

# Identification and Instance Segmentation of Oil Spills Using Deep Neural Networks

Zahra Ghorbani<sup>1</sup>, Amir H. Behzadan<sup>1</sup>

<sup>1</sup>Texas A&M University  
3137 TAMU, College Station, TX, USA  
zahraghorabani@tamu.edu; abehzadan@tamu.edu

**Abstract** - Oil spills impact natural and built environments, people and communities, food chain, and wildlife, and the road to full recovery is often long and costly to businesses, contractors, communities, and local governments. Existing oil spill detection methods including in-situ measurements and remote sensing primarily depend on involving skilled personnel in data collection, processing, and analysis, which could be expensive, slow, and subjective (influenced by prior experience, and judgment of problem parameters or solution space). In addition, oil pipelines and platforms can be located in remote and harsh areas, making it difficult and even hazardous for engineers to conduct timely inspections. Applying artificial intelligence (AI) can streamline this process and create more objective measures of oil spill and leakage detection. In this research, deep learning models, namely VGG-16 and mask R-CNN (mask region-based convolutional neural network) are employed to identify and locate oil spills. These models represent state-of-the-art object recognition algorithms in computer vision. Red-green-blue (RGB) training images are collected using semi-supervised learning (i.e., keyword search) from the web. This initial visual dataset consists of a diverse set of photos taken by unmanned aerial vehicles (UAVs) or first-person cameras from previous oil spill accidents. The methodology consists of model training and validation, image classification, object detection, and semantic segmentation. The VGG16 model is used for image classification (to predict the existence of oil spill in an image) and yields an accuracy of ~93%. The mask R-CNN model is used for instance segmentation (to detect oil spills and marking their boundaries at pixel-level) and yields average precision and recall of 61% and 70%, respectively. Results can create opportunities for advancing the current practice of integrating AI and data analytics into downstream and upstream operations in the oil and gas industry, as well as enabling non-intrusive techniques for detection of environmental pollutants.

**Keywords:** Deep neural networks; oil spills; classification; instance segmentation; drones.

## 1. Introduction

Recent advancements in data sensing and artificial intelligence (AI) have created new opportunities to tackle challenging problems in environmental monitoring such as air, solid waste, and wastewater pollution [1]. In the area of air pollution, past work has developed artificial neural networks (ANNs) to predict daily or hourly values of critical pollutants such as nitrogen dioxide (NO<sub>2</sub>), carbon monoxide (CO), and Ozone (O<sub>3</sub>) in the atmosphere [2]. In solid waste management, researchers have investigated waste generation and waste composition generation, and for instance, proposed AI algorithms based on support vector machine (SVM), k-nearest neighbour (KNN), ANN, and adaptive neuro-fuzzy inference system (ANFIS) to forecast waste generation in Queensland, Australia [3]. Other work includes the use of genetic algorithms to optimize the type of vehicle and length of waste collection route, and a fuzzy logic to represent customer satisfaction [4]. ANNs and fuzzy logic models have been also used to predict leachate penetration into groundwater and assess its environmental impacts [5]. In wastewater and water pollution control, researchers have used AI models to estimate water treatment processes and control pollutant flows. Multivariate adaptive regression splines (MARS), M5 model tree (M5Tree), and least square support vector machine (LSSVM) were used to model river water pollution and predict monthly chemical oxygen demand (COD) [6]. In another study, back propagation neural network models (BPNNs) and radial basis function neural network (RBFNN) were used to forecast water quality index (WQI) based on variables such as pH, dissolved oxygen (DO), total suspended solids (TSS), biological oxygen demand (BOD), and COD [7].

Despite previous work, there is still a dearth of research in using AI for monitoring oil spills, as a major source of environmental pollution. Oil spills can negatively affect plant growth [8] and soil nutrient levels [9], and lead to soil infertility [10] and contamination [11]. With increasing global oil consumption [12, 13] especially in the developing world, oil pollution is also expected to be on the rise [14]. Traditional in-situ methods of detecting oil pollution such as pressure-point-analysis

(PPA) [15], acoustic methods [16], vapor sampling [17], and negative pressure wave (NPW) [18] involve large-array sensor installation and deployment of skilled personnel, making them costly, resource-intensive, and subjective. More recent techniques that take advantage of satellite remote sensing focus mainly on major oil spills, and rely on data that cannot be obtained frequently due to user restrictions. Smaller scale oil spills such as those resulting from local offshore drilling operations or leaking tankers or oil rigs [19] are by far more frequent and underreported [20]. To monitor this kind of spills, frequent observations of the target area from relatively lower altitudes are needed, which cannot be achieved by satellite flyovers.

