12th International Seminar on Inland Waterways and Waterborne Transportation

Virtual, 19th to 21st October de 2021

A Review of Deep Learning Application for Computational Vision within the Maritime Industry

Ramiro Fernandes Ramos, UFRJ/COPPE, Rio de Janeiro/Brazil**,** ramirofr@oceanica.ufrj.br
Mikael Strauhs, UFRJ, Rio de Janeiro/Brazil, mikaelstrauhs@poli.ufrj.br
Severino Virgínio Martins Neto, UFRJ, Rio de Janeiro/Brazil, severino_virginio@poli.ufrj.br
Antonio Rafael Paulino de Lira, UFRJ, Rio de Janeiro/Brazil, rafaellira@poli.ufrj.br
Maricruz A. Fun Sang Cepeda, UFRJ/COPPE, Rio de Janeiro/Brazil**,** maricruzcepeda@oceanica.ufrj.br
Luiz Felipe Guaycuru de Carvalho, UFRJ/COPPE, Rio de Janeiro/Brazil**,** luiz_salomon@oceanica.ufrj.br
João Vitor Marques de Oliveira Moita, UFRJ/COPPE, Rio de Janeiro/Brazil**,** joaov@oceanica.ufrj.br
Jean-David Caprace, UFRJ/COPPE, Rio de Janeiro/Brazil**,** jdcaprace@oceanica.ufrj.br

Abstract

Computational vision is the ability of a machine to process images or videos. Many technologies are used to develop this ability. Deep learning is a technique inside the machine learning area, where a neural network is used but with a more profound complexity than the usual neural networks. Deep learning is changing the scenario for computational vision since it is more efficient than the others and is becoming more popular. Given this popularity, deep learning for computational vision is applied for multiple fields of study. In the maritime industry, this technique is being used by different researchers to help them identify and classify ships automatically through images and videos. Also, it is being applied in subsea inspections to simplify the task of detecting objects in the seabed. This paper reviews the current state of the art of using deep learning techniques for computational vision within the Maritime Industry. The target for each study is analyzed. A comparison of the dataset used and the type of deep learning are made. The most common target is to identify ships using surveillance cameras or satellite images. Subsea equipment, corals, underwater mines, marine organisms, and oil spills are other targets. The most used deep learning technique is the convolutional networks. This result is not only observed in the maritime domain but for any computational vision problem. The neighbors' data influence the convolutional network result, and the pixel of an image is very similar to its neighbor pixel, evidencing its advantage. There are at least ten different types of convolutional networks being used in the reviewed papers, making it clear that there is space for innovation in this matter. This review and comparison can help future research, giving information on which deep learning technique to choose and how to evaluate its target.

1. Introduction

Computational vision is understood as a machine's ability to process images or videos by classifying different objects, finding and counting them on the image, or identifying each pixel that composes the target. The computational vision is usually divided into four groups, as shown in Figure 1: classification, localization, detection, and segmentation. All can be objectives for a computational vision algorithm. It will depend on the purpose of the algorithm.

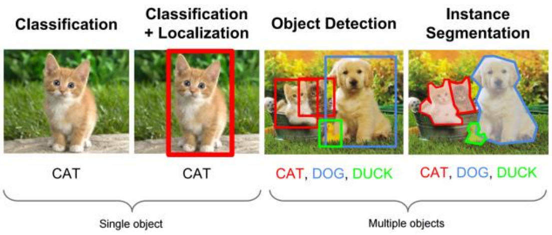

Figure 1 - Difference between classification, localization, detection, and segmentation (Jaiswal et al., 2021)

Deep learning is a technology inside the machine learning technologies that aims to get even closer to how the human brain works. It is very similar to the already known neural network algorithms. However, it goes deeper in complexity. It usually has more hidden layers or an elaborate structure (Benuwa et al., 2016). One of the first uses of deep learning in computational vision is described in LeCun et al. (1998). It shows the creation of the convolutional neural network (CNN) algorithm and applies it to a letter identification problem. The same author published another paper years later, demonstrating that deep learning is already consolidated in both industrial and academic communities (LeCun et al., 2015).

