ARTICLE IN PRESS

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Underwater human detection using faster R-CNN with data augmentation

Uma N. Dulhare *, Mohd Hussam Ali

Muffakham Jah College of Engineering & Technology, Hyderabad, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Computer vision
Deep learning
Data augmentation
Faster R-CNN
Human detection
Underwater

ABSTRACT

According to WHO, Drowning is a very much serious and every year 3,74,000 people claiming the lives as a public health threat. Underwater human body detection through live video is a very challenging task because of movement of water currents continuously and changes the intensity of water. Detecting people from an underwater video is a very complex and challenging problem as the video can be affected by various factors such as undesired artifacts (e.g. noise), monitoring limitation of cameras, illumination variation etc. It is very important to build the appropriate model to address this problem. The major goal of this paper is to examine the possibility of detect humans in an underwater Ambient using one of the Faster R-Convolution neural network (Faster R-CNN) and data enhancement algorithms and acutely, recall and precise evaluation of the proposed model results.

Selection and peer-review under responsibility of the scientific committee of the International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology.

1. Introduction

Computer vision is exorbitant due to extraction of information like object detection and recognition for content searching from the images automatically. During past few years, computer vision with deep learning plays a major role for detection of object in any environment. Deep learning-based framework is more powerful for detection and recognition of underwater human object as compared with the traditional algorithms due to robust in various environmental attributes like changes in illumination, blurry movements and perspective distortion. One of the recent deep learning algorithms, Faster R-CNN (Convolution Neural Network) is considered as inexpensive and efficient to detect humans in an underwater environment. It is very difficult to detect the human from some underwater images which are affected by various conditions viz. noisy, distortion, hazy. To overcome this problem proposed method is used the Data augmentation (DA) with Faster R-CNN. Three data augmentation techniques are used in this proposed system as Image Degradation, Perspective transformation, Illumination changes.

2. Related work

Author built the model with background subtraction, thresholding, inter-frame differencing, foreground computer vision based algorithms for extracting the features of underwater human and then subtracting white pixels with the reference background image. Moving object can be detected easily but it's a non-learning approach where in certain environmental conditions may not be detecting the moving object [1].Author [2] highlighted about feature mapping of moving objects from one frame to other frame with optimal solution. A vision-based model is used to detect the drowning incidents by placing the overhead cameras to properly monitor the swimmers in swimming pools by analyzing several key features like shape, motion [3]. For early stage drowning detection in outdoor swimming pool background subtraction, blob splitting, denoising and data fusion technique are used. If the object is not clear while tracking with the foreground detection, it could not be able to detect [4].

A moving object detection in underwater video using subtraction of background model with the five methods which are preprocessing, initialization normal distribution, decision the pixels as objects or background, background modeling, the object detection. Where the moving object can be detecting easily but used only for the indoor environment [5]. In this approac [6] compared the background modelling methods in video analysis using

* Corresponding author.
E-mail address: uma.dulhare@mjcollege.ac.in (U.N. Dulhare).

https://doi.org/10.1016/j.matpr.2021.05.653
2214-7853/
Selection and peer-review under responsibility of the scientific committee of the International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology.

Please cite this article as: U.N. Dulhare and M. Hussam Ali, Underwater human detection using faster R-CNN with data augmentation, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.05.653

Gausian mixture model,code book.consensus based method. Mixture-of-Gaussians (MoG) method used for object detection in an outdoor environment with less computationally [7]. Author compared the study of statistical background and substraction by GMM & estimated parameters by EM algorithm [8]. In this approach [9] author proposed Underwater & Satellite Image Enancement using auto threshold method by histogram clipping, sub-division and equlization. A novel multi-feature algorithm pipelined for underwater object detection and pose estimation to find human-made artifacts [10].In other approach, two-step process is used for edge detection. In first step, apply neural network to segment the various regions corresponding to sea bottom or edges of the underwater images then analyzed all region the pairs to select the correct one. The presence of obstacle can be determined by no. of pixels classified as edge in image [11].In approach [12] described underwater image processing for object detection using Laplacian of Gaussian(LoG).

