





Article

Leveraging Deep Neural Networks to Map Caribou Lichen in High-Resolution Satellite Images Based on a Small-Scale, Noisy UAV-Derived Map

Shahab Jozdani ¹, Dongmei Chen ¹, Wenjun Chen ^{2,*}, Sylvain G. Leblanc ², Christian Prévost ², Julie Lovitt ², Liming He ² and Brian A. Johnson ³

¹ Department of Geography and Planning, Queen's University, Kingston, ON K7L 3N6, Canada; shahab.jozdani@queensu.ca (S.J.); chendm@queensu.ca (D.C.)

² Canada Centre for Remote Sensing, Natural Resources Canada, 560 Rochester Street, Ottawa, ON K1S 5K2, Canada; sylvain.leblanc@canada.ca (S.G.L.); christian.prevost@canada.ca (C.P.); julie.lovitt@canada.ca (J.L.); liming.he@canada.ca (L.H.)

³ Natural Resources and Ecosystem Services Area, Institute for Global Environmental Strategies, 2108-1 Kamiyamaguchi, Hayama, Kanagawa 240-0115, Japan; johnson@iges.or.jp

* Correspondence: wenjun.chen@canada.ca



Citation: Jozdani, S.; Chen, D.; Chen, W.; Leblanc, S.G.; Prévost, C.; Lovitt, J.; He, L.; Johnson, B.A. Leveraging Deep Neural Networks to Map Caribou Lichen in High-Resolution Satellite Images Based on a Small-Scale, Noisy UAV-Derived Map. *Remote Sens.* **2021**, *13*, 2658. <https://doi.org/10.3390/rs13142658>

Academic Editor: Izaya Numata

Received: 8 May 2021

Accepted: 2 July 2021

Published: 6 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Lichen is an important food source for caribou in Canada. Lichen mapping using remote sensing (RS) images could be a challenging task, however, as lichens generally appear in unevenly distributed, small patches, and could resemble surficial features. Moreover, collecting lichen labeled data (reference data) is expensive, which restricts the application of many robust supervised classification models that generally demand a large quantity of labeled data. The goal of this study was to investigate the potential of using a very-high-spatial resolution (1-cm) lichen map of a small sample site (e.g., generated based on a single UAV scene and using field data) to train a subsequent classifier to map caribou lichen over a much larger area (~0.04 km² vs. ~195 km²) and a lower spatial resolution image (in this case, a 50-cm WorldView-2 image). The limited labeled data from the sample site were also partially noisy due to spatial and temporal mismatching issues. For this, we deployed a recently proposed *Teacher-Student* semi-supervised learning (SSL) approach (based on U-Net and U-Net++ networks) involving unlabeled data to assist with improving the model performance. Our experiments showed that it was possible to scale-up the UAV-derived lichen map to the WorldView-2 scale with reasonable accuracy (overall accuracy of 85.28% and F1-score of 84.38%) without collecting any samples directly in the WorldView-2 scene. We also found that our noisy labels were partially beneficial to the SSL robustness because they improved the false positive rate compared to the use of a cleaner training set directly collected within the same area in the WorldView-2 image. As a result, this research opens new insights into how current very high-resolution, small-scale caribou lichen maps can be used for generating more accurate large-scale caribou lichen maps from high-resolution satellite imagery.

Keywords: remote sensing; lichen mapping; deep learning; semi-supervised learning; teacher-student learning; WorldView-2; unmanned aerial vehicle



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Global caribou herd populations have been reportedly declining in recent decades because of multiple factors ranging from climate change to hunting and disease, according to “International Union for the Conservation of Nature” [1]. Although the main driver of this population decline is still ambiguous, most factors thought to be contributing are directly or indirectly related to human activities [2,3], such as land-cover changes that can affect resource availability for caribou [4] and can cause them to change their distribution, migration, and timing patterns when foraging for food [5]. Lichen is an important source of food for caribou especially during winters [6,7].

Since the base data for caribou lichen mapping is generally UAV images covering small areas, one of the primary challenges is to extend the maps generated based on UAV images to much larger areas. Therefore, along with the type of data used, the method for lichen mapping is also of crucial importance to improve the results. This is due to the fact that the limited amount of data provided by UAV images may not help generalize results to larger scales. Given recent advances in the field of artificial intelligence (AI), this, therefore, demands for deploying advanced algorithmic approaches (such as semi-supervised learning (SSL)) that can partially compensate for the lack of sufficient labeled data.