The maritime industry comprehends many different activities: maritime transportation, offshore oil & gas exploration, shipbuilding, coast surveillance, offshore wind turbines, and others. The naval sector involves every activity connected to the sea or waterways.

The maritime industry is present in different countries and different ways. In Brazil, for instance, offshore exploration is one of the most significant Brazilian activities. Subsea systems are current in most of the oil fields in Brazil and have massive relevance for the oil & gas industry. Inspecting the equipment of those systems can be very tricky. show that this inspection can benefit from deep learning techniques applied for computational vision in AUV images.

Subsea images taken from AUV and ROV are also used for many other purposes and not necessarily with deep learning techniques. In Costa et al. (2019), an AUV is used to explore underwater archaeology in Portugal, where somebody found a shipwreck of a German boat from World War II. In Walther et al. (2004), it is described how they used the video captured by an ROV to detect and track objects, such as marine animals. Also, a robot was used by Lee et al. (2012) to take videos from subsea targets and follow them. They also evaluated the impact of the object's geometry on the result.

This activity is present almost everywhere and has a significant impact on the international economy. Ramos et al. (2019)pointed out that maritime transportation was responsible for 80% of the global trade, being the most crucial transportation mode. One opportunity for computational vision in marine transportation is to use a surveillance camera to identify the ships arriving or leaving a port. It is also a good objective for safety and military authorities. As demonstrated by Strauhs et al. (2020), a deep learning network is completely capable of identifying ships and classify them by their correspondent type.

For those examples and many other applications, there are many studies, papers, and practical applications already publicly available. However, as pointed by Prasad et al. (2019), maritime computer vision is in a nascent stage at this moment. It is essential to assemble in one paper the results obtained by others to help its development. Also, a correlation between target and methodology can make it easy for future researchers to choose a method between all available.

With this in mind, this paper aims to review published scientific work that uses a deep learning application for computational vision purposes in the maritime industry. It will discuss and compare the targets of each study, the dataset used, and the chosen deep learning algorithms. Where it is appropriated, the results will be compared, and a correlation between technique and target will be presented.

2. Targets

The target for deep learning studies is the variable or feature that the algorithm predicts or finds.

Machine learning algorithms are generally divided into two main types: the supervised and the unsupervised algorithms. The difference between them is either if the training dataset contains the target variable or not. In supervised, during the training phase, the algorithm will be trying to fit the known input data to find a known target, and then, on the test phase and in an actual application, it will be using a known input data to discover an unknown target. Examples of supervised algorithms are Classification and Regression. Deep Learning neural networks applied to computational vision are typically used as supervised methods. However, as Wang et al. (2020) pointed out, some unsupervised algorithms of deep learning can also be found in this application. Unsupervised algorithms do not see the target either in the training phase or later. It will try to predict the result based on reinforcements chosen by the developer. Examples of unsupervised are Clustering and Association (Love, 2002).

And the target will be chosen by the necessity of the person employing the algorithm. As computational vision is the focus of this review, the target will be the classification of images and videos or the detection of objects. Here, the paper will segregate for the target chosen.

2.1. Ships

A popular target is a ship. Since it is one of the most common objects at sea and there are more cameras on the surface than subsea cameras, identifying or classifying vessels would be a good target for computational vision.

An example of this is the work demonstrated by Strauhs et al. (2020). In this paper, it is possible to see that the target of the algorithm was to classify an image between several ship types. Ore carrier, bulk cargo, and fishing boat are 3 of the six classes that this neural network can classify. And the algorithm is not only to classify the images. It can detect multiple ships in a single image, localizing them all.

A very similar study is presented in Shao et al. (2018). The paper shows different classes of ships being classified and identified in images. The most significant difference is the method used to determine the occlusion of a boat. So, if a vessel is not entirely shown in the picture, the algorithm identifies the size of occlusion.

The work presented by Shan et al. (2021) aimed to detect different ships in video frames. Many ship classes were used, but the neural network does not classify the frame. It only detects each vessel in the image.

Another study that used a deep learning algorithm to identify ships is described in (Gallego et al., 2018). This paper uses aerial images to train a neural network to classify the pictures between 7 different groups. Ship, Detail, Coast, Sea, and Land are some of them. In this case, no localization was made. It can say if there are multiple ships in the same image, using a multi-group. It is an excellent way to separate pictures that have only one ship.

