

Underwater Fish Identification in Real-Time using Convolutional Neural Network

S.K.Aruna

Research Scholar, Department of
Computer Science and Engineering,
Saveetha School of Engineering,
Saveetha Institute of Medical and
Technical Sciences,
Chennai
arunask9001.sse@saveetha.com

N.Deepa

Associate Professor, Department of
Computer Science and Engineering,
Saveetha School of Engineering,
Saveetha Institute of Medical and
Technical Sciences,
Chennai
deepa23narayan@gmail.com

Devi.T

Professor, Department of Computer
Science and Engineering,
Saveetha School of Engineering,
Saveetha Institute of Medical and
Technical Sciences,
Chennai
devi.janu@gmail.com

Abstract— Artificial Intelligence (AI) is the wide application that learns the problem and features by given data and processes the data like the human brain. When a computer program imitates a characteristic of the human brain that is considered "innovator." Among the methods are statistical methods, methods for artificial intelligence, and traditional order to verify the validity. The expansion of AI is also related to virtually infinite storage and an abundance of data, including exchanges, geospatial information, video files, photos, text messages, and audio files. Machine learning is divided into deep learning and deep learning is primarily divided into numerous layers of neural networks. This pattern gives it the ability to learn a lot of information and attempt to replicate the brain function. Increasing the efficiency by attaching more covert layers can be beneficial. It is used to gather the data and transfer the data. Aquaculture production has grown into a barrier to the growth of fish culture and the counting operation represents one of the problems experienced during the spawning process. Previous studies have primarily relied on the application of manual and automated counting techniques, which has prevented it from producing accurate results. The proposed method offers a promising method for enhancing image detection by combining the IoT techniques. The image data were divided into three categories: low frequency, intermediate density, as well as high frequency. The proposed method has used 8200 images to train and 2500 images for verification. Only the data relevant data sources were used during the train and verification phase in order to find the proper parameters and create a better VGG19 parameter calibration strategy. Consequently, the improved VGG19 model can achieve an accuracy of 98%.

Keywords— Artificial Intelligence, Underwater object detection, Machine Learning, Tracking Deep Learning.

I. INTRODUCTION

A growing shortage of renewable resources and the growth of the world economy have made the exploration of the underwater acoustic popular recently. Additionally, a number of studies as well as activities relating to ocean designing now rely increasingly on underwater images taken from underwater vehicles.

A. Artificial Intelligence

Computers that display intelligence are said to have artificial intelligence (AI). The study of "autonomous algorithms," or any handset that interprets its atmosphere and acts in a way that maximizes its chances of success in achieving a goal, is what the field of artificial intelligence (AI) study in computer science makes reference. Whenever

a machine imitates "intellectual" processes that people typically connect with other human thought, such as "having to learn" and "decision making," the term "neural network" is used. At the moment, AI is capable of comprehending speech sounds, excelling at strategic games consoles (like board game and then go), self-driving vehicles, smart routing in data centers, military simulation models, and decoding complex data.

Reasoning, understanding, making plans, having to learn, communication using natural speech, perspective, and the capacity to move and movement of objects are the main issues of AI technologies. One of the long-term objectives of the field is intelligence. Techniques include integrated AI, intelligent systems, and statistical methods. Computer engineering, arithmetic, cognitive science, philology, ideology, cognitive science, artificial cognitive science, among other fields are all used in the field of AI. When activated, IoT devices produce and collect data, which AI after which analyzes to produce insights and increase productivity. AI gains insights by employing techniques like data learning.

B. Convolution Neural Network

A CNN is a machine learning subset. It is one of several types of neural networks that are used for various applications and types of data. A CNN is a type of network infrastructure for algorithms for deep learning that is particularly used for image processing and pixel data processing activities. There are other kinds of neural network systems in DL, but CNNs are the neural network of choice for identifying and recognizing objects. As a result, they are ideal for Computer Vision (CV) tasks and applications requiring object recognition, such as self-driving cars and face recognition.

(1) Deep Learning Model Layers

CNN has three layers: one is convolutional, second is pooling, and third one is FC Layer. The first layer is the convolutional layer, and the last layer is the FC layer.

The difficulty of CNN intensifies convolution layers to the fully connected layer. This increases the difficulty of enabling the identification of bigger and more complicated portions of a picture till it gets the final result to recognize the component in the entire length of the data.

