3-D Hybrid CNN Combined With 3-D Generative Adversarial Network for Wetland Classification With Limited Training Data

Ali Jamali, Masoud Mahdianpari, Fariba Mohammadimanesh, Brian Brisco, and Bahram Salehi, *Senior Member, IEEE*

***Abstract*—Recently, deep learning algorithms, specifically convolutional neural networks (CNNs), have played an important role in remote sensing image classification, including wetland mapping. However, one limitation of deep CNN for classification is its requirement for a great number of training samples. This limitation is particularly enhanced when the classes of interest are spectrally similar, such as that of wetland types, and the training samples are limited. This article presents a novel approach named 3-D hybrid generative adversarial network (3-D hybrid GAN) that addresses the limited training sample issue in the classification of remote sensing imagery with a focus on complex wetland classification. We used a conditional map unit that generates synthetic training samples for only classes with a lower number of training samples to improve the per-class accuracy of wetlands. This procedure overcomes the issue of imbalanced data in conventional wetland mapping. Based on the achieved results, better classification accuracy is obtained by integrating a 3-D generative adversarial network (3-D GAN) and the CNN network of a 3-D hybrid CNN using both 3-D and 2-D convolutional filters. Experimental results on the avalon pilot site located in eastern Newfoundland, Canada, and covering five wetland types of bog, fen, marsh, swamp, and shallow water demonstrate that our model significantly outperforms other CNN models, including the HybridSN, SpectralNet, MLP-mixer, as well as a conventional algorithm of random forest for complex wetland classification by approximately 1% to 51% in terms of F-1 score.**

***Index Terms*—Convolutional neural network (CNN), deep convolutional neural network (DCNN), generative adversarial network (GAN), random forest (RF), wetland mapping.**

I. Introduction

WETLANDS provide various ecological services to the fauna, flora, and humans [1], including water storage, carbon sequestration, as well as flood and shoreline protection. Wetland's contributions to the environment originated from their biogeochemical and hydrological processes. These processes are regarded as wetland functions, including biomass production, nutrient transformation, biogeochemical transformation, hydraulic recharge and storage, and improving water quality [2]. Despite these benefits, they are threatened by anthropogenic activities, such as urban and agricultural developments and industrialization [3], as well as by the ever-changing climate. It is reported that recent global droughts can alter wetlands from carbon sinks to carbon sources by a higher rate of decay [4]. As such, accurate and up-to-date systematic mapping, and monitoring of these vital yet in danger natural ecosystems are crucial for their protection and preservation.

Currently, remote sensing is the state-of-the-art technology for the accurate and systematic method of collecting coarse- to very high-resolution images of the earth's surface. Complexities such as wetlands' spectral similarity and vegetation composition and dynamics (i.e., phenological changes) limit the capability of conventional machine learning approaches for wetland classification [3]. On the other hand, the existence of water near or on the ground surface or below the vegetation canopy is used to define various classes of wetlands rather than a common type of land cover or vegetation that increases the complexity of wetland classification [5]. In recent years many advanced machine learning algorithms have been developed for the classification of wetlands [6], [7], [8], [9], [10], [11]. Currently, deep learning (DL) algorithms, specifically the convolutional neural networks (CNNs), are effectively used for complex wetland classification [12], [13], [14], [15].

The issue of the DL methods arises from their need for many training data for their training to reach their full potential of highly accurate image classification. Collecting training data is costly, time-consuming, and labor-intensive in remote sensing. Two techniques can address the issue of the limited number of wetlands training data: 1) transfer learning [16], [17]; 2) generating synthetic samples using generative adversarial networks (GANs) [18], [19], [20], [21]. Various research on hyperspectral image classification has utilized a limited number of training data [22], [23], [24], [25], [26]. For example, Liu et al. [24] proposed a methodology based on a metric space to address the issue of limited training data in the classification of hyperspectral imagery. Moreover, an unsupervised cross-data set model based

Manuscript received 15 February 2022; revised 1 July 2022 and 10 August 2022; accepted 4 September 2022. Date of current version 23 September 2022. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) Discovery under Grant RGPIN-2022-04766. *(Corresponding author: Masoud Mahdianpari.)*

Ali Jamali is with the Civil Engineering Department, Faculty of Engineering, University of Karabük, 78050 Karabük, Turkey (e-mail: alijamali@karabuk.edu.tr).

Masoud Mahdianpari is with the C-CORE, St. John's, NL A1B 3X5, Canada, and also with the Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada (e-mail: m.mahdianpari@mun.ca).

Fariba Mohammadimanesh and Brian Brisco are with the Canada Centre for Mapping and Earth Observation, Ottawa, ON K1S 5K2, Canada (e-mail: fm7550@mun.ca; brian.brisco@Canada.ca).

Bahram Salehi is with the Department of Environmental Resources Engineering, State University of New York College of Environmental Science and Forestry (SUNY ESF), Syracuse, NY 13210 USA (e-mail: bsalehi@esf.edu).

Digital Object Identifier 10.1109/JSTARS.2022.3206143

on the deep domain adaptation was developed by Ma et al. [26] to achieve a high level of classification accuracy in hyperspectral images. Moreover, Qu et al. [25] proposed a transfer learning methodology using shared abundance space in the classification of hyperspectral imagery with a few training data.

Although there have been various research on the use of GANs for the classification of remote sensing images, specifically for hyperspectral data, to the best of our knowledge, GANs have not been fully investigated for wetland classification. Due to the complex characteristics of wetlands, generating synthetic samples through GANs may cause errors. The reason is that generating a fake wetland image may increase the complexity of the classification task, resulting in a lower wetland classification accuracy. In this article, to address the issue of limited training samples in remote sensing, specifically complex wetland mapping, we propose a wetland classification method that integrates and enhances the power of GANs and CNN models. Our method is called 3-D hybrid Generative Adversarial Network (3-D hybrid GAN). Our motivation behind the integration of a hybrid network and GANs is to establish a robust CNN algorithm to classify wetlands with a limited number of training samples. More precisely, our contributions are summarized as follows.

1) We propose a CNN algorithm that optimizes the classification model and improves the quality and the number of wetlands training data. Moreover, compared to conventional (or solo) CNNs, the 3-D hybrid GAN overcomes the problem of limited training samples for complex wetland classification.
2) We propose using a conditional map unit focusing on the generation of high-quality wetland samples to address the issue of imbalanced data in the current complex wetland mapping. As such, we achieved a high overall accuracy (OA) level and significantly improved the per-class accuracy of wetlands compared to the current CNN algorithms.
3) The generation of wetland samples is costly and requires expert-level field data acquisition. At the same time, nonexperts can recognize and produce samples of nonwetlands through the utilization of very high remote sensing images such as RapidEye satellite data and Google Earth images. This research studies the generation of high-quality wetland samples using advanced and state-of-the-art algorithms (i.e., GANs).

The rest of this article is organized as follows. Related works on image classification, CNN, and GAN are discussed in Section II. The proposed CNN method of 3-D hybrid GAN is illustrated and formulated in Section III. Experiments, including the settings, study area, and results, are provided in Section IV. Finally, Section V concludes this article and summarizes the findings.

II. Related Work

A. Image Classification

Traditional classifiers, such as maximum likelihood, cannot effectively interpret the complicated high-dimensional data of remote sensing [27]. The normal distribution of input data is one of these algorithms' assumptions, which is not applicable to many remote sensing samples. Random forest (RF) [28], decision trees [29], and support vector machine (SVM) [30] classifiers, on the other hand, are unaffected by the input data distribution. Procedures for the classification of remote sensing images using shallow classifiers, such as RF and SVM include feature extraction and classification. Spatial, temporal, and spectral data of satellite images are transformed into the feature vectors in the feature extraction process, while these extracted features are classified into different land cover types in the classification stage [31], [32], [33]. It is worth highlighting that the manual feature engineering of the conventional classification methods (i.e., the selection of the most appropriate extracted features in the classification of satellite data) significantly affects the final classification results.

B. Convolutional Neural Networks

The issue of manual feature engineering was overcome by introducing DL techniques such as CNNs [9], [34], [35], [36], [37]. CNNs are specifically used for image classification, semantic image segmentation, and object detection. It is worth mentioning that CNNs extract low-, intermediate-, and high-level features from images. In each layer (l), based on the weights (W) and biases (B) of the previous layers, features are extracted. Then, they are updated in the next iteration (1), (2)

$$\Delta W_l\ (t+1) = -\frac{x\lambda}{r} W_l - \frac{x}{n}\frac{\partial C}{\partial W_l} + m\,\Delta W_l\ (t) \quad (1)$$

$$\Delta\ B_l\ (t+1) = -\frac{x}{n}\frac{\partial C}{\partial B_l} + m\,\Delta\ B_l\ (t) \quad (2)$$

where λ, n, and x present a regularization parameter, the total number of training samples, and the learning rate. In addition, t, m, and C are updating step, momentum, and cost function, respectively. Based on the dataset, to obtain optimal performance, the learning rate (x), regularizing parameter (λ), and momentum (m) are fine-tuned. The DL methods use and learn from representation rather than using an empirical manual feature engineering in the traditional classifiers (e.g., RF), achieving much higher accuracies for high-resolution earth observation data classification [38], [39], [40], [41], [42], [43]. It is reported that the DL classifiers usually recognize a more generalized pattern in the satellite data that is regarded as their superiority over the traditional algorithms [3], [44]. In addition, DL methods, including the CNNs, use spatial and spectral data in a patch of pixels (e.g., a patch of 15 by 15 pixels) rather than spectral information of a pixel utilized by conventional shallow classifiers, such as RF. The issue of deep convolutional neural networks (DCCNs) arises from their need for a large number of training data as they have a relatively significantly high number of parameters to be fine-tuned.