The work presented in Hordiiuk et al. (2019) also uses satellite images of ships. The main contribution of the authors lies in the segmentation of the targets. As shown in Hordiiuk et al. (2019), one single image can contain multiple ships with different formats and pointing toward different directions. To do this segmentation, they found each pixel composing each vessel. It allows not only to segregate but also to count the number of ships in each image.

An image from Strauhs et al. (2020) was chosen in Figure 2 to illustrate the abovementioned description. It shows the neural network classifying and localizing a fishing boat with a 94% of confidence score.

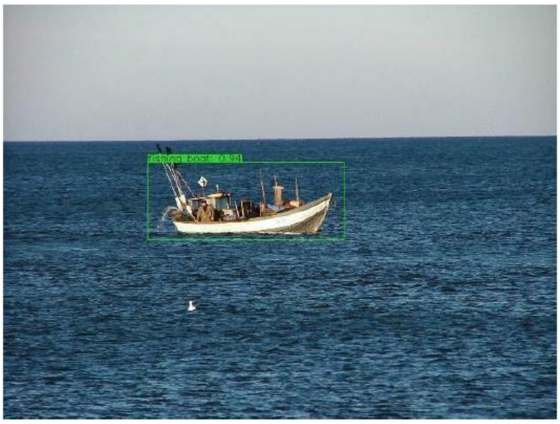

Figure 2 - Fishing boat detected with a confidence score of 94%. (Strauhs et al., 2020)

2.2. Subsea Equipment

On offshore oil fields, lots of equipment are laid on the seabed. Since it is hard to make a visual inspection on those places, autonomous underwater vehicles (AUV), remotely operated vehicles (ROV), or autonomous surface vehicles (ASV) are typically used to do these inspections. They can work together with the computational vision.

The first example of a deep learning model with a subsea target is described in Moita et al. (2020). In this case, the target was finding oil pipes crossings that lay on the seabed. The authors analyzed both the classification of the images having a crossing event and the localization of the event. Note that in this case, only the crossing was the target. Therefore, the algorithm's answer for each image was binary, meaning either if the image had a crossing or not.

A very similar target was described in Martin-Abadal et al. (2021). They focused on underwater pipes and valves. Unlike Moita et al. (2020), they made a 3D recognition, identifying each pixel corresponding to a pipeline or a valve in the image presented, making a segmentation. As well, in this case, two classes were the target. The algorithm could identify and distinguish the valves from the pipes.

Another study that goes on this target is the Inzartsev et al. (2019). The main objective of this study is to make an AUV follow a pipeline automatically. To this, the target of the used algorithm is to identify the pipe. Them, with post-processing, it is possible to make a vector between the identifications that will point the direction of the pipeline. It is only possible to work with a series of images or a video in this specific case.

n image of the result of Moita et al. (2020) is presented in Figure 3 to illustrate the subsea target. It demonstrates two crossings of subsea pipes and a rectangle localizing them with a 99% confidence score.

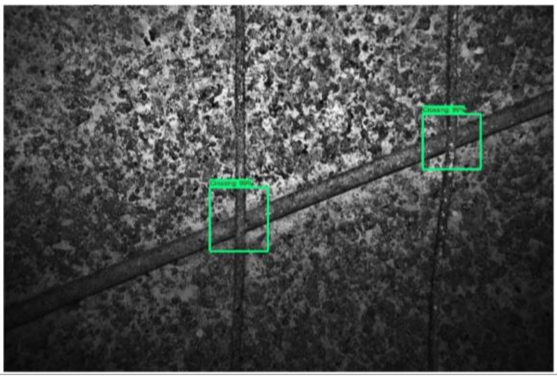

Figure 3 - Detection of crossing pipes (Moita et al., 2020)

2.3. Corals

Corals may not be the most common application of this review, but a significant one. The coral's location is essential for dredging or offshore operations since it has a high significance for the environment.