(1.1) Layer of convolution:

The vast bulk of calculations take place inside the convolutional layer, which serves as the foundation of a CNN. A 2nd convolutional layer can be added after the first. Convolution occurs when a core or funnel within this layer moves across the picture's receptive fields, checking for the presence of a feature.

(1.2) Pooling Layer:

The layer, just like a convoluted layer, reaches a core or funnel throughout the source target image. However, the convolution layer of the max minimizes the pooling of the amount of parameters in the input while also causing a few information loss. On the plus side, this layer reduces the complexity and enhances CNN's efficiency.

(1.3) Fully connected layer:

In this layer of the CNN is where classification of images occurs according to the characteristics retrieved in the preceding layer. FC means that all of the input data or endpoints of one layer are linked to every initiation of the next layer.

Padding: total number of input is calculated by

$$I=m \times m \text{ ---Eq(1)}$$

Where as m is input size

$$\text{And filter size is } f_t \times f_t \text{ ---Eq(2)}$$

when we use padding as pd, the result will be

$$(m-f_t+1) \times (m-f_t+1) \text{ ---Eq(3)}$$

when padding is applied the result image will be the same as the input image.

the result of the padding will be calculated in this equation

$$(m+2pd-f_t+1) \times (m+2pd-f_t+1) \text{ ---Eq(4)}$$

C. Deep Learning-Based Image Recognition

Even before deep learning, recognition software is not always at its best because image characteristics are gathered and demonstrated using a handcrafted feature, which is an algorithm created according to the expertise of researchers. A method for gaining knowledge categorization and extraction of features from training images is the CNN, one form of deep learning, as shown in Fig. 1. and discusses trends in its use for image processing.

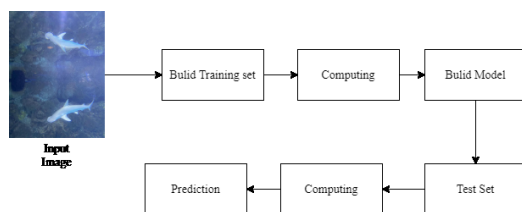


Fig.1.Work Flow of Machine Learning

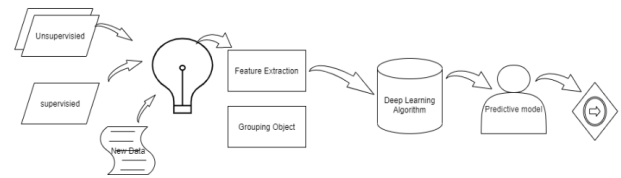


Fig.2.Work Flow of Deep Learning

D. Underwater object detection

Techniques for detecting underwater objects have long been used in research on marine ecology. It is used based on color definitions to identify carp that were being moved on a conveyor system and being watched by a camera. It demonstrated a vision device that employs video edge detection, object recognition, and monitoring to find, and count carp through real-time videos. The aforementioned techniques, however, rely primarily on custom features, which has restricted the ability to represent the crap. It utilized CNN to classify the different species of fishes in the area and a foreground detection technique to extract the fish boundary in the underwater. In this article, samples of underwater objects are primarily thought to have been found. Fishes, debris, and a diver.

II. RELATED WORK

The following reasons are discussed about the existing limitation in speech recognition. The most difficult area in the study of computer vision is submerged anomaly detection. Due to the complex submerged surroundings, submerged object tracking tasks are significantly hampered by submerged images' loud noises, poor accessibility, blurred corners, contrast adjustment, and color variance. Traditional object recognition techniques frequently perform inadequately in terms of precision and generalization skills when used for submerged object detection methods. Many investigators have suggested different submerged item detection methods that utilize deep learning for the precise, steady, generally applicable, actual, and lighter weight detection models required for submerged object detection. Even though submerged object detection has seen many notable successes over years, there is still a lack of truly united induction in the field's research, and some current problems must be addressed. This study can be used as a reference for subsequent studies on underwater object recognition.

Most of the trickiest areas in the study of computer vision today is submerged anomaly detection. Oceanographers must perform all manual tasks that are relevant to submerged goal recognition and classification since traditional marine environment exploration mainly relies on living thing diving operations. In addition to being ineffective and challenging, obtaining more rich information does have additional demands for diving. It is unrealistic to explore for an extended period of time in the deep sea. With the development of computer vision technology, it is now possible for people to use underwater detection tools like ROV and autonomous submarine vehicles to use marine ecological resources without harming them.

