCF-101-24 dataset : Challenging dataset with multiple action instances per video
2. J-HMBD-21 : Challenging dataset with single action instance per video

Table 1. ST action localization results (v-mAP) on UCF-101-24 dataset. Last two columns compare the f-mAP and FPS.

Method	v-mAP 0.2	v-mAP 0.5	v-mAP 0.75	v-mAP 0.5:0.95	f-mAP @0.5	FPS
Saha *et al.*[1]$^{\diamond}$	66.6	36.4	7.9	14.4	-	4
Peng(w/ MR) *et al.*[2]$^{\diamond}$	72.9	-	-	-	65.7	-
Zhang *et al.*[8]$^{\diamond\star}$	74.8	46.6	16.7	21.9	67.7	37.8
ROAD (w/ AF) [9]‡	73.5	46.3	15.0	20.4	-	7
ROAD (w/ RTF) [9]‡*	70.2	43.0	14.5	19.2	-	28
Our (A+AF)‡	72.9	**46.7**	**16.2**	**20.9**	**70.8**	7.7
Our (A+RTF)‡*	69.6	42.1	**15.5**	**19.3**	**69.6**	**37.9**
ROAD (A) [9]†*	69.8	40.9	15.5	18.7	-	40
Ours (A)†*	**70.2**	**44.3**	**16.6**	**20.6**	**71.8**	**52.9**
Ours†*	**71.6**	**44.1**	**17.0**	**20.7**	**74.4**	**52.9**

$^{\diamond}$ Offline * Real-time † Online with no OF ‡ Online with OF

Table 2. ST action localization results (v-mAP) on J-HMDB-21 dataset. Last two columns compare the f-mAP and FPS.

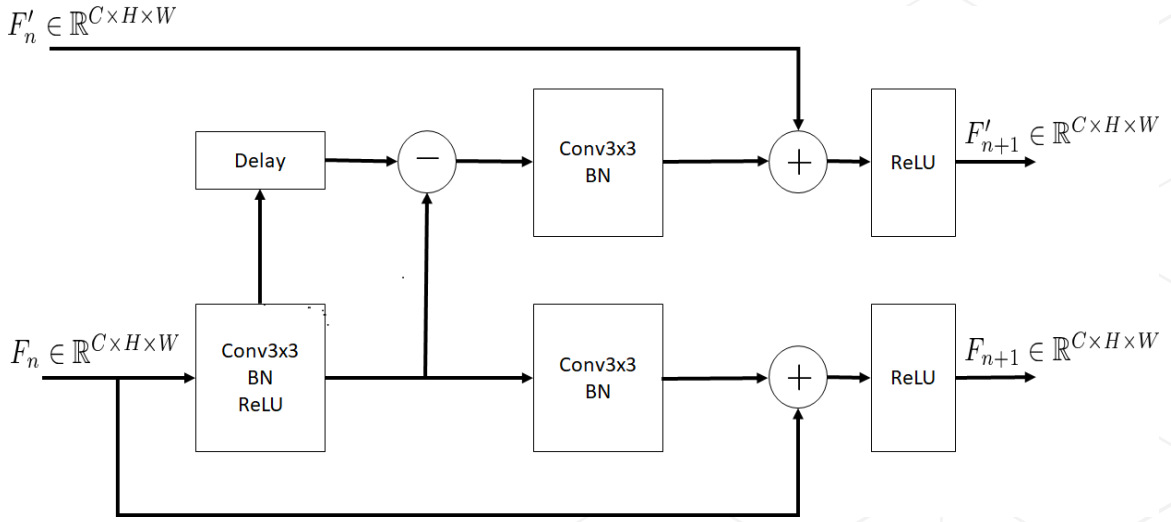
Method	v-mAP 0.2	v-mAP 0.5	v-mAP 0.75	v-mAP 0.5:0.95	f-mAP @0.5	FPS
Saha *et al.*[1]$^{\diamond}$	72.6	71.5	43.3	40.0	-	4
Peng(w/ MR) *et al.*[2]$^{\diamond}$	74.3	73.1	-	-	58.5	-
Zhang *et al.*[8]$^{\diamond\star}$	-	-	-	-	37.4	37.8
ROAD (w/ AF) [9]‡	70.8	70.1	43.7	39.7	-	7
ROAD (w/ RTF) [9]‡*	66.0	63.9	35.1	34.4	-	28
Our (A+AF)‡	68.8	67.6	**49.9**	**43.7**	46.9	7.7
ROAD (A) [9]†*	60.8	59.7	37.5	33.9	-	40
Ours (A)†*	59.3	59.2	**48.2**	**41.2**	**51.2**	**52.9**
Ours†*	57.2	55.9	48.1	39.9	47.9	**52.9**

