IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 18, 2022, accepted February 1, 2022, date of publication February 10, 2022, date of current version March 2, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3150969

Improved Sea-Ice Identification Using Semantic Segmentation With Raindrop Removal

NAHED M. ALSHARAY[1], (Student Member, IEEE), YUANZHU CHEN[2], (Member, IEEE), OCTAVIA A. DOBRE[3], (Fellow, IEEE), AND OSCAR DE SILVA[3], (Member, IEEE)

[1]Computer Science Department, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada
[2]School of Computing, Queen's University, Kingston, ON K7L 3N6, Canada
[3]Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada

Corresponding author: Nahed M. Alsharay (nmmqalsharay@mun.ca)

This work was supported in part by the Memorial University of Newfoundland's Research Chair, and in part by the Mitacs Lab2Market Program.

ABSTRACT Sea-ice identification is an essential process for safety critical navigation support of surface vessels in polar waters. Semantic segmentation has drawn much attention as an enabling technique for fast detection of objects in a scene including sea-ice conditions. Identifying sea-ice is a challenging problem, especially in the presence of raindrops. The raindrop alters the boundaries of the objects in the scene, and thus, degrades the identification performance. In this work, a raindrop removing framework is developed to enhance the classification performance. Three deep-learning semantic segmentation networks are trained to classify the scene of sea-ice images into ice, water, ship, and sky. The deep-learning networks are VGG-16, fully convolutional network, and pyramid scene parsing network. Transfer learning along with data augmentation operations have been implemented to improve the training process. Results illustrate that data augmentation operations enhance the performance of the three models. Moreover, the raindrop removing framework improves the models' performance, e.g. the average intersection over union of the VGG-16 model is improved from 85.91% to 91.70%.

INDEX TERMS Convolutional neural networks, raindrop removing, sea-ice, semantic segmentation.

I. INTRODUCTION

Recently, navigation through sea-ice has attracted more research efforts as the need for operating surface vessels and offshore operations in polar waters have increased. Unlike in land navigation where road networks are deterministic, navigable area in polar waters is continuous and the ice thickness, shape, and concentration are changing over time.

Navigation under sea-ice conditions requires highly trained and experienced ice navigators to find a safe route through rapidly changing environment and to avoid the hazardous ice conditions following established polar operation risk assessment standards [1]. Satellite-based synthetic aperture radar systems provide large-scale and high-spatial resolution of the sea-ice floes, which is useful for operational planning through polar waters [2]. However, relying on aerial air-borne or space-borne imagery requires a reliable and sustainable communication channel between the aerial scanner and the earth. Moreover, the quality and temporal frequency of aerially captured images degrades significantly in bad weather conditions such as clouds. One advantageous alternative is to obtain the navigation information in situ, without the need for external platforms. Onboard systems are considered to gather and process information in real-time, which is important for time-critical decisions in maneuvering, control, and monitoring systems [3], [4]. An example of onboard system is the marine radar, which is an in-situ sensing device that operates in the X-band. It has the capability of measuring the backscatter from the polar water surface in space and time, independent of lighting conditions and under different weather conditions. However, marine radars have limited spatial resolution and it is quite difficult to establish form and type of ice based on the radar information alone. Currently, for in-situ navigational risk assessment, the detailed information about ice is captured by ice navigators using primarily visual observations [5].

Computer-aided scene analysis techniques such as automated image processing and segmentation have paved the

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Gaggero.

way for autonomous navigation systems, which reduces costs, processing time, and human bias in the navigation process. For this application at least, a trained human is considered an expert over any onboard system due to the experience required. Moreover, the resurgence of deep neural networks (DNNs) has dramatically improved the performance of many computer-aided scene analysis techniques such as image classification [6], object detection and localization [7], and semantic segmentation [8].

In contrast to image prediction, semantic segmentation generates a fine-grained delineation of objects that embeds their spatial information, which makes it a key enabling technique to address diverse remote sensing problems [9]. In context of sea-ice monitoring, DNNs have been utilized in image-based ice detection techniques such as sea-ice classification in synthetic aperture radar (SAR) images [10], [11], ice objects classification in optical close-range images [12], lake ice monitoring algorithm [13], and river ice classification in images collected by an unmanned aerial vehicle (UAV) [14].

A. RELATED WORK

Various classification and identification problems are of interest in the ocean environments, such as underwater target classification [15], maritime targets classification on high-resolution image [16], classification of coral reef images [17], and sea-ice condition identification and assessment [4]. Research efforts have been devoted to develop techniques for sea-ice classification [2], [3], [18], [19]. In [2], the authors proposed a remote sensing algorithm that utilizes radar images to estimate the ice-drift velocity vector in a region around a moving ship. Two Kalman filters were integrated with radar image processing to estimate the local drift vector from the vessel motion. In [3], the authors developed an algorithm to quantify ice concentration and to estimate ice thickness. The global Otsu method and the K-means method were utilized to implement the ice concentration analysis. In [18], a sea-ice monitoring model using SAR was designed. This model performs SAR segmentation and classification using the Markov random-field theory such that a region-growing technique keeps refining the segmentation iteratively. In [19], an ice navigation system was presented. This system implements ship-based ice awareness by utilizing a combination of radar, lidar, and video processing for ice detection and classification.

NNs have attracted considerable research attention recently as a promising tool to avail automated sea-ice monitoring solutions [10]–[14], [20]. In [10], a NN algorithm was designed to classify sea-ice in SAR images of central arctic. The algorithm classified the images into open water, and deformed ice based on third and fourth central statistical moments, inertia, cluster prominence, energy, and homogeneity of image brightness. In [12], ice objects in optical close-range images were classified into several categories using convolutional neural networks (CNNs). In [11], an algorithm was proposed to classify the scene of SAR images into several classes using CNNs. In [14], a data set was collected by a UAV of river ice and used to train a semantic segmentation deep network, which classifies the scene into ice, water, and other class. In [13], a lake ice monitoring algorithm based on semantic segmentation was proposed. The authors utilized video streams acquired by a webcam to generate the data set with nomenclature classes of water, ice, and clutter. In [20], two data sets were used to train DNNs. The scene in the first data set captures four classes, namely ice, vessel, ocean, and sky; while the scene in the second data set captures more ice classes.

