Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1734-1740 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Machine Vision Approach for Underwater Remote Operated Vehicle to Detect Drowning Humans

Yaswanthkumar S K*, Keerthana M, Vishnu Prasath M S

Department of Electrical & Electronics Engineering, Kumaraguru College of Technology, Coimbatore, 641049, India

ARTICLE INFO

Article history:
Received: 15 September, 2020
Accepted: 29 December, 2020
Online: 30 December, 2020

Keywords:
Remote Operated Vehicle
Image Processing
Drowning People
Semi-Autonomous

ABSTRACT

Today, Drowning is the 3rd major cause for unintentional injury death accounting for 7% of deaths of all injury deaths. Drowning is a state of suffocation when water or other fluids accumulate the lungs, resulting in respiratory impairment, ultimately leading to death. The predominant problem during rescue operations of such accidental drowning is to locate or track the person underwater, facilitated with the help of our bare eye sight, which makes the process too cumbersome. Also, in case of moving water bodies such as rivers and sea, the process of tracking the person becomes too difficult. Thus, there is pressing need for development of an engineered solution to solve this problem. This research involves development of such a robotic technology which will facilitate the work of locating the position of the drowning person underwater. This method comprises of Image processing along with GPS tagging embedded on a Remote Operated Vehicle maneuvering underwater at the site of accident to detect the exact location of the drowning person along with the GPS coordinates in order to ease the process of rescue. This robotic system was tested on a real time environment in order to demonstrate the efficacy of the system under practical circumstances. Thus, this system can be used at places where emergency rescue operations are needed.

1. Introduction

Drowning accounts for 7% of all deaths that occur due to injury. It is also the 3rd major cause for unintentional injury death. As drowning is an unexpected event, there are neither any precautionary steps to prevent it, nor any existing technology to initiate rescue operations rapidly. According to stats from 2016 nearly 350000 people are reported dead due to drowning. In India, Drowning is the second largest cause of death for children below age of 15. The situation is even worse at Russia. The drowning mortality rate is of more than 3.9 per 10000 individuals.

Nowadays, few assistive technologies are available in the market to provide safety for humans in or under water. But these solutions are based on scenarios where people must use this technology before nearing any water bodies. But since Accidental Drowning are unpredictable and the underwater environment varies from place to place a solution adaptable to all water bodies is the main focus of this work. Development of this technology can be used not only on rescue operations of accidental drowning, but also to mitigate the risk of loosing track of people during natural calamities like heavy floods or Tsunamis. Thus, there is a urgent need to focus on development of a solution for the above mentioned scenario.

Our research work is to embedded Image Processing methodologies to detect drowning people underwater and to use GPS coordinates of their exact location with the use of a Robotic Remote Operated Vehicle which can by controlled autonomously or semi autonomously to navigate through water bodies during the rescue process. Also, this system uses a on board camera, live camera feed can also be surfaced over radio communications to the rescue team, which would prove to be great advantage to the people on the field for the rescue operations. All the existing methods are either not cost-effective or have installation issues at underwater environments.

2. Survey on Previous Researches

This paper is an extension of work [1] originally presented in 2019 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), where a different method for the detection of humans drowning underwater was illustrated.

In [2], is a research work that involves development of an automated system that detects and tracks objects that are of potential interest for human video annotators. By pre-selecting salient targets for track initiation using a selective attention algorithm, we reduce the complexity of multi-target tracking, in

*Corresponding Author: Yaswanthkumar S K, Email: yaswanthsk4@gmail.com

particular of the assignment problem. Detection of low-contrast translucent targets is difficult due to variable lighting conditions and the presence of ubiquitous noise from high-contrast organic debris ("marine snow") particles.

In [3], the paper presents an attentional selection system for processing video streams from remotely operated underwater vehicles (ROVs). The system identifies potentially interesting visual events spanning multiple frames based on low-level spatial properties of salient tokens, which are associated with those events and tracked over time. Using video cameras, it is possible to make quantitative video transects (QVTs) through the water, providing high-resolution data at the scale of the individual animals and their natural aggregation patterns.