C. Generative Adversarial Network

As discussed, generating new data in remote sensing, specifically wetland mapping, is time-consuming, labor-intensive, and costly. On the other hand, although the DL methods have been successfully employed in different fields of remote sensing, such as object detection [45], [46], [47], [48], [49] and classification

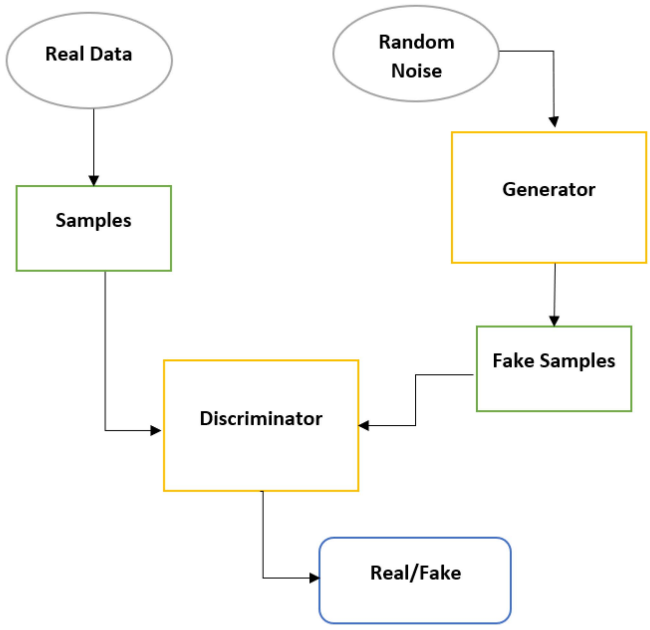


Fig. 1. Generator and discriminator networks of GANs.

[50], [51], [52], [53], [54], [55], they have a need for a large set of training data [3]. This issue can be addressed with the utilization of the new architecture of GANs that revolutionized the DL field introduced by Goodfellow et al. [56]. There are two generator and discriminator networks in the GANs architecture, as presented in Fig. 1.

As shown in Fig. 1, in the 3-D GAN, using a random noise vector, the generator network creates new synthetic samples while the discriminator tries to differentiate between the synthetic and real image data. Thus, the 3-D GAN is trained while the generator (G) produces more realistic synthetic samples, and the discriminator network (D) tries to differentiate the real samples from the synthetic ones [57], [58], [59], [60]. The main benefit of GAN models in remote sensing image classification is the possibility of creating high-quality ground-truth data without any field data collection. This is vital as producing new ground-truth data for wetland classifications is significantly costly and time-consuming. Moreover, there is a need for expert-level knowledge for the recognition of different wetlands in the field data collection as well. Addressing such issues will substantially improve the large-scale wetland classification in the near future and will improve the current wetland classification accuracies.

III. Proposed CNN Network

In this section, we describe the details of the proposed CNN algorithm of 3-D hybrid GAN, which can generate and classify complex wetland data through the integration of a 3-D GAN and a modified version of the hybridSN network [61]. It should be noted that generating samples for the nonwetland classes such as urban regions does not require expert knowledge, while producing wetland samples necessitates the assistance of biologist experts and field data acquisition. The reason is that, unlike the nonwetlands, wetlands do not have regular (i.e., clear-cut) boundaries, and some of them have similar vegetation types and structures. Thus, we propose a model that mainly focuses on the generation of synthetic samples of high-resolution wetlands. The main idea is to generate samples for the classes with minority training data. As such, we propose the 3-D hybrid GAN for the generation and classification of wetlands and nonwetlands based on the generative adversarial minority oversampling (GAMO) [62] and 3-D-HyperGAMO [60] models.

The overall architecture of the proposed 3-D hybrid GAN is illustrated in Fig. 2. The proposed 3-D GAN model is composed of three major parts, including the generator (G), the discriminator (D), and the classifier (C), while the generator has two parts of the conditional map unit and the 3-D RapidEye patch generator as described in the next sections. In the 3-D GAN model, the generator (G) tries to create high-quality wetland samples, while the discriminator (D) tries to recognize the real samples from the synthetic ones. On the other hand, the classifier (C) in the 3-D GAN model uses the synthetic wetland and nonwetland samples produced by the generator (G) and the real RapidEye image patches to classify the wetland and nonwetland classes (part I).

It should be noted that the major class is chosen based on the number of training data in part I. The reason is that training the generator network (G) is a costly and time-consuming procedure. After obtaining a high level of wetland classification accuracy by the 3-D classifier (C) in the 3-D GAN network, based on the number of test data, a much higher number of synthetic data is produced. The reason is that after the proper training of the generator network (G), a much higher number of new synthetic RapidEye samples can be generated in just a few minutes. Then, the real and synthetic data generated by the 3-D generator (G) are fed to the 3-D hybrid CNN classifier for precise wetland classification (part II).

A. 3-D Tensor

In this study, we consider a RapidEye image as a 3-D tensor ($X_{\text{original}} \in R^{W\times H\times B}$) defined by width ($W$), height ($H$), and the number of bands (B). The wetland classification is done by classifying the RapidEye image pixels into l_c wetland and nonwetland classes, denoted by $Y = y_1, y_2, \ldots, y_{l_c}$. As such, a pixel ($x_{i,j} \in X_{\text{original}}$), is defined as a spectral vector ($x_{i,j} = x_{i,j,1}, \ldots, x_{i,j,B} \in R^B$, where $i = 1, 2, \ldots, W$ and $j = 1, 2, \ldots, H$). In this study, we are interested in the extraction of 3-D patches of the wetland and nonwetland samples defined by $Ra_i \in R^{S\times S\times B}$, where the $S \times S$ is the size of the image patches and B is the number of spectral bands. In our model, a sliding window of size $S \times S$ is used to extract the spectral and spatial information of the RapidEye imagery, which is then used as the input of the 3-D patch generator and the classifier and discriminator networks. To use the developed 3-D DL model, we stack all generated 3-D RapidEye image patches into a 3-D

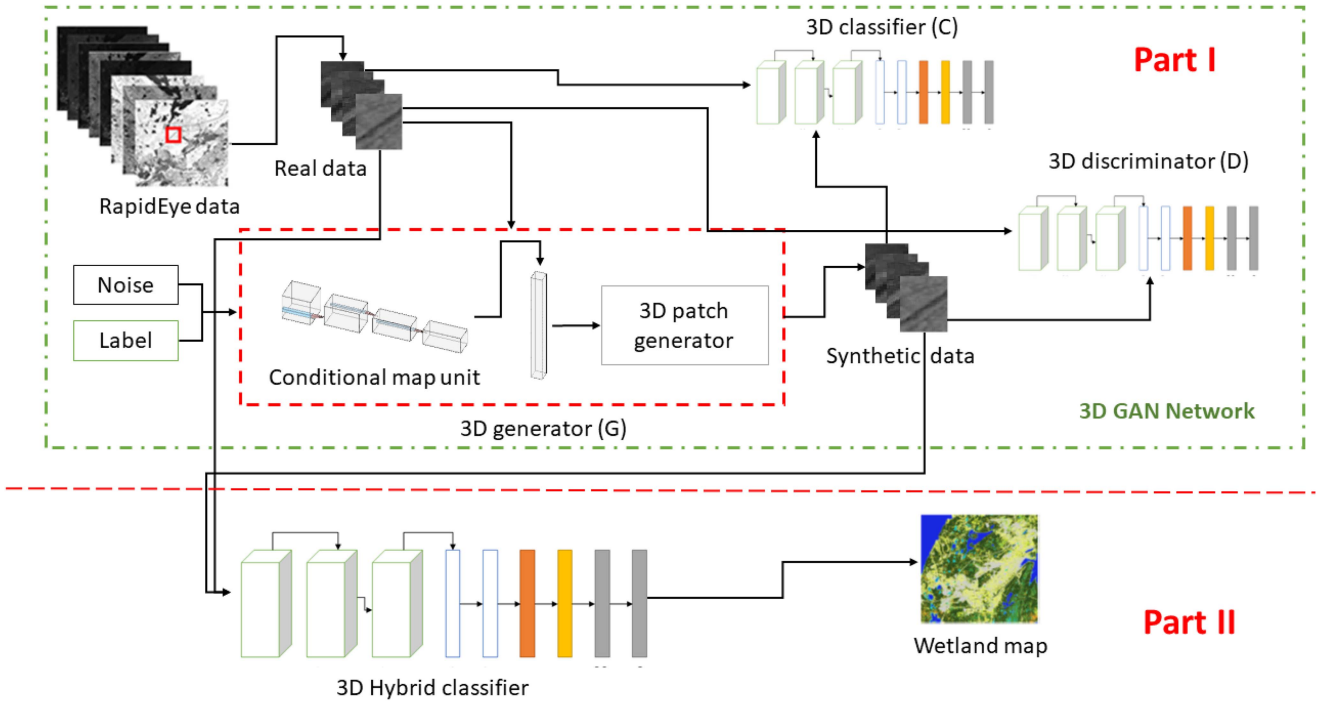


Fig. 2. Architecture of the proposed 3-D Hybrid GAN.