In this paper, we focus on studying the effect of rain droplets on the sea-ice identification and develop a framework to remedy this effect. Moreover, we study the effect of changes in the camera location and mounting angle on the sea-ice identification performance by data augmentation operations such as cropping and rotating the training set. DNNs are utilized to classify images of sea-ice scenes captured onboard of a ship using semantic segmentation. Each sea-ice scene image is classified into four classes, namely ice, water, ship, and sky. The initial dataset of this work consists of 428 and 23 training and evaluation images, respectively; this is constructed from images taken by the Nathaniel B. Palmer icebreaker during its expedition through the Ross Sea.[1] Data augmentation operations are applied to enhance the size of the dataset, such that the total data set of this work consists of 10,700 and 575 training and evaluation images, respectively. Moreover, transfer learning is implemented using the Cityscapes and CoCo datasets, which are widely used in academia in the context of training semantic segmentation models. Three DNN models are trained, namely VGG-16, fully convolutional network (FCN-8), and pyramid scene parsing network (PSPNet-50). The classification performance of the models is measured using the precision and intersection over union (IoU) of the predicted and the ground truth images.

The rest of this paper is organized as follows. Section II introduces the adopted DNN models. The data augmentation operations are discussed in Section III. Section IV introduces the raindrop removing framework. Results are discussed in Section V and Section VI concludes the paper.

II. DNN MODELS

In this work, we consider three models, each representing different DNN architecture, namely CNN, fully convolutional network, and encoder-decoder network architecture. This section introduces the adopted DNN models.

A. VGG-16

VGG-16 is a CNN model proposed by the University of Oxford [6]. The VGG-16 model includes convolution layers and max pool layers consistently throughout its architecture. In the end it has 2 fully connected layers (FCS) followed by a

[1]A video footage of the dataset can be seen in the following link: https://youtu.be/BNZu1uxNvlo.

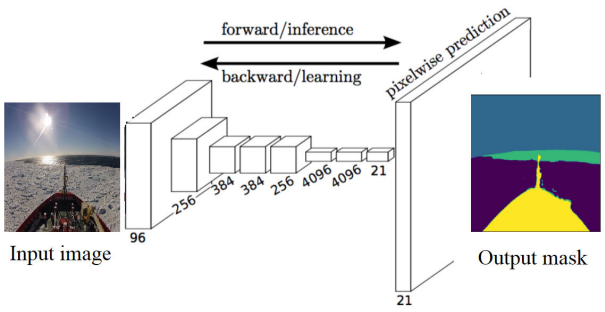

FIGURE 1. The architecture of the FCN model [8].

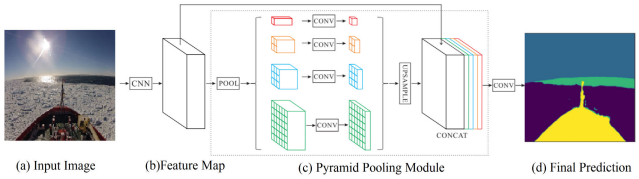

FIGURE 2. The architecture of the PSPNet model [21].

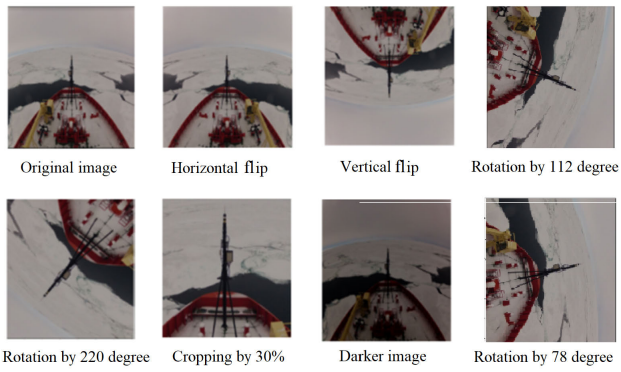

FIGURE 3. Sample of data augmentation operations.

soft-max for output. The number 16 in VGG-16 refers to the fact that it has 16 layers.

B. FULLY CONVOLUTIONAL NETWORK (FCN-8)

FCN-8 is the first model to train a network end-to-end for semantic segmentation that gained this name from its architecture, which is built from fully connected layers [8]. FCN-8 can work regardless of the image size. Figure 1 shows the architecture of the FCN-8 model.

C. PYRAMID SCENE PARSING NETWORK (PSPNet-50)

PSPNet-50 model takes into account the global context of the image to perform the local level predictions [21]; hence, it achieves better performance in comparison with the FCN model which classifies pixels without capturing the context of the whole image. Figure 2 shows the structure of the PSPNet model. It starts with an input image, then first uses the CNN to obtain the feature map of the last convolutional layer as in part (b), after which a pyramid parsing module is applied to harvest different sub-region representations, followed by upsampling and concatenation layers to form the final feature representation; this carries both local and global context information as in part (c). Finally, the representation is fed into a convolution layer to get the final per-pixel prediction as in part (d).

III. DATASET ENHANCEMENT USING DATA AUGMENTATION OPERATIONS

A. IMAGES SOURCE

The dataset is constructed from images taken from the Nathaniel B. Palmer expedition through the Ross Sea, Antarctica [20].[1] The images were captured in different light conditions encountered in the voyage ranging from midday sun to gray skies and setting sun. In addition, some images present precipitation of rain. The original data set consists of 428 training images (380 clear weather images and 48 rainy weather images) and 23 evaluation images that comprise different weather conditions such as sunny, cloudy, rainy, and clear weather. The scene in the images consists of ice, water, ship, and sky classes. As the ship moves, the surrounding ice, water, and to some extent the sky change. However, the ship class does not change much because the images are taken from a fixed location on the ship. Consequently, data augmentation operations are performed to increase the data diversity for a more robust model.