Early research on lichen cover mapping using RS data was based on non-machine-learning (ML) approaches. In one of the earliest studies in this field, Petzold and Goward [8] found that Normalized Difference Vegetation Index (NDVI) may not be a reliable measure for lichen cover estimation (especially where the surface is completely lichen covered) as the values of this index may be misinterpreted as sparse green vascular plants. The Normalized Difference Lichen Index (NDLI) proposed by Nordberg [9] is an index specifically for lichen detection that is mainly used in conjunction with other vegetation indices. In a more recent study, Théau and Duguay [10] improved the mapping of lichen abundance using Landsat TM data by employing spectral mixing analysis (SMA). The authors reported that the SMA overestimated (underestimated) lichen fractional coverage over sites with low (high) lichen presence.

Orthogonal to the aforementioned studies, in 2011, Gilichinsky et al. [11] tested three classifiers for lichen mapping using SPOT-5 and Landsat ETM+ images. They found that *Mahalanobis* distance classifier generated the highest accuracy (84.3%) for mapping lichen based on SPOT-5 images in their study. On the other hand, it turned out that a maximum likelihood classifier reached an overall accuracy (OA) of 76.8% on Landsat ETM+ imagery. However, the authors reported that the model did not perform well over areas where lichen cover was <50%. Later in 2014, Falldorf et al. [12] developed a lichen volume estimator (LVE) based on Landsat TM images using a 2D Gaussian regression model. To train the model, the authors used the NDLI and Normalized Difference Moisture Index (NDMI) as the independent variables, and ground-based volumetric measurements of lichen cover as the dependent variable. Using a 10-fold cross-validation, the authors reported that the model achieved an R^2 of 0.67. In 2020, a study by Macander et al. [13] employed different data sources, including plot data, UAV images, and airborne data, to map the fraction of lichen cover in Landsat imagery using a Random Forest (RF) regressor. However, the authors reported that they observed an overestimation for low lichen cover sites, and an underestimation for high lichen cover sites.

Although there have recently been seminal studies on lichen mapping using powerful ML models (e.g., deep neural networks (DNNs) [14]), there are some gaps that have not yet been approached in this field. An important issue in this field is the lack of sufficient labeled, ground-truth lichen samples. As with any other environmental RS studies, lichen mapping suffers from “the curse of insufficient training data”. It is now clear that insufficient training data deteriorates the generalization power of a classifier, especially DNNs which can easily memorize training data [15].

In case of limited labeled samples, a workaround is to use unlabeled samples that could be available in large volumes. In fact, learning from unlabeled samples could implicitly help improve the generalization performance of the model trained on limited labeled samples. There are a wide variety of methods for involving unlabeled data in a classification setup. Perhaps, one of the most commonly used approaches is unsupervised pre-training of a network and reusing its intermediate layers for training another network on the labeled set of interest [16]. If unlabeled samples are not abundant, another method could be based on transfer learning [17]. This approach is based on using a network trained on a large data set in another domain (e.g., natural images) or in the same domain but for a different application (e.g., urban mapping). This network can be then either fine-tuned on the limited labeled data of interest or used as a feature extractor for another classifier or

other downstream tasks. The potential of transfer learning has also been proven in the RS community [18].

Another class of methods that can take advantage of unlabeled data is semi-supervised learning (SSL). Some SSL methods, known as wrapper methods [19], generally explicitly incorporate unlabeled samples into the training process by automatically labelling them to improve the performance of the model [20]. The concept of *Teacher-Student* SSL is an example of such a method that has shown to improve the generalization when limited labeled data are available [20,21]. In this regard, *Naïve-Student* [22] has been recently proposed as a potentially versatile SSL approach for the pixel-wise classification of street-level urban video sequences. This approach integrates SSL and self-training iteratively through a *Teacher* and *Student* network, where the *Student* network can be the same as or larger than the *Teacher* network. In contrast to a similar approach known as *Noisy-Student* proposed earlier [21], *Naïve-Student* does not inject data-based noise (i.e., data augmentation) directly to the training process to improve the robustness of the *Student* network. Instead, it uses test-time augmentation (TTA) functions when predicting pseudo-labels (i.e., labeling unlabeled samples) on which later the *Student* network is trained. The premise of such an approach is to train a *Student* network that outperforms the *Teacher* network. Besides, the iterative nature of both *Noisy-Student* and *Naïve-Student* approaches helps improve the *Student* network further.

