CONEM 2022
CONGRESSO NACIONAL DE ENGENHARIA MECÂNICA
TERESINA - PI

ABCM
XI CONGRESSO NACIONAL DE ENGENHARIA MECÂNICA
DE 07 A 11 DE AGOSTO DE 2022, TERESINA-PI, BRASIL

AUTOMATIC IDENTIFICATION OF FLANGED JOINTS IN PIPELINE SYSTEMS IN THE OIL AND GAS INDUSTRY

Nathalia S. Nascimento, nathalia.nascimento@cefet-rj.br[1]
Alessandro R. L. Zachi, alessandro.zachi@cefet-rj.br[1]
Milena F. Pinto, milena.pinto@cefet-rj.br[1]
Aurelio G. Melo, aurelio.melo@engenharia.ufjf.br[2]
Aline G. Manhães, aline.manhaes@cefet-rj.br[1]
Luciana A. Faletti, luciana.almeida@cefet-rj.br[1]

[1]Federal Center for Technological Education of Rio de Janeiro (CEFET-RJ), Rio de Janeiro, Brazil
[2]Federal University of Juiz de Fora (UFJF), Juiz de Fora, Brazil

Abstract: *In recent years, the offshore Oil & Gas industry has increased the search for automated processes to ensure efficiency and quality in manufacturing. With new robotic technologies, traditional operations were moved towards more autonomous and advanced processes, reducing operational costs and human effort, and increasing safety and reliability. It is well-known from the literature that regular underwater inspections can sometimes be difficult to accomplish, very expensive to maintain, and also be time-consuming. Moreover, such kind of inspection is generally carried out by Remotely Operated Vehicles (ROVs) equipped with end-effect tools, spotlights, sensors, and appropriate cameras. However, in such an approach, the human operator must guide the vehicle safely through the structure while performing the visual inspection to detect possible structural faults and movements. In such a context, computer vision technology provides a suitable non-contact alternative to the real-time detection problem and has emerged as a potential in robotics sensing. Therefore, this work aims to detect flanged joints from image scenes collected during the inspection services of piping systems operated both onshore and offshore. It presents and discusses the construction of a computational intelligence tool setting parallelism between two state-of-art detectors: Yolov4 and Faster-RCNN. The captured images are used to validate both and provide a conclusion of which is better suited for decision making. By characterizing the presence of a connection region in the underwater structure, the purpose of searching for flanged joints is justified as means of streamlining the work of the operator who would previously perform a manual search. Some preliminary tests have been carried out using the image dataset to feed the proposed algorithms. The results obtained showed that the use of machine learning-based techniques have achieved satisfactory metrics that show certain interesting and attractive efficiency levels that would characterize it as a potential tool for carrying out this type of task.*
Keywords: *Subsea Mechanical Systems, Offshore Industry, Inspection, Convolutional Neural Network, Computer Vision.*

1. INTRODUCTION

The offshore Oil & Gas industry is extremely important to the Brazilian economy and global energy supply, comprising exploration, transport, and distribution of natural gas and deep-sea crude oil Bogue (2019); Tan *et al.* (2020). In recent years, this industry has sought to increase both the efficiency and quality of manufacturing, using robots to inspect and maintain underwater structures through Remotely Operated Underwater Vehicles (ROVs) Yu *et al.* (2019).

Furthermore, new robotic technologies were developed to move traditional operations towards more autonomous and advanced processes, reducing operational costs and human effort and increasing safety and reliability. As an example of the use of advanced technologies in industrial processes, in Durdevic *et al.* (2019), the authors proposed using an Unmanned Aerial Vehicle (UAV) to inspect a wind farm. Another example is seen in Pinto et al. Pinto *et al.* (2020b), where UAVs are proposed to inspect and monitor slope slides and dams. Robotic technologies are already being widely used in the offshore industry, as discussed in Yu *et al.* (2019).

Part of these intelligent robotic systems uses optical sensors to guide them. A reason explained both by the cost-benefit and by the fact that they approach human behavior Pinto *et al.* (2020a). Regarding ROVs, these robots collect a variety of sensor data and can make underwater maps or investigate specific features. Many researchers continuously study underwater robot control theory, navigation methods, and detection techniques Kim and Yu (2016); Yan *et al.* (2005).