$^{\diamond}$ Offline * Real-time † Online with no OF ‡ Online with OF

Action Detection

Short-Term Motion Module

- Feature Extractor Structure: ResNet-50
- Detection Head: CenterNet
- Pretrained Weights on: ImageNet



STM Module Architecture

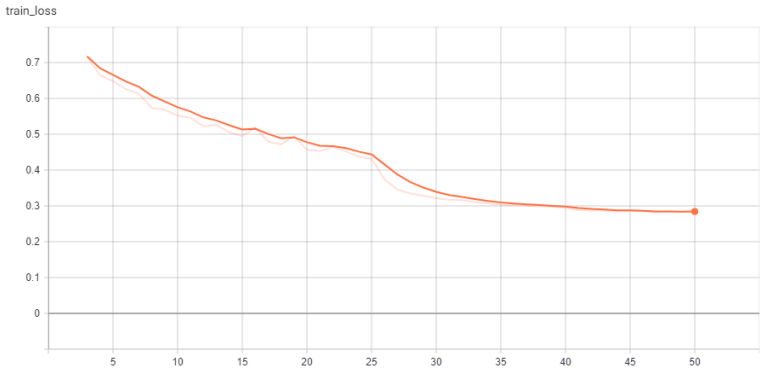
STMTrack

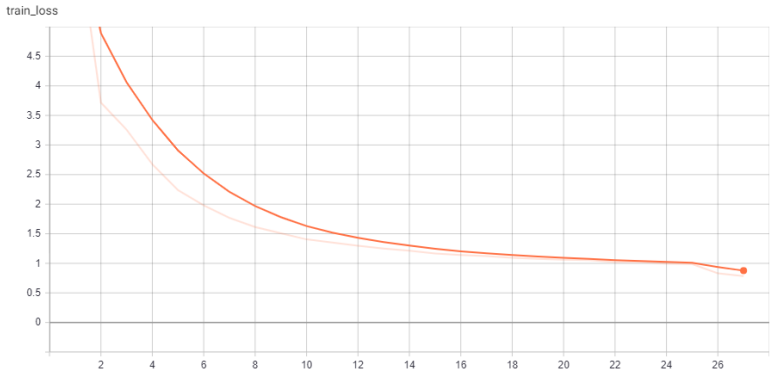
Zhou, Xingyi, Vladlen Koltun, and Philipp Krähenbühl. "Tracking objects as points." *European Conference on Computer Vision*. Springer, Cham, 2020.