Deep Learning efficiently recognized the fishes in different resolution than the SVM classifier with extraction of Histogram of Oriented Gradients (HOG) features. Strength and weaknesses of supervised machine learning two methods elaborated for automatic detect and recognize underwater coral reef fishes from videos [13].

Author [14] discussed the detecting and classifying various underwater marine objects using deep learning. Deep Learning based underwater pose estimation & detection of object detection [15], author developed synthetic training image set with annotation by using 3D CAD model and proved that it is very much potential for detection.

3. Proposed approach for underwater human detection

Our proposed work divided into training phase and testing phase and the procedure as follows -

- Collect the video data
- Convert video data into to the images or frames.
- Resize the images into the one particular size
- Divide these images into training and testing images & annotate to each image
- Convert into the XML files to get the classes of those images as labeled images
- Create TF Records (test and train) which can be served input data for the training.
- Apply three Data augmentation technique to detect the human in any mis-focus of camera or any environmental conditions
- Apply Faster R-CNN (Object detection) Algorithm for training the model and testing the model.
- Evaluate the performance of the model with PASCAL VOC metrics for this human detection.

Fig. 1 shows a brief overview of design and implementation of complete scenario is performed.

4. Result analysis

For object detection challenges the PASCAL VOC evaluation metrics are popular. In this proposed method PASCAL VOC used to evaluate the model in terms of precision, recall, average precision (AP), Precision × Recall

4.1. Performance metrics

The important performance metrics are discussed as follows-

- True Positive (TP) – if a human is present, the algorithm is still recognized as a human.
- False Positive (FP) – If human is not present, except for human algorithms.
- False Negative (FN) – if human, it does not recognize human algorithms.
- Real Negative (TN) – if no humans are around, but nothing is discovered.
- Precision—The no. of objects expected to be human and to be human in the underwater is described as precision. Precision

Recall: Recall is defined as the fraction of humans that are predicted as humans.

Accuracy × Curvature recall:
Reminder = (TP + FN) Reminder.

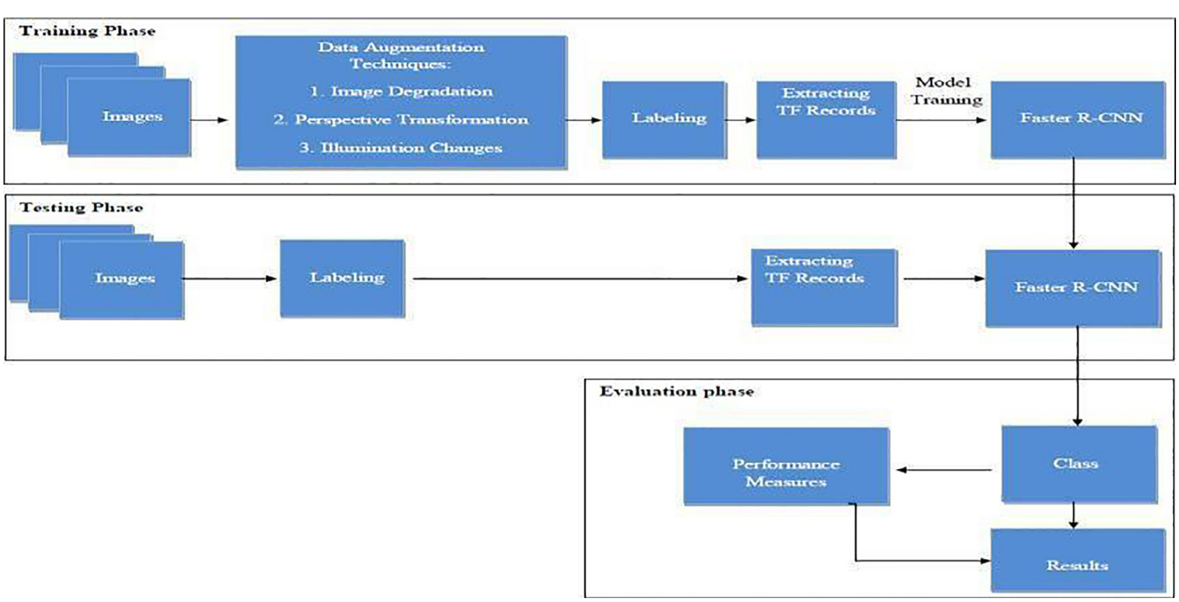

Fig. 1. Proposed architecture.