B. DATA AUGMENTATION OPERATIONS

Data augmentation encompasses a suite of operations that enhances the size of training and evaluation datasets such that better deep learning models can be built without the need to collect new data [22]. Two main categories of data augmentation operations are considered in this work, namely geometrical data augmentation and image effect data augmentation. In the following, we briefly describe the data augmentation operations.

- Vertical and horizontal flip: The original images are flipped along vertical (top to bottom) or horizontal (left to right) axis.
- Rotation: Rotation augmentation is done by rotating the images right on an axis with angle θ; we consider different value of $\theta = 10, 45, 65, 85, 90, 110, 130, 150, 180, 240, 270, 300, 320$, and 340 degrees.
- Cropping: We adopt a random cropping approach, in which we create random segments of the original image. Each cropping operation is associated with an appropriate scaling operation to maintain the size of each image before and after cropping.
- Adding noise: Adding noise augmentation consists of adding a random value to each pixel; the random value is usually drawn from a zero-mean normal distribution with variance σ^2, $\mathcal{N}(0, \sigma^2)$. In this work, we consider $\sigma^2 = 10$ and $\sigma^2 = 20$.
- Changing lighting condition: Changing lighting condition augmentation is implemented by increasing or decreasing the brightness of the images.

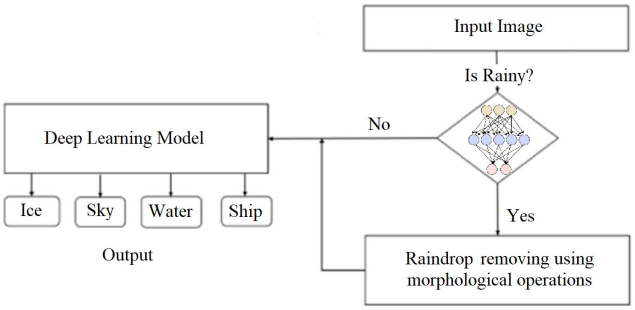

FIGURE 4. Raindrop removing framework.

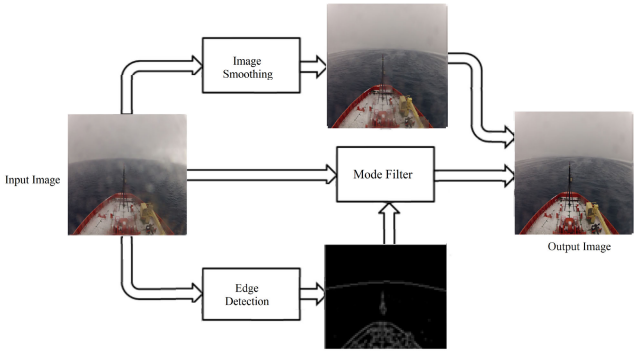

FIGURE 5. Illustration of the morphological operations.

The original training set consists of 428 images; we add 2×428 images through noise augmentation, 3×428 images by applying cropping operation, and 18×428 images by applying the rotation, flipping, and changing lighting condition operations. Figure 3 illustrates a sample example of some of the augmentation operations.

IV. RAINDROP REMOVING FRAMEWORK

Images captured onboard a ship are subject to weather conditions such as rain. The raindrop degrades the image quality and alter the boundaries of the objects in the image, which reduces the performance of the semantic segmentation models. Including rainy weather images in the training dataset does not improve their performance. To cope with this issue, removal of raindrop effects is essential to improve the performance of DNN. Figure 4 illustrates the raindrop removing framework, which consists of the following operations:

- The first step is to test whether there is any raindrop in the image or not. This step is important to apply the morphological operations to the images with raindrop. We trained a deep CNN model as binary classifier to classify the images. This model consists of a convolutional 2D layer with 16 filters, a kernel of 3×3 pixels, the input size as our image dimensions, i.e., $713 \times 713 \times 3$. After that, a stack of 5 max pooling layers is added. Finally, the output is flatten and feed into a fully-connected layer, and then to a sigmoid layer for binary classification. A data set of 10,700 images is utilized to train the binary classifier, which achieves 99.5% accuracy.
- Morphological operations: Images with raindrop undergo morphological operations to reduce the corresponding effects; these operations are image smoothing and object edge detection.
  - Image smoothing is the process of capturing important objects in the image while leaving out fine-scale structures/rapid patterns. In this work, we adopt a bilateral filter which smooths out the raindrop and preserves the edges of the object [23], [24]. The bilateral filter takes a weighted sum of the pixels in a local neighborhood; the weights depend on both the spatial distance and the intensity distance. In this way, edges are preserved well while noise is averaged out. Mathematically, at a pixel location $\boldsymbol{x}$, the output of the bilateral filter is calculated as follows

$$\hat{I}(x) = \frac{1}{C} \sum_{\boldsymbol{y} \in N(\boldsymbol{x})} \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{2\sigma_d^2}\right) \times \exp\left(-\frac{|I(\boldsymbol{y})-I(\boldsymbol{x})|^2}{2\sigma_r^2}\right) I(\boldsymbol{y}),$$

    where I and $\hat{I}$ are the input and filtered images, respectively, σ_d^2 and σ_r^2 are parameters controlling the fall-off of the weights in spatial and intensity domains, respectively, $N(\boldsymbol{x})$ is a spatial neighborhood of $\boldsymbol{x}$, and C is a normalization constant such that

$$C = \exp\left(\frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{2\sigma_d^2}\right) \exp\left(\frac{|I(\boldsymbol{y})-I(\boldsymbol{x})|^2}{2\sigma_r^2}\right),$$