In [4], this paper explains an agent system is strategy to enhance the underwater manipulation. If the location of the agent can be measured, the end effector is able to be place to any position. To implement this system, the method of an agent vehicle localization is proposed. The method uses the sonar images of moving agent obtained by forward-looking sonar. To detect the location of the agent in the sonar images, the convolutional neural network is applied. Through field experiment, we confirm the proposed method can detect and track the agent in the successive sonar images.

In [5], This paper presents methods applied for automated detection of fish based on cascade classifiers of Haar-like features created using underwater images from a remotely operated vehicle under ocean survey conditions. The images are unconstrained, and the imaging environment is highly variable due to the moving imaging platform, a complex rocky seabed background, and still and moving cryptic fish targets. Several Haar cascades are developed from the training set and applied to the validation and test video images for evaluation.

In [6], This paper presents a robotics vision-based system for an underwater pipeline and cable tracker. This system is introduced to improve the level of automation of underwater remote operated vehicles (ROVs) operations which combines computer vision with an underwater robotics system to perform target tracking and intelligent navigation. This study focuses on developing image processing algorithms and nonlinear control method for tracking the pipeline or cable. By using the adaptive sliding mode controller, the ROV can easily track the cable or pipeline.

In [7], is a proposed system of a multi-step and all-round underwater image processing specially for the underwater images taken in succession to improve the image quality, remove the dynamic interference and reconstruct the image. For images with complex texture, the inpainting result brings out much smoothness, at the cost of massive details. For images with low resolution, the transition of color near the edge of the vacancy is abrupt

In [8], The paper presents a wavelet-based fusion method to enhance the hazy underwater images by addressing the low contrast and color alteration issues. The publicly available hazy underwater images are enhanced and analyzed qualitatively with some state-of-the-art methods. Green color dominates the Image in the evaluated set of images. In the results, it can be observed that the natural color profiles of Image are preserved by correcting the dominated color and it was hard before to distinguish the color of the water and objects

In [9], it shows that man-made objects in the image can be distinguished from natural objects. This paper proposes that a control strategy to adjust image processing parameters, and sequence multiple applications of tools, is especially important for digital processing of underwater images. This paper highlights the key difficulties in processing underwater images such as obscuration of objects of interest by fish / sand / coral / algae in the water and change of illumination, color balance and scale of fish images due to refraction, absorption in water and large range of distances of observation and different sizes of fish

3. Schematic & Working of ROV

As shown in Figure 1. The devised system consists of a on board camera embedded to the ROV to communicate the live video feed to the operator and also for the video to be processed by the on-board Processor to detect for humans underwater. The main system consists of sub units such as:

- Transmission & Receiving Unit
- Processor Unit
- Sensing Unit
- Maneuvering Unit
- Powering Unit

Each of which will be discussed on the following sections.

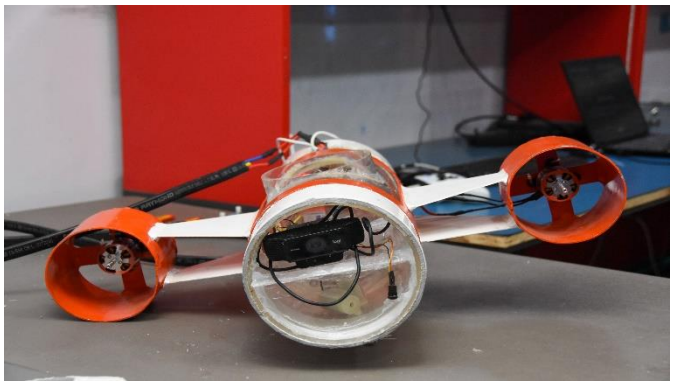

Figure 1: Hardware assembly of ROV

3.1. Transmission & Receiving Unit

A). Introduction

The transmission Unit consists of a Two axis controller to control the movement of the ROV along with a Screen or a monitor to view the live feed from the camera on-board which is interfaced to a Raspberry Pi through USB communication. A PC Joystick controller with TWO axis was used to provide the control signals as a input ls to the motor to control the movement of the ROV. Common PC monitor was used along with the raspberry PI to view the live feed from the ROV's on-board camera. The Receiver Unit Consists of a Raspberry PI which is on-board in the ROV. The Video data from the camera is given as an input to this raspberry pi from which is then transferred to the other raspberry pi through wireless communication.