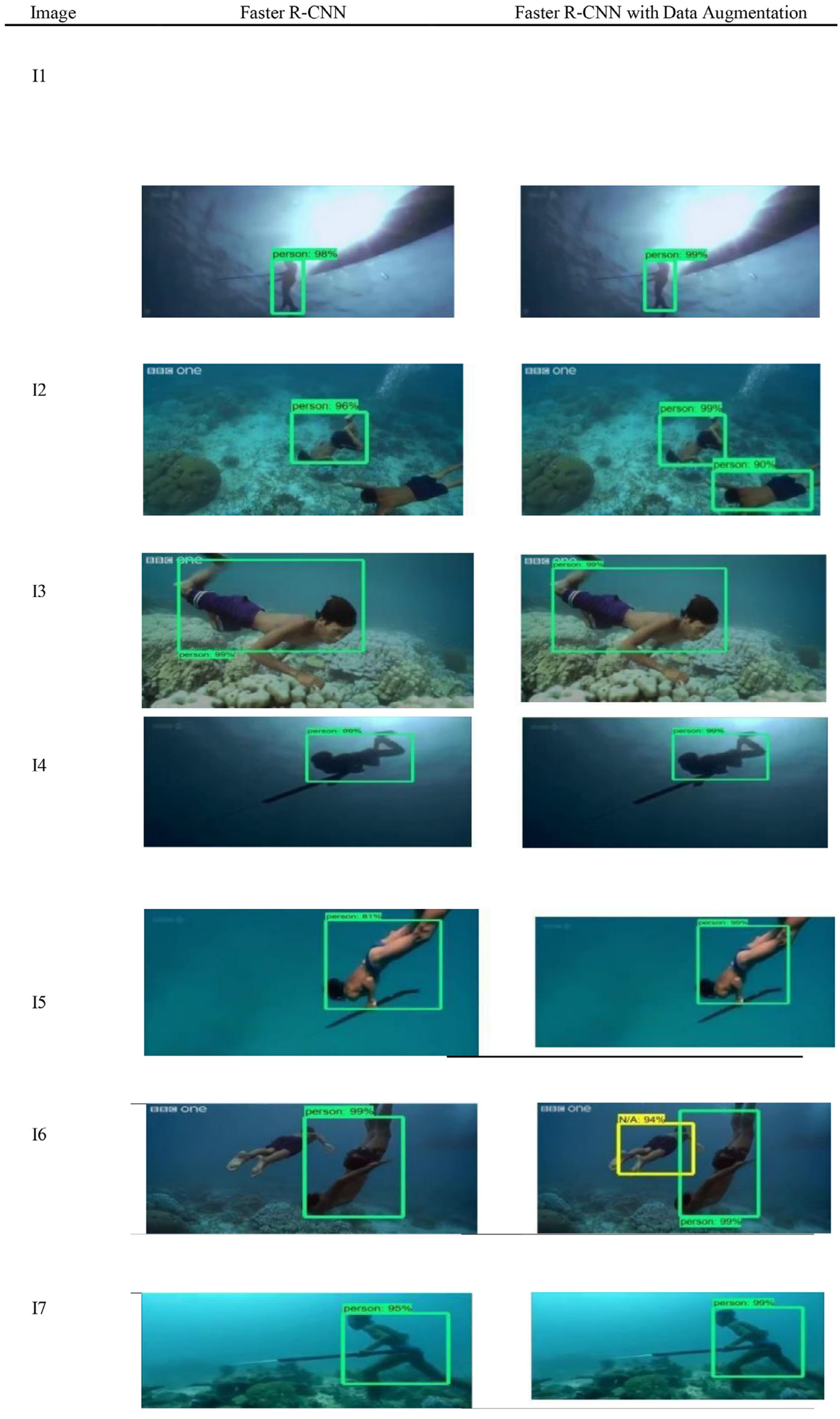

Fig. 2. Detection of humans in an underwater environment for sample images.