    with $|\cdot|$ and $\|\cdot\|$ as the first and second norms, respectively. The selection of σ_d^2 and σ_r^2 depends on the image intensity and the size of the object that needs to be smoothed out [24], [25]. In this work, we set $\sigma_d^2 = 45$ and $\sigma_r^2 = 150$.
  - Objects edge detection: Edge detection is a morphological operation for finding the boundaries of objects within an image. The edge detection algorithm identifies points in the image at which the image brightness changes sharply or, more formally, has discontinuities. The points at which the image brightness changes sharply are typically organized into a set of curved line segments termed edges. In this work, we adopt the Canny edge detection algorithm which represents one of the most efficient edge detection algorithms [26]. It consists of the following steps: (1) Finds the intensity gradients of the image; (2) Apply non-maximum suppression to remove the spurious response to edge detection. Non-maximum suppression means that edge points are defined as points where the gradient magnitude assumes a local maximum in the gradient direction; (3) Apply a threshold to determine potential edges such that the pixels with gradient

values greater than this threshold will be considered as edge; we set the threshold as 100, which is obtained heuristically based on sensitivity analysis.

Clear boundaries between the objects in the image enables the DNN models to accurately apply the semantic segmentation. The raindrops alter objects' boundaries and the smoothing filter cannot obtain these boundaries. The raindrops do not alter all the boundary pixels of an object in the original image because their size is smaller than the object. Consequently, the boundary pixels affected by the raindrops can be reconstructed using the neighboring boundary pixels. The mode filter[2] is an edge-preserving filter, in which the value of the output pixel is obtained by the mode over all pixels within the filter's window. The locations of boundary pixels are determined using the edge detector. The input of the mode filter is the intensity of the boundary pixels in the original image and its output determines the intensity of the boundary pixels in the resulted image. Figure 5 illustrates a sample of the morphological operations.

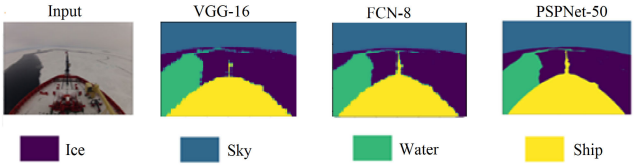

FIGURE 6. A sample input image and the results of VGG-16, FCN-8, and PSPNet-50 models.

A. COMPLEXITY ANALYSIS

Table 1 illustrates the number of training epochs, estimated time of arrival (ETA) which in the context of Keras is the estimated time before the model finishes one epoch, time of each training step, and testing time per image of each deep-learning network. The training epochs of each network are selected by evaluating the performance after each iteration to determine the optimal number of training epochs and to avoid over-fitting. It is noticed that the PSPNet-50 performs more epochs. It is worth mentioning that the training is an off-line process and can be performed using powerful computers.

The proposed raindrop removing framework consists of three morphological operations: (1) A bilateral image smoothing filter, which has a complexity on the order of $\mathcal{O}(N + \frac{N}{\sigma_d^2}\frac{R}{\sigma_r})$, where N is the number of pixels and R is the extent of the intensity scale [28, Chapter 4.5]. (2) Canny edge detection, which has a complexity on the order of $\mathcal{O}(N^2 \log N)$. (3) Mode filter, which has a complexity on the order of $\mathcal{O}(N^2 \log N)$. Consequently, the computational complexity of the proposed framework is $\mathcal{O}(N + \frac{N}{\sigma_d^2}\frac{R}{\sigma_r} + N^2 \log N + N^2 \log N) = \mathcal{O}(N^2 \log N + \frac{N}{\sigma_d^2}\frac{R}{\sigma_r})$. It is worth mentioning that the execution time of the proposed raindrop removing framework is 55 ms per image and it is applied to only the images with raindrop.

[2]Each output pixel of a mode filter is assigned the intensity of the most frequently occurring pixel in the input pixels [27].

TABLE 1. Training and testing time of the considered deep-learning networks.

DNN model	Number of epochs	ETA	Time per step	Testing time per image
VGG-16	30	185 s	361 ms	550 ms
FCN-8	39	299 s	583 ms	692 ms
PSPNet-50	43	373 s	728 ms	992 ms

V. RESULTS

In this section, we evaluate the three DNNs using the original and augmented datasets, with and without the raindrop removal.

A. TRAINING SETTINGS

The three DNN models are trained using the original dataset (428 images) and the augmented dataset (10,700 images); the size of each image is 713 × 713 pixels. Transfer learning has been implemented to train the models with both original and augmented datasets. To train the DNN models with the original dataset, pre-trained models with Cityscapes and CoCo datasets have been used as starting points. The starting points to train the DNN models with the augmented dataset are the resulted models of the original dataset. The training has been performed using a Lenovo ThinkStation-P920 server running on Linux version Ubuntu 18.04 LTS; the central processor unit is Intel Xeon 24 cores and 64 GB RAM is used. The workstation has a GeForce RTX 2080 Ti graphics card with 11 GB memory.

The classes in each image represent the following:

- Ice: ice visible in the image including ice pans;
- Water: open water of the ocean that appears in the image
- Ship: sections of the ship that appear in the image
- Sky: visible sky in the image.

We evaluate the performance of the deep learning networks using the intersection over union (IoU), precision, recall, and F1-score [29]. IoU is a similarity coefficient representing the ratio of the overlapping area of ground truth and predicted area to the total area, and can be expressed as

$$\text{IoU} = \frac{T \cap P}{T \cup P} \times 100\%, \tag{1}$$

where T and P represent the ground truth image and image produced by the deep learning model, respectively.

The precision and recall are defined as

$$\text{Pr} = \frac{P_T}{P_T + P_F} \times 100\%, \tag{2}$$

$$\text{R} = \frac{P_T}{P_T + N_F} \times 100\%, \tag{3}$$

respectively, where P_T is the number of true positive pixels, P_F is the number of false positive pixels, and N_F is the number of false negative pixels.