One of the primary motivations for using such alternatives is to perform exhausting tasks or tasks that demand a high level of performance. Massive undersea exploration of oil and gas has made the inspection and maintenance of subsea pipelines one of those situations. However, it is observed that in the current scenario, the visual inspection of these structures is generally done manually. A tedious and challenging task, especially in the case of large plants and due to the limitation of images to low visibility captures Antich and Ortiz (2003). Thus, regular operations for these systems can sometimes be difficult and expensive.

Therefore, computer vision technology provides a suitable non-contact approach to real-time detection problem Wang *et al.* (2019). It is important to note that these algorithms can be valuable for task autonomy, where recognizing and detecting specific patterns allows the identification of objects and textures.

It should also be remembered that initiatives that implement computer vision already have successful applications in other fields of study, including finance François-Lavet *et al.* (2018), biology Rundo *et al.* (2018); Chen *et al.* (2018), agriculture Kamilaris and Prenafeta-Boldú (2018); Patrício and Rieder (2018), accessibility Huang *et al.* (2018); Cihan Camgoz *et al.* (2018) and industrial automation Grilo and Figueiredo (2018); Li *et al.* (2018).

Convolutional Neural Networks (CNNs) became popular during the 1990s, losing the podium to systems based on support vector machines afterward. In 2014, a convolutional neural network based on the selection of regions of interest for investigation was introduced to the scientific community by Girshick Girshick *et al.* (2014). The article describes the Region-based Convolutional Neural Network (R-CNN), a network based on the selective selection algorithm of up to 2000 regions of interest, which presented good results for both detection and semantic segmentation.

Girshick Girshick *et al.* (2014) demonstrated in his work the superiority of R-CNN over the OverFeat Sermanet *et al.* (2014), a similar architectural model, in detection with the ILSVRC2013 dataset. The slowness of R-CNN is justifiable because each proposed region is treated independently by the convolutional network without sharing the feature maps obtained from these portions of the image. Based on R-CNN's continuous improvement, Fast R-CNN Girshick (2015) was created. Some of its features were higher detection quality, single-stage training, updating of all network layers at once, and no need for disk storage to cache feature maps.

Another CNN based on selecting regions of interest is You Only Look Once (YOLO). The objective behind YOLOv4 Bochkovskiy *et al.* (2020) was to have in a one-step detector a fast and accurate model, which could operate in real-time on a conventional GPU and require a single GPU during training.

YOLOv4 retains the character of its predecessors. The image is considered in its entirety in the detection of bounding boxes, that is, the detection is faced directly, from the image's pixels to the bounding boxes, so that it is only necessary to look at the image once. This causes considerable divergence from YOLO networks to systems based on R-CNN, for example, as it dispenses region classifiers such as Selective Selection Uijlings *et al.* (2013), Region Proposal Network (RPN) Girshick *et al.* (2016), Spatial Pyramid Pooling (SPP) He *et al.* (2015), Feature Pyramid Network (FPN) Kim *et al.* (2018), in addition to promoting agility in the response.

Over the years, CNNs were applied in a wide range of applications, such as in Oil & Gas industry Li *et al.* (2021); Xu *et al.* (2020). This technique has been applied to salt bodies identifications Waldeland *et al.* (2018), prediction Li *et al.* (2021), fault detection Cunha *et al.* (2020), among others.

1.1 Main Contributions

This work introduces a solution to help the user carry out an inspection of the system in the neighborhood of the connectors, essentially, in the search for end regions via visual analysis through the use of CNNs. For this challenge, it compared the performance of the networks Faster R-CNN and YOLOv4. Its contributions can be summarized as:

- Speed up the work of identifying end regions of structures characterized by the presence of a flanged connection;
- Development of a database containing images of flanged connections;
- Overcoming the challenge of analyzing images captured in low visibility environment or/and populated with visual pollution;

The database and scripts consolidated in this work were made available on the GitHub repository **Flange Annotation Tool** Nascimento (2021).