In environmental studies, there are cases where current maps were generated based on very high-spatial resolution aerial images covering small areas, but later the aim will be to reuse those maps to generate much larger-scale maps using another image source without conducting any further labeling. Scaling-up local maps is restricted by many factors (e.g., spatial/temporal misalignments, classification errors, etc.) that may prevent generating large-scale maps with appropriate accuracy. The accuracy of such a scale-up is further degraded in cases where the current high-resolution maps cover a very small, homogeneous area. This is due to differences in soil/substrate or climate conditions that could lead to variability in the spectral signature of the vegetation of interest [23,24]. Although multi-scale/sensor mapping in the field of RS is a common application, it has been mostly performed either using conventional ML approaches or using data-fusion frameworks [25–27]. In this study, we approached this problem in a different way: To what degree is it possible to scale-up very-high-resolution lichen maps produced over small, homogenous areas to a much larger scale without labeling any new samples?

Taking into account the above question in this research, we aimed to scale-up a very high-resolution lichen map, generated from a small UAV scene, to a high-resolution satellite level, namely WorldView-2 (hereafter, WorldView). This scaling-up task posed several challenges. First, training a fully supervised approach in such a case would result in poor accuracy because training samples were not sufficient and diverse enough. Second, we expected the classifier performance to be negatively impacted by errors associated with image misalignments. These errors, caused by various spatial and temporal offsets, can lead to undesirable errors in the model training process. These sources of error caused noisy labels (i.e., wrong labels) in the WorldView image when the UAV-derived lichen map was scaled-up. We, therefore, investigated whether a small, noisy UAV-derived lichen map could be used to train a highly generalized model capable of producing accurate results over a much larger area. To improve classification performance, we adapted *Naïve-Student* [22] to the lichen mapping task in this study to leverage unlabeled data in the training process.

2. Methodology

2.1. Study Area and Data

The study area (~405 km²) is located south of the Manicouagan Reservoir, Québec, Canada, around the Manic-5 dam (Figure 1). In undisturbed portions, this area is comprised primarily of dense boreal forests (black spruce dominated) and open lichen woodlands, or forests with lichen. This region features diverse vegetation species including various

types of bright caribou ground forage lichens within the *Cladonia* genus (predominantly *C. stellaris*, *C. rangiferina*, *C. stygia*, and *C. uncialis*) and *Stereocaulon* genus.