Finally, the F1-score is defined as

$$\text{F1} = 2 \times \frac{\text{Pr} \times \text{R}}{\text{Pr} + \text{R}} \times 100\%. \tag{4}$$

B. RESULTS OF A SAMPLE IMAGE

Figure 6 illustrates a sample image and the results of VGG-16, FCN-8, and PSPNet-50 models. It is clear that the models are able to classify the classes in the image. To get detailed insight, Table 2 summarizes the performance of the three models evaluated using the image in Figure 6. It is noticed that the best performance is achieved by the PSPNet-50, and the FCN-8 model outperforms the VGG-16 model. This is in line with the results in [8] where FCN-8 outperforms the VGG-16, as the former was designed by upgrading the latter through including FCNs and transferring its learned representations by fine-tuning. The PSPNet architecture achieves state-of-the-art performance on both the original dataset and the augmented dataset because it takes into account the global context of the image, hence gives better performance. It is worth mentioning that the values in Table 2 are the average of the four classes ice, water, ship, and sky.

TABLE 2. The performance of the three models evaluated using the image in Figure 6.

DNN model	Av. IoU	Av. Precision	Av. Recall	Av. F1-score
VGG-16	93.00	95.08	94.70	94.88
FCN-8	93.80	95.65	94.12	94.86
PSPNet-50	97.50	98.04	97.84	97.93

C. PERFORMANCE OF THE MODELS USING THE ORIGINAL DATASET

Figure 7 illustrates some sample images and the corresponding results of the images based on training the models using the original dataset.

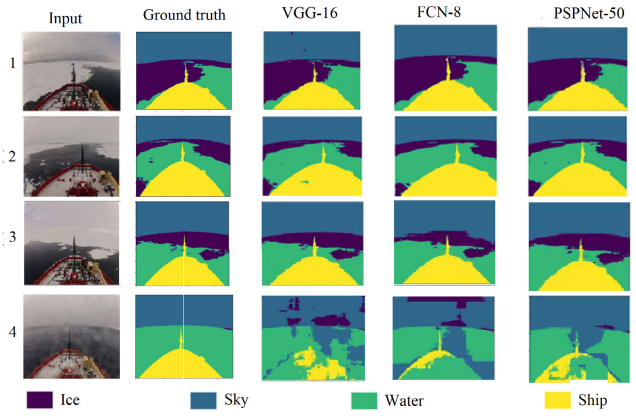

FIGURE 7. Sample original images and the corresponding results of VGG-16, FCN-8, and PSPNet-50 models.

Figure 8 summarizes the performance of the three models with the original dataset. It is noticed that the performance of the three models in the water and ice classes is lower in comparison with the other two classes. This is attributed to the fact that the location of the sky and ship do not change in the images of the original dataset. Consequentially, the location of the sky and ship enables the models to identify these classes more accurately. On the other hand, the location, shape, and size of the water and ice classes are changing in the original dataset, which reduces the models' capability to classify these classes. Motivated by this observation, geometrical data augmentation operations are applied to increase the diversity in the dataset.

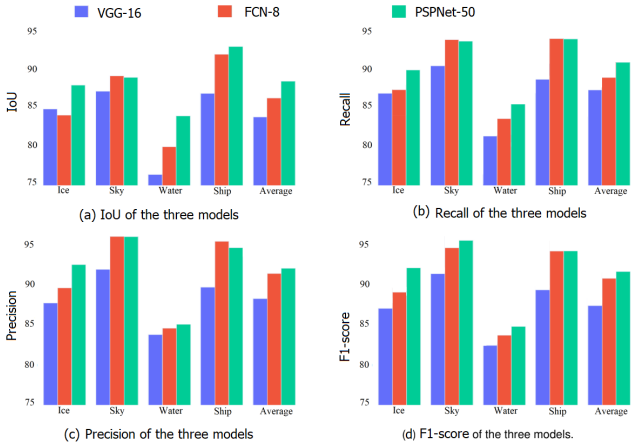

FIGURE 8. The performance of the three models with the original dataset.

D. PERFORMANCE OF THE MODELS WITH DATA AUGMENTATION

Figure 9 illustrate an image with different augmentation operations in which image 1 is the original image and images 2, 3, 4, and 5 represent the following augmentation operations: rotation by 45°, rotation by 150°, darker image, and cropped image, respectively. It is noticed that the models can classify the ice, sea, ship, and sky classes.

To study the effect of each augmentation operation, Table 3 illustrates the average IoU of the three DNN models using the original dataset, the dataset after applying each augmentation operation, and the entire augmented dataset. It is noticed that each operation improves the performance of the three DNN models and the operation of adding noise provides minor improvement. The performance of the three models improve when all augmentation operations are performed.

TABLE 3. Average IoU of the data augmentation operations.

DNN model	Without augmentation	Vertical and horizontal flip	Rotation	Cropping	Adding noise	Changing lighting condition	All operations
VGG-16	83.01	85.33	86.24	88.66	83.09	83.21	89.51
FCN-8	85.55	87.06	87.10	88.50	85.71	86.00	90.08
PSPNet-50	88.13	89.12	90.78	91.20	88.70	88.89	92.25

To get deep insight into the models performance with the augmented dataset, Figure 10 summarizes the performance of the three models. It is clear that data augmentation remarkably improves the performance of the PSPNet-50 model, which takes into account the global context of the image while the other models perform pixel per pixel classification. It is worth noting that the classification performance of the models is improved with the augmented dataset and the classification performance of the water and ice classes is close to that of the ship and sky classes. This result indicates that the augmented dataset enables the models to gain spatial diversity of the classes.