1.2 Organization

Section 2. presents the proposed methodology. It contains all the architectures conceptualization and details, as well as the information of the database. The results obtained and the necessary discussions are presented in section 3. Finally, section 4. presents the final considerations, in addition to the difficulties faced in the course of development, and ideas for subsequent works.

2. PROPOSED METHODOLOGY

2.1 Overall Idea

In this work, two models popularly recognized for showing promising results in computer vision challenges were tested. The proposed methodology addresses the use of Faster R-CNN and YOLOv4 networks to be trained with the flange database and submitted for validation. At that moment, a comparison was made between the performance of the networks, considering metrics such as accuracy and speed response, in order to consolidate the one with the best performance to integrate the flange annotation tool definitively. Figure 1 illustrates the overall idea of the proposed methodology.

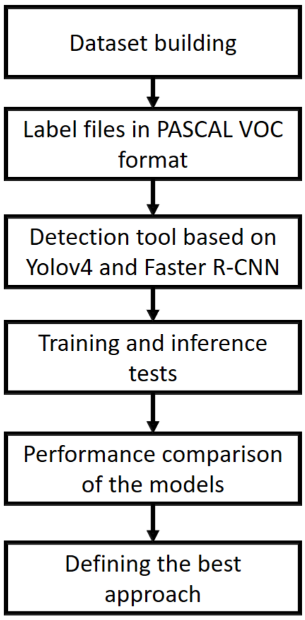

Figura 1: **Flowchart of the proposed methodology.**

2.2 CNN Models

2.2.1 Faster R-CNN

Faster R-CNN is subdivided into a deep convolutional network, the RPN, which proposes the regions of interest, and the reused detector module of the Fast R-CNN. Figure 2 illustrates how the network works.

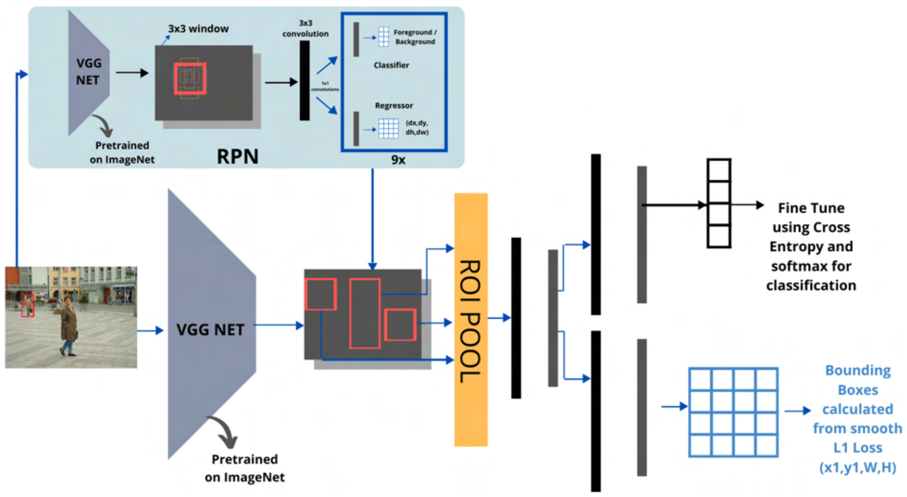

Figura 2: **Faster R-CNN Ren *et al.* (2015).**

The RPN Girshick *et al.* (2016) is a fully convolutional network (FCN) Long *et al.* (2015), which uses an input image of any dimension to generate regional proposals, each with a score that correlates it with a given object class. RPN is expected to share some of its convolutional layers with the detection network of Fast R-CNN. The RPN works in a sliding window scheme of dimensions nxn, in which a small network slides over the map of features obtained from the last convolutional layer. Each time the window passes through a portion, k rectangular regions of interest with different scales and aspect ratios are generated, using the central point of the window as an anchor. In the article, the authors mention adopting factors n=3 and k=9 and two sister layers in the next stage, one of classification and the other of regression.