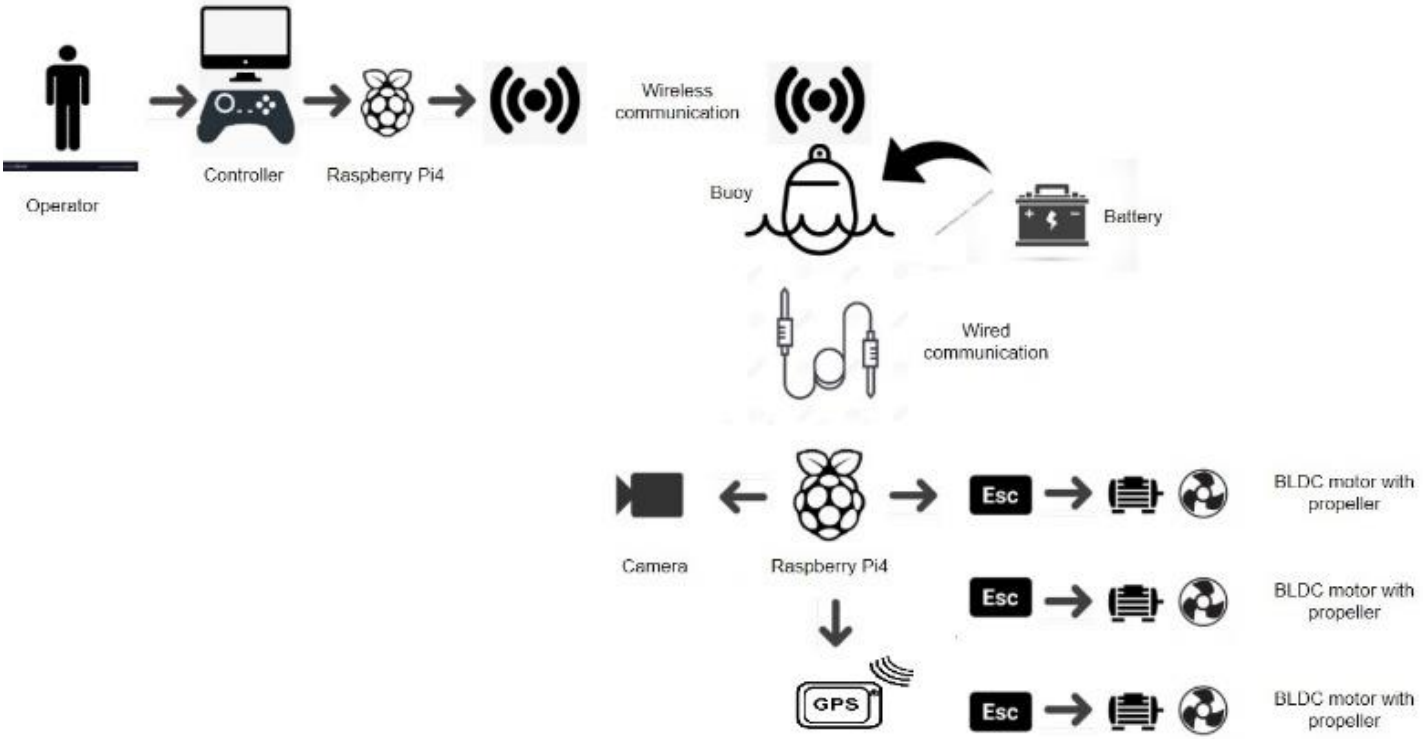

Figure 2: Schematic Diagram

B). Working

The control signals from the operator are provided as joystick input through USB communication to Raspberry PI on the Transmitter side. These signals are then transmitted to the other Raspberry PI on the Receiver side through WIFI Communication. This transfer of data can also be performed using RF transmission or MQTT via internet which will provide further distance range between the operator and the ROV. The operator can view the live feed in the monitor, to control the ROV effectively. The video feed from the USB Camera is also transmitted wirelessly through WIFI Communication from the Receiver Raspberry Pi to Transmitter Raspberry Pi.