This paper proposes the use of low-cost and easy-to-access unmanned aerial vehicle (UAV) (a.k.a., drone) imagery and AI models to detect oil spills in smaller-scale offshore pollutions and prevent further damage to the environment. Recently, drones have been increasingly used in the oil and gas (O&G) industry for various applications [21] such as monitoring oil and gas air emission [22, 23], pipeline inspection [24, 25, 26], safety management [27], and logistics [28]. Given the ubiquity of drones and the decreasing technology cost, if drones were to be deployed for oil spill detection, the process could be conveniently crowdsourced by splitting the data collection and monitoring tasks among various stakeholders and offshore contractors who operate in a large geographical area. The following sections describe the scientific methodology to train convolutional neural network (CNN) models on what an oil spill looks like as described by parameters such as color palette, pixel density, appearance, geometry, and progression pattern. These fully trained models can then predict the presence and location of oil spills in new footage.

## 2. Convolutional Neural Networks (CNNs) for Visual Recognition

A CNN architecture contains layers each encapsulating a feature map [29]. The input layer feature map is a 3D matrix of pixel intensities for various color channels such as RGB [30]. Several types of transformation can be applied to feature maps, including pooling and filtering [31]. Pooling functions (e.g., average-pooling and max-pooling), condense multiple cells (i.e., field) of a feature map into one, and gradually generate more vigorous feature descriptions [32]. Filtering (i.e., convolution) function, on the other hand, convolutes a filter matrix (learned weights) with the values of a receptive field of neurons and takes a nonlinear function to obtain final responses [33]. A feature hierarchy is created by alternating between pooling and convolution operations. By adding various fully connected (FC) layers with specific activation functions, this hierarchy is fine-tuned and adapted for various tasks [34].

Object recognition in an image can be done in the form of image classification (i.e., determining the mere presence of an object) or object detection (i.e., determining the location and marking the boundaries of an object) [35]. In order to train a CNN model for object recognition, large amount of training data and optimal model parameters (a.k.a., hyperparameters) are desired [36]. However, the scarcity of annotated data for oil spill and marine object detection poses a challenge to training high-performing CNN models. A potential solution to this problem is to pre-train the model on data from another domain, and then retrain the network on the primary domain data. In deep learning (DL), this process is termed transfer learning [37]. As shown in Table 1, in this research, two pre-trained models are used, i.e., VGG16 and mask R-CNN (mask region-based CNN), which are trained on ImageNet and COCO datasets, respectively. VGG16 is used for image classification (i.e., predicting the existence of oil spill) while mask R-CNN is used for object detection (i.e., detecting the location of oil spill) and semantic segmentation (i.e., marking the oil spills at pixel-level).

Table 1: Overview of deep learning models used in this study.

Model	Network Structure	Pre-trained Dataset	Application
VGG16 [30]	16 Convolutional Layers	ImageNet [38]	Image classification
Mask R-CNN [39]	Resnet-50 [40]	COCO [41]	Instance segmentation

### 2.1. Oil Spill Image Classification

For image classification, a well-established DL network, called VGG16 [30], is adopted and fine-tuned (by transfer learning). This network takes an RGB image as input, generates intermediate features through convolution and max-

pooling operations, passes the features to the FC layers, and outputs the probabilities of the image belonging to either of the two classes of “spill” or “no spill.” As shown in Figure 1, the model is comprised of one input, 18 intermediate, two FC, and one output layers. The intermediate layers consist of a series of convolutional and max-pooling operations with 14,714,688 pre-trained weights. The last of these layers is connected to a flattened layer including 8,192 nodes, which is fully connected to the next layer of 256 nodes. In this layer, a dropout operation is performed with 50% probability, i.e., during each training iteration, 50% of the nodes are randomly excluded from weight updating. Together, the two FC layers contain 2,097,408 (i.e., 8,192 multiplied by 256) weights. The training process consists of two steps. First, the model learns how to classify new images using filters from the pre-trained dataset (i.e., ImageNet). Here, only the weights of the FC layers are updated, and the weights of the intermediate layers remain unchanged (i.e., frozen). Next, fine-tuning is performed during which all previously frozen layers adapt to the domain data [42]. In this step, the model is fed with training data, and the weights of the last three convolutional layers and two FC layers are updated again. Hyperparameters (number of epochs: 30, batch size: 20, and learning rate:  $10^{-4}$ ) are selected empirically, and the trained model is validated using the validation set.