These last layers filter the number of regions of interest that will be supplied to the detector. At this point, it is simple to imagine that a significant number of proposed areas will be generated while scanning the nxn window, especially for large images. It is expected that most of these proposals do not contain any trace of the object, being mere portions of the image's background. The classifier, at this point, makes a binary selection to exclude the background regions. Afterward, the rest goes through a softmax transformation that generates a score for the presence of some object class within the rectangle's boundaries. Again, a filter is performed between the regions with the highest score, configuring the RPN result.

This work used the implementation of Faster R-CNN in *python* available in the repository Detectron2 Wu *et al.* (2019).

2.2.2 Yolo

The detection architecture of YOLOv4 (Figure 3) can be distributed into three subdivisions, presented in Bochkovskiy *et al.* (2020) as a metaphor of the human body corresponding to the spine, neck, and head, respectively.

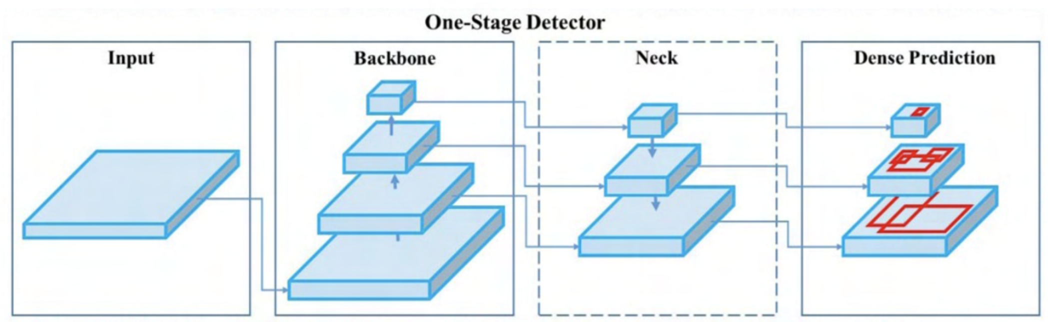

Figura 3: **YOLOv4 Bochkovskiy *et al.* (2020).**

CSPDarknet53 Wang *et al.* (2020) is the dense convolutional neural network, with 29 convolutional layers and capacity of up to 27.6 million parameters. It was defined as the backbone of YOLO because it has a large receptive field and presents good detection results. The SPP He *et al.* (2015) module comes next as a device to increase the receptive field and filter out the most relevant *features* of the CSPDarknet53 response.

Further on, Path Integral Based Convolution for Deep Graph Neural Networks (PAN) Liu *et al.* (2018) was used as a method of aggregating parameters from different levels of the initial dense network. Finally, YOLOv3 is a representative structure of the head of this model and is responsible for detection itself. It produces the coordinates of the predicted bounding boxes.

3. RESULTS AND DISCUSSION

3.1 Dataset

The calibration of neural network parameters is closely linked to the database. In this sense, consolidating compelling information that reflects the problem's reality is essential for the model's accuracy. Figure 4 presents a few images of the dataset.