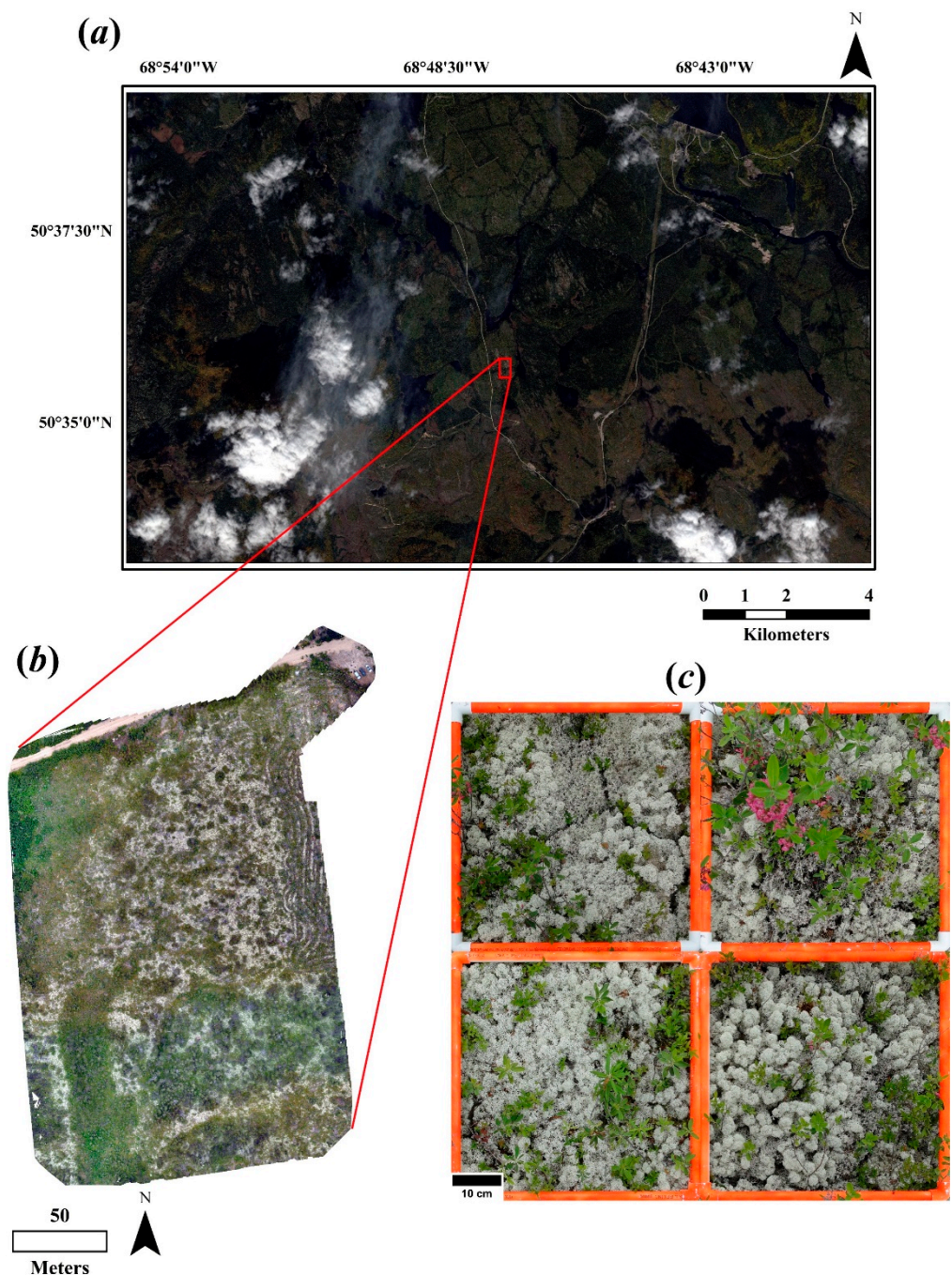


Figure 1. (a) southern part of the WorldView scene (used for model training), (b) the UAV scene, and (c) four examples of the micro-plots collected within the UAV survey extent.

We had three main sources of image data to accurately map lichen cover over the study area:

- In-situ digital photographs of ground conditions within vegetation survey micro-plots (Figure 1c)
- Very-high-resolution aerial data: UAV image with a spatial resolution of 1 cm (Figure 1b)
- High-resolution satellite data: Pansharpened WorldView image with a spatial resolution of 50 cm (Figure 1a)

The field data and aerial image used in this study were part of a field campaign conducted between 22 July and 2 August 2019, in Québec and Labrador, Canada. The very-high-resolution aerial scene was acquired using a Sentra Double 4K NDRE camera onboard a DJI Inspire-1 UAV at a spatial resolution of 1 cm. The UAV was flown at 35 m above ground level at 17 km/h. The camera onboard was set to acquire one frame per second. To improve the spatial accuracy of the final mosaicked UAV scene, we distributed three ground control points in visible areas across the site and recorded their center points using a high-precision Global Navigation Satellite System (GNSS). The image originally had four spectral bands, namely RGB and NIR. We used Pix4D software to produce a georeferenced orthomosaic over the UAV surveyed site. The final mosaicked UAV scene covered approximately 0.04 km² (4 ha). Along with the UAV survey, we collected digital photographs of 11 vegetation survey micro-plots on 24 July, 2019. The second RS image source was an 8-band WorldView scene acquired on 9 September 2017 covering the whole study area. To improve the spatial resolution of the multi-spectral bands, we pansharpened the WorldView image to the spatial resolution of the panchromatic band (50 cm) using the proprietary pansharpening algorithm in Geomatica Banff software. The WorldView image was divided into two parts, namely a Northern part and a Southern part (Figure 1a). The Southern part contained the UAV survey location and was used exclusively for model training purposes. We reserved the Northern part for model testing. To mitigate the effect of spatial misalignments between the datasets, we co-registered the UAV and WorldView images using an Affine model with 14 manually identified tie points. This resulted in an RMSE of 0.68 m.

2.2. Classification Framework

A high-level overview of the scaling-up framework used in this study is presented in Figure 2. We started the procedure by classifying the UAV image based on a labeled set collected using in-situ plot images and the UAV image itself. The quality of the scale-up was highly dependent on the UAV map.