A study pointing to this type of target can be found in Raphael et al. (2020). It uses images in shallow water to identify points that can have coral on them. Then, transform the pixels around this point into sub-images and classify them by the type of coral on them. To cite a few types: Acronora, Favia, and Platygyra are between them. This approach is interesting because it detects a point in the first image and classifies the second. But since the second image is a part of the original image, the result is detection based on the type of coral.

Another work with this target is presented in Mahmood et al. (2017). They also classify corals based on the species shown in the image. In this case, only the classification is made.

To better illustrate the coral's target, Figure 4 brings a detection of 4 different species of coral. The intermediate result presented in Raphael et al. (2020) is the intermediate result, where its sub-images will be classified later.

Figure 4 - Coral species detection (Raphael et al., 2020)

2.4. Other targets

Many maritime objects can be the target of a deep learning network applied for computational vision. Ships and subsea equipment are frequently more used, and corals have been gaining the scientific community's attention.

However, there are a few more targets that are worth the mention. Finding underwater mine warfare was the target in Denos et al. (2017), where they applied deep learning in AUV data. In Lu et al. (2018), the computational vision was able to identify marine organisms such as shrimps and crabs. In Temitope Yekeen et al. (2020), a deep learning model is used in satellite images to identify oil spills.

3. Datasets

Datasets assemble data that will be used to train, test, and validate the machine-learning algorithm created. The data can be presented in multiple formats: text, number, audio, images, and videos. In the context of this paper, pictures and videos are the formats present in the datasets since it is exploring computational vision problems.

The dataset is essential for the result of the algorithm since it will learn from it. Because of this, the data has to be representative of the problem approached. If the resulting algorithm will be applied in surveillance cameras, the dataset is expected to be composed of surveillance images and not satellite images.

Also, if it is wanted to classify ships in select types, the dataset must have a significant number of images for each class. If one type has few pictures, it can be possible that the algorithm will not have high accuracy for this type. Those considerations are essential to have in mind when evaluating the dataset used in different papers.

For instance, in Strauhs et al. (2020), there are six types of ships studied. Some groups contain more than one known ship type. It is, for instance, the case for the group of the bulk cargo carrier. It could be divided into oil cargo, chemical cargo, and edible cargo. But since this group represents only 25.87% of the dataset, dividing it could low the total accuracy.

The dataset used by Strauhs et al. (2020) is a subset containing 7000 images of the SeaShip dataset (Shao et al., 2018), a publicly available dataset formed by lateral photos of ships in China waters. This dataset has an excellent example of representativeness. Figure 5 shows the ore carrier type considered in this dataset. It is specific from the location where the dataset was made, and it has many differences compared to a common ore carrier, shown in Figure 6.

At Shan et al. (2021), instead of images, videos were used. Those videos were separated into frames, and those frames were treated as images. The dataset was composed of 224 videos, and a total of 33018 frames were extracted.

Figure 5 - Ore carrier and bulk carrier identified (Strauhs et al., 2020)

Figure 6 - A common ore carrier (available at https://gcaptain.com/vale-brasil-worlds-largest-carrier/, accessed in 28/08/2021)

As mentioned, the objective can be to use satellite images for composing the dataset. For this, many papers can be cited. Liu et al. (2017) use a dataset called HRSC2016 containing more than 1500 satellite images of ships moored at ports, and Gallego et al. (2018) used a freely available dataset containing 6000 optical aerial photos (MASATI). The dataset called SPOT-5 contains 1000 satellite images, and it is used in Corbane et al. (2008). Also, Liu et al. (2017) use a dataset called QuickBird containing 150 images from satellites. All these works cited for satellite images have vessels as targets. Figure 7 brings an example of satellite images.

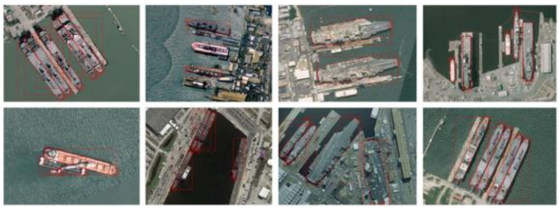

Figure 7 - Satellite images from the HRSC2016 dataset (Liu et al., 2017)

As mentioned, the datasets used for subsea equipment are mainly formed by AUV, ROV or ASV images. images are taken from an AUV inspection in an offshore oil field. The dataset is not publicly available, and the number of images is not mentioned. However, for Martin-Abadal et al. (2021), the photos are taken from an ROV. They transformed the images into 262 cloud points that were used in the training of the neural network. This dataset of images is publicly available. Figure 3 is a good example of a photo taken from an AUV dataset.