3.2. Processor Unit

A). Introduction

The Processor unit consists of a Raspberry Pi (receiving unit) on-board in the ROV. This Receives the control signals from the Operator which is then used to control the movement of the ROV. This also receives data from the on-board camera and on-board GPS for further transmission.

B). Working

As the PWM from the Transmitter unit are received by the Processor Unit (receiving unit), these signals are the compared with the on-board accelerometer to measure the difference in the values. These values are again corresponded to respective PWM signals and then is given as an input to the respective Electronic Speed Controllers (ESC's) which control the spin of the motors. The Processor also performs the function of receiving the input data from camera and GPS on the ROV and transfer them to the operator.

3.3. Sensing Unit

A). Introduction:

The sensing unit consists of a IP68 waterproof high megapixel camera to be interfaced with the Processor. This will provide the video data for Image Processing and as well as send the live data from underwater to the operator. It also consists of a waterproof GPS chip to be enclosed in an air tight chamber of ROV. This provides the GPS coordinates of the ROV underwater.

B). Working

The camera transmits the data to the processor through USB communication, which is then used as the input for Image Processing algorithm and is also transmitted to the operator. The GPS module sends the signal to raspberry pi through SPI protocol which is also sent to the operator through wireless Communication.

3.4. Maneuvering Unit

A). Introduction

Maneuvering unit consists of 60 Amps ESC's, each individual for respective 500 kV BLDC motors. As the PWM control signals are given as an input into ESC, the speed of the motor varies accordingly the position and moment of the ROV also varies.

B). Working

As the operator provides the input signals through the JOYSTICK controller, this input data from the Transmitter Raspberry pi, reaches the Processor Raspberry pi (receiver unit). Those input signals are then provided as an input to the respective ESC after being compared with the current accelerometer value &

then the required amount of power will be delivered to the respective motor which is now being controlled by ESC.

3.5. Powering Unit

ROV is completely powered through wires as it is being tested for its buoyancy when battery pack is being placed inside the air tight enclosure of the ROV.

4. Imaging Methodology to Detect Drowning Humans in a - underwater environment

Image Processing has overcome the limits of its own discipline. Today, it is possible to process a live video feed right after few trained models. Due to this rapid increase in its uses, this technology is pressured to produce strong methodologies which would offer a wide variety of solutions for real time problems. Our methodology involves usage of Faster Region Convoluted Neural Network (FRCNN) algorithm.

4.1. Hardware Specification

The hardware specifications are illustrated in table 1.

Table 1: Hardware Specification

System	Dell G3 3579
GPU	Nvidia GeForce GTX 1050Ti
CPU	Intel Core i7-87507H
Memory (RAM)	8192 MB
Display Memory (VRAM)	4018 MB
Operating System	Windows 10 64-bit

4.2. Dataset

As only humans are to be detected underwater a dataset consisting of 800 images is used as dataset. These 800 images comprise a variety of underwater images with varying light intensity, water color, and humans in it. This dataset also contains images of halfway human pictures along with overlapping images too.

A). Data Preprocessing:

Reading

Using the matplotlib library on a python script is used to reading all the images from the dataset. This helps us to know whether all the all the images from the dataset are loadable and could be read by the algorithm.

Re-sizing

As larger images would take longer time to process, all the images collected are res-sized to be less than 200KB in size.