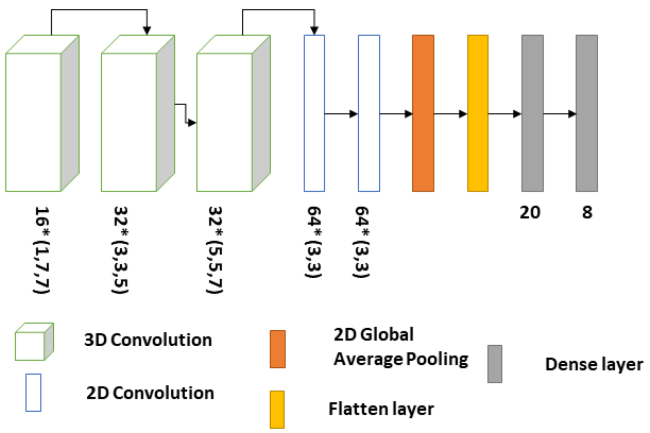


Fig. 3. Architecture of the implemented 3-D Hybrid CNN classifier.

tensor ($X = \{Ra_1, Ra_2, \ldots, Ra_n\}$, where $Ra_i \in R^{S\times S\times B}$ and n is the number of training data).

B. Generator Network

For the generation of the synthetic wetland samples, like both models of GAMO and 3-D-HyperGAMO, we used a conditional map unit to generate samples from a random noise vector, only for classes with a lower number of training data (Table I). The advantage of this model is solving the issue of imbalanced data that is common in wetland mapping. Then, the output of the conditional map unit is flattened to form a conditional feature vector. Afterward, the output of the conditional map unit is linked to a 3-D patch generator according to the label information. The

TABLE I
ARCHITECTURE OF THE CONDITIONAL FEATURE MAPPING UNIT

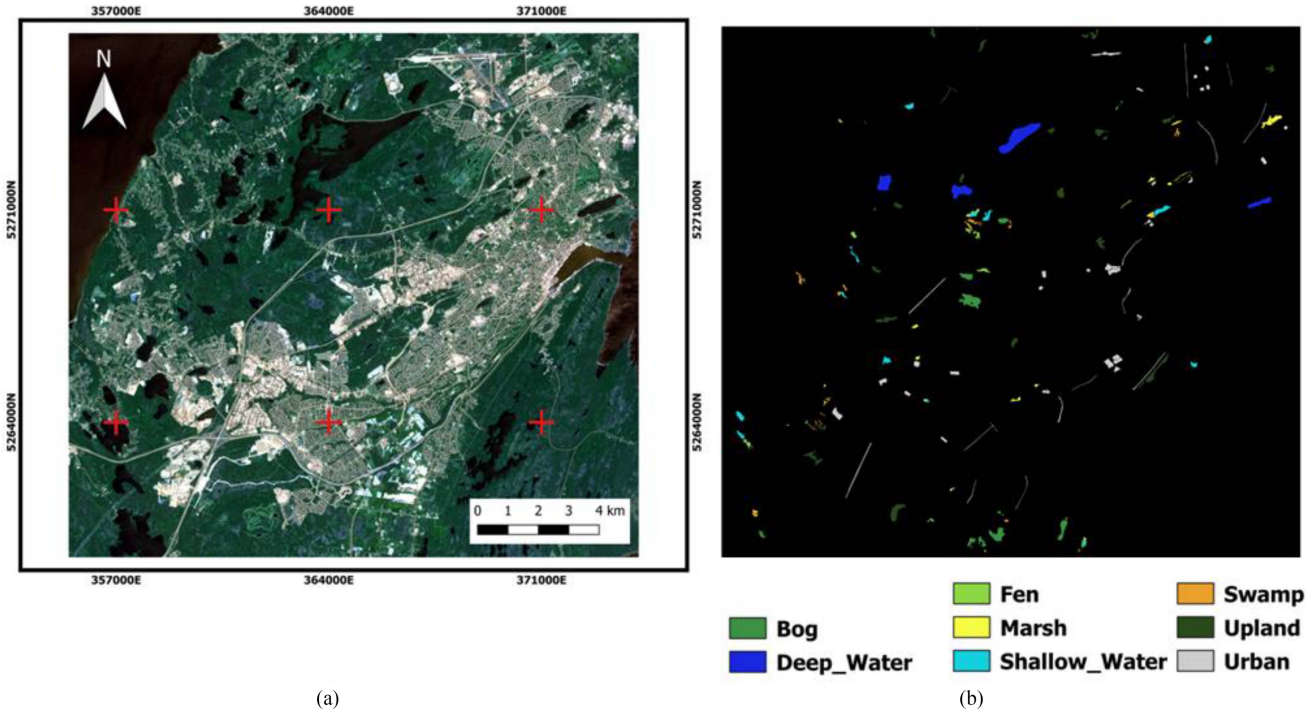
Layer	Filters/Kernel size	Batch normalization	Activation function
Dense	256*7*7	Yes	LeakyRelu
Reshape	256,7,7	-	-
Conv2DTranspose	128,5,5	Yes	LeakyRelu
Conv2DTranspose	1,5,5	Yes	LeakyRelu

patch generator creates a different number of 3-D patches for classes with the minor number of training samples that were sent by the conditional mapping unit. The process is described as follows.

In the proposed 3-D patch generator, the urban class with the highest number of training data is considered the majority class, while the other wetland and nonwetland classes are the minority classes. As a result, the 3-D patch generator uses seven ($c-1$, c presents the number of classes) units, one unit for each of the minor classes. The unit U_i generates γ_i^g samples as defined by

$$\gamma_i^g = \gamma_m - \gamma_i \tag{3}$$

where γ_m is equal to 1899 and γ_i presents the number of training data belonging to the minor classes. Each U_i uses two inputs of the 3-D patches of RapidEye image samples with the dimension of $\gamma_i \times S \times S \times B$, as well as the output of the conditional unit map. In our model, a dense layer is used to transform the output of the conditional unit map (intermediate feature presented by I_f) to a feature with the length of γ_i, followed by a softmax layer. The feature vector is repeated by $n = S \times S \times B$ times



(a) (b)

Fig. 4. Map of the study area. (a) Location of the pilot site of the Avalon, Newfoundland, Canada. (b) Ground-truth wetland samples acquired by the biologist team across the study area.

for the generation of the class-specific random feature (I_m) with the dimension of $n \times \gamma$. The 3-D RapidEye patches with the dimension of $\gamma_i \times S \times S \times B$ are then reshaped into a matrix of P_m with a dimension of $n \times \gamma$. Then, the wetland and nonwetland class-specific feature matrix (F_m) with the dimension of $n \times \gamma$ is calculated as shown in

$$F_m = I_m \,.(F_m)^T \tag{4}$$

where $(F_m)^T$ is the transpose of the matrix F_m and $(.)$ is the dot product. It is followed by a flattened vector (F_v) with a dimension of n that is calculated through the sum of column-wise of each row belonging to the F_m matrix. Afterward, F_v is reshaped into the dimension of $S \times S \times B$ for the generation of the synthetic wetland and nonwetland samples for the classes with a minor number of training data.

C. 3-D Classifiers and Discriminator Networks

For our proposed 3-D GAN classifier and discriminator networks, we used a modified architecture of the hybrid SpectralNet (i.e., HybridSN) proposed by Roy et al. (2020) [61]. The architecture of the proposed model is presented in Fig. 3.

As seen in Fig. 3, the proposed classifier and discriminator have three 3-D convolutional layers, followed by two 2-D convolutional layers. It is followed by a 2-D global average pooling and a flatten layer. We included two dense layers with sizes of 20 and 8 at the end of the proposed 3-D hybrid classifiers. It should be noted that the last dense layer for the discriminator network has a size of 1.

D. Objective Function

Considering our generator (G), discriminator (D), and classifier (C), the objective function of the 3-D GAN model is formulated as

$$\min_G \max_D \max_C \; O\,(G, D, C) \; = \sum_{i \in c} O_i \tag{5}$$

where $O_i = (O_{i1} + O_{i2} + O_{i3} + O_{i4} + O_{i5} + O_{i6})$ and c is the number of classes. As such, the objective of the generator network (G) is described as fooling the discriminator network (O_{i6}), while the discriminator network (D) should classify the extracted features of the RapidEye image as real O_{i5} and the generated synthetic samples as fake samples (O_{i6}). The classifier (C) should maximize the wetland and nonwetlands class score for the ith class with respect to the real and synthetic samples of class ith (O_{i1} and O_{i3}), while trying to minimize the class score with respect to the other wetland and nonwetland classes (O_{i2} and O_{i4}), simultaneously. In other words, the probability of belonging to the real and synthetic sample for class ith should be high (i.e., a high-class score), while the probability of that real or synthetic sample belonging to the other classes should be low (i.e., a low score for the rest of the classes).

IV. Experiments

In the experiments, we validate the effectiveness of the proposed 3-D hybrid GAN model for the generation and classification of complex wetland and nonwetland regions.

Fig. 5. Confusion matrices of test data based on the prediction of each shallow and CNN algorithms of (a) 3-D hybrid, (b) HybridSN, (c) MLP-mixer, (d) RF, (e) SpectralNet, and (f) the proposed 3-D hybrid GAN.