Denos et al. (2017) was also using AUV images. But a significant difference in this dataset is that the authors combined natural and synthetic images. It was developed for the dataset to reach 5000 images.

Another worth pointing dataset is the one used by Lu et al. (2018). Called by the name of Kyutech-10K, this dataset is composed not only of images but also videos. It contains more than 10000 data (image + video) that were captured by subsea cameras. Figure 8 brings some examples of this data.

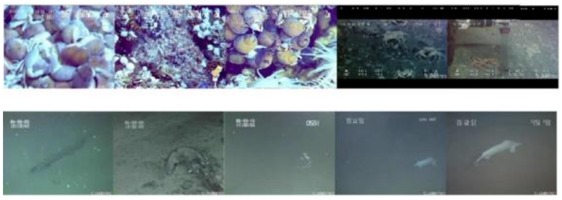

Figure 8 - Dataset of images and videos used to identify marine organisms (Lu et al., 2018)

4. Algorithms

Several types of algorithms can be used in deep learning. As pointed in the introduction, one of them is called Convolutional Neural Network (CNN), first described by LeCun et al. (1998). This algorithm is viral for computational vision problems due to its architecture that considers neighbor points in the analysis. It has an advantage on images and videos, where a pixel usually is similar to its neighbor pixels. However, as expected of a deep learning algorithm, the CNN architecture is not simple. Figure 9 shows the architecture of LeNet-5, the algorithm described by LeCun et al. (1998).

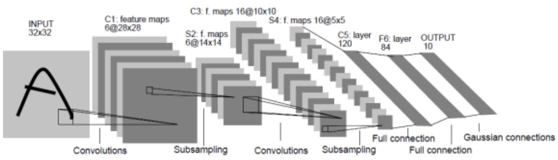

Figure 9 - LeNet-5 architecture (LeCun et al., 1998)

The algorithm used by Strauhs et al. (2020)is known as YOLO, initials for "You Only Look Once". The version used by them for ship detection and classification was the YOLOv4 Bochkovskiy et al. (2020), but the YOLOv5 is already publicly available. All versions of YOLO are based on CNN architecture with different modifications.

The work presented in Shao et al. (2018) brings three different architectures applied to the ship detection and classification problem. It was used a Fast R-CNN, a YOLOv2, and an SSD algorithm. As can be inferred by the name Fast R-CNN, better described in He et al. (2015), it is based on CNN architecture and the other two algorithms used. However, Shao et al. (2018) changed those algorithms differently, creating 12 different applications. The results show that for every class of ship they studied, the better architecture was the Fast R-CNN.

Still, in the neural networks where ships are the target, Shan et al. (2021) used a ResNet-50 with a Feature Pyramid Networks (FPN) structure. This architecture is also based on CNN. Another paper that used ResNet is Gallego et al. (2018). They experimented with five different architectures, all based on CNN: VGG-16, VGG-19 (Simonyan & Zisserman, 2014), Inception Szegedy et al. (2016), ResNet (He et al., 2016), and Xception (Chollet, 2017). In this case, the results vary depending on the set used, but in general, VGG performed better.

For the subsea-focused papers, the architectures are similar. In study, the architecture used to create the neural network was an R-CNN, described in Ren et al. (2016). In Martin-Abadal et al. (2021), the neural network used is called PointNet (Qi et al., 2017). It is a little different from the others above since the objective in this study was the segmentation of pipes.

Jumping to Raphael et al. (2020), where Corals were the target, the author states that CCN and Deep Belief net (DBN) (Elawady, 2015) are the most common type of architecture used in this problem. However, the author himself used CNN architectures, such as VGG-16 and ResNet-50, in his work. On Mahmood et al. (2017) the architecture type used to identify corals was a CNN model.