B). Data Labelling

The datasets must be labelled, since labelled images must be provided as input to train the training algorithm. In order to label the datasets, the images are first loaded onto the LabelImg tool. Now, humans in each image are to be bounded by drawing bounding boxes around them as shown in Figure 3. Then, the labelled data of all these images are to be stored in an XML format. Each image file would create individual XML files containing the File name, path, the label and coordinates for the bounding boxes. These are also called annotation files which will be used to train the algorithm.

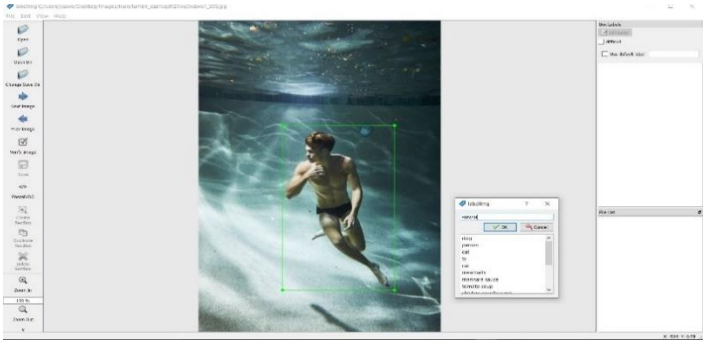

Figure 3: Labelling the datasets

4.3. Implementation

In order to bring out the best outcome, two algorithms were used 1). Faster RCNN and Yolo V3. Both these algorithms were implemented and trained using the TensorFlow Object Detection API. This makes the training process faster as TensorFlow object detection API comes with the model called ZOO which is a collection of models and algorithms.

A). Configuring the training algorithm

In order to detect the best algorithm which can be used, both the Faster RCNN and YOLO algorithms were configured to the same values.

Data Slicing

Both the algorithms were trained and tested using the same data sets. So, for both the algorithms 75% of the images and 25% of the images were used as training datasets and testing datasets respectively.

Classes

Since, humans were the only object to be detected by the algorithms, the class for these algorithms was set to 1.

Learning Rate

Learning rate is a parameter which allows us to control the model's response to estimated error each time the model weights are updated. The learning rate was set to a default value of 0.0002 while training both the algorithms.

Batch Size

The batch size corresponds to the consumption of VRAM, after looking into Table 1. The value set of batch size of the algorithms is set as 4.

4.4. Training Procedure

Figure 4 shows the training process of Faster RCNN algorithm along with the classification loss and the time taken after each iteration.

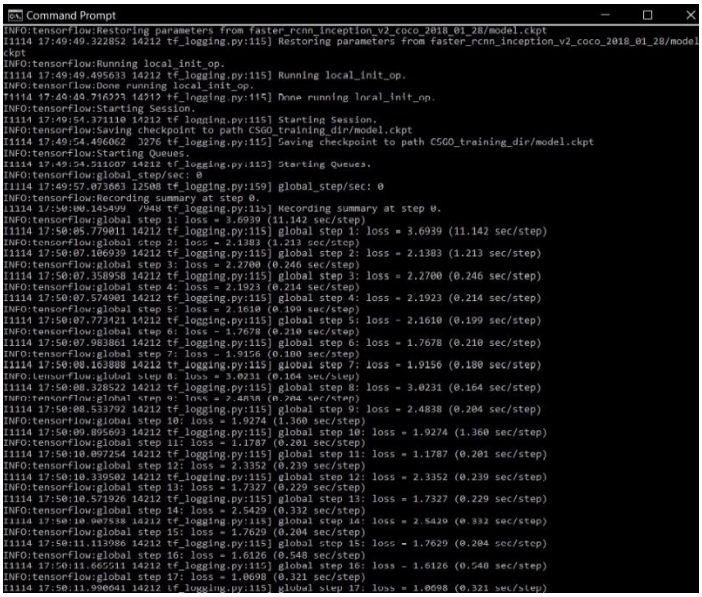

Figure 4: Training Process of Faster RCNN

As you can see, Classification loss is high when the training processes started. But after each step, there is a gradual decrease in the classification loss as the algorithm kept learning in each step. Figure 4. shows the exact decrease in classification loss during training process compared along with the testing process.