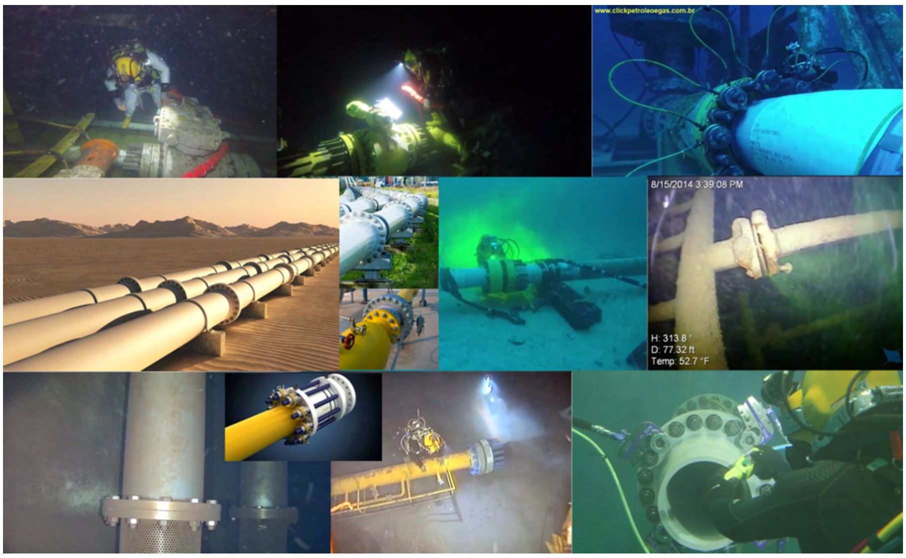

Figura 4: **A few images of the dataset.**

Some images from the public dataset Trash Can v1.0 Hong *et al.* (2020) were also used, consisting of over 7000 captures of garbage deposited on the seabed along with underwater fauna and flora. This material was selected to integrate the base of negative cases for the presence of flanged joints.

The database includes captures of flanges at different stages of life, considering examples of intact and deteriorated structures. In some cases, flanges are presented partially or entirely covered by encrustations like marine life. In addition, certain catches demonstrate severe damage to the gasket fasteners or seal failure, setting a leak itself.

The flange captures were extracted from platforms of shared data such as Google Images and YouTube videos. The video playlist used on this dataset is available on YouTube: **Subsea and Surface Pipeline Systems** Nascimento (2020).

After establishing a reasonable amount of data for the collection, it was necessary to perform manual filtering to remove very low-resolution images. In addition, some of them needed to be cut to avoid excessive graphic effects such as text, margins, vignette, etc. All procedures at this point were taken to avoid compromising the training with data with no information value for the adjustment of *kernels* on the convolutional layers.

3.2 Results

The first result concerns accuracy since it is essential to identify adequate detections in low-light and visually polluted environments. By using the dataset and the trained model, it was possible to calculate the confusion matrix. The basic terminology used in the confusion matrix is given below.

- **Positive Condition (CP):** Total number of positive real cases in the data;
- **Negative Condition (CN):** Total number of negative real cases in the data;
- **True positive (TP):** Positive condition detected as positive;
- **False Positive (FP):** Negative condition detected as positive and equivalent to false alarm;
- **False negative (FN):** Positive condition detected as negative and equivalent to error;
- **True negative (TN):** Negative condition detected as negative. This metric does not apply to object detection, as there are infinite possibilities for predictions that should not be detected in an image;

Due to the ambiguity of the concept of true negatives (TN) in detecting objects, some recurrent parameters from the confusion matrix are not applicable, such as the *True Negative Rate* (TNR), and *False Positive Rate* (FPR). In this case, the concepts of precision and *recall*, presented below, are used:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

The networks were also evaluated on a city dataset with two performance metrics: intersection over union (IoU), Equation 3, and mean average precision (mAP), Equation 4. These are concepts widely used as comparative criteria for convolutional neural networks and were used in this work to verify the accuracy of the models. The IoU varies between 0 and 1, with 1 indicating a perfect overlap between the prediction and the ground truth.

$$IoU = \frac{Intersection\ Area}{Union\ Area} \tag{3}$$

mAP is a metric that calculates precision by class considering all classes in the database and a given IoU value. Equation 4 is the mAP equation for a database of n object classes.

$$mAP = \frac{1}{n} \sum_{i=1}^{n} AP_i \tag{4}$$

where the average precision (AP_i) is in the Equation 5.

$$AP = \int_0^1 p(r)\, dr \tag{5}$$

p(r) is a curve defined by the relation *Precision* x *Recall* from outputs of the detection model. In essence, the average precision evaluates to the area beneath this curve.

Figure 5 demonstrates the comparison between the predicted bounding box, in blue, and the ground truth, in red for Yolov4. IoU was calculated and displayed in the same figure. Tables 1 and 2 compile the recorded results of the confusion matrix and metrics for the flange annotation tool with Yolov4 as detector module.