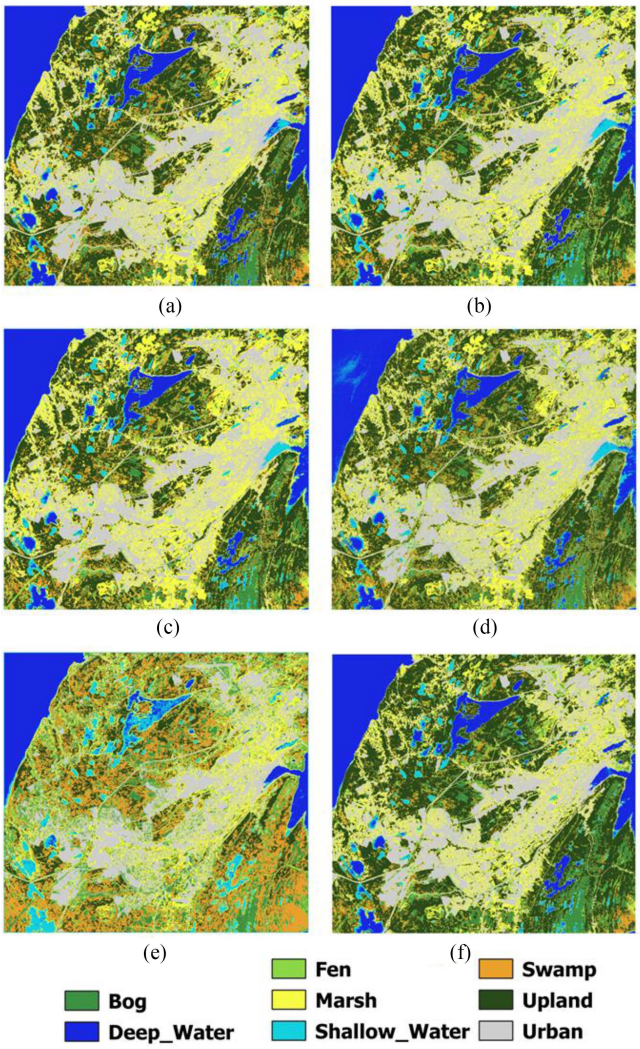


Fig. 6. Wetland classification maps using the shallow and DCNN models of (a) 3-D hybrid, (b) HybridSN, (c) MLP-mixer, (d) RF, (e) SpectralNet, and (f) the proposed 3-D hybrid GAN.

TABLE II
NUMBER OF WETLANDS AND NONWETLANDS TRAINING AND TESTING SAMPLES

Class	Number of training pixels	Number of testing pixels
Bog	1152	21883
Fen	309	5865
Marsh	551	10475
Swamp	333	6332
Shallow water	789	14985
Urban	1899	36079
Deep water	1712	32530
Upland	1683	31980

A. Study Area and Data Descriptions

This article uses the RapidEye Image of the Avalon pilot site to evaluate the proposed method of 3-D hybrid GAN. As shown in Fig. 4, the pilot site is on the most eastern side of Newfoundland, Canada. In the pilot site of Avalon, five wetland classes of shallow water, bog, marsh, fen, and swamp defined by the canadian wetland classification system (CWSC) are present, where the bog and fen classes (i.e., peatlands) are the most dominant. In the study area, the capital and the largest city of Newfoundland, the St. John's town, with a population of approximately 226 000, is situated. The ground truth data have been collected from the years 2015 to 2017. The wetlands reference data has been collected and labeled by a team of experts, including wetland specialists and biologists in the field, and high-resolution imagery has also been employed as a reference to minimize the uncertainty of reference data labeling. It is worth mentioning that the possible wetland areas were recognized with the utilization of very high-resolution images of RapidEye and Google Earth prior to the ground truth data acquisition. Moreover, GPS points, notes, and photographs were taken to improve the delineation of the wetland regions. Fig. 4 presents several samples of wetlands acquired by the wetland experts (i.e., biologists).

The number of wetlands training and test samples is shown in Table II. We used a random sampling strategy where 5% of the ground truth data were used as training and the other 95% as our test data.

For the complex wetland classification, we used five bands of blue, green, red, red edge, and near-infrared of two level-3a RapidEye images with spatial resolutions of 5 m acquired on 18 June and 22 October 2015. Moreover, several spectral indices, including the ratio-vegetation index, the red edge normalized difference vegetation index (RENDVI), and green NDVI (GNDVI), were used to improve the final wetland classification accuracy (Table III).

B. Experimental Settings

In this research, we selected RapidEye image patch sizes of 6×6 experimentally. The reason was that although selecting a larger image patch size would result in higher classification accuracy (e.g., we achieved a high classification accuracy of 97% with RapidEye patch sizes of 11×11), linear features such as buildings and road areas will lose their sharp and linear shape. Moreover, smaller objects inside larger features will be lost as well. As such, we selected a smaller RapidEye patch size to preserve the linear objects' shape and maintain information on smaller ground features.

Hyper-parameters: We used Adam optimizer to train our 3-D hybrid GAN network with a learning rate of 0.0002. We set the maximum training iteration to 30 000 epochs in the 3-D GAN network. Moreover, we set the noise dimension and training batch size as 100 and 16, respectively. The training iteration in the other CNN networks was set to 100 epochs.

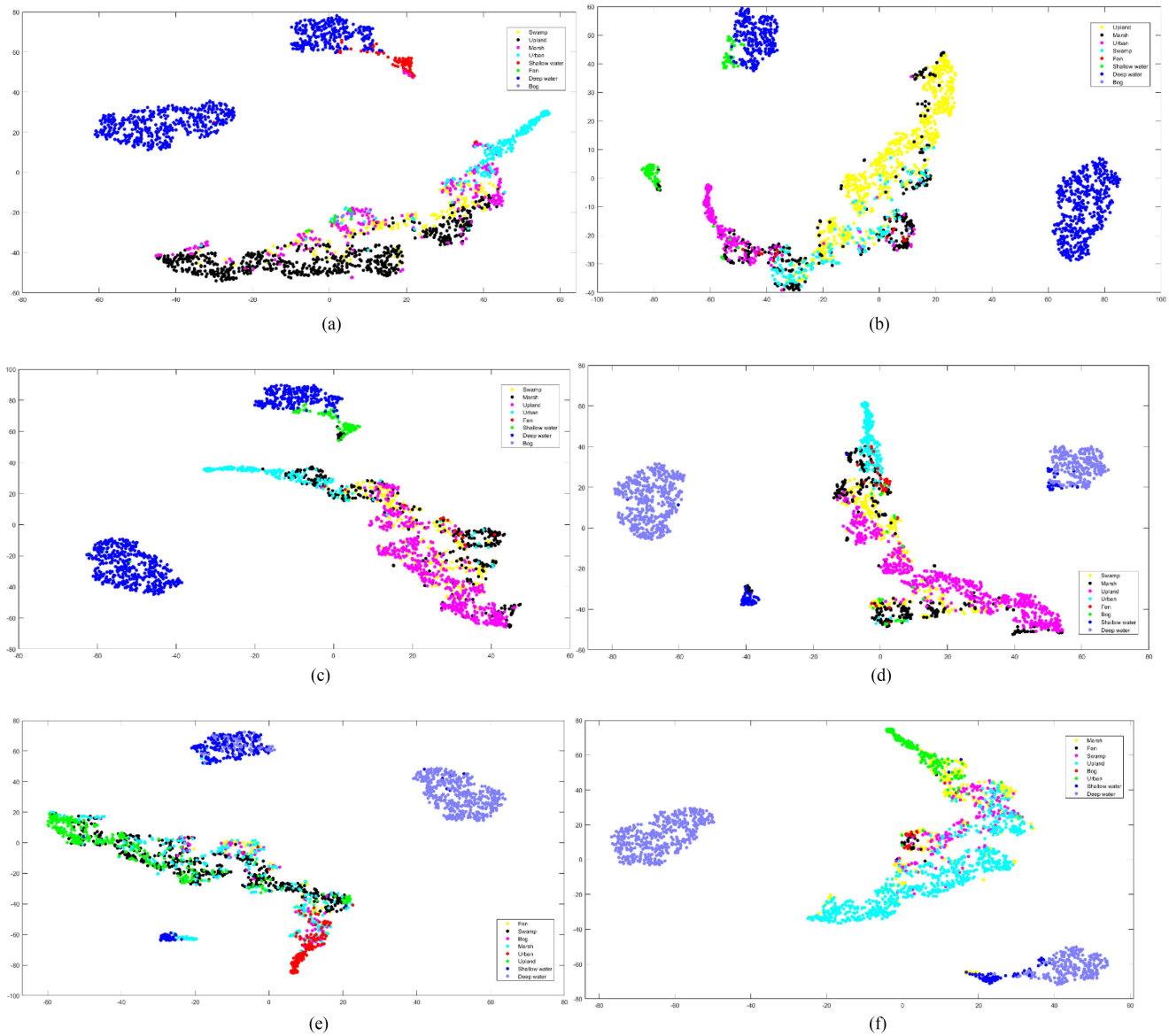


Fig. 7. Feature visualization using the t-SNE algorithm for (a) 3-D hybrid, (b) HybridSN, (c) MLP-mixer, (d) RF, (e) SpectralNet, and (f) the proposed 3-D hybrid GAN.