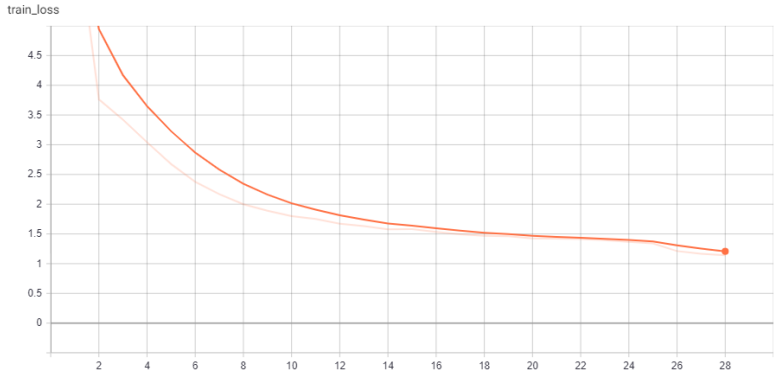
Action Detection

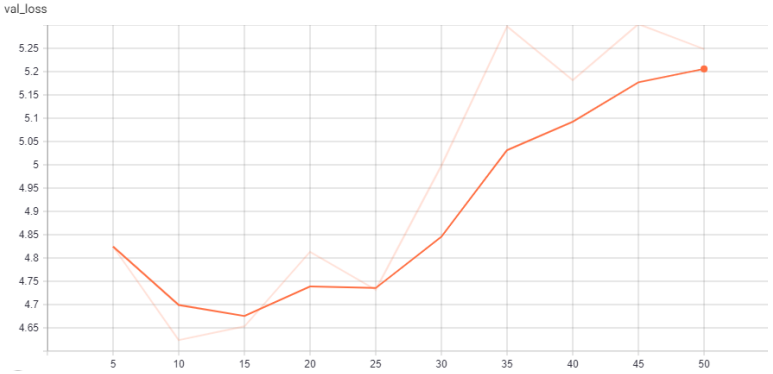
STMNet vs. ResNet vs STMTrack on Action Classification on HMDB-21

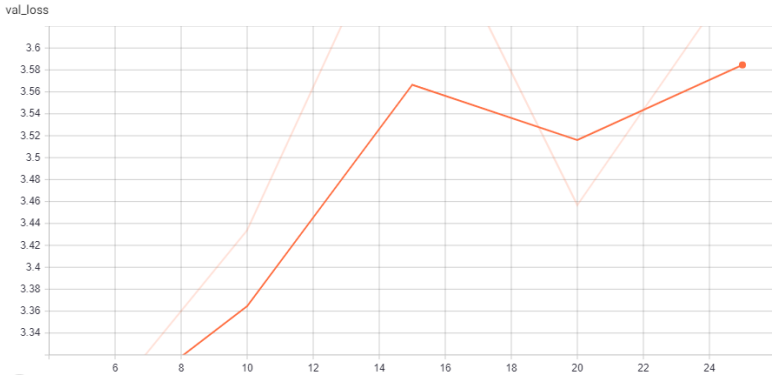
ResNet	STMNet	STMTrack
39.0%	33.4%	27.5%

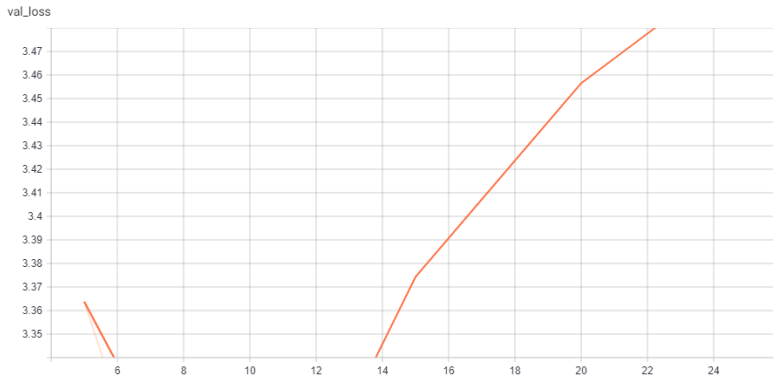












Action Detection

Thermal dataset creation for Action Detection

- Experimental Technique.
- Used Generative Adversarial Network (GAN) based frameworks to generate synthetic images.
- Used Pix2Pix[17] framework.

- Challenges :
    1. High Noise.
    2. Different thermal signatures.