Some of the cited papers, such as Strauhs et al. (2020) and Moita et al. (2020), used a technique called transfer learning. Transfer learning is a machine learning technique where the calibration of a previously developed learning is reused as a starting point for developing new learning. In this case, the second learning does not start from zero and benefit itself from previous works, even if the target on both works is different Tan et al. (2018).

5. Conclusion

This paper did a review for deep learning application for computational vision within the maritime industry. It was observed multiples scientific works with different targets, datasets, and algorithms.

The targets observed were ships, subsea equipment, and corals. It was also mentioned underwater mines, marine organisms, and oil spills as examples of other types of targets. It was noted that stating the target of a paper is very important to analyze the other aspects. It has a heavy influence on the dataset chosen and can also impact the architecture of the algorithm.

When analyzing the dataset used in each paper, it was clear that it will depend on the final objective of the study. As expected, satellite images are used primarily to form the dataset of neural networks that identify ships because ships are large enough to appear in those images. Subsea equipment and corals are underwater, hidden from those cameras. As pointed, another target that uses satellite is the oil spill.

On the other hand, AUV, ROV, and ASV images are widely used for the dataset of subsea targets. The first reason is that those vehicles are already used for this kind of inspection, and also the fact that AUV and ROV can reach deep waters, being possible to evaluate deep subsea equipment.

Every paper reviewed for this work used at least one CNN architecture to perform the task for algorithms. The conclusion that CNN algorithms are better for computational vision is not exclusively for the maritime domain. This architecture is more effective than the others when treating images or videos. But, inside CNN there is still room for innovations and improvements. At least ten different CNN algorithms were cited in this paper. Results point to pre-built CNN, such as VGG, as a better solution that is highly connected to the target and the objective of the neural network.

With this paper, many authors can discover more about the application of deep learning in the maritime industry. It can help a new project understand where the state of art is and see which architecture and type of dataset should be used for a given target. Since this type of application is still in its early stages, many possibilities are yet to be discovered. Advances can be made and published focusing on a new target or a new algorithm architecture.

6. Acknowledgments

The present work was realized with the partial support of ANP, National agency for Oil, Natural Gas and Biofuels, FINEP, Financier of Studies and Projects and MCTI, Ministry of science, technology, and innovation through PRH18-ANP/MCTI, ANP Program for Oil, Natural Gas and Biofuels, ongoing at the Ocean Engineering Department (https://oceanica.ufrj.br/).

Furthermore, the UFRJ, Federal university of Rio de Janeiro, and the CNPq, Institutional Program for Scientific Initiation Scholarships, also partially support the development of the present work through Institutional Program for Scientific Initiation Scholarships (PIBIC), ongoing at the Ocean Engineering Department.

7. References

Benuwa, B. B., Zhan, Y. Z., Ghansah, B., Wornyo, D. K., & Banaseka Kataka, F. (2016). A review of deep machine learning. *International Journal of Engineering Research in Africa*, *24*, 124-136.

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition,

Corbane, C., Pecoul, E., Demagistri, L., & Petit, M. (2008). Fully automated procedure for ship detection using optical satellite imagery. Remote Sensing of Inland, Coastal, and Oceanic Waters,

Costa, M., Pinto, J., Ribeiro, M., Lima, K., Monteiro, A., Kowalczyk, P., & Sousa, J. (2019). Underwater Archaeology with Light AUVs. OCEANS 2019-Marseille,

Denos, K., Ravaut, M., Fagette, A., & Lim, H.-S. (2017). Deep learning applied to underwater mine warfare. OCEANS 2017-Aberdeen,

Elawady, M. (2015). Sparse coral classification using deep convolutional neural networks. *arXiv preprint arXiv:1511.09067*.

Gallego, A.-J., Pertusa, A., & Gil, P. (2018). Automatic ship classification from optical aerial images with convolutional neural networks. *Remote Sensing*, *10*(4), 511.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, *37*(9), 1904-1916.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition,

Hordiiuk, D., Oliinyk, I., Hnatushenko, V., & Maksymov, K. (2019). Semantic segmentation for ships detection from satellite imagery. 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO),

Inzartsev, A., Eliseenko, G., Panin, M., Pavin, A., Bobkov, V., & Morozov, M. (2019). Underwater pipeline inspection method for AUV based on laser line recognition: Simulation results. 2019 IEEE Underwater Technology (UT),

Jaiswal, A., Babu, A. R., Zadeh, M. Z., Banerjee, D., & Makedon, F. (2021). A survey on contrastive self-supervised learning. *Technologies*, *9*(1), 2.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324.