4.5. Performance Metrics

In order to compare the results of these two models, accuracy of these two models are to be contrasted. To calculate the accuracy few parameters are to be considered as metrics:

1. True Positive TP: A human is present in the Image and, the algorithm detects the human.
2. False Positive FP: No human is present in the Image, but the algorithm detects human.
3. False Negative FN: A human is present in the Image, but the algorithm does not detect the human.
4. True Negative TN: No human is present in the Image, and nothing is being detected.

A). Accuracy:

It is a measure to check whether the model or the algorithm is trained correctly to detect humans in an underwater environment or not. The formula to calculate the accuracy as follows:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

B). Precision

It is a ratio of positively predicted images that contain humans in it to that of which actually contains humans in it. The formulae to calculate precision as follows:

$$\text{Precision} = \frac{TP}{(TP+FP)}$$

C). Recall

It is the ratio of images that actually contain human in it to that of which were predicted positively. The formula to calculate recall as follows:

$$\text{Recall} = \frac{TP}{(TP+FN)}$$

D). F_1 Score

It is the measure of accuracy of the model combining both precision and recall. A model or algorithm is considered to be perfect if it's F_1 Score is 1. The formula to calculate F1 Score as follows:

$$F_1 \text{ Score} = 2 * \left(\frac{\text{Percision} * \text{Recall}}{\text{Persion} + \text{Recall}}\right)$$

5. Results

The results for both the algorithms are summarized in this section. The analysis of results as follows:

5.1. Faster RCNN Results

Figure 5, Figure 6 & Figure 7 are a collection of test data from 3 different samples under varying light conditions and other environmental factors for faster RCNN algorithm.

Figure 5: RCNN Output 1

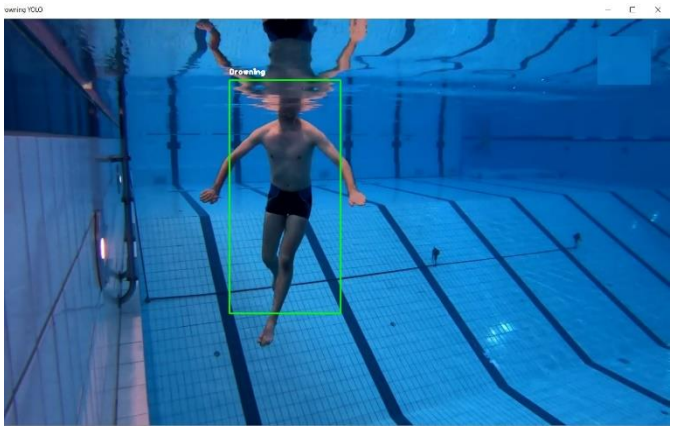

Figure 6: RCNN Output 2

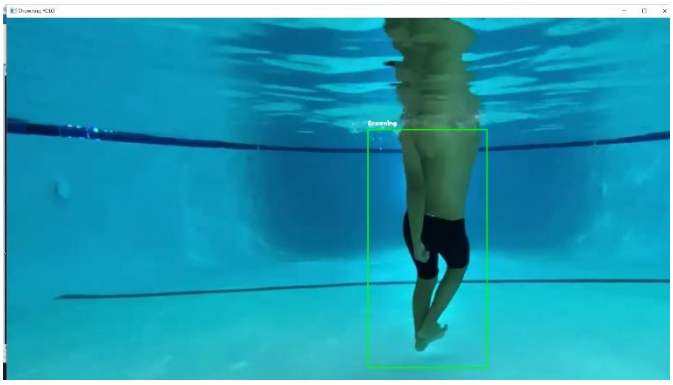

Figure 7: RCNN Output 3

The number of TP, TN, FP, FN from RCNN is tabulated in the table 2. These are further used to analysis the algorithms.