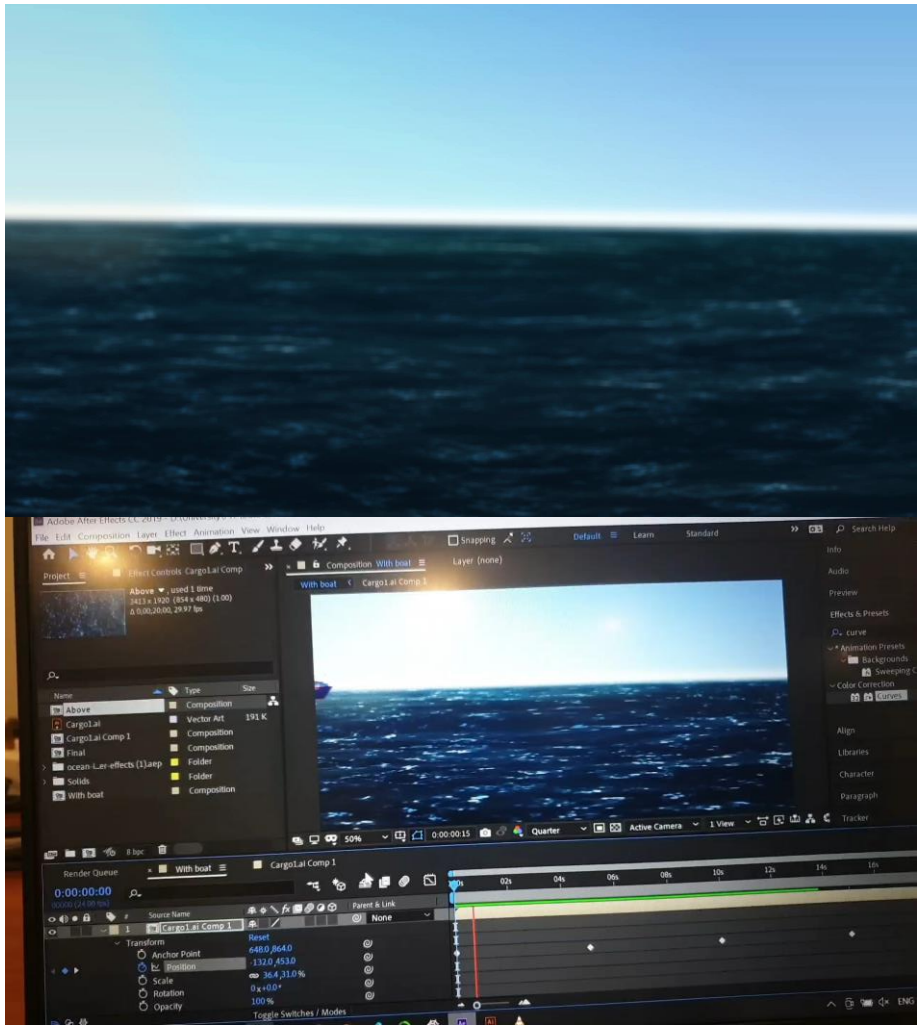
Converted RGB video using Pix2Pix

Action Detection

Maritime dataset creation for Action Detection

- Experimental Technique with the intention of creating a maritime dataset with action instances.
- Used Adobe After Effects to develop videos from scratch.

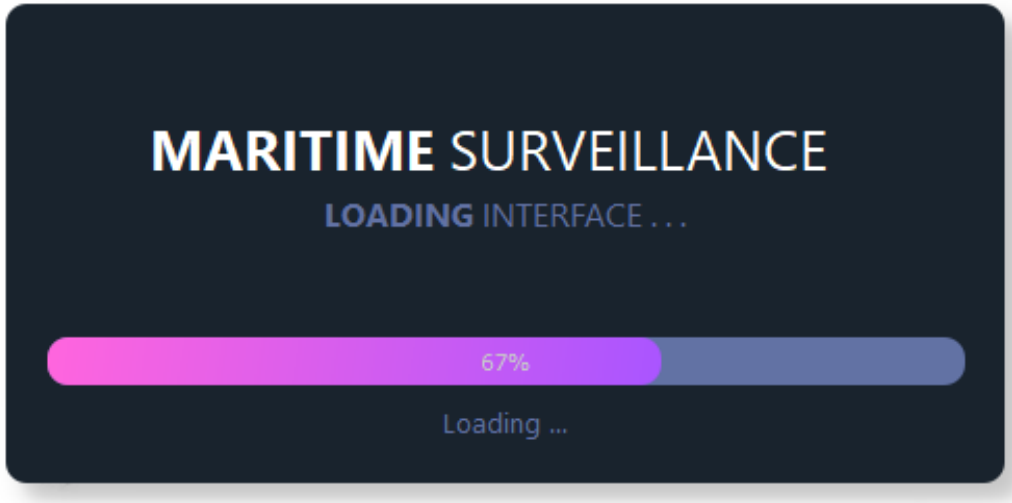
- Challenges:
    1. High rendering/creating time for each video.
    2. Repetition of patterns in the sea waves, which might affect the deep learning algorithms.