TABLE III
SPECTRAL BANDS AND INDICES USED IN THIS STUDY

RapidEye bands in micrometers	Band ratios and spectral Indices
Blue (0.44-051)	$RENDVI = \frac{NIR-RE}{NIR+RE}$
Green (0.52-0.6)	$NDVI = \frac{NIR-R}{NIR+R}$
Red (0.63-0.69)	$GNDVI = \frac{NIR-G}{NIR+G}$
Red Edge (0.69-0.73)	
Near Infra-Red (0.76-85)	

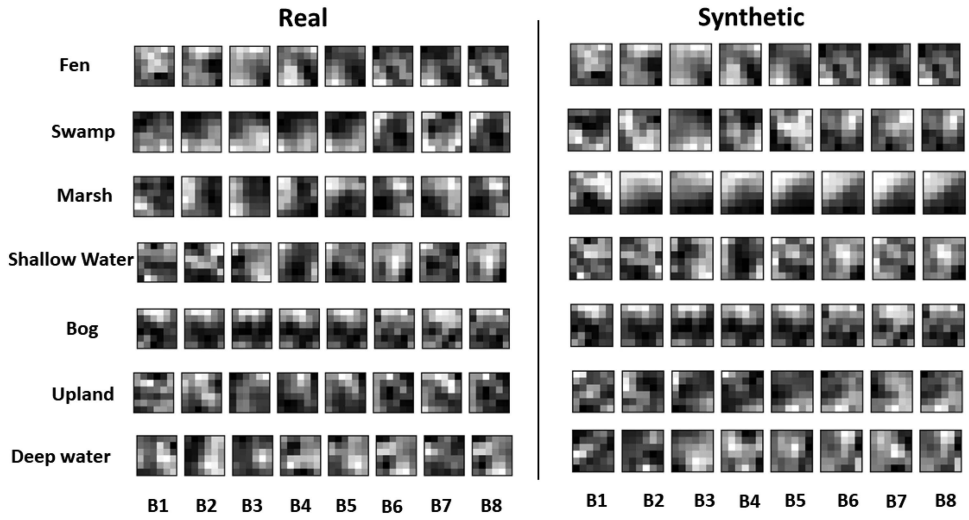


Fig. 8. Example of real and synthetic wetland samples generated by the proposed 3-D GAN network.

C. Evaluation Metrics

To assess the quantitative performance of the developed models, we adopted F1-score, as well as the OA, average accuracy (AA), and Kappa index (KI) as our evaluation metrics (6–9)

$$F1-\text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

$$\text{Overall Accuracy} = \frac{(TP+TN)}{\text{Total number of pixels}} \times 100 \tag{7}$$

$$\text{Average Accuracy} = \frac{\sum_{i=1}^{n} \text{Recall}_i}{n},$$

$$\text{Recall} = \frac{\text{True positive}}{(\text{True positive} + \text{False negative})} \tag{8}$$

$$\text{Kappa} = \frac{p_0 - p_e}{1 - p_e}, p_0 = \frac{\sum x_{ii}}{N}, p_e = \frac{\sum x_{i+} x_{i+}}{N^2} \tag{9}$$

where x_{i+} is the marginal total of row *I*, the total number of observations is shown by *N*, and x_{ii} is observation in row i and column i. TP, FN, and FP are true positive, false negative, and false positive values, respectively.

D. Comparison With Other Deep Convolutional Neural Networks

For the evaluation of the efficiency of the proposed 3-D hybrid GAN model, the wetland classification results were compared with several state-of-the-art algorithms in remote sensing and computer vision, including the HybridSN [61], SpectralNet [63], multi-layer perceptron (MLP)-mixer [64], and RF [65] classifiers.

E. Results

Comparison results on the pilot site of Avalon are shown in Table IV, including the results of the HybridSN, SpectralNet, MLP-mixer, and RF. Table IV shows that the proposed 3-D hybrid GAN significantly outperformed other CNN techniques and the RF classifier in terms of F-1 score for wetland classification by 1%–51%. It should be noted that we fed only the real data to other implemented classifiers, while for the training of the proposed wetland classifier, both real and synthetic data generated by the 3-D GAN generator were utilized. The F-1 score values of swamp, marsh, bog, fen, and shallow water wetland classes obtained by the 3-D hybrid GAN methodology were 0.92, 0.93, 0.94, 0.94, and 0.97, respectively, as shown in Table IV. It is worth highlighting that for a better understanding, the accuracy obtained by the 3-D GAN classifier, with an AA of 87.38%, is also shown in Table IV. By generating high-quality synthetic wetland samples, in terms of F-1 score, the 3-D hybrid GAN method improved classification results of the original 3-D hybrid CNN network for the recognition of bog, shallow water, marsh, swamp, and fen wetlands by 2%, 5%, 16%, 31%, and 38%, respectively.

Table IV shows that although the shallow RF classifier achieved a high OA of 90.242%, there is a low level of agreement between the ground truth data and the predicted class values for wetlands. The reason is that for the RF and other CNNs, we are dealing with the imbalanced data issue where the OA is relatively high, but the final predicted wetlands are not as accurate as nonwetland regions due to the lower number of training data for wetland classes. For example, the RF classifier obtained F-1 scores of 0.45, 0.58, 0.67, 0.88, and 0.89 for the prediction of fen, swamp, marsh, bog, and shallow water wetland classes, respectively, much lower compared to the proposed classification algorithm. In terms of F-1 score, the 3-D hybrid GAN method achieved a much higher agreement between the predicted wetland classes and ground truth data than that of the RF classifier by 6%, 8%, 26%, 34%, and 49% for

TABLE IV
RESULTS OF THE MACHINE LEARNING ALGORITHMS OF THE HYBRIDSN, SPECTRALNET, MLP-MIXER, RF, AND THE PROPOSED 3-D HYBRID GAN IN TERMS OF F1-SCORE, AA, KI, AND OA BASED ON THE TEST DATA

Class	3D Hybrid Classifier (Using only real data, ours)	HybridSN [58]	Spectral Net [60] (Using only real data)	MLP-Mixer [61]	Random Forest [62] (Using only real data)	3D GAN Classifier (ours)	3D Hybrid Classifier (Using synthetic and real data, ours)
Bog	0.92	0.93	0.86	0.92	0.88	0.93	**0.94**
Fen	0.56	0.74	0.27	0.71	0.45	0.68	**0.94**
Marsh	0.77	0.82	0.63	0.81	0.67	0.79	**0.93**
Swamp	0.61	0.69	0.53	0.73	0.58	0.66	**0.92**
Shallow water	0.92	0.95	0.46	0.92	0.89	0.9	**0.97**
Urban	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Deep water	**0.99**	1	0.47	0.98	0.97	0.97	**0.99**
Upland	0.96	0.96	0.88	**0.97**	0.95	0.95	0.94
KI (%)	91.18	93.51	68.05	92.43	88.25	90.91	**94.91**
OA (%)	92.67	94.61	72.71	93.68	90.24	92.39	**95.55**
AA (%)	82.76	87.97	68.55	88.45	79.02	87.38	**95.49**

Best results are presented in bold.

the recognition of bog, shallow water, marsh, swamp, and fen wetlands, respectively. In addition, although the implemented CNN models similar to the RF classifier obtained a relatively high level of agreement between the ground truth nonwetland values and their predicted classes, they demonstrated relatively much lower per-class classification accuracy for the prediction of wetlands in terms of F-1 score, as seen by Table IV.

It should also be noticed that the MLP-mixer algorithm, with an average accuracy of 88.45%, outperformed other algorithms of RF (79.02%), HybridSN (87.97%), and SpectralNet (68.55%), demonstrating the better performance of state-of-the-art computer vision algorithms over the current remote sensing techniques.

F. Confusion Matrices

To further investigate which wetland and nonwetland classes are more confused with each other, the confusion matrices of testing results on the Avalon pilot site are presented in Fig. 5. As seen in Fig. 5, it is clear that there is a high level of confusion between wetland classes predicted by the CNN models of HybridSN, SpectralNet, and MLP-mixer, as well as the shallow classifier of RF. Nevertheless, the wetland classes are far less distinguishable by the RF algorithm and other CNN models compared to the proposed methodology. It can be explained as follows. First and foremost, the number of training data for wetland classes compared to the nonwetlands was relatively low. Second, there is a high level of similarity shared between fen, bog, and marsh regions in terms of vegetation structure and pattern (i.e., saturated vegetation, wet soils, and some emergent vegetation), besides the similarity between the upland forest and swamp in terms of tree dominance, resulting in similar spectral reflectance in the RapidEye optical imagery. We can see from Fig. 5 that the general results obtained by the CNN models are similar; a high level of per-class accuracy for non-wetlands and a much lower level of per-class accuracy for the wetland samples that is resulted due to the intrinsic complexity of wetlands, as well as the limited number of training data. Although generating high-quality synthetic samples for the classes with a minor number of training and test data by the proposed 3-D hybrid GAN model, the confusion between wetland classes substantially decreased (Fig. 5).

G. Classification Maps

To better understand the performance of the shallow algorithm of RF and other CNN models, namely HybridSN, SpectralNet, and MLP-mixer for wetland classification, their predicted wetland maps are presented in Fig. 6. In order to support the presented visual results, Fig. 7 presents the distributions of features for the developed classifiers, as well as the proposed 3-D hybrid GAN model in 2-D space using the t-distributed stochastic neighbor embedding (t-SNE) algorithm [66]. As seen in Fig. 6, almost all shallow and CNN models presented high separability in the recognition of shallow and deep waters. Based on the visual interpretation and results of the t-SNE algorithm, there is a high level of over-classification for recognizing the swamp wetland class by the SpectralNet model. Based on the visual and statistical metrics, the proposed 3-D hybrid GAN model showed the best results compared to the other classifiers (Figs. 6 and 7).