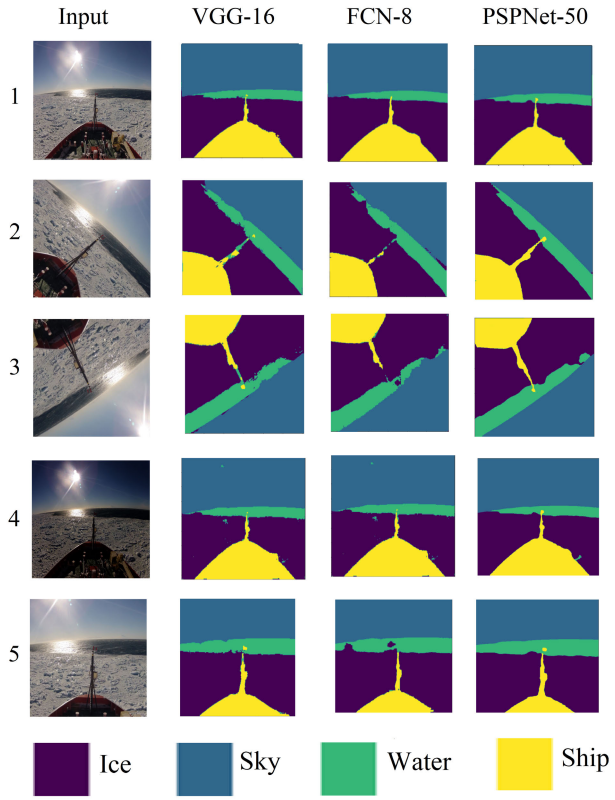

FIGURE 9. Sample augmented images and the corresponding results of VGG-16, FCN-8, and PSPNet-50 models.

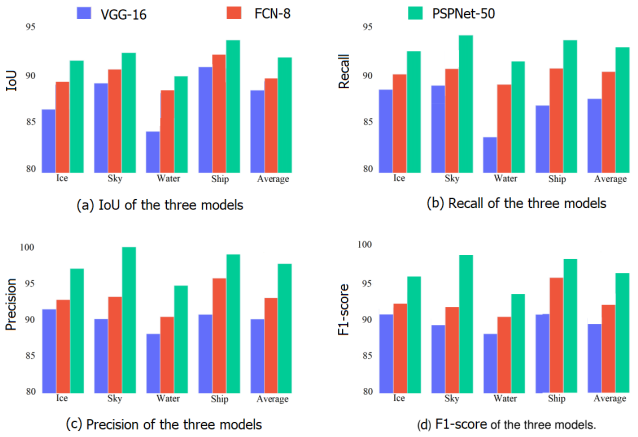

FIGURE 10. The performance of the three models with data augmentation.

E. RAINDROP EFFECT

In Figure 7, it is noticed that the three models are incapable to predict the classes properly in image 4, which is an image with raindrop. It is clear that such images degrade the performance of the three models dramatically. This motivates the application of a removing model. Figure 11 shows a sample of an image with raindrop before and after applying the removing process using the proposed framework. It is noticed that the performance of the models is improved with this framework.

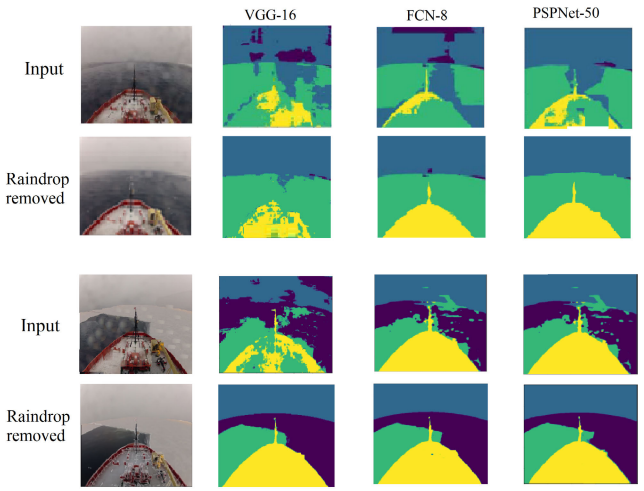

FIGURE 11. A sample of the models performance with rainy images and the results of the raindrop removing model.

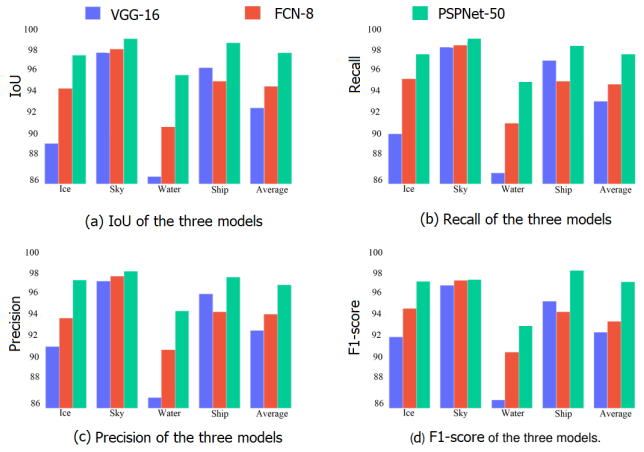

FIGURE 12. Models' performance with good weather images.

To gain deep insight on the effect of the raindrop, Figure 12 and Figure 13 illustrate the average IoU and precision of the three models when evaluated using images of good and rainy weather condition (without raindrop removal), respectively. It is clear that rainy weather images significantly degrade the performance of the three models.

Table 4 illustrates the performance of the three models before and after removing the raindrop using the proposed framework, the ℓ_0-gradient minimization approach proposed in [30], and the conditional generative adversarial network developed in [31]. It is noticed that removing the raindrop enhances the performance of the three models. Furthermore, the proposed framework outperforms the approaches in [30] and [31]. The three approaches are also compared in terms of the peak signal-to-noise ratio (PSNR), which is expressed as:

$$\text{PSNR} = 10 \times \log\left(\frac{M^2}{\text{MSE}}\right), \tag{5}$$

where M is the maximum pixel score and MSE is the mean square error between the input and resulted images. The average PSNR of the proposed framework, the approach proposed

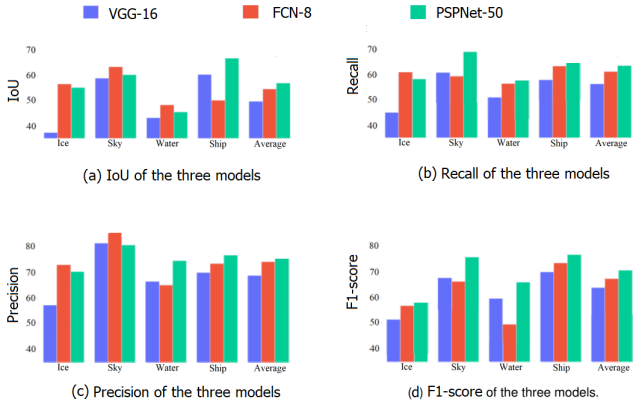

FIGURE 13. Models' performance with rainy weather images.