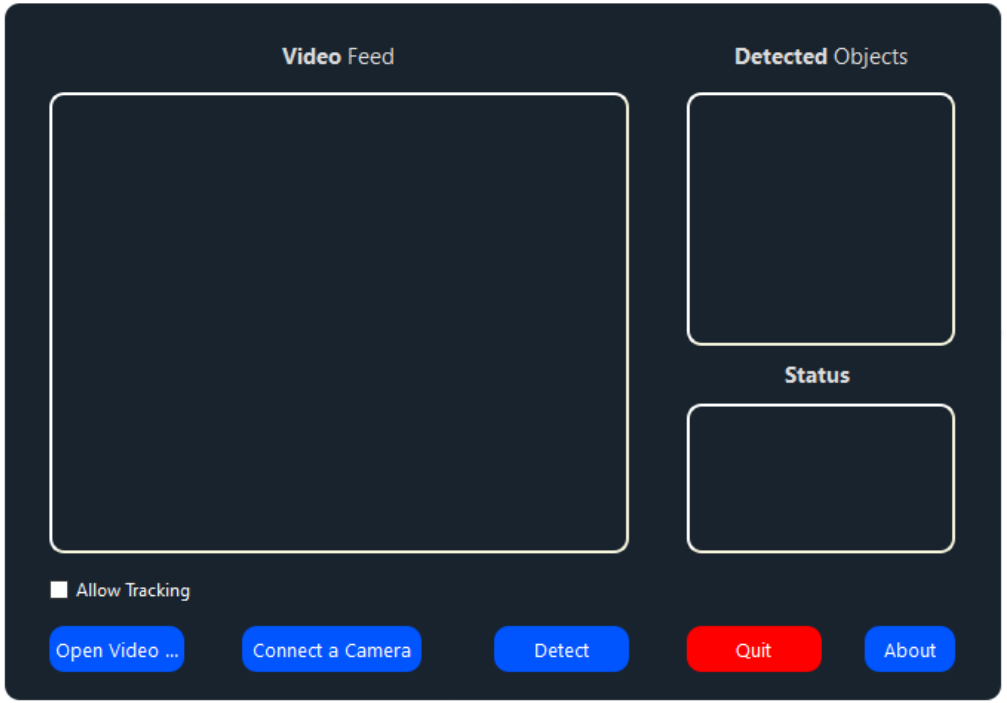


Custom made maritime action classification data using After Effects.

User Interface

- Developing using PyQt5.
- Python based interface with the ability to display detected objects, activities and track those in real-time.
- Displays detected objects on the side.