Table 2: RCNN Output Tabulation

#	True Positives	True Negative	False Positive	False Negative
Total	148	14	9	29

5.2. Yolo V3 Results

Figure 8, Figure 9 & Figure 10 are a collection of test data from 3 different samples under varying light conditions and other environmental factors for faster RCNN algorithm.

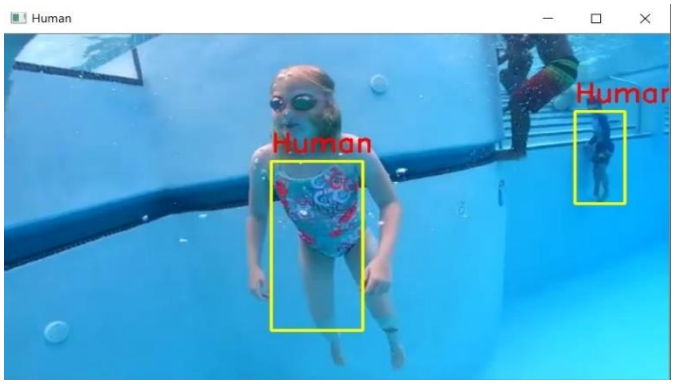

Figure 8: Yolo V3 Output 1

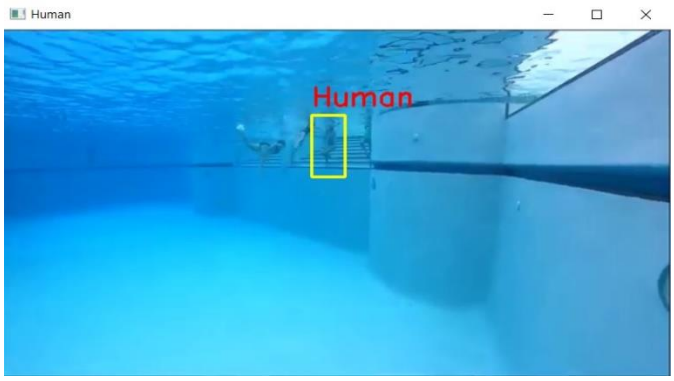

Figure 9: Yolo V3 Output 2

The number of TP, TN, FP, FN from YoloV3 is tabulated in the table 3. These are further used to analysis the algorithms.

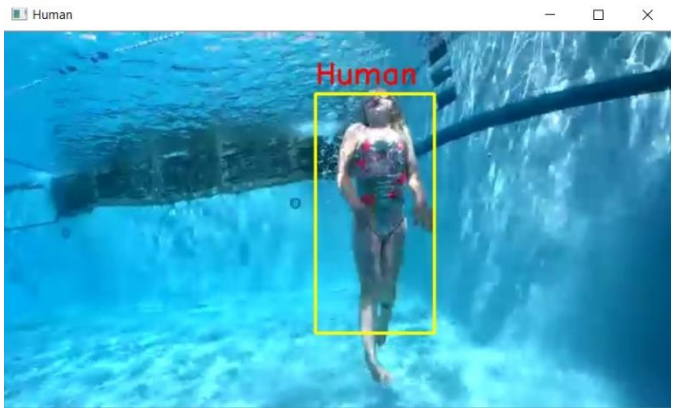

Figure 10: Yolo V3 Output 3

Table 3: YoloV3 Output Tabulation

#	True Positives	True Negative	False Positive	False Negative
Total	125	12	3	60

5.3. Analysis of Results

Table 4, provides us the collection of results containing the number of TP, TN, FP, FN parameters obtained by Faster RCNN and Yolo V3 algorithms.

Table 4

#	True Positives	True Negative	False Positive	False Negative
F - RCNN	148	14	9	29
YOLO V3	125	12	3	60

From this table, the accuracy of Faster RCNN and YOLO V3 algorithms are 81% and 68.5% respectively. Clearly, Faster RCNN is more Accurate than YOLO V3 in detecting humans in a underwater environment. It is also to note that Faster RCNN was able to consistently detect Humans if overlapping occurs. While, YOLO V3 failed to detect humans during overlapping frames.