TABLE 4. Performance of the raindrop removing framework.

		VGG-16	FCN-8	PSPNet-50
Without raindrop removal	Av. IoU	85.91	88.97	91.83
	Av. Precision	89.98	92.85	97.44
	Av. Recall	88.01	90.10	95.10
	Av. F1-score	89.07	91.45	93.29
Proposed raindrop removal framework	Av. IoU	**91.70**	**92.32**	**97.64**
	Av. Precision	**94.44**	**95.87**	**98.99**
	Av. Recall	**92.20**	**94.02**	**97.90**
	Av. F1-score	**93.30**	**94.93**	**98.44**
Raindrop removal approach proposed in [30]	Av. IoU	88.12	90.22	94.32
	Av. Precision	92.41	94.12	98.34
	Av. Recall	90.22	92.30	96.95
	Av. F1-score	91.30	93.20	97.64
Raindrop removal approach proposed in [31]	Av. IoU	87.20	89.01	92.24
	Av. Precision	91.52 1	93.08	98.03
	Av. Recall	89.12	91.66	96.06
	Av. F1-score	90.00	92.70	95.01

in [30], and the model developed in [31] is 31.91 dB, 31.02 dB, and 30.30 dB, respectively.

VI. CONCLUSION

In this work, three deep-learning semantic segmentation networks (VGG-16, FCN-8, PSNet-50) were applied to identify sea-ice in a scene of ice, sky, water, and ship. The performance of the models has been evaluated using the IoU and precision metrics. Data augmentation operations have been implemented to increase the diversity of the dataset and a raindrop removing framework has been applied to improve the performance of the models under rainy weather conditions. Results have showed that the data augmentation operations enhance the performance of the three models. Moreover, results have illustrated that the raindrop removing framework improves the performance of the three models. Future research involves increasing the training dataset to include different sea-ice types such as new ice, first-year ice, and multi-year ice.

REFERENCES

[1] D. Snider, *Polar Ship Operations: A Practical Guide*. London, U.K.: Nautical Institute, 2018.
[2] O. K. Kjerstad, S. Loset, R. Skjetne, and R. A. Skarbo, "An ice-drift estimation algorithm using radar and ship motion measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3007–3019, Jun. 2018.
[3] W. Lu, Q. Zhang, R. Lubbad, S. Løset, and R. Skjetne, "A shipborne measurement system to acquire sea ice thickness and concentration at engineering scale," in *Proc. All Days*, Oct. 2016, pp. 1–5.
[4] H.-M. Heyn, M. Blanke, and R. Skjetne, "Ice condition assessment using onboard accelerometers and statistical change detection," *IEEE J. Ocean. Eng.*, vol. 45, no. 3, pp. 898–914, Jul. 2020.
[5] B. O'Connell, "Marine radar for improved ice detection," in *Proc. 8th Int. Conf. Exhib. Ships Struct. ICE (ICETECH)*. Princeton, NJ, USA: Citeseer, 2008, pp. 1–6.
[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.
[8] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 3431–3440.
[9] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community," *J. Appl. Remote Sens.*, vol. 11, no. 4, 2017, Art. no. 042609.
[10] N. Y. Zakhvatkina, V. Y. Alexandrov, O. M. Johannessen, S. Sandven, and I. Y. Frolov, "Classification of sea ice types in ENVISAT synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2587–2600, May 2013.
[11] H. Boulze, A. Korosov, and J. Brajard, "Classification of sea ice types in sentinel-1 SAR data using convolutional neural networks," *Remote Sens.*, vol. 12, no. 13, pp. 1–20, Jul. 2020.
[12] E. Kim, G. S. Dahiya, S. Løset, and R. Skjetne, "Can a computer see what an ice expert sees? Multilabel ice objects classification with convolutional neural networks," *Results Eng.*, vol. 4, pp. 1–13, Dec. 2019.
[13] M. Xiao, M. Rothermel, M. Tom, S. Galliani, E. Baltsavias, and K. Schindler, "Lake ice monitoring with webcams," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. IV-2, pp. 311–317, May 2018.
[14] X. Zhang, J. Jin, Z. Lan, C. Li, M. Fan, Y. Wang, X. Yu, and Y. Zhang, "ICENET: A semantic segmentation deep network for river ice by fusing positional and channel-wise attentive features," *Remote Sens.*, vol. 12, no. 2, pp. 1–22, Jan. 2020.
[15] D. P. Williams, "On the use of tiny convolutional neural networks for human-expert-level classification performance in sonar imagery," *IEEE J. Ocean. Eng.*, vol. 46, no. 1, pp. 236–260, Jan. 2021.
[16] C. Bentes, D. Velotto, and B. Tings, "Ship classification in TerraSAR-X images with convolutional neural networks," *IEEE J. Ocean. Eng.*, vol. 43, no. 1, pp. 258–266, Jan. 2018.
[17] A. Mahmood, M. Bennamoun, S. An, F. A. Sohel, F. Boussaid, R. Hovey, G. A. Kendrick, and R. B. Fisher, "Deep image representations for coral image classification," *IEEE J. Ocean. Eng.*, vol. 44, no. 1, pp. 121–131, Jan. 2019.
[18] Q. Yu and D. A. Clausi, "SAR sea-ice image analysis based on iterative region growing using semantics," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3919–3931, Dec. 2007.
[19] T. Reid, T. Walter, P. Enge, and A. Fowler, "Crowdsourcing Arctic navigation using multispectral ice classification & GNSS," in *Proc. 27th Int. Tech. Meeting Satell. Division Inst. Navigat.*, Sep. 2014, pp. 707–721.
[20] B. Dowden, O. De Silva, W. Huang, and D. Oldford, "Sea ice classification via deep neural network semantic segmentation," *IEEE Sensors J.*, vol. 21, no. 10, pp. 11879–11888, May 15, 2021.
[21] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2881–2890.
[22] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, p. 60, Jul. 2019.
[23] N. S. Pal, S. Lal, and K. Shinghal, "A visibility restoration framework for rainy images by using L_o gradient minimization and bilateral filtering," in *Proc. Int. Conf. Adv. Comput., Commun. Control Netw. (ICACCCN)*, Oct. 2018, pp. 848–852.
[24] M. Zhang and B. K. Gunturk, "Multiresolution bilateral filtering for image denoising," *IEEE Trans. Image Process.*, vol. 17, no. 12, pp. 2324–2333, Dec. 2008.
[25] J. Yang, F.-L. Luo, and A. Nehorai, "Spectral contrast enhancement: Algorithms and comparisons," *Speech Commun.*, vol. 39, nos. 1–2, pp. 33–46, Jan. 2003.
[26] M. Basu, "Gaussian-based edge-detection methods—A survey," *IEEE Trans. Syst., Man, Cybern.*, vol. 32, no. 3, pp. 252–260, Aug. 2002.