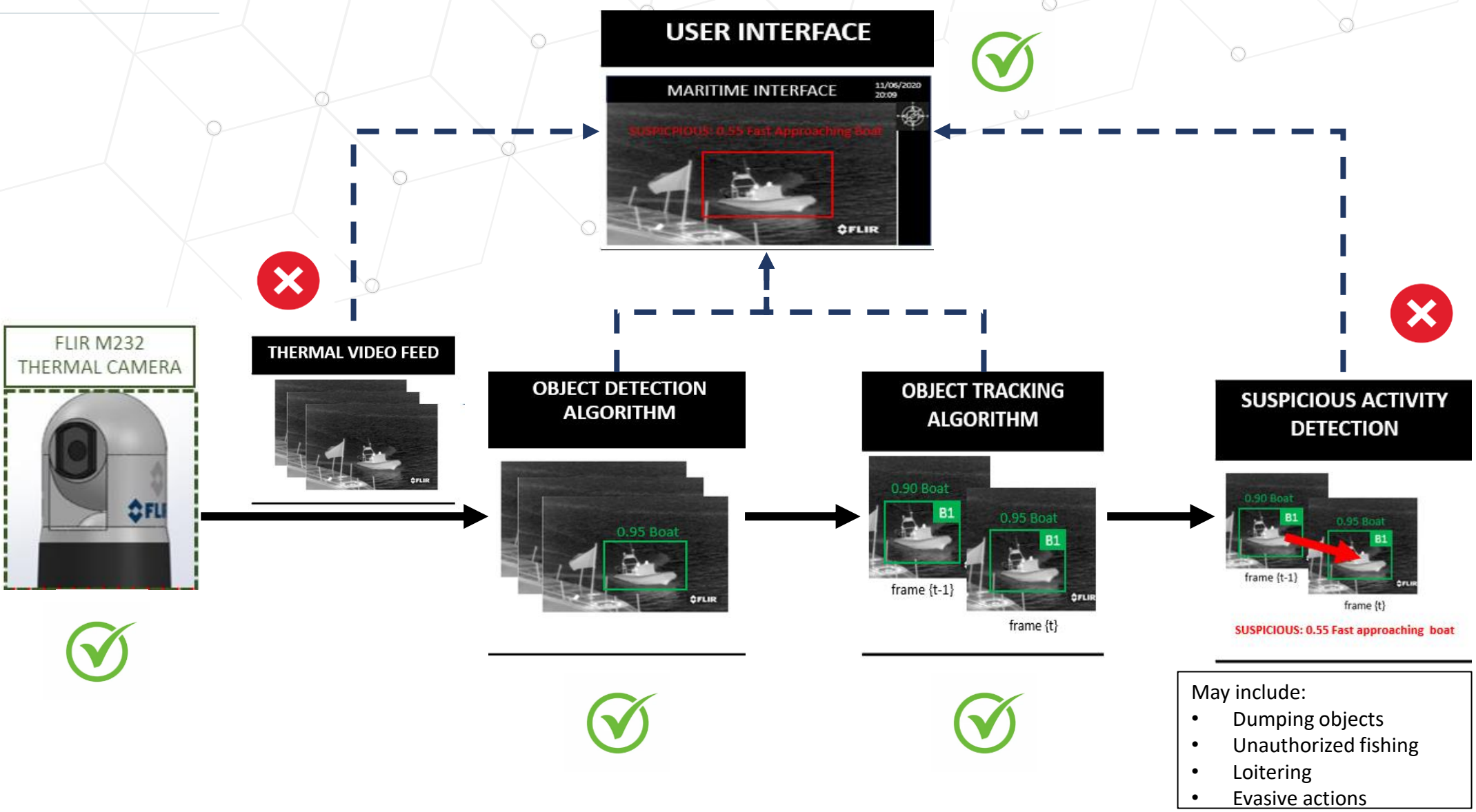
Interface created using PyQt5

Challenges

- Obtaining the Thermal Camera.
- Collaboration with the Sri Lanka Navy.
- Suitable, publicly available action detection datasets.

Meeting conducted with Sri Lanka Navy on obtaining the maritime dataset and planning the next steps.