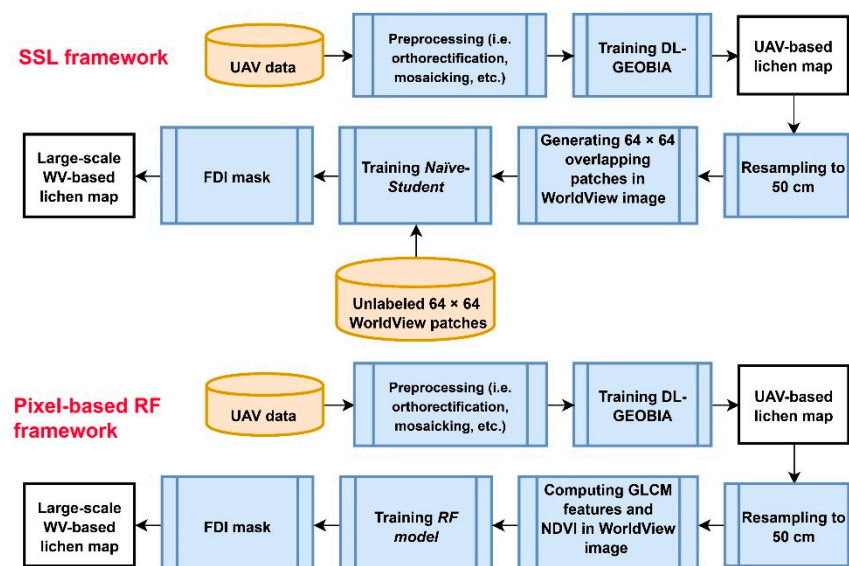


Figure 2. Overall scaling-up procedure employed in this study.

We first used an integration of deep learning and geographic object-based image analysis (DL-GEOBIA) approach [28] to detect caribou lichen cover within the UAV scene. The first step in this approach was image segmentation. We used the multi-resolution image segmentation (MRS) algorithm [29,30] with a Scale parameter (SP) of 30 and a neutral value (i.e., 0.5) for the Compactness and Smoothness parameters, which were selected by performing a trial-and-error process. The SP was chosen in such a way that it

would generate pure but not very over-segmented image objects, preventing the extraction of useful features. Afterwards, 1825 point samples were randomly selected in the UAV image. It was made sure that none of the point samples fell inside the same segments. For training, 80% (1460) of the segments inside which the corresponding point samples fell were selected. Of these 1460 segments, 20% (292) were used as a validation set to fine-tune the model. To mitigate bias/optimism in our accuracy assessment, we performed a pixel-based accuracy assessment rather than an object-based one, and thus chose the corresponding pixels of the remaining 20% points (365) as a test set. To identify the true labels of the image segments/pixels, the micro-plot images and direct image interpretation of the UAV image were used. Based on an exhaustive cross-validation and grid-search, we then concluded that a four-layer multi-layer perceptron (MLP) (Table 1) in the DL-GEOBIA setup with 30 neurons in each hidden layer, ReLU activation for all layers (except for the output layer with a sigmoid function), Adam optimizer with a learning rate of 0.001, and dropout layers with a probability of 0.5 resulted in the best accuracy.

Table 1. Details of the networks used in this study.

Model	Learning Rate	Optimizer	# of Neurons/Filters	# of Trainable Parameters
DL-GEOBIA	0.001	Adam	30 (in each layer)	3691
U-Net, U-Net++	Training: 0.001 Fine-tuning: 0.0001	Adam	32-64-128-256-512	U-Net: 7,767,425 U-Net++: 9,164,769

After classifying the UAV image, we resampled it to the spatial resolution of the pansharpened WorldView image, namely 50 cm. The rescaling was based on a majority-voting approach; that is, if more than 50% of the pixels of the UAV-derived lichen map within a pixel footprint of WorldView image were classified as lichen, we considered that pixel as lichen. This implies that because of different types of errors, there were some pixels not corresponding to a true lichen patch (or background) in the WorldView image. The resampled lichen/non-lichen map was then used as the label map for training another classifier on the WorldView image. We employed *Naïve-Student* [22] as an SSL approach to take advantage of unlabeled data for training a more generalizable model. The use of SSL would indirectly help learn about the area missing within our original training data. In fact, our main goal was to improve the robustness of the classifier against misleading samples our classifier had not seen in the labeled data.