H. Synthetic Wetland Generation

To better understand and evaluate how the generation of high-quality wetland samples can improve the per-class and overall wetland classification accuracy, we included the proposed 3-D hybrid classifier's results with and without synthetic samples, as shown in Table IV. It can be seen that the overall and per-class wetland classification accuracies of the 3-D hybrid classifier were considerably improved with the inclusion of synthetic RapidEye samples. The OA, KI, and AA of the 3-D hybrid classifier improved by 2.88%, 3.73%, and 12.73%, respectively,

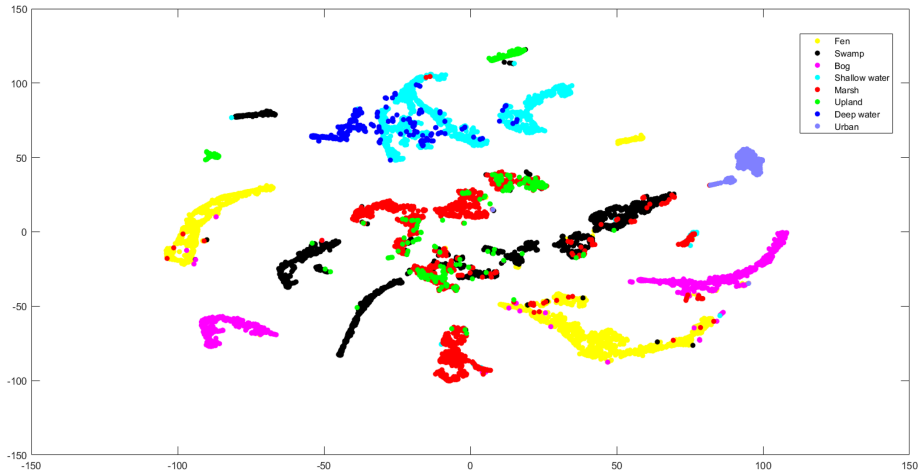


Fig. 9. Visualization of the generated synthetic RapidEye data using the t-SNE algorithm.

with the inclusion of the synthetic RapidEye image data. Moreover, Fig.8 illustrates several synthetic wetland and nonwetland samples generated by the 3-D GAN generator. Additionally, to visually present the quality of synthetic RapidEye samples, we used the t-SNE algorithm, as presented in Fig. 9.

I. Computation Cost

The computation cost of the proposed CNN model of the 3-D hybrid GAN model in terms of time is relatively high compared to the other classifiers. The required time for the RF classifier with 200 trees was about 2 min, achieving the least computation cost in terms of time as it uses only spectral information. The CNN models of SpectralNet, MLP-mixer, and hybridSN required approximately 18, 10, and 5 min, respectively, for their training. On the other hand, the 3-D hybrid GAN model required approximately 120 min for its training, which is much higher than that of the other classifiers. The reason is that unlike other CNN models, the 3-D hybrid GAN model is composed of three main networks: 1) the classifier; 2) the generator; 3) the discriminator. It is worth mentioning that the experiments were done with an Intel processor (i.e., i7-10750H Central Processing Unit (CPU) of 2.60 GHz), a Graphical Processing Unit (GPU) of NVIDIA GeForce RTX 2070, and a 16 GB Random Access Memory (RAM) operating on 64-bit Windows 10.

V. Conclusion

The CNN models require a large number of training samples to fine-tune their many parameters compared to the shallow classifiers, such as RF. Collecting and generating training samples in remote sensing (i.e., ground truth data) is costly, time-consuming, and labor-intensive. To address the issue of limited training samples, we proposed the 3-D hybrid GAN model for the classification of complex wetlands and applied it to a pilot site located in the Avalon Peninsula of Newfoundland, Canada. The proposed CNN model uses a modified architecture of the hybridSN model. It is important to take full advantage of the recently developed state-of-the-art algorithms such as GANs to solve the issues that we are dealing with in remote sensing, including a limited number of high-quality ground truth data for ecological mapping. As such, to tackle the usual imbalance data issue, we utilized the idea of generating synthetic samples only for the classes with a lower number of training data using a conditional mapping unit. Based on the results, the 3-D hybrid GAN with an average accuracy of 95.49% significantly outperformed the RF (79.02%), HybridSN (87.97%), SpectralNet (68.55%), and MLP-mixer (88.45%) classifiers in terms of both average accuracies and per-class F-1 score. Although other CNN algorithms reached a high level of overall accuracy, they demonstrated relatively poor results for the wetlands per-class accuracies. However, with the utilization of a 3-D Hybrid classifier, as well as a 3-D synthetic sample generator, the 3-D hybrid GAN method achieved the F-1 scores of 0.92, 0.93, 0.94, 0.94, and 0.97 for the classification of swamp, marsh, bog, fen, and shallow water wetland regions, respectively; significantly improving the wetland classification accuracies compared to the other CNN techniques, and the RF classifier by approximately 1%–51%. As such, this article presented the power and high capabilities of the current advanced machine learning models for complex ecological classification with the integration of GANs and DCNN models.

References

[1] N. C. Davidson, "Ramsar convention on wetlands: Scope and implementation," in *The Wetland Book I: Structure and Function, Management and Methods*, C. M. Finlayson, M. Everard, K. Irvine, R. J. McInnes, B. A. Middleton, A. A. van Dam, and N. C. Davidson, Eds. Berlin, Germany: Springer, 2016, pp. 451–458, doi: 10.1007/978-90-481-9659-3_113.

[2] J. M. Marton et al., "Geographically isolated wetlands are important biogeochemical reactors on the landscape," *BioScience*, vol. 65, no. 4, pp. 408–418, Apr. 2015, doi: 10.1093/biosci/biv009.

[3] A. Jamali, M. Mahdianpari, B. Brisco, J. Granger, F. Mohammadimanesh, and B. Salehi, "Comparing solo versus ensemble convolutional neural networks for wetland classification using multi-spectral satellite imagery," *Remote Sens.*, vol. 13, no. 11, pp. 2046, 2021, doi: 10.3390/rs13112046.

[4] S. Salimi, S. A. Almuktar, and M. Scholz, "Impact of climate change on wetland ecosystems: A critical review of experimental wetlands," *J. Environ. Manage.*, vol. 286, May 2021, Art. no. 112160, doi: 10.1016/j.jenvman.2021.112160.

[5] B. Slagter, N. E. Tsendbazar, A. Vollrath, and J. Reiche, "Mapping wetland characteristics using temporally dense sentinel-1 and sentinel-2 data: A case study in the St. Lucia Wetlands, South Africa," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 86, 2020, Art. no. 102009.

[6] S. Bhatnagar, L. Gill, S. Regan, S. Waldren, and B. Ghosh, "A nested drone-satellite approach to monitoring the ecological conditions of wetlands," *ISPRS J. Photogrammetry Remote Sens.*, vol. 174, pp. 151–165, Apr. 2021, doi: 10.1016/j.isprsjprs.2021.01.012.

[7] E. R. DeLancey, J. F. Simms, M. Mahdianpari, B. Brisco, C. Mahoney, and J. Kariyeva, "Comparing deep learning and shallow learning for large-scale wetland classification in Alberta, Canada," *Remote Sens.*, vol. 12, no. 1, pp. 2, 2020, doi: 10.3390/rs12010002.

[8] A. Ghorbanian, M. Kakooei, M. Amani, S. Mahdavi, A. Mohammadzadeh, and M. Hasanlou, "Improved land cover map of Iran using sentinel imagery within google earth engine and a novel automatic workflow for land cover classification using migrated training samples," *ISPRS J. Photogrammetry Remote Sens.*, vol. 167, pp. 276–288, Sep. 2020, doi: 10.1016/j.isprsjprs.2020.07.013.

[9] M. Mahdianpari, B. Salehi, M. Rezaee, F. Mohammadimanesh, and Y. Zhang, "Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery," *Remote Sens.*, vol. 10, no. 7, pp. 1119, 2018.

[10] M. Merchant, C. Haas, J. Schroder, R. Warren, and R. Edwards, "High-latitude wetland mapping using multidate and multisensor Earth observation data: A case study in the Northwest Territories," *J. Appl. Remote Sens.*, vol. 14, no. 3, Aug. 2020, Art. no. 034511, doi: 10.1117/1.JRS.14.034511.

[11] M. Rezaee, M. Mahdianpari, Y. Zhang, and B. Salehi, "Deep convolutional neural network for complex wetland classification using optical remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3030–3039, Sep. 2018.

[12] S. Kentsch et al., "Analysis of UAV-Acquired wetland orthomosaics using GIS, computer vision, computational topology and deep learning," *Sensors*, vol. 21, no. 2, pp. 471, 2021, doi: 10.3390/s21020471.

[13] G. L. O'Neil, J. L. Goodall, M. Behl, and L. Saby, "Deep learning using physically-informed input data for wetland identification," *Environ. Model. Softw.*, vol. 126, Apr. 2020, Art. no. 104665, doi: 10.1016/j.envsoft.2020.104665.

[14] M. Pashaei, H. Kamangir, M. J. Starek, and P. Tissot, "Review and evaluation of deep learning architectures for efficient land cover mapping with UAS hyper-spatial imagery: A case study over a wetland," *Remote Sens.*, vol. 12, no. 6, pp. 959, 2020, doi: 10.3390/rs12060959.