Table 1
Comparison of performance parameters of models.

Images	Proposed system(Faster R-CNN)			Proposed (Faster R-CNN + Data Augmentation)		
	Precision	Recall	Confidence	Precision	Recall	Confidence
I1	1.00	0.14	98%	1.00	0.10	99%
I2	1.00	0.29	96%	1.00,1.00	0.20,0.40	99%,90%
I3	0.67	0.29	99%	1.00	0.40	99%
I4	0.75	0.43	99%	1.00	0.50	99%
I5	0.60	0.43	81%	0.83,0.86	0.50,0.60	94%,99%
I6	0.50	0.43	99%	0.88,0.78	0.70,0.70	94%,99%
I7	0.57	0.57	95%	0.70	0.70	99%
	AP = 47.45%					
				AP = 67.50%		

In order to determine the trust shifted by plotting the curve for each item class for an object detector, the precision × recall curve [1] is used. If precise reminder remains strong, a specific object detector class is considered fine.

Average accuracy (AP): Average accuracy curves are zigzagged. The region under the curve (AUC) of the precision × recall curve [1] for an object sensor was measured also using this metric

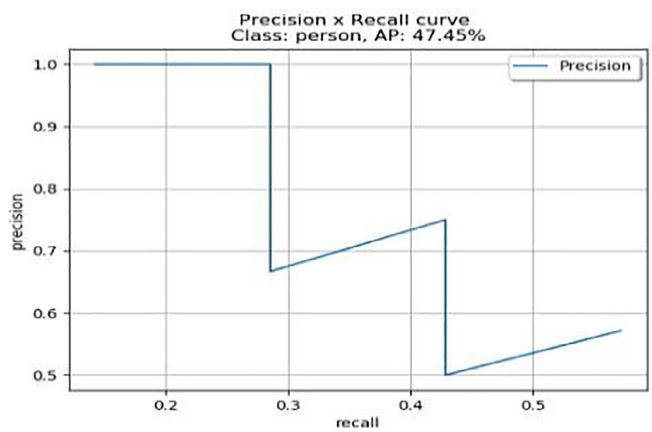

Fig. 3. Precision × Recall by Faster R-CNN.

4.2. Dataset information

Two videos are taken from the internet. One views a moving person and camera is fixed. The other views a moving person and the camera is not fixed. One video is of 1 min and 51 s and other one is 4 min and 56 s. These videos are further divided into frames for the detection.

Video Dataset
Video 1–
Video 2 –
Frame Width and Height: 640 × 360 Video Length: 04:56 sec

https://www.youtube.com/watch?v=YDspP4BhlTw
Frame Width and Frame Height: 640 × 360 Video Length: 01:51 sec
https://www.youtube.com/watch?v=YIKm3Pq9U8M

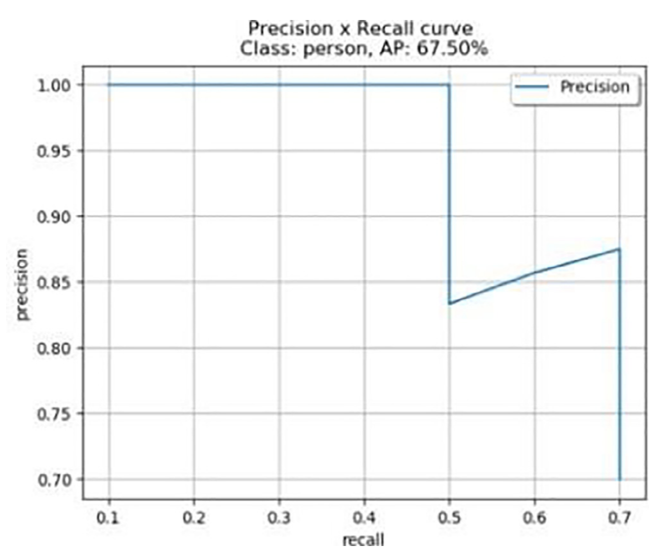

Fig. 4. Precision × Recall by Faster R-CNN with Data Augmentation.