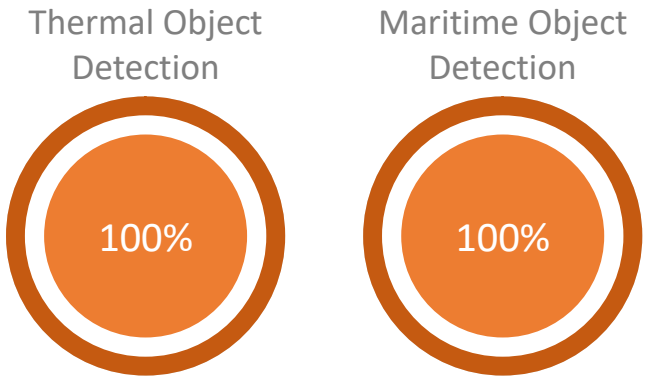
Modified Architecture



Percentage Completed

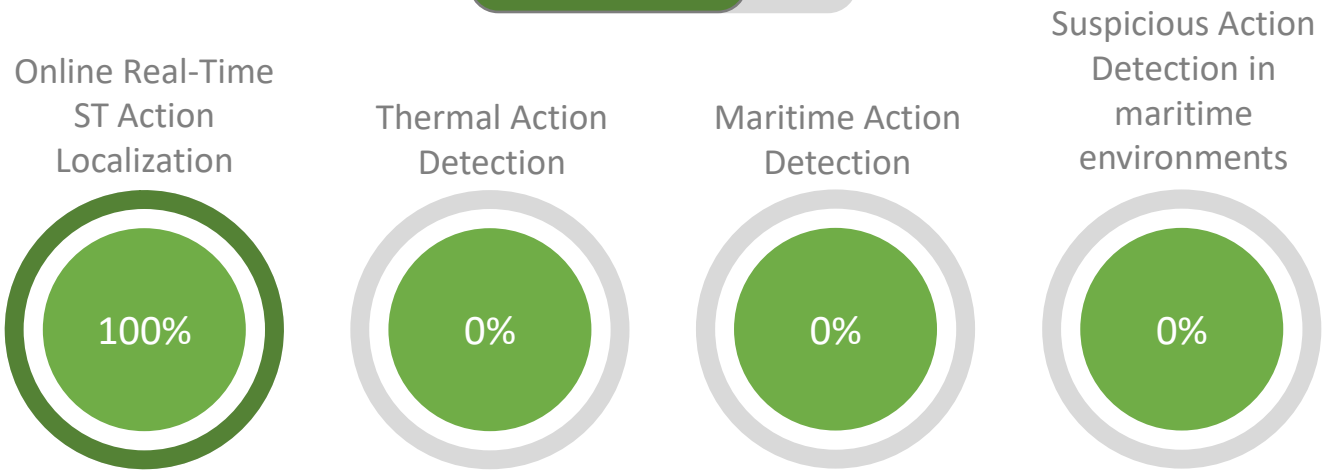
Object Detection

100%



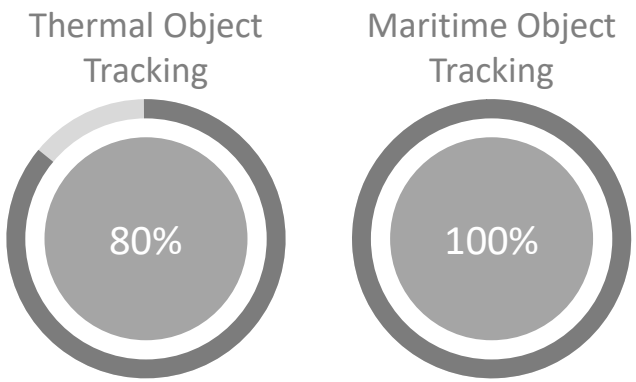
Action Detection

70%



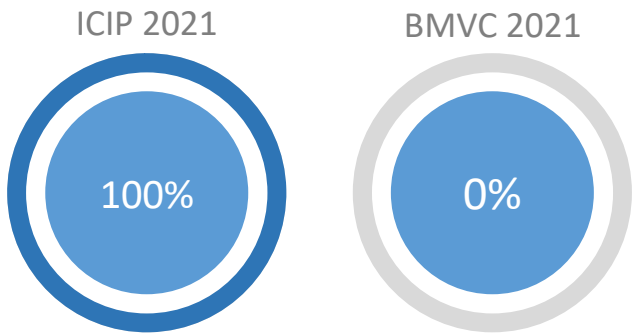
Object Tracking

90%



Paper Submission

100%



User Interface

60%

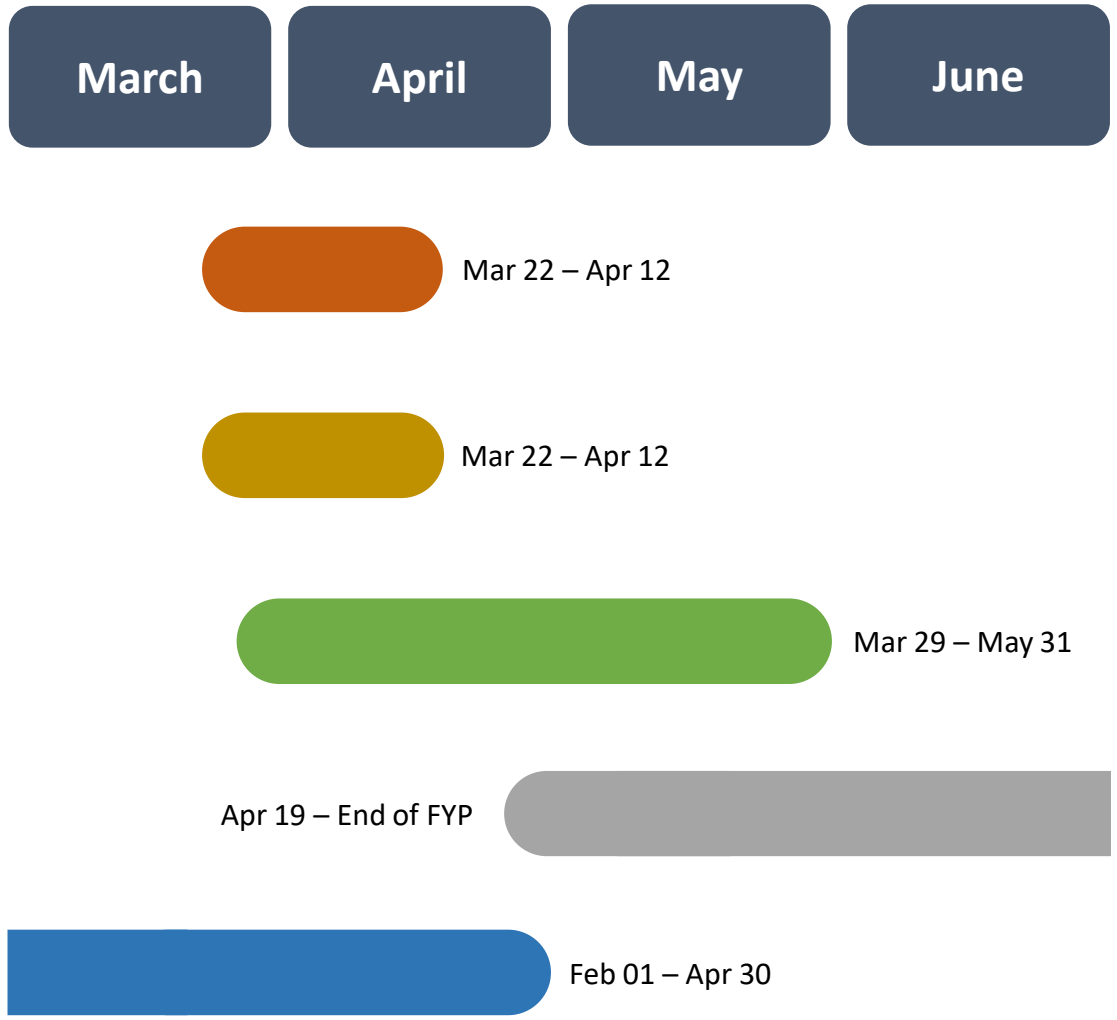
Thermal Camera Demonstration Plan

- Training the object detector algorithm with FLIR ADAS dataset
- Use our FLIR M232 thermal camera to collect our own thermal urban driving videos both in day and nighttime.
- Use trained object detector for inference on the collected videos.

FLIR M232 Thermal Camera

Final Completion Plan

Task	Task Delegation			
	Kalana	Sakuna	Sachira	Shechem
• End to end Object Detection and Tracking	✓		✓	
• Custom made interface for monitoring and Controlling		✓		✓
• Object Detection demo using FLIR Thermal Camera	✓	✓	✓	✓
• Documentation	✓	✓	✓	✓
• BMVC 2021 Paper Submission	✓	✓	✓	✓



Thank you.

References

[1] SSD: Single Shot MultiBox Detector - Liu et al.

[2] CornerNet-Lite: Efficient Keypoint Based Object Detection - Law et al.

[3] Objects as Points – Zhou et al.

[4] Real-time tracking via on-line boosting – Grabner et al.

[5] High-speed tracking with kernelized correlation filters – Henriques et al.

[6] Forward-backward error: Automatic detection of tracking failures – Kalal et al.

[7] Visual tracking with online multiple instance learning – Babenko et al.

[8] Visual object tracking using adaptive correlation filters – Bolme et al.

[9] Tracking-learning-detection – Kalal et al.

[10] Discriminative correlation filter with channel and spatial reliability – Lukezic et al.

[11] Temporal Recurrent Networks for Online Action Detection – Xu et al.

[12] Learning to Discriminate Information for Online Action Detection – Eun et al.

[13] STM: SpatioTemporal and Motion Encoding for Action Recognition – Jiang et al.

[14] A Structured Model For Action Detection – Zhang et al.

[15] STEP: Spatio-Temporal Progressive Learning for Video Action Detection – Yang et al.

[16] Online Real-time Multiple Spatiotemporal Action Localisation and Prediction – Singh et al.

[17] Image-to-image translation with conditional adversarial networks. - Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017).