[15] A. Jamali and M. Mahdianpari, "Swin transformer and deep convolutional neural networks for coastal wetland classification using sentinel-1, sentinel-2, and LiDAR data," *Remote Sens.*, vol. 14, no. 2, pp. 359, 2022, doi: 10.3390/rs14020359.

[16] Z. Huang, Z. Pan, and B. Lei, "Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data," *Remote Sens.*, vol. 9, no. 9, pp. 907, 2017, doi: 10.3390/rs9090907.

[17] J. S. Smith et al., "Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning," *Nature Commun.*, vol. 10, no. 1, pp. 2903, Jul. 2019, doi: 10.1038/s41467-019-10827-4.

[18] J. Wang, S. Guo, R. Huang, L. Li, X. Zhang, and L. Jiao, "Dual-channel capsule generation adversarial network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5501016, doi: 10.1109/TGRS.2020.3044312.

[19] W. Xie, J. Zhang, J. Lei, Y. Li, and X. Jia, "Self-spectral learning with GAN based spectral–spatial target detection for hyperspectral image," *Neural Netw.*, vol. 142, pp. 375–387, Oct. 2021, doi: 10.1016/j.neunet.2021.05.029.

[20] R. Zhang, S. Newsam, Z. Shao, X. Huang, J. Wang, and D. Li, "Multi-scale adversarial network for vehicle detection in UAV imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 180, pp. 283–295, Oct. 2021, doi: 10.1016/j.isprsjprs.2021.08.002.

[21] S. Zhao, S. Yang, J. Gu, Z. Liu, and Z. Feng, "Symmetrical lattice generative adversarial network for remote sensing images compression," *ISPRS J. Photogrammetry Remote Sens.*, vol. 176, pp. 169–181, Jun. 2021, doi: 10.1016/j.isprsjprs.2021.03.009.

[22] M. Rao, P. Tang, and Z. Zhang, "Spatial–spectral relation network for hyperspectral image classification with limited training samples," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5086–5100, Dec. 2019, doi: 10.1109/JSTARS.2019.2957047.

[23] Y. Qu, R. K. Baghbaderani, and H. Qi, "Few-shot hyperspectral image classification through multitask transfer learning," in *Proc. 10th Workshop Hyperspectral Imag. Signal Process.: Evol. Remote Sens.*, 2019, pp. 1–5. doi: 10.1109/WHISPERS.2019.8920992.

[24] B. Liu, X. Yu, A. Yu, P. Zhang, G. Wan, and R. Wang, "Deep few-shot learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2290–2304, Apr. 2019, doi: 10.1109/TGRS.2018.2872830.

[25] Y. Qu, R. K. Baghbaderani, W. Li, L. Gao, Y. Zhang, and H. Qi, "Physically constrained transfer learning through shared abundance space for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10455–10472, Dec. 2021, doi: 10.1109/TGRS.2020.3045790.

[26] X. Ma, X. Mou, J. Wang, X. Liu, H. Wang, and B. Yin, "Cross-data set hyperspectral image classification based on deep domain adaptation," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10164–10174, Dec. 2019, doi: 10.1109/TGRS.2019.2931730.

[27] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, S. Homayouni, and E. Gill, "The first wetland inventory map of newfoundland at a spatial resolution of 10 m using sentinel-1 and sentinel-2 data on the google earth engine cloud computing platform," *Remote Sens.*, vol. 11, no. 1, pp. 43, 2019.

[28] A. Jamali, M. Mahdianpari, B. Brisco, J. Granger, F. Mohammadimanesh, and B. Salehi, "Deep forest classifier for wetland mapping using the combination of sentinel-1 and sentinel-2 data," *GIScience Remote Sens.*, vol. 58, no. 7, pp. 1072–1089, 2021, doi: 10.1080/15481603.2021.1965399.

[29] R. M. Vieira et al., "Land degradation mapping in the MATOPIBA region (Brazil) using remote sensing data and decision-tree analysis," *Sci. Total Environ.*, vol. 782, Aug. 2021, Art. no. 146900, doi: 10.1016/j.scitotenv.2021.146900.

[30] A. Razaque, M. Ben Haj Frej, M. Almi'ani, M. Alotaibi, and B. Alotaibi, "Improved support vector machine enabled radial basis function and linear variants for remote sensing image classification," *Sensors*, vol. 21, no. 13, pp. 4431, 2021, doi: 10.3390/s21134431.

[31] H. Ebrahimy, B. Mirbagheri, A. A. Matkan, and M. Azadbakht, "Per-pixel land cover accuracy prediction: A random forest-based method with limited reference sample data," *ISPRS J. Photogrammetry Remote Sens.*, vol. 172, pp. 17–27, Feb. 2021, doi: 10.1016/j.isprsjprs.2020.11.024.

[32] J. Magidi, L. Nhamo, S. Mpandeli, and T. Mabhaudhi, "Application of the random forest classifier to map irrigated areas using google earth engine," *Remote Sens.*, vol. 13, no. 5, pp. 876, 2021, doi: 10.3390/rs13050876.

[33] S. S. Paul, J. Li, Y. Li, and L. Shen, "Assessing land use–land cover change and soil erosion potential using a combined approach through remote sensing, RUSLE and random forest algorithm," *Geocarto Int.*, vol. 36, no. 4, pp. 361–375, 2021, doi: 10.1080/10106049.2019.1614099.

[34] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, "Classification of remote sensing images using efficientnet-B3 CNN model with attention," *IEEE Access*, vol. 9, pp. 14078–14094, 2021, doi: 10.1109/ACCESS.2021.3051085.

[35] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogrammetry Remote Sens.*, vol. 173, pp. 24–49, Mar. 2021, doi: 10.1016/j.isprsjprs.2020.12.010.

[36] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639–3655, Jul. 2017, doi: 10.1109/TGRS.2016.2636241.

[37] A. Jamali, M. Mahdianpari, F. Mohammadimanesh, A. Bhattacharya, and S. Homayouni, "PolSAR image classification based on deep convolutional neural networks using wavelet transformation," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4510105, doi: 10.1109/LGRS.2022.3185118.

[38] S. Bera and V. K. Shrivastava, "Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification," *Int. J. Remote Sens.*, vol. 41, no. 7, pp. 2664–2683, Apr. 2020, doi: 10.1080/01431161.2019.1694725.

[39] C. Shi, T. Wang, and L. Wang, "Branch feature fusion convolution network for remote sensing scene classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5194–5210, 2020, doi: 10.1109/JSTARS.2020.3018307.

[40] J. Liang, Y. Deng, and D. Zeng, "A deep neural network combined CNN and GCN for remote sensing scene classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4325–4338, 2020, doi: 10.1109/JSTARS.2020.3011333.