GOD and UIE methods are the main sources of inspiration for current submerged object detection techniques. While the latter seeks to improve the appearance of image data with poor quality, the former seeks to more accurately find and recognize targets. Numerous investigators have done in-depth studies on such two activities in recent years. The methods for processing submerged images have been widely summarized. It gave an evaluation and contrast of DL based image improving methods. In-depth research on the evolution of item detection methods over a 20-year period was conducted. The identification and classification processes for naval ground and submerged objects were compiled. The pertinent uses of submerged marine object recognition. The present and past surveys on the detection of submerged objects. Despite this, a deal with various research directions in submerged object recognition is still lacking, especially in regards to the interaction between underwater image processing and submerged object detection. Methods for detecting submerged objects have long been used in research on marine science. Fish were identified using color and shape descriptive terms while being transferred on a conveyor and observed. It presented a visual acuity system that uses youtube clip feature extraction, object recognition, and track detecting, tracking, and number count fish in real-time videos. The lead to the onset, however, heavily relies on palm features, which has restricted the ability. Due to the complexity of the anomaly and the messy backgrounds and objects in the scenes, the problem of anomalous happenings being identified in actual video sequences is difficult. While there have been numerous studies on using deep networks of neurons to solve this issue, very little of the literature focuses on fish anomalous behavior real-time detection.

This study combines fish tracking, directed cycle graph (DCG), learning algorithms, object recognition, and DTW to present a method for identifying fish anomalous behavior underwater Dynamic Time Warping. The technique is helpful in anticipating biologically unusual submerged fish so that quick preventative measures could be prepared for and carried out. Additionally, through thread, the cause of illnesses or deaths can be determined, preventing needless loss. Assist in the precise spawning selection process and promote conserve the environment education. A smart aquaponics system that combines the suggested method and Internet - of - things detectors enables extensive collection of data during the program's operation in various farming fields, allowing for the creation of ideal culturing conditions. Both of these are especially beneficial to scientists and the aquaculture sector. Table 1: Illustrates the challenges faced in underwater objects.

Table 1: Identifying underwater objects is a difficult task

Endeavour	Techniques
Loss of image quality	Image processing before enhancement
	Interconnected image acquisition in the detector
Detection of tiny items	recognition on various scales
	Knowledge about the context
	Task of taking input of a low resolution
	The ratio of constructive and detrimental examples
Inadequately aggregated	Data improvement
	Transforming the web address

The following are the key distinctions between this study and previous surveys:

Systematic evaluation of submerged object recognition research challenges: This paper conducts a thorough literature review of related works on submerged object recognition. The primary difficulties in object recognition in the underwater image are systematically identified. We have extensively discussed three issues: picture quality degeneration, small object tracking, and poor generalization.

Thorough review in light of noted difficulties: An absolutely superb review of the literature must provide data about what the survey method is, what obstacles researchers have been challenged to, and what remedies are suggested. This literature survey is more valuable and functional for the understanding of submerged marine object recognition thanks to an in-depth investigation of core technologies and cutting-edge approaches within the structure of three research challenges.

A deeper understanding of upcoming trends: Any survey should include a discussion of how submerged object recognition will develop. This paper provides a deeper understanding of upcoming trends in underwater object detection based on cutting-edge techniques in submerged object recognition and forethought of the next artificial intelligence.

III. PROPOSED WORK

In this proposed method an algorithm to identify species from underwater using CNN. A fixed shape contour and a color pattern are two instances of low-level characteristics that are matched in the conventional approach to object monitoring. Due to poor lighting, loud noises, and picture haze, this method failed to work well in underwater images. A CNN, on the various sides, uses a number of layers of structure to simultaneously learn high-level characteristics and the clustering algorithm that uses these characteristics. As long as the test is used, the learned high-level characteristics in this situation are more probable to succeed than generic low-level characteristics. The main objectives of this paper are to increase computation speed and accuracy. Numerous object detection techniques have been created

with a focus on processing data more quickly for real-time applications. Only Look Once VGG19 algorithm, among these, for the real-time fish detection process. VGG19 is an organization that implements that period and knows the place, length, and learning algorithm of an item, in contrast to other Cnn models that detect objects by sliding trained classification model. As a result, VGG19 performs quickly. We implemented a CNN with 24 convolutional layers and two layers that are completely linked using the VGG19 structure. We then used our unique dataset to train the network. Fig 3 illustrates the block diagram for fish detection using VGG19 architecture. The object detection is constructed as,

$$H = (X,D) \quad \text{---Eq(5)}$$

Where X and D denote the vertex and edge of the fish

where X and D denote vertex and edge of the fish and H denotes the graph. When a fish is identified, VGG19 will sketch a boundary to cover the whole fish, while also allocating a completely separate vertices to each clearly recognized body part that corresponds to the center of the frame. Only four of the eight essential body parts—the body, rostral, pec, glute, head, eye, and mouth—are actually required to define a stance (see Fig.3).