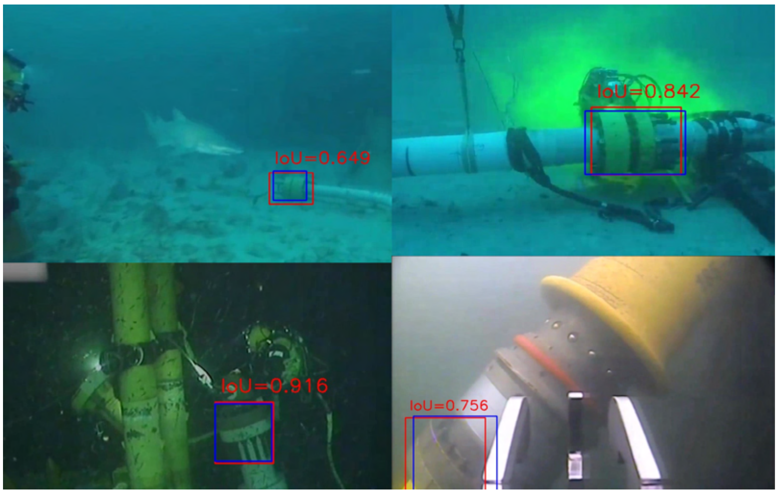

Figura 5: **IoU Results for Yolov4.**

Tabela 1: **Confusion Matrix for YOLOv4.**

	Positive Condition (CP)	Negative Condition (CN)
Expected Positive Condition	275 (TP)	4 (FN)
Expected Negative Condition	14 (FP)	N/A (TN)

Tabela 2: **Results of the Confusion Matrix for YOLOv4.**

Precision	*Recall*	IoU_{mdio}	mAP@50 (class *flange joint*)
95.16%	98.57%	86.20%	90.87%

Figura 6: **IoU Results for Faster-RCNN.**

Figure 6 gives the same results for the flange annotations with Faster R-CNN. An analysis in these results demonstrated that this model appeared less accurate than YOLOv4, especially in images populated by multiple occurrences of flanges. The training of this model showed signs of overfitting on several occasions, even with changes in network parameters such as learning rate and the number of training iterations.

Note that the YOLOv4 network stood out concerning Faster R-CNN in terms of the accuracy of the results, concluding that this is the best choice to integrate the annotation tool definitively. In terms of underwater imaging, its performance was remarkable in detecting flanges in adverse conditions, such as in low-light environments and full of sand particles, in

Tabela 3: **Confusion Matrix for Faster R-CNN.**

	Positive Condition (CP)	Negative Condition (CN)
Expected Positive Condition	243 (TP)	7 (FN)
Expected Negative Condition	248 (FP)	N/A (TN)

Tabela 4: **Confusion Matrix Results for Faster R-CNN.**

Precision	*Recall*	IoU_{mdio}	mAP@50 (class *flange joint*)
49.49%	97.20%	85.37%	89.33%

addition to cases where there is partial or total concealment of the flange by marine life. The surface predictions were also assertive, demonstrating reasonable indications for multiple occurrences, regardless of the proximity of the image capture location.

4. CONCLUSIONS AND FUTURE WORK

This work used machine learning techniques based on convolutional neural networks to create a tool capable of managing the detection of flanged connections in captures collected during the inspection of piping systems. It proposed parallelism between the famous detection models Yolov4 and Faster-CNN to verify which would be better suited to predict flanges in underwater and surface environments. Later results showed that Yolov4 yielded more accurate outputs, being chosen as the final detector module for the flange annotation software.

It is expected that this development speeds up the specialist's activity in fault analysis, efficiently signaling the capture of one of these structures in several conditions. It is important to highlight that the frames of underwater pipelines presented inferior quality. Still, the methods of deep neural networks employed proved to be efficient and robust in overcoming this issue.

This research opens up the possibility of several subsequent works. Thus, in the future, it is intended to apply deep neural networks not only in detecting flanges but also in monitoring specific structural failures such as deformations and sealing failure.

The computer vision employed can be expanded to detect, in fact, hydrocarbon leaks in the flange region and its surroundings. The automatism of this type of task is crucial in-plant inspection and maintenance applications since regular monitoring can predict and prevent damage that could compromise the entire system.

5. ACKNOWLEDGMENT

The authors would like to thank the following Brazilian Federal Agencies CEFET-RJ, FAPERJ, CAPES, and CNPq for supporting this research.