[41] B. Petrovska, E. Zdravevski, P. Lameski, R. Corizzo, I. Štajduhar, and J. Lerga, "Deep learning for feature extraction in remote sensing: A case-study of aerial scene classification," *Sensors*, vol. 20, no. 14, pp. 3906, 2020, doi: 10.3390/s20143906.
[42] X. Yang, Y. Ye, X. Li, R. Y. K. Lau, X. Zhang, and X. Huang, "Hyperspectral image classification with deep learning models," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5408–5423, Sep. 2018, doi: 10.1109/TGRS.2018.2815613.
[43] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral–Spatial residual network for hyperspectral image classification: A 3-D deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, Feb. 2018, doi: 10.1109/TGRS.2017.2755542.
[44] X. Pan and J. Zhao, "A central-point-enhanced convolutional neural network for high-resolution remote-sensing image classification," *Int. J. Remote Sens.*, vol. 38, no. 23, pp. 6554–6581, 2017.
[45] E. Liu, Y. Zheng, B. Pan, X. Xu, and Z. Shi, "DCL-Net: Augmenting the capability of classification and localization for remote sensing object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7933–7944, Sep. 2021, doi: 10.1109/TGRS.2020.3048384.
[46] G. Cheng, Y. Si, H. Hong, X. Yao, and L. Guo, "Cross-scale feature fusion for object detection in optical remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 431–435, Mar. 2021, doi: 10.1109/LGRS.2020.2975541.
[47] K. Li, G. Wan, G. Cheng, L. Meng, and J. Han, "Object detection in optical remote sensing images: A survey and a new benchmark," *ISPRS J. Photogrammetry Remote Sens.*, vol. 159, pp. 296–307, Jan. 2020, doi: 10.1016/j.isprsjprs.2019.11.023.
[48] X. Sun, P. Wang, C. Wang, Y. Liu, and K. Fu, "PBNet: Part-based convolutional neural network for complex composite object detection in remote sensing imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 173, pp. 50–65, Mar. 2021, doi: 10.1016/j.isprsjprs.2020.12.015.
[49] X. Zhang, L. Han, L. Han, and L. Zhu, "How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery?," *Remote Sens.*, vol. 12, no. 3, pp. 417, 2020, doi: 10.3390/rs12030417.
[50] D. Hong et al., "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021, doi: 10.1109/TGRS.2020.3016820.
[51] G. Cheng, X. Xie, J. Han, L. Guo, and G. -S. Xia, "Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3735–3756, 2020, doi: 10.1109/JSTARS.2020.3005403.
[52] A. Ma, Y. Wan, Y. Zhong, J. Wang, and L. Zhang, "SceneNet: Remote sensing scene classification deep learning network using multi-objective neural evolution architecture search," *ISPRS J. Photogrammetry Remote Sens.*, vol. 172, pp. 171–188, Feb. 2021, doi: 10.1016/j.isprsjprs.2020.11.025.
[53] P. Dou, H. Shen, Z. Li, X. Guan, and W. Huang, "Remote sensing image classification using deep–shallow learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3070–3083, 2021, doi: 10.1109/JSTARS.2021.3062635.
[54] Y. Li, Y. Zhang, and Z. Zhu, "Error-tolerant deep learning for remote sensing image scene classification," *IEEE Trans. Cybern.*, vol. 51, no. 4, pp. 1756–1768, Apr. 2021, doi: 10.1109/TCYB.2020.2989241.
[55] A. Jamali and M. Mahdianpari, "A cloud-based framework for large-scale monitoring of ocean plastics using multi-spectral satellite imagery and generative adversarial network," *Water*, vol. 13, no. 18, pp. 2553, 2021, doi: 10.3390/w13182553.
[56] I. Goodfellow et al., "Generative adversarial nets," in *Adv. Neural Inf. Process. Syst.*, vol. 27, pp. 2672–2680, 2014.
[57] H. Zhang, Y. Song, C. Han, and L. Zhang, "Remote sensing image spatiotemporal fusion using a generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4273–4286, May 2021, doi: 10.1109/TGRS.2020.3010530.
[58] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Generative adversarial networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5046–5063, Sep. 2018, doi: 10.1109/TGRS.2018.2805286.
[59] S. Ji, D. Wang, and M. Luo, "Generative adversarial network-based full-space domain adaptation for land cover classification from multiple-source remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3816–3828, May 2021, doi: 10.1109/TGRS.2020.3020804.
[60] S. K. Roy, J. M. Haut, M. E. Paoletti, S. R. Dubey, and A. Plaza, "Generative adversarial minority oversampling for spectral-spatial hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 7, 2022, Art. no. 5500615, doi: 10.1109/TGRS.2021.3052048.
[61] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 277–281, Feb. 2020, doi: 10.1109/LGRS.2019.2918719.
[62] S. Subhra Mullick, S. Datta, and S. Das, "Generative adversarial minority oversampling," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 1695–1704.
[63] T. Chakraborty and U. Trehan, "SpectralNET: Exploring spatial-spectral WaveletCNN for hyperspectral image classification." arXiv.org, 2021. [Online]. Available: https://arxiv.org/abs/2104.00341
[64] I. O. Tolstikhin et al., "MLP-mixer: An all-MLP architecture for vision," *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 24261–24272, 2021, [Online]. Available: https://proceedings.neurips.cc/paper/2021/file/cba0a4ee5ccd02fda0fe3f9a3e7b89fe-Paper.pdf
[65] L. Breiman, "Random forests," *Mach. Learn.*, vol. 54, no. 1, pp. 5–32, 2001.
[66] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.

Ali Jamali received the doctor of philosophy (Ph.D.) degree in geoinformatics from Universiti Teknologi Malaysia (UTM), Iskandar Puteri, Malaysia, in 2017.

He is an experienced researcher with a demonstrated history of working in higher education, skilled in statistics, mathematics, GIS, remote sensing, image processing, machine learning, and algorithm optimization. He has authored or coauthored more than 30 papers in Geoinformatics and remote sensing fields. His current research interest includes remote sensing image processing based on advanced mathematical and statistical algorithms.

Masoud Mahdianpari received the B.S. degree in surveying and geomatics engineering and the M.Sc. degree in remote sensing engineering from the School of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran, in 2010 and 2013, respectively, and the Ph.D. degree in electrical engineering from the Department of Engineering and Applied Science, Memorial University, St. John's, NL, Canada, in 2019.

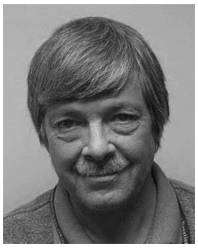
In 2019, he was an Ocean Frontier Institute (OFI) Postdoctoral Fellow with Memorial University and C-CORE. He is currently a Remote Sensing Technical Lead with C-CORE and a Cross Appointed Professor with the Faculty of Engineering and Applied Science, Memorial University. He is the author or coauthor of more than 100 publications, including peer-reviewed journal articles, conference papers, books, and book chapters. His research interests include remote sensing and image analysis, with a special focus on PolSAR image processing, multisensor data classification, machine learning, geo big data, and deep learning.

Dr. Mahdianpari is also an Editorial Team Member for the *Remote Sensing Journal* and the *Canadian Journal of Remote Sensing*. He was a recipient of the Research and Development Corporation Ocean Industries Student Research Award, organized by Newfoundland Industry and Innovation Center, amongst more than 400 submissions, in 2016, the T. David Collett Best Industry Paper Award organized by IEEE in 2016, the Com-Adv Devices, Inc., Scholarship for Innovation, Creativity, and Entrepreneurship from the Memorial University in 2017, the Artificial Intelligence for Earth Grant organized by Microsoft in 2018, and the Graduate Academic Excellence Award organized by Memorial University in 2019. In 2020, he received the Best Industry Paper Award organized by IEEE Newfoundland Electrical and Computer Engineering Conference (NECEC).

Fariba Mohammadimanesh received the B.Sc. degree in surveying and geomatics engineering and the M.Sc. degree in remote sensing engineering from the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran, in 2010 and 2014, respectively, and the Ph.D. degeree in electrical engineering from the Department of Engineering and Applied Science, Memorial University, St. John's, NL, Canada, in 2019.

She is currently a Remote Sensing Scientist with C-CORE's Earth Observation Team, St. John's, NL, Canada. Her research interests include image processing, machine learning, segmentation, and classification of satellite remote sensing data, including SAR, PolSAR, and optical, in environmental studies, with a special interest in wetland mapping and monitoring, as well as geo-hazard mapping using the interferometric SAR (InSAR) technique.

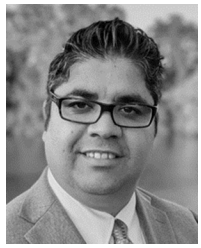
Dr. Mohammadimanesh was the recipient of several awards, including the EmeraGraduate Scholarship (2016–2019), CIG-NLAward (2018), and the IEEE NECEC Best Industry Paper Award (2016 and 2020).



Brian Brisco received the B.Sc. degree in ecology and the M.Sc. degree in soil science from the University of Guelph, Guelph, ON, Canada, and the Ph.D. degree in remote sensing/physical geography from the University of Kansas, Lawrence, KS, USA, in 1985.

He is an internationally recognized authority on Synthetic Aperture Radar (SAR) and its application to a wide range of environmental monitoring applications. He has been involved in remote sensing since 1975 and participated in the SURSAT project from 1977–1980 before spending four years with the Remote Sensing Laboratory, University of Kansas under the supervision of Dr. F.T. Ulaby, widely recognized as one of the world's leading authorities on radar. He worked for Intera from 1989 until 1997 as a Research Associate after completion of an NSERC Postdoctoral Fellowship served with the Canada Centre for Remote Sensing. Since 1997 he has worked for Noetix Research Inc. where he is the Director of Research and Applications Development. His work has included experience with interferometry, polarimetry, and radar backscatter modeling including software development and operational implementation. He provides peer review services to all the major remote sensing journals and participates as an external examiner for many graduate students at various universities in Canada and abroad. Brian has been consulted or contracted by government and nongovernment organizations on a wide range of SAR applications and system development including NRCan, CSA, DND, AAFC, EC, NASA, ESA, NASDA, NOAA, USDA, CAS, etc. He has extensive contacts in the SAR community worldwide and has worked in China, Vietnam, Malaysia, Thailand, Indonesia, South Africa, Argentina, Uruguay, Chile, Brazil, Columbia, and Costa Rica. Brian has authored or coauthored more than 200 publications including over 50 peer reviewed journal publications and is the author of two chapters in the Manual of Remote Sensing volume on radar applications published by ASPRS. His research interests include using remote sensing, particularly SAR, for mapping and managing renewable resources.

Dr. Brisco is a Past-President of the Canadian Remote Sensing Society (CRSS) and the Canadian Aeronautics and Space Institute (CASI). His skills contribute to ESS Programs on Remote Sensing Science, Earth Science for Environment and Health, Enhancing Resilience in a Changing Climate and Northern Resources and Development.



Bahram Salehi (Senior Member, IEEE) received the B.Sc. degree in 2002 and the M.Sc. degree in 2005 in geomatics engineering from K. N. Toosi University of Technology, Tehran, Iran, and University of Tehran, Tehran, Iran, respectively, and the Ph.D. degree in remote sensing engineering from the Department of Geomatics Engineering, University of New Brunswick, Canada in 2012.

He is an Assistant Professor of Remote Sensing Engineering with the State University of New York College of Environmental Science and Forestry (SUNY-ESF), Syracuse, NY, USA. He has over 20 years of industrial and academic experiences in applications and algorithm development in Remote Sensing and its environmental applications. His research has resulted in over 50 journal publications and book chapters. At SUNY ESF, his research focus areas are remote sensing of wetlands, forest biomass, and water quality using Optical, RADAR, and UAV Sensors.

Dr Salehi has received the 2019 Early Career Medal for his achievements and services in remote sensing from the Canadian Remote Sensing Society. He is a Professional Engineer (P.Eng.), and a licensed project management professional (PMP). He is a member of ASPRS, AGU, and CRSS.