5. Results

In this section, the performance of the proposed Faster R-CNN with Data augmentation better than Faster R-CNN as shown in Fig. 2 and Table 1. Fig. 3 shows the AP (average precision) by Faster R-CNN is 47.45% and Fig. 4 shows the AP (average precision) by Faster R-CNN with Data Augmentation is 67.5% of sample images. Fig. 6 shows the AP (average precision) for video1 by proposed system is 78.98% and Fig. 7 shows the AP (average precision) for video2 is 51.64% (See Fig. 5) shows Detection of human in various underwater conditions with Data Augmentation.

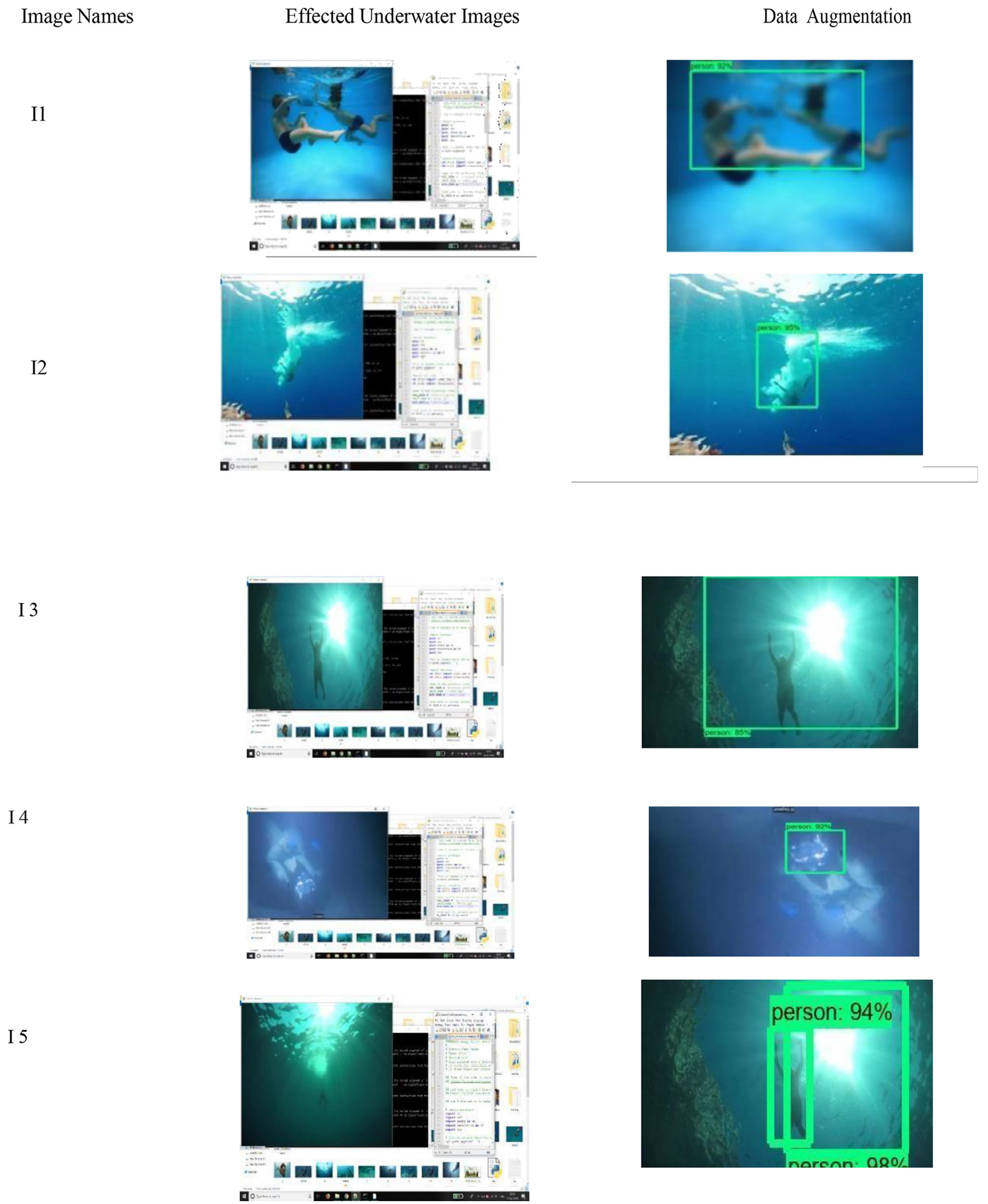

Fig. 5. Detection of human in various underwater conditions after Data Augmentation.

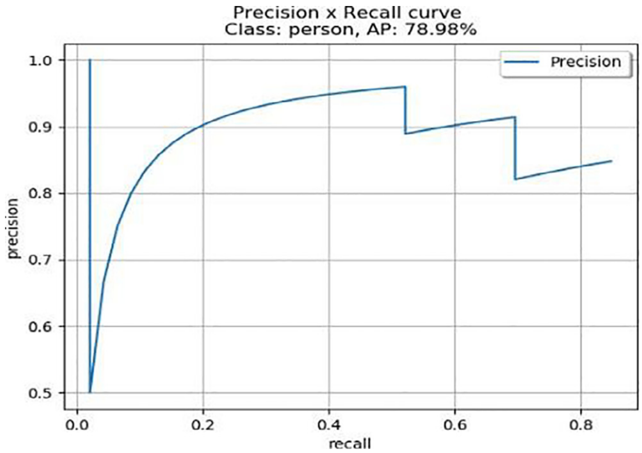

Fig. 6. Precision × Recall by Faster R-CNN with DA for video 1.

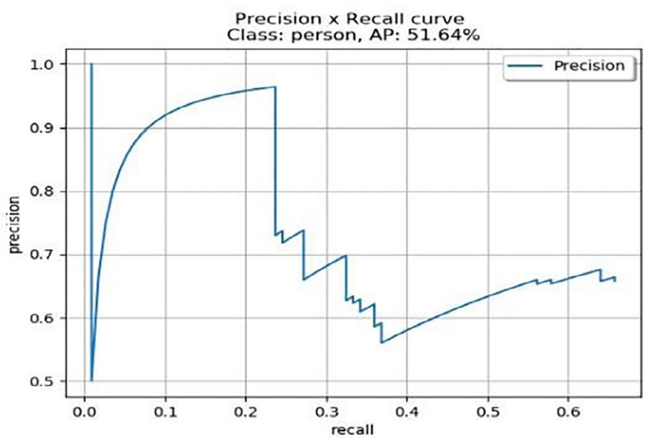

Fig. 7. Precision × Recall by Faster R-CNN with DA for video 2.

5.1. Results of underwater effected images after Data augmentation

As mentioned above that randomly collected images from the internet couldn't able to detect the human body by the model. To overcome this, Data augmentation techniques applied and then model is re-trained then got the best results and also model could able to detect the human body as shown in.

6. Conclusion

The experimental results confirmed that the proposed method perform the best results with the data augmentation methods in underwater environment variations. Faster R-CNN with data augmentation gives the better results with the Average precision. The key idea is to save the life by early detection of the suffocating person who stays underwater and have risk of death. In future work, we plan to develop the algorithm to improve the average precision for longer video and also develop IOT based gadget which can be able to detect the actual problem of underwater human by various emotions like anxiety, panic etc.

CRediT authorship contribution statement

Uma N. Dulhare: Conceptualization, Methodology, Software, Visualization. **Mohd Hussam Ali:** Data curation, Supervision, Software, Validation, Writing - original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Fatma Günseli YaŞar ,Hüseyin KusetoĞullar, "Underwater human body detection using computer vision algorithms" IEEE 10.1109/SIU.2018.8404305, Year2018.