[27] L. D. Griffin, "Mean, median and mode filtering of images," *Proc. Roy. Soc. London. Ser. A, Math., Phys. Eng. Sci.*, vol. 456, no. 2004, pp. 2995–3004, Dec. 2000.
[28] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, *Bilateral Filtering: Theory and Applications*. Norwell, MA, USA: Now Publishers Inc, 2009.
[29] W. Wang, Y. Fu, F. Dong, and F. Li, "Semantic segmentation of remote sensing ship image via a convolutional neural networks model," *IET Image Process.*, vol. 13, no. 6, pp. 1016–1022, Apr. 2019.
[30] B. N. Manu, "Rain removal from still images using L_0 gradient minimization technique," in *Proc. 7th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Oct. 2015, pp. 263–268.
[31] H. Zhang, V. Sindagi, and V. M. Patel, "Image de-raining using a conditional generative adversarial network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 11, pp. 3943–3956, Nov. 2020.

NAHED M. ALSHARAY (Student Member, IEEE) received the B.Sc. degree in information technology from Taiz University, Yemen, in 2009, and the M.Sc. degree in computer science from King Faisal University, Saudi Arabia, in 2017. She is currently pursuing the Ph.D. degree with the Computer Science Department, Memorial University of Newfoundland, Canada. Her current research interests include image processing, autonomous navigation systems design, and machine learning.

YUANZHU CHEN (Member, IEEE) received the B.Sc. degree from Peking University, China, in 1999, and the Ph.D. degree from Simon Fraser University, Canada, in 2004. From 2004 to 2005, he was a Postdoctoral Researcher with Simon Fraser University. In 2005, he joined the Memorial University of Newfoundland (Memorial University) as a tenure-track Assistant Professor. While at Memorial University, he was the Deputy Head of Undergraduate Studies, from 2012 to 2015, the Deputy Head of Graduate Studies, from 2016 to 2019, and the Department Head, from 2019 to 2021. He then joined the Queen's School of Computing, in 2021. He has been a Professor of computing science, since 2005, and is currently affiliated with the School of Computing, Queen's University. His research interests include complex networks, computer networking, online social networks, mobile computing, graph theory, web information retrieval, and evolutionary computation, with funding from national agencies and various university programs and awards. He was a recipient of the President's Award for Distinguished Teaching, in 2018.

OCTAVIA A. DOBRE (Fellow, IEEE) received the Dipl. Ing. and Ph.D. degrees from the Politehnica University of Bucharest, Romania, in 1991 and 2000, respectively.

Between 2002 and 2005, she was with the New Jersey Institute of Technology, USA. She was a Visiting Professor with the Massachusetts Institute of Technology, USA, and the Université de Bretagne Occidentale, France. In 2005, she joined Memorial University, Canada, where she is currently a Professor and the Research Chair. Her research interests include wireless technologies, such as non-orthogonal multiple access and full duplex, as well as optical and underwater communications, and machine learning for communications. She has (co)authored over 350 refereed papers in these areas.

Dr. Dobre was a Fulbright Scholar, a Royal Society Scholar, and a Distinguished Lecturer of the IEEE Communications Society. She is a fellow of the Engineering Institute of Canada and the Canadian Academy of Engineering. She obtained the Best Paper Awards at various conferences, including IEEE ICC, IEEE Globecom, IEEE WCNC, and IEEE PIMRC. She also served as the general chair, the technical program co-chair, the tutorial co-chair, and the technical co-chair of symposia at numerous conferences. She was the Editor-in-Chief (EiC) of the IEEE COMMUNICATIONS LETTERS, a senior editor, an editor, and a guest editor for various prestigious journals and magazines. She serves as the EiC for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY.

OSCAR DE SILVA (Member, IEEE) received the B.Sc. degree in engineering from the University of Moratuwa, Sri Lanka, in 2009, and the Ph.D. degree from the Memorial University of Newfoundland (MUN), St. John's, NL, Canada, in 2015. Following his postdoctoral research work with the ABS Harsh Environment Technology Center, St. John's, he joined MUN as a Faculty Member, in 2016, where he is currently an Assistant Professor with the Faculty of Engineering and Applied Science. His main research interests include autonomous robotics, navigation systems design, and machine learning.