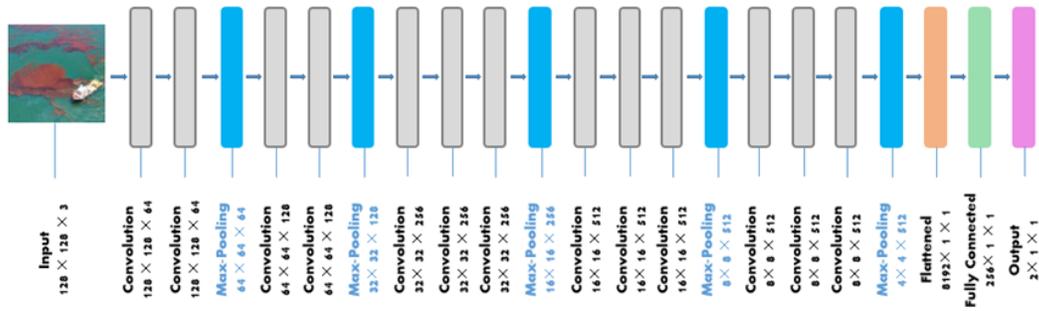


Fig. 1: VGG16 architecture for image classification.

## 2.2. Oil Spill Instance Segmentation

Instance segmentation comprises of detecting and segmenting objects in an image (i.e., semantic segmentation). The mask R-CNN algorithm [39] is used in this research for pixel-level instance segmentation. As illustrated in Figure 3, mask R-CNN uses a layer, called region of interest (RoI) align, that implements bilinear interpolation [43] to compute the precise values of the input features at four sampling locations in each RoI bin. This layer preserves the pixel-level spatial correspondence, thus enhancing mask accuracy [34].

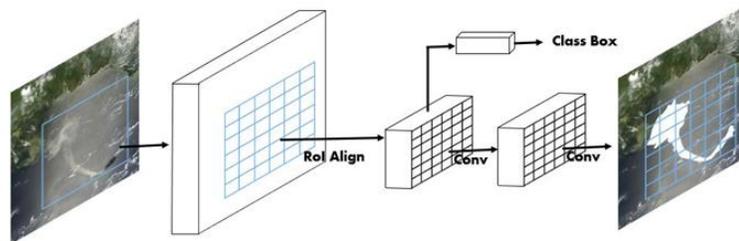


Fig. 2: Mask R-CNN architecture for instance segmentation.

Eq. (1) calculates the total loss of the mask R-CNN model as the sum of three individual losses, namely the classification loss ( $L_c$ ), the bounding box regression loss ( $L_b$ ), and the mask loss ( $L_m$ ) [34]. Eq. (2) denotes the classification loss, in which,  $p_i$  is the predicted probability of anchor  $i$  being an object, and  $p_i^*$  is the ground truth anchor label. The bounding box regression loss function is shown in Eq. (3), in which  $t_i$  is the predicted parametrized coordinate, and  $\lambda$  is a balancing value used so that  $L_c$  and  $L_b$  are assigned approximately the same weight. The mask loss is defined as the average binary cross-

entropy loss, as shown in Eq. (4), in which,  $y_{ij}$  represents a cell  $(i, j)$  that belongs to the ground truth mask for the region of size  $m \times m$ , and  $\hat{y}_{ij}^k$  is the predicted value of the same cell in the mask learned for the ground truth class  $k$ .

$$L = L_c + L_b + L_m \quad (1)$$

$$L_c(p_i, p_i^*) = -p_i^* \log p_i - (1 - p_i^*) \log(1 - p_i) \quad (2)$$

$$L_b = \sum_{i \in \{x, y, \omega, h\}} (t_i - d_i(p))^2 + \lambda \|w\|^2 \quad (3)$$

$$L_m = -\frac{1}{m^2} \sum_{1 \leq i, j \leq m} [y_{ij} \log \hat{y}_{ij}^k + (1 - y_{ij}) \log(1 - \hat{y}_{ij}^k)] \quad (4)$$