6. REFERENCES

Antich, J. and Ortiz, A., 2003. "Underwater cable tracking by visual feedback". In *Iberian Conference on Pattern Recognition and Image Analysis*. Springer, pp. 53–61.

Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M., 2020. "Yolov4: Optimal speed and accuracy of object detection".

Bogue, R., 2019. "Robots in the offshore oil and gas industries: a review of recent developments". *Industrial Robot: the international journal of robotics research and application*.

Chen, C.C., Juan, H.H., Tsai, M.Y. and Lu, H.H.S., 2018. "Unsupervised learning and pattern recognition of biological data structures with density functional theory and machine learning". *Scientific reports*, Vol. 8, No. 1, p. 557.

Cihan Camgoz, N., Hadfield, S., Koller, O., Ney, H. and Bowden, R., 2018. "Neural sign language translation". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 7784–7793.

Cunha, A., Pochet, A., Lopes, H. and Gattass, M., 2020. "Seismic fault detection in real data using transfer learning from a convolutional neural network pre-trained with synthetic seismic data". *Computers & Geosciences*, Vol. 135, p. 104344.

Durdevic, P., Ortiz-Arroyo, D., Li, S. and Yang, Z., 2019. "Vision aided navigation of a quad-rotor for autonomous wind-farm inspection". *IFAC-PapersOnLine*, Vol. 52, No. 8, pp. 61–66.

François-Lavet, V., Henderson, P., Islam, R., Bellemare, M.G., Pineau, J. *et al.*, 2018. "An introduction to deep reinforcement learning". *Foundations and Trends® in Machine Learning*, Vol. 11, No. 3-4, pp. 219–354.

Girshick, R., 2015. "Fast r-cnn". In *Proceedings of the IEEE international conference on computer vision*. pp. 1440–1448.

Girshick, R., Donahue, J., Darrell, T. and Malik, J., 2014. "Rich feature hierarchies for accurate object detection and semantic segmentation". In *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 580–587.

Girshick, R., Donahue, J., Darrell, T. and Malik, J., 2016. "Region-based convolutional networks for accurate object detection and segmentation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 38, No. 1, pp. 142–158. doi:10.1109/TPAMI.2015.2437384.

Grilo, F. and Figueiredo, J., 2018. "Computer vision in industrial automation and mobile robots". In *Introduction to Mechanical Engineering*, Springer, pp. 241–266.

He, K., Zhang, X., Ren, S. and Sun, J., 2015. "Spatial pyramid pooling in deep convolutional networks for visual recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*.

Hong, J., Fulton, M.S. and Sattar, J., 2020. "Trashcan 1.0 an instance-segmentation labeled dataset of trash observations". URL `https://doi.org/10.13020/g1gx-y834`.

Huang, J., Zhou, W., Zhang, Q., Li, H. and Li, W., 2018. "Video-based sign language recognition without temporal segmentation". In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Kamilaris, A. and Prenafeta-Boldú, F., 2018. "A review of the use of convolutional neural networks in agriculture". *The Journal of Agricultural Science*, Vol. 156, No. 3, pp. 312–322.

Kim, J. and Yu, S.C., 2016. "Convolutional neural network-based real-time rov detection using forward-looking sonar image". In *2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*. IEEE, pp. 396–400.

Kim, S.W., Kook, H.K., Sun, J.Y., Kang, M.C. and Ko, S.J., 2018. "Parallel feature pyramid network for object detection". In *Proceedings of the European Conference on Computer Vision (ECCV)*. pp. 234–250.

Li, J., Hu, D., Chen, W., Li, Y., Zhang, M. and Peng, L., 2021. "Cnn-based volume flow rate prediction of oil–gas–water three-phase intermittent flow from multiple sensors". *Sensors*, Vol. 21, No. 4, p. 1245.

Li, P., Niggemann, O. and Hammer, B., 2018. "A geometric approach to clustering based anomaly detection for industrial applications". In *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, pp. 5345–5352.