Table 5

Algorithms	Precision	Recall	F_1 Score
Faster RCNN	0.94	0.83	0.881
YOLO V3	0.97	0.67	0.792

Table 5, Shows us that YOLO V3 is more precise compared to that of Faster RCNN, but then Faster RCNN has a high recall value than that of YOLO V3. To compare both the models, F_1 Score was calculated which also points out that Faster RCNN algorithm is the Best fitted algorithm to detect humans underwater. The results contribute to some preliminary findings to the growing field of underwater robotics in real time environments.

6. Future Work

The experimentation has been tested only on clear water bodies. The implementation of this ROV under real time circumstances such as lakes and river where the opacity will be translucent or opaque, it is better to use the thermal imaging as the

video input to this image processing models. As the body temperature of the aquatic species in that water body will be the similar, it will be easy to differentiate human body using thermal sensors and video feed of thermal imaging which will be performed in the near future.

7. Conclusion

This paper illustrates, an idea to develop a robotic system which will detect drowning people underwater at the times of an emergency. It also concludes that Faster RCNN is the best fit algorithm which can be used for this practical application. In our further research we plan to our work from a single ROV into a swarm of ROV's collectively working at a site of emergency to detect every drowning human underwater. We are also planning to enable full autonomous nature of the ROV in order to minimize the reaction time and communication time between the ROV and the Operator.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We would like to thank Forge Accelerator for providing the necessary support to complete this research work. We also thank Mr. Ravi Shankar, Mr. Koushal & Miss. Poovika for contributing to this research work by building the ROV from Scratch.

References

[1] S.K. Yaswanthkumar, O.K. Praveen, R. V. Rohit, "Autonomous Utility Vehicle (Auvs) Based Emergency Human Drowning Detection System Using Sonar and Thermal Detection Methods," in 2019 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics, DISCOVER 2019 - Proceedings, 2019, doi:10.1109/DISCOVER47552.2019.9007992.

[2] D. Walther, D.R. Edgington, C. Koch, "Detection and tracking of objects in underwater video," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, doi:10.1109/cvpr.2004.1315079.

[3] D.R. Edgington, K.A. Salamy, M. Risi, R.E. Sherlock, D. Walther, C. Koch, "Automated event detection in underwater video," in Oceans Conference Record (IEEE), 2003, doi:10.1109/oceans.2003.178344.

[4] J. Kim, S.C. Yu, "Convolutional neural network-based real- Time ROV detection using forward-looking sonar image," in Autonomous Underwater Vehicles 2016, AUV 2016, 2016, doi:10.1109/AUV.2016.7778702.

[5] G. Cutter, K. Stierhoff, J. Zeng, "Automated detection of rockfish in unconstrained underwater videos using haar cascades and a new image dataset: Labeled fishes in the wild," in Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2015, 2015, doi:10.1109/WACVW.2015.11.

[6] R. Trahan, R. Kiang, "An analysis of the die testing process using taguchi techniques and circuit diagnostics," in Proceedings - International Test Conference, 1992, doi:10.1109/TEST.1992.527832.

[7] C. Cai, Y. Zhang, T. Liu, "Underwater image processing system for image enhancement and restoration," in 2019 IEEE 11th International Conference on Communication Software and Networks, ICCSN 2019, 2019, doi:10.1109/ICCSN.2019.8905310.

[8] A. Khan, S.S.A. Ali, A.S. Malik, A. Anwer, F. Meriaudeau, "Underwater image enhancement by wavelet based fusion," in USYS 2016 - 2016 IEEE 6th International Conference on Underwater System Technology: Theory and Applications, 2017, doi:10.1109/USYS.2016.7893927.

[9] R.E. Blake, "Digital image processing of underwater images," in 2006 IEEE US/EU Baltic International Symposium, BALTIC 2006, 2006, doi:10.1109/BALTIC.2006.7266195.