### 3. Data Preparation

To create an image dataset in this research, relevant images are retrieved from Google using web mining using keywords such as “oil spill”, “ocean aerial imagery”, and “sea aerial view”. Next, a web-based labeling toolbox is used for image labeling and annotation. The generated in-house dataset, named Nafta 2019, has 1,292 images that are used to train, validate, and test the CNN models. One half of these images contains oil spills and the other half does not contain oil spills. To test model robustness, different viewpoints are considered; 16% of the images are taken at high altitude (i.e., satellite), 14% are from first-person view, and the remaining 70% are captured at low altitude (i.e., drone). Each image is segmented by marking the boundaries of oil spills. In total, 1,431 instances of oil spills are marked. Segmentation is done at pixel-level to extract information such as shape, size, and the local position.

## 4. Results and Discussion

### 4.1. VGG16 Model for Image Classification

The performance of the VGG16 model in classifying images based on the presence of oil spills is calculated using precision, accuracy, and recall metrics [44]. Accuracy is defined as the ratio of number of images that are correctly classified to the total number of images, and is calculated by Eq. (5), in which true positive (TP) and true negative (TN) are to the number of correctly predicted values. TP indicates positive predictions that are correct, while TN indicates negative predictions that are correct. false positive (FP) and false negative (FN), on the other hand, are to the number of incorrectly predicted values. FP indicates positive predictions that are incorrect, while FN indicates negative predictions that are incorrect. The denominator of Eq. (5), therefore, is the total number of all predictions made by the model. Knowing TP, TN, FP, and FN, precision and recall values can be calculated by Eqs. (6) and (7).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

Overall, the model achieves an accuracy of %92.77 and a test loss of %23.71. The outcome of classification in the form of a confusion matrix is shown in Figure 3, which indicates that in 83 test images, oil spill is classified correctly, while in only 1 test image it is classified incorrectly (i.e., the model predicts an oil spill when there is no oil spill). Similarly, in 90 of the test images, the absence of oil spill is classified correctly, while in 10 test images this is classified incorrectly (i.e. the model predicts that the image does not contain an oil spill when there is an oil spill).

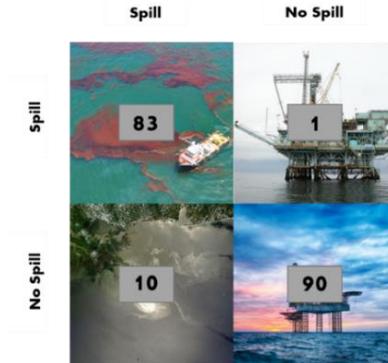


Fig. 3: Confusion matrix for oil spill classification.

#### 4.2. Mask R-CNN Model for Instance Segmentation

In the mask R-CNN model, features are extracted from four feature maps of the Resnet-50 architecture. Each layer reduces the size of the feature map by half and doubles the number of feature maps. To generate final feature maps (P2, P3, P4, and P5), the top-bottom pathway approach is used, which starts from the smallest feature map and performs upscaling to get to the largest one. Then, by performing max-pooling on P5, the last feature map (P6) is generated. As shown in Figure 4, these feature maps are then passed through a  $3 \times 3$  convolution layer and the output is passed through two branches, for bounding box and object class prediction. To evaluate the performance of this model, each test image is assessed at pixel-level using precision and recall, as well as intersection over union (IoU), F-1 score, and dice coefficient (DC) [45]. Results show that the average precision, recall, IoU, F-1 score, and DC for all the images are 61%, 70%, 48%, 60%, and 60%, respectively. Table 2 summarizes measures of central tendency and variability of all evaluation metrics calculated for the test images at pixel-level.