The subscapular fin, gluteal fin, visual acuity, and mouth in particular serve as auxiliary components in the construction of a DCG. Eye or mouth can take the place of the head as a node if VGG19 is unable to detect, for example, a head part as long as one of the pec and pelvic blades is correctly identified, it can also be used as a fin component. This tactic essentially improves the accuracy of object recognition, improving system performance as a whole.

Fish swimming directions differ from those of cars, making it challenging to track them because they can change suddenly and unexpectedly. The fact that species are all alike is even worse. We once more use the bounding box's coordinates to solve these issues. To ascertain whether 2 packages contain the same fish, we calculate the intersect area between the box's exact location of the consecutive frames as well as the story it states. Identify Fit as that of the fish at the gth frame as the set of fish being tracked. The fish with TH:5 corresponds to the fish with TH:1 at the previous frame

Algorithm 1 CNN using VGG19

Training images xyz train as input

Test images xyz test while training underlying data objects $Z = z_1, \dots, z_x$. Results of the detection: W.

1 = a1 ,B = 1, ..., A.

1: starting point the item weights I_j in order

2: for $g = 1$ to G do

3: starting point the item weights I_j in order Use to calculate the weights of samples tested (15)

4: Use (1)- to train the gth VGG19 $H_m(3)$.5: compute the gth VGG19

5: determine the gth VGG19 error rate E_m (6).

6: Calculate the ensemble model's gth VGG19 weight using (7).

7: Reduce the weights of objects that aren't being detected and boost the weights of those that are by using (8).

8: Finish up

9: Get the final object detections U using (9).

10 : detection output U

IV. RESULTS AND DISCUSSION

In addition to our generated image data set, we also gathered a number of underwater pictures for training, verifying, and experimenting the recommended deep submerged object detection network, as shown in Table 2. Only video data from some data sets can be found in Table 2, and we chose a few related video frames to include in our data set.

A. Dataset Description

Table.2. Description of dataset

Dataset Description	
Object Types	Number of Images
Fishes	5,300
Underwater debris	1,100
Divers	750
fishes/divers/debris	2,000
Total images of underwater	9,150
Training and Verification	8,200/2,200
Testing	2,500

As shown in Table 2, we chose 8200 pictures to train and 2200 images for verification to generate our training given dataset. Only the data from the relevant data source were used in our model during the training and validating phase. Additionally, the percentages of fish, debris, and divers in our data set represent roughly 25%, 50%, and 25% of the maximum image collection, respectively. The sample size is 32, and each picture fix is 320 x 320 in shape. In our experimental studies, the learning rate is originally set to 0.002 and decomposes by 0.2 for every of the ten epochs if the loss also isn't further reduced. If the damage is not further lowered for a period of twenty epochs, the slightly earlier stop is triggered. Additionally, the Leaky ReLU function is the perceptron implemented in the proposed deep model and the model parameters were prefixed based on the weight initialization technique. Fig 4: illustrate about the training and verification of the underwater pictures of fishes. Fig 4 and 5 shows the accuracy range and loss rate for the VGG19 model.

Table.4. Comparison Table

Comparison Table with neural network architecture					
Methodology	Rachis	Sea cucumber	scallop	map	sea
ssd300	VGG16	30.5	69	45	70
YOLOv3	DarkNet53	70	71	40	75
Faster RCNN	ResNet50	78	67	85	90
CNN	VGG19	90	75	95	98