2.2.1. Semi-Supervised Learning: Naïve-Student

Building on [20,21], *Naïve-Student* is an approach to training neural networks in a semi-supervised fashion when limited labeled data are available. This approach is based on a *Teacher-Student* framework where the goal is to improve the performance of the *Student* network iteratively using both labeled and unlabeled samples. In contrast to Knowledge Distillation [31], *Naïve-Student* aims to expand knowledge (or the capacity of the *Student*). In other words, the *student* network will eventually generalize better than the *Teacher* network. The premise of such an SSL framework basically is to improve robustness [21].

Since we had limited data available, we first split the UAV and WorldView images into equally sized image patches before inputting them to the networks in this study. Given this, to formalize the *Naïve-Student*, let $X = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ be a set of paired n image patches and respective label maps. In addition, let $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_m\}$ be a set of m unlabeled image patches. The size of both the labeled and unlabeled patches in this study was chosen to be 64 by 64 pixels, with the labeled image patches having an overlap of 75%. The reason for using a large overlap was two-fold: (1) minimizing the edge-effect problem in predictions, and (2) increasing the number of training data, which can be considered as a data augmentation process [32]. Given these two data sets, *Naïve-Student* procedure in the context of this research can be summarized as follows:

1. Train a *Teacher* network θ_*^t minimizing binary cross entropy on the labeled set:

$$\frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i, \theta_*^t)) \quad (1)$$

2. Generate hard pseudo-labels for the unlabeled set with test-time augmentation (TTA) functions (in this study, vertical and horizontal flips):

$$\tilde{y}_i = f(TTA(\tilde{x}_i), \theta_*^t), \forall i = 1, \dots, m \quad (2)$$

3. Train an equal or larger *Student* network θ_*^s minimizing binary cross entropy on the pseudo-labeled image set:

$$\frac{1}{m} \sum_{i=1}^m L(\tilde{y}_i, f(\tilde{x}_i, \theta_*^s)) \quad (3)$$

4. Fine-tune the trained *Student* network θ_*^s on the labeled set:

$$\frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i, \theta_*^s)) \quad (4)$$

5. Replace the *Student* network θ_{**}^s with the *Teacher* network and start over from step 2.

These steps, starting from the second step after the first iteration, are repeated for a certain number of iterations until desired results are achieved based on smallest validation loss.

The choice of the networks used as the *Teacher* and *Student* is highly dependent on the application at hand. Choosing very deep models for easy tasks may lead to extreme overfitting after a few epochs. We used a U-Net model [33] as the initial *Teacher* model. This network has been frequently proven to be an accurate model for different mapping tasks in RS, and thus it was also a suitable choice for lichen mapping. For the *Student* model, we selected a U-Net++ [34], which is a nested version of the plain U-Net and generally outperforms U-Net. The reason for choosing U-Net++ was two-fold: (1) we needed a model that would be more accurate than U-Net, thus more likely helping improve the performance of the *Student* model; and (2) we needed a consistent model to U-Net which was not much deeper, causing early overfitting. In fact, since U-Net++ is an enhanced version of U-Net, its architecture is consistent with U-Net, so that hyperparameter tuning of U-Net++ would not be different from U-Net. It should be noted that the original U-Net++ uses a custom loss function which is a combination of the standard binary cross entropy and Dice function. However, in this study, we instead used the standard binary cross entropy, as we found it more suitable and stable for the mapping task in this study. In addition, neither data over-/under-sampling nor class weighting was employed while training the *Student* network on pseudo-labels. In fact, there were more complex patterns in non-lichen features that the model needed to decipher to improve its robustness against misclassifying background pixels as lichen.

We used a U-Net with 32 initial convolutional filters and the convolutional blocks of 32-64-128-256 convolutional filters, and the bottleneck with 512 convolutional filters. The schematic of the U-Net model applied in this study can be seen in Figure 3. We also used a U-Net++ (Figure 4) with the same 32 initial convolutional filters and convolutional blocks, without the deep supervision mode. We used the Adam optimizer to train both the networks. As mentioned earlier, rather than using the single training image as direct input to the SSL, we split it into equally sized, overlapping patches, which is a common practice when limited data are available and/or when the entire image cannot fit in the GPU memory [33,35]. Since we had limited labeled samples, we initialized our *Teacher* and *Student* networks using supervised pretrained U-Net and U-Net++. To make the models more adaptable to the domain of this study, we pre-trained them on an urban pixel-wise

classification task on a large WorldView scene and reused the intermediate weights in the *Naïve-Student* setup.