Lee, D., Kim, G., Kim, D., Myung, H., & Choi, H.-T. (2012). Vision-based object detection and tracking for autonomous navigation of underwater robots. *Ocean Engineering*, *48*, 59-68.

Liu, Z., Yuan, L., Weng, L., & Yang, Y. (2017). A high resolution optical satellite image dataset for ship recognition and some new baselines. International conference on pattern recognition applications and methods,

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic bulletin & review*, *9*(4), 829-835.

Lu, H., Li, Y., Uemura, T., Ge, Z., Xu, X., He, L., Serikawa, S., & Kim, H. (2018). FDCNet: filtering deep convolutional network for marine organism classification. *Multimedia tools and applications*, *77*(17), 21847-21860.

Mahmood, A., Bennamoun, M., An, S., Sohel, F., Boussaid, F., Hovey, R., Kendrick, G., & Fisher, R. B. (2017). Deep learning for coral classification. In *Handbook of neural computation* (pp. 383-401). Elsevier.

Martin-Abadal, M., Piñar-Molina, M., Martorell-Torres, A., Oliver-Codina, G., & Gonzalez-Cid, Y. (2021). Underwater Pipe and Valve 3D Recognition Using Deep Learning Segmentation. *Journal of Marine Science and Engineering*, *9*(1), 5.

Moita, J. V. M. d. O., Ramos, R. F., Soares, L. N. d. A. C., Alvarenga, I. D., Rodrigues, G. P., de Souza, M. I. L., & Caprace, J. D. J. E. M. (2020). A deep learning approach for automatic identification of subsea events using AUV data.

Prasad, D. K., Dong, H., Rajan, D., & Quek, C. (2019). Are object detection assessment criteria ready for maritime computer vision? *IEEE Transactions on Intelligent Transportation Systems*, *21*(12), 5295-5304.

Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition,

Ramos, R. F., F. S. Cepeda, M., Moita, J. V., Monteiro, G., Fiksdahl, O., & Caprace, J. (2019). Detecting Possible Near Miss Collisions in Santos Bay from AIS Big Data. 11th International Seminar on Inland Waterways and Waterborne Transportation, Brasília.

Raphael, A., Dubinsky, Z., Iluz, D., Benichou, J. I., & Netanyahu, N. S. (2020). Deep neural network recognition of shallow water corals in the Gulf of Eilat (Aqaba). *Scientific reports*, *10*(1), 1-11.

Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, *39*(6), 1137-1149.

Shan, Y., Zhou, X., Liu, S., Zhang, Y., & Huang, K. (2021). SiamFPN: A Deep Learning Method for Accurate and Real-Time Maritime Ship Tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, *31*(1), 315-325. https://doi.org/10.1109/TCSVT.2020.2978194

Shao, Z., Wu, W., Wang, Z., Du, W., & Li, C. (2018). Seaships: A large-scale precisely annotated dataset for ship detection. *IEEE transactions on multimedia*, *20*(10), 2593-2604.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Strauhs, M., Lira, A. R., Fernandes Ramos, R., Moita, J. V., Virgínio Martins Neto, S., & Caprace, J. (2020). A Deep Convolutional Neural Network Applied to Ship Detection and Classification. 28th International congress on waterborne transportation, shipbuilding and offshore constructions, Rio de Janeiro.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. Proceedings of the IEEE conference on computer vision and pattern recognition,

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. International conference on artificial neural networks, Temitope Yekeen, S., Balogun, A. L., & Wan Yusof, K. B. (2020). A novel deep learning instance segmentation model for automated marine oil spill detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, *167*, 190-200. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2020.07.011

Walther, D., Edgington, D. R., & Koch, C. (2004). Detection and tracking of objects in underwater video. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.,

Wang, N., Wang, Y., & Er, M. J. (2020). Review on deep learning techniques for marine object recognition: Architectures and algorithms. *Control Engineering Practice*, 104458.