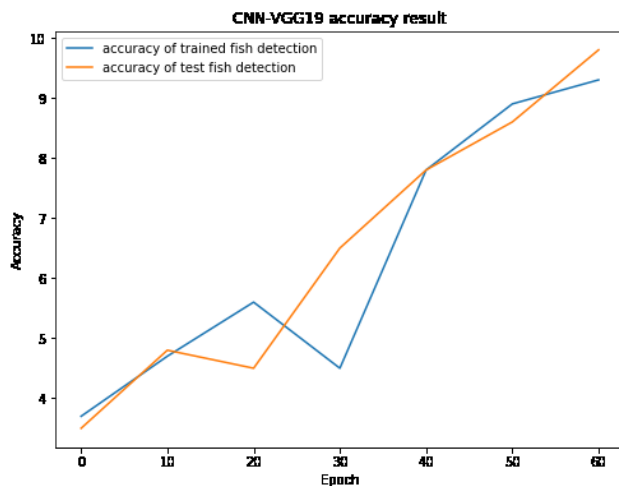


Fig.4.Accuracy Range of Proposed Model

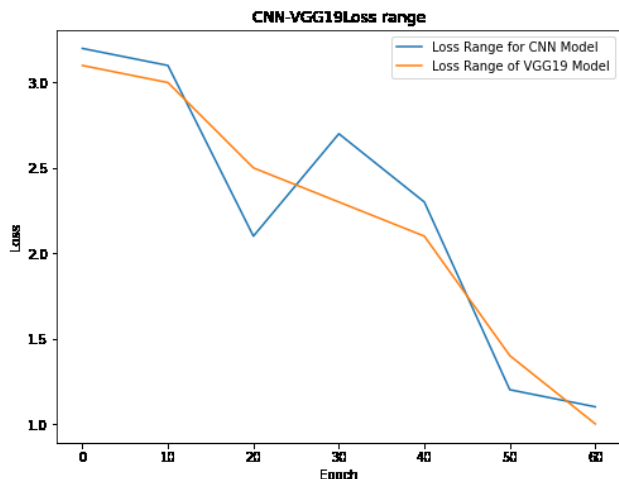


Fig.5.Loss Range of VGG19 Model

V. CONCLUSION AND FUTURE WORK

The proposed study provides a new approach to solve the problem of submerged fish intrusion detection. Three things constitute the study's primary contributions: (1) using data from feature extraction algorithms to enforce a chart classification for ening a fish stance, (2) realizing true unusual behavior rec _ ition by treating fish behavior as an

encoded time - series data, each of which is made up of a number of body posture encrypted during an identity object tracking, and (3) supplying a structure for the development of AI training datasets and open datasets for fish behaviors. Proposed VGG19 approach is more cost-effective than LSTM and RNN, because underground having to learn object recognition requires much less labelled data and, consequently, far less computational power to develop the object recognition of fish parts of the body.

REFERENCES

- [1] Alice, K., Deepa, N., Devi, T., BeenaRani, B. B., & Nagaraju, V. (2023). Effect of multi filters in glucoma detection using random forest classifier. *Measurement: Sensors*, 25, 100566.
- [2] Prasad, B. R., & Deepa, N. (2021). Classification of analyzed text in speech recognition using RNN-LSTM in comparison with convolutional neural network to improve precision for identification of keywords. *REVISTA GEINTEC-GESTAO INOVACAO E TECNOLOGIAS*, 11(2), 1097-1108.
- [3] Tulasi, K. S., & Deepa, N. Sentimental Analysis To Improve Teaching And Learning European. *Journal of Molecular & Clinical Medicine*, 7(3), 2194-2200.
- [4] N. Deepa and D. T, "E-TLCNN Classification using DenseNet on Various Features of Hypertensive Retinopathy (HR) for Predicting the Accuracy," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1648-1652
- [5] Wang, Jung-Hua, et al. "Anomalous behaviors detection for underwater fish using ai techniques." *IEEE access* 8 (2020): 224372-224382.
- [6] Y.-C. Lai, R.-J. Huang, Y.-P. Kuo, C.-Y. Tsao, J.-H. Wang, and C.-C. Chang, "Underwater target tracking via 3D convolutional networks," in *Proc. IEEE 6th Int. Conf. Ind. Eng. Appl. (ICIEA)*, Apr. 2019, pp. 485–490.
- [7] R.-J.Huang,C.-Y.Tsao,Y.-P.Kuo,Y.-C.Lai,C.LiuZ.-W.Tu,J.-H.Wang, and C.-C. Chang, "Fast visual tracking based on convolutional networks," *Sensors*, vol. 18, no. 8, p. 2405, Jul. 2018.
- [8] C. C. Liu, "Incorporating object detection and stereo vision for real-time underwater fish range measurement," M.S. thesis, Nat. Taiwan Ocean Univ., Keelung, Taiwan, 2018.
- [9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.
- [10] Y. C. Lai, R. J. Huang, Y. P. Kuo, C. Y. Tsao, J. H. Wang, and C. C. Chang, "Underwater target tracking via 3d convolutional networks," in *2019 IEEE 6th Int Conf. on Ind. Eng. and Appl.*, pp. 485–490.
- [11] Muwei Jian, Qiang Qi, et al. "The OUC-Vision Large-Scale Underwater Image Database". *Proceedings 2017 IEEE International Conference on Multimedia and Expo (ICME 2017)*, Hong Kong, July, 2017.
- [12] Jian, M., Qi, Q., Dong, J., Yin, Y., Lam, K. M. Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection. *Journal of visual communication and image representation*, 53, 31-41, 2018.
- [13] Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. Learning to reweight examples for robust deep learning. *arXiv preprint arXiv:1803.09050*, 2018.
- [14] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. In *Proceedings of the IEEE' international conference on computer vision*, pages 2980–2988, 2017.
- [15] Tomasz Malisiewicz, Abhinav Gupta, Alexei A Efros, et al. Ensemble of exemplarsvms for object detection and beyond. In *Iccv*, volume 1, page 6. Citeseer, 2011.
- [16] M Pawan Kumar, Benjamin Packer, and Daphne Koller. Self-paced learning for latent variable models. In *Advances in Neural Information Processing Systems*, pages 1189–1197, 2010.
- [17] Rathi, D., Jain, S., & Indu, D. S. (2018). Underwater Fish Species Classification using Convolutional Neural Network and Deep Learning. *arXiv preprint arXiv:1805.10106*