[2] Ambeshwar Kumar, R. Manikandan and C. Thaventhiran, **Pattern Recognition Concepts**, in: Dulhare Uma N. (Ed.), Machine Learning and Big Data: Concepts, Algorithms, Tools and Applications, Wiley Scrivener publishing LLC, USA, July 2020.

[3] Wenmiao Lu and Yap-Peng Tan, "A vision-based approach to early detection of drowning incidents in swimming pools," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 2, pp. 159-178, Feb. 2004 doi: 10.1109/TCSVT.2003.821980. Year2004.

[4] Eng H.L., Toh K.A., Yeu W.Y. and Wang J., "DEWS: A live visual surveillance system for early drowning", IEEE Transactions on Circuits and Systems for Video Technology, 18:196-210, Year 2008.

[5] M. R. Prabowo, N. Hudayani, S. Purwiyanti, S. R. Sulistiyanti and F. X. A. Setyawan, "A moving objects detection in underwater video using subtraction of the background model," 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI),Yogyakarta,2017,pp.1-4. doi: 10.1109/EECSI.2017.8239148, Year2017.

[6] Yong, Xu, Ji xiang Donga, Bob Zhang, Dao yun Xu "Background modeling methods in video analysis: A review and comparative evaluation" 2468-2322/2016. Year2016.

[7] Shaik Mohamed, Shahrizat & Tahir, Noorita & Adnan, Rahmi. (2010). Background modelling and background subtraction performance for object detection. 1 - 6. 10.1109/CSPA.2010.5545291, Year2010.

[8] Mourad Moussa, Comparative Study of Statistical background Modeling and Subtraction, Indones. J. Electr. Eng. Comput. Sci. 8 (2) (2017) 287–295, https://doi.org/10.11591/ijeecs.v8.i2.pp287-295.

[9] Singh, S., & Gaur, P., Mrs. (2018). Matlab based underwater and satellite image enhancement using auto threshold method. International journal of engineering sciences & research technology, 7(5), 423-428.Year2018

[10] TY - JOUR AU - Lodi Rizzini, Dario AU - Kallasi, Fabjan AU - Oleari, Fabio AU - Caselli, Stefano PY - 2015/06/24 SP - 1 EP - 13 T1 - Investigation of Vision-based Underwater Object Detection with Multiple Datasets VL - 12 DO - 10.5772/60526 JO - International Journal of Advanced Robotic Systems ER . Year2015.

[11] TY - JOUR AU - Foresti, G.L. AU - Gentili, Stefania PY - 2000/03/01 SP - 167 EP - 188 T1 - A Vision Based System for Object Detection in Underwater Images. VL - 14 DO - 10.1142/S021800140000012X JO - IJPRAI ER . Year2000.

[12] Niranjan Anand Hule "Underwater Image Processing For Object Detection" International Journal of Innovative and Emerging Research in Engineering Volume 2, Issue 2, Year2015.

[13] Villon, Sebastien & Chaumont, Marc & Subsol, Gérard & Villéger, Sébastien & Claverie, Thomas & Mouillot, David. (2016). Coral Reef Fish Detection and Recognition in Underwater Videos by Supervised Machine Learning: Comparison Between Deep Learning and HOG+SVM Methods. 10016. 160-171. 10.1007/978-3-319-48680-2_15 Year2016.

[14] TY - BOOK AU - Moniruzzaman, Md AU - Islam, Syed AU - Bennamoun, Mohammed AU - Lavery, Paul PY - 2017/11/23 SP - 150 EP - 160 SN - 978-3-319-70352-7 T1 - Deep Learning on Underwater Marine Object Detection: A Survey DO - 10.1007/978- 3-319-70353 4_13 ER , Year2017.

[15] Eng, How-Lung & Toh, K.A. & Kam, A.H. & Wang, J & Yau, W.-Y. (2003). An automatic drowning detection surveillance system for challenging outdoor pool environments. Proceedings of the IEEE International Conference on Computer Vision. 1. 532-539 vol.1. 10.1109/ICCV.2003.1238393. Year2003.