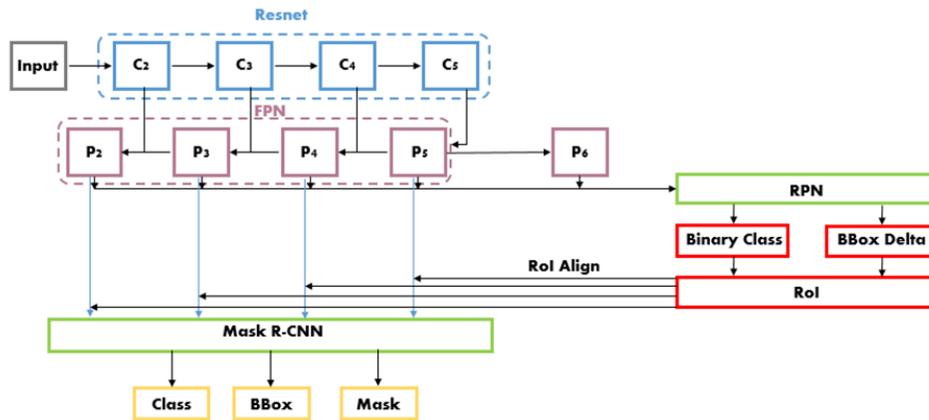


Fig. 4: Process of bounding box and object class prediction by mask R-CNN model.

Table 2: Measures of central tendency and variability for the testing set images at pixel level.

Metric	Max	Min	1 <sup>st</sup> Quartile	Median	3 <sup>rd</sup> Quartile	Mean	SD
Precision	0.99	0.00	0.35	0.68	0.86	0.61	0.30
Recall	0.97	0.00	0.54	0.79	0.89	0.70	0.26
IoU	0.92	0.00	0.22	0.48	0.73	0.48	0.28
F1	0.96	0.00	0.37	0.65	0.84	0.60	0.27
DC	0.96	0.00	0.37	0.65	0.84	0.60	0.27

Figure 5 illustrates examples of ground truth and predictions made by the mask R-CNN model. Since the model predicts each mask independent from predictions in the neighbouring areas, for a given test image, generated masks may be fragmented (i.e., separated from each other). As such, post-processing is performed to connect individual masks prior to measuring model performance. Figure 6 shows examples of how detected masks are merged to create a single mask.



Fig. 5: Sample images from the testing set (left) and corresponding predictions by mask R-CNN (right).

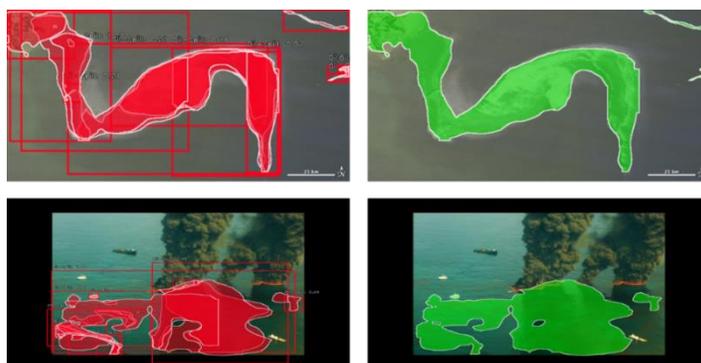


Fig. 6: Examples of merging separate masks (left) into one single mask (right).

## 5. Summary and Conclusion

Timely and accurate detection of oil spills can help authorities take proper actions immediately and effectively, thus minimizing cascading damages. Most existing oil spill detection techniques have elements of in-situ measurements or remote sensing, which require the involvement of skilled personnel in data collection and analysis, rendering them expensive, slow, and subjective. In some cases, the location of oil spill is an impediment due to being remote or hazardous, thus preventing physical access for proper inspection. In this paper, two DL models (VGG16 and mask R-CNN) were used to identify and localize oil spills in aerial imagery. The VGG16 model was used for classification and yielded an accuracy of ~93%. The mask R-CNN model was used for instance segmentation and yielded average precision and recall of 61% and 70%, respectively. An in-house image dataset was created through web mining of publicly available photos of previous oil spill incidents, and used to train, validate, and test the developed models. This dataset contained 1,292 images taken from high altitude (i.e., satellite), first-person view, and low-altitude (i.e., drone), and comprised an equal number of images with and without instances of oil spill. Model development was performed on Texas A&M University High Performance Research Computing (HPRC) clusters. To minimize the implementation cost and increase the likelihood of technology adoption by the O&G industry, this research used red-green-blue (RGB) imagery as input since consumer-grade drones are already equipped with RGB camera. With some modifications, the same method can be applied to other input types (e.g., infrared or thermal imagery).