Liu, S., Qi, L., Qin, H., Shi, J. and Jia, J., 2018. "Path aggregation network for instance segmentation". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Long, J., Shelhamer, E. and Darrell, T., 2015. "Fully convolutional networks for semantic segmentation". In *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 3431–3440.

Nascimento, N., 2020. "Subsea and surface pipeline systems". URL `https://youtube.com/playlist?list=PLvofOEUnfw04W3652VkIUcnI6nQmN- eHq`.

Nascimento, N., 2021. "Flange Annotation Tool". URL `https://github.com/nathnascimentto2/FlangeAnnotationTool`.

Patrício, D.I. and Rieder, R., 2018. "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review". *Computers and Electronics in Agriculture*, Vol. 153, pp. 69–81.

Pinto, M.F., Honório, L.M., Marcato, A.L., Dantas, M.A., Melo, A.G., Capretz, M. and Urdiales, C., 2020a. "Arcog: An aerial robotics cognitive architecture". *Robotica*, pp. 1–20.

Pinto, M.F., Honorio, L.M., Melo, A. and Marcato, A.L., 2020b. "A robotic cognitive architecture for slope and dam inspections". *Sensors*, Vol. 20, No. 16, p. 4579.

Ren, S., He, K., Girshick, R. and Sun, J., 2015. "Faster r-cnn: Towards real-time object detection with region proposal networks". In *Advances in neural information processing systems*. pp. 91–99.

Rundo, F., Conoci, S., Ortis, A. and Battiato, S., 2018. "An advanced bio-inspired photoplethysmography (ppg) and ecg pattern recognition system for medical assessment". *Sensors*, Vol. 18, No. 2, p. 405.

Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R. and LeCun, Y., 2014. "Overfeat: Integrated recognition, localization and detection using convolutional networks."

Tan, J.H., Roberts, B., Sundararaju, P., Sintive, C., Facheris, L., Vanden Bosch, J., Lehning, V., Pegg, M. *et al.*, 2020. "Transforming offshore oil and gas production platforms into smart unmanned installations". In *Offshore Technology Conference Asia*. Offshore Technology Conference.

Uijlings, J., van de Sande, K., Gevers, T. and Smeulders, A., 2013. "Selective search for object recognition". *International Journal of Computer Vision*, pp. 154–171. doi:https://doi.org/10.1007/s11263-013-0620-5.

Waldeland, A.U., Jensen, A.C., Gelius, L.J. and Solberg, A.H.S., 2018. "Convolutional neural networks for automated seismic interpretation". *The Leading Edge*, Vol. 37, No. 7, pp. 529–537.

Wang, C., Wang, N., Ho, S.C., Chen, X. and Song, G., 2019. "Design of a new vision-based method for the bolts looseness detection in flange connections". *IEEE Transactions on Industrial Electronics*, Vol. 67, No. 2, pp. 1366–1375.

Wang, C.Y., Liao, H.Y.M., Wu, Y.H., Chen, P.Y., Hsieh, J.W. and Yeh, I.H., 2020. "Cspnet: A new backbone that can enhance learning capability of cnn". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPR Workshop)*.

Wu, Y., Kirillov, A., Massa, F., Lo, W.Y. and Girshick, R., 2019. "Detectron2: A pytorch-based modular object detection library". URL `https://ai.facebook.com/blog/-detectron2-a-pytorch-based-modular-object-detection-library-`

Xu, Z., Wu, F., Yang, X. and Li, Y., 2020. "Measurement of gas-oil two-phase flow patterns by using cnn algorithm based on dual ect sensors with venturi tube". *Sensors*, Vol. 20, No. 4, p. 1200.

Yan, Y., Ma, P., Wang, D. and Gao, X.g., 2005. "Development of deep sea rov and its working system". *Robot*, Vol. 27, No. 1, pp. 82–89.

Yu, L., Yang, E., Ren, P., Luo, C., Dobie, G., Gu, D. and Yan, X., 2019. "Inspection robots in oil and gas industry: a review of current solutions and future trends". In *2019 25th International Conference on Automation and Computing (ICAC)*. IEEE, pp. 1–6.

7. RESPONSIBILITIES

The authors are the only responsible for this work's content.