- [18] Kratzert, F., & Mader, H. (2018). Fish species classification in underwater video monitoring using Convolutional Neural Networks. OpenKratzert, Frederik, and Helmut Mader. "Fish Species Classification in Underwater Video Monitoring Using Convolutional Neural Networks". EarthArXiv, 15
- [19] P., & Müller, H. (2017, September). Lifeclef 2017 lab overview: multimedia species identification challenges. In International Conference of the Cross-Language Evaluation Forum for European Languages (pp. 255-274). Springer, Cham.
- [20] Matzner, S., Hull, R. E., Harker-Klimes, G., & Cullinan, V. I. (2017, September). Studying fish near ocean energy devices using underwater video. In OCEANS-Anchorage, 2017 (pp. 1-7). IEEE.
- [21] Sungbin Choi. Fish identification in underwater video with deep convolutional neural network: Snmedinfo at lifeclef fish task 2015. In CLEF (Working Notes), 2015.
- [22] Sebastien Villon, Marc Chaumont, G' erard Subsol, S' ebastien Vill' eger, Thomas Claverie, and David Mouillot. Coral reef fish detection and recognition in underwater videos by supervised machine learning: Comparison between deep learning and hog+ svm methods. In International Conference on Advanced Concepts for Intelligent Vision Systems, pages 160–171. Springer, 2016.
- [23] Xiu Li, Min Shang, Hongwei Qin, and Liansheng Chen. Fast accurate fish detection and recognition of underwater images with fast r-cnn. In OCEANS2015-MTS/IEEE Washington, pages 1–5. IEEE, 2015.
- [24] Ross Girshick. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision, pages 1440–1448, 2015.
- [25] Xiu Li, Min Shang, Jing Hao, and Zhixiong Yang. Accelerating fish detection and recognition by sharing cnns with objectness learning. In OCEANS2016-Shanghai, pages 1–5. IEEE, 2016.
- [26] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems, pages 91–99, 2015.
- [27] Muwei Jian, Qiang Qi, et al. "The OUC-Vision Large-Scale Underwater Image Database". Proceedings 2017 IEEE International Conference on Multimedia and Expo (ICME 2017), Hong Kong, July, 2017.
- [28] Jian, M., Qi, Q., Dong, J., Yin, Y., Lam, K. M. Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection. Journal of visual communication and image representation, 53, 31-41, 2018.
- [29] Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. Learning to reweight examples for robust deep learning. arXiv preprint arXiv:1803.09050, 2018.
- [30] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision, pages 2980–2988, 2017.
- [31] Tomasz Malisiewicz, Abhinav Gupta, Alexei A Efros, et al. Ensemble of exemplar svm for object detection and beyond. In lccv, volume 1, page 6. Citeseer, 2011.