## References

- [1] K. Yetilmezsoy, B. Ozkaya, and M. Cakmakci, "Artificial intelligence-based prediction models for environmental engineering," *Neural Network World*, vol. 21, no. 3, pp. 193-218, 2011.
- [2] E. Agirre-Basurko, G. Ibarra-Berastegi, and I. Madariaga, "Regression and multilayer perceptron-based models to forecast hourly O<sub>3</sub> and NO<sub>2</sub> levels in the Bilbao area," *Environ. Model. Softw.*, vol. 21, pp. 403-446, 2006.
- [3] M. Abbasi and A. E. Hanandeh, "Forecasting municipal solid waste generation using artificial intelligence modelling approaches," *J. Waste Manag.*, vol. 56, pp. 13-22, 2016.
- [4] A. Krol, P. Nowakowski, and B. Mrowczynska, "How to improve WEEE management? Novel approach in mobile collection with application of artificial intelligence," *J. Waste Manag.*, vol. 50, pp. 222-233, 2016.
- [5] M. Bagheri, A. Bazvand, and M. Ehteshami, "Application of artificial intelligence for the management of landfill leachate penetration into groundwater, and assessment of its environmental impacts," *J. Clean. Prod.*, vol. 149, pp. 784-796, 2017.
- [6] O. Kisi, and K. S. Parmar, "Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution," *J. Hydrol.*, vol. 534, pp. 104-112, 2016.
- [7] M. Hameed, S. S. Sharqi, Z. M. Yaseen, H. A. Afan, A. Hussain, and A. Elshafie, "Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia," *Neural Comput. & Applic.*, vol. 28, pp. S893-S905, 2017.
- [8] G. Malallah, M. Afzal, M. Kurian, S. Gulshan, and M.S.I. Dhmi, "Impact of oil pollution on some desert plants," *Environ. Int.*, vol. 24, no. 8., pp. 919-924, 1998.
- [9] T. A. Abii and P. C. Nwosu, "The effect of oil spillage on the soil of Eleme in Rivers State of the Niger-Delta area of Nigeria," *Res. J. Environ. Sci.*, vol. 3, no. 3, pp. 316-320, 2009.
- [10] O. E. Essien and I. A. John, "Impact of crude-oil spillage pollution and chemical remediation on agricultural soil properties and crop growth," *J. Appl. SCI. Environ. Manage.*, vol. 14, no. 4, pp. 147-154, 2010.
- [11] G. Nicolotti and S. Egli, "Soil contamination by crude oil: impact on the mycorrhizosphere and on the revegetation potential of forest trees," *Environ. Pollut.*, vol. 99, no. 1, pp. 37-43, 1998.
- [12] D. Gately, N. Al-Yousef, and H. M. H. Al-Sheikh, "The rapid growth of domestic oil consumption in Saudi Arabia and the opportunity cost of oil exports foregone," *Energ. Policy*, vol. 47, no. 1, pp. 57-68, 2012.
- [13] G. Zou and K. W. Chau, "Short- and long-run effects between oil consumption and economic growth in China," *Energ. Policy*, vol. 34, no. 18, pp. 3644-3655, 2006.
- [14] H. Bloch, S. Rafiq, and R. Salim, "Economic growth with coal, oil and renewable energy consumption in China: prospects for fuel substitution," *Econ. Model.*, vol. 44, pp. 104-115, 2015.
- [15] G. Geiger, *Principles of leak detection*, 1<sup>st</sup> ed. Breda, the Netherlands, Krohne Oil & Gas, 2005.
- [16] L. Meng, L. Yuxing, W. Wuchang, and F. Juntao, "Experimental study on leak detection and location for gas pipeline based on acoustic method," *J. Loss Prevent. Proc.*, vol. 25, no. 1, pp. 90-102, 2012.
- [17] P. S. Murvay and L. Silea, "A survey on gas leak detection and localization techniques," *J. Loss Prevent. Proc.*, vol. 25, no. 6, pp. 966-973, 2012.
- [18] G. Chuanhu, W. Guizeng, and Y. Hao, "Analysis of the smallest detectable leakage flow rate of negative pressure wave-based leak detection systems for liquid pipelines," *Comput. Chem. Eng.*, vol. 32, no. 8, pp. 1669-1680, 2008.
- [19] A. B. Alexopoulos and G. Dounias, "An assessment of vessel-source oil pollution incidents in the Mediterranean Sea using inductive machine learning methodologies," University of the Aegean, Department of Shipping Trade and Transport, 2003.
- [20] M. Fingas, *Oil spill science and technology: prevention, response, and clean-up*. Burlington, MA, Elsevier, 2011.
- [21] P. Sow, "The use of drones in the oil and gas industry: A 4.0 contract," *PM World J.*, vol. 7, no. 12, pp. 1-14, 2018.
- [22] L. Satterlee, "Climate drones: A new tool for oil and gas air emission monitoring," *Envtl. L. Rep. News & Analysis*, vol. 46, pp. 11069-11083, 2016.
- [23] I. Uyanik and A. W. Halliburton, "Next generation gas emission monitoring system," *SPE Middle East Oil and Gas Show and Conference*, Society of Petroleum Engineers, 2019.

- [24] A. Shukla and H. Karki “Application of robotics in onshore oil and gas industry – a review part I,” *Robot. Auton. Syst.*, vol. 75, pp. 490-507, 2016.
- [25] A. Shukla, H. Xiaoqian, and H. Karki, “Autonomous tracking and navigation controller for an unmanned aerial vehicle based on visual data for inspection of oil and gas pipeline,” *16<sup>th</sup> International Conference on Control, Automation and Systems (ICCAS)*, Gyeongju, Korea, 2016, pp. 194-200.
- [26] L. Yu, E. Yang, P. Ren, C. Luo, G. Dobie, D. Gu, and X. Yan, “Inspection robots in oil and gas industry: a review of current solutions and future trends,” in *Proceedings of the IEEE 25<sup>th</sup> International Conference on Automation & Computing (ICAC)*, Lancaster, UK, 2019, pp. 1-6.
- [27] J. Cho, G. Lim, T. Biobaku, S. Kim, and H. Parsaei, “Safety and security management with unmanned aerial vehicle (UAV) in oil and gas industry,” *Procedia Manufacturing*, vol. 3, pp. 1343-1349, 2015.
- [28] J. Baghirova, “The use of drones in oil and gas logistics,” M.S. thesis, College of Logistics, Molde Univ., Molde, Norway.
- [29] Y. LeCun, Y. Bengio, and G. E. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *NIPS*, pp.1097-1105, 2012.
- [31] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, “Learning and transferring mid-level image representations using convolutional neural networks,” *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, 2014, pp. 1717-1724.
- [32] K. Kavukcuoglu, R. Fergus, and Y. LeCun, “Learning invariant features through topographic filter maps,” *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, 2009, pp. 1605-1612.
- [33] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, and E. Shelhamer, “cuDNN: efficient primitives for deep learning,” *arXiv:1410.0759 [cs.NE]*, 2014.
- [34] Z.Q. Zhao, P. Zheng, S. Xu, and X. Wu, “Object detection with deep learning: a review,” in *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1-21, 2019.
- [35] J. R. Uijlings, K. E. Van De Sande, T. Gevers, A. W. Smeulders, “Selective search for object recognition,” *Int. J. Comput. Vision*, vol. 104, no. 2, pp. 154-171, 2013.
- [36] Z. Kolar, H. Chen, and X. Luo, “Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images,” *Automat. Constr.*, vol. 89, pp. 58-70, 2018.
- [37] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345-1359, 2009.
- [38] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, “ImageNet: a large-scale hierarchical image database,” *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, 2009, pp. 248-255.
- [39] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask R-CNN,” *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, 2018, pp. 2961-2969.
- [40] T. Akiba, S. Suzuki, and K. Fukuda, “Extremely large minibatch SGD: training resnet-50 on Imagenet in 15 minutes,” *arXiv preprint arXiv:1711.04325*, 2017.
- [41] T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and L. Zitnick, “Microsoft COCO: common objects in context,” *European Conference on Computer Vision*, Zurich, Switzerland, 2014, pp. 740-755.
- [42] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv:1409.1556*, 2014.
- [43] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, “Spatial transformer networks,” *NIPS*, pp.2017-2025, 2015.
- [44] M. H. Dunham, *Data mining: introductory and advanced concepts*. Upper Saddle River, NJ, Pearson Education, 2006.
- [45] Z. Ghorbani, “Oil spill detection using deep neural networks,” M.S. thesis, Dept. Cons. Sci., Texas A&M Univ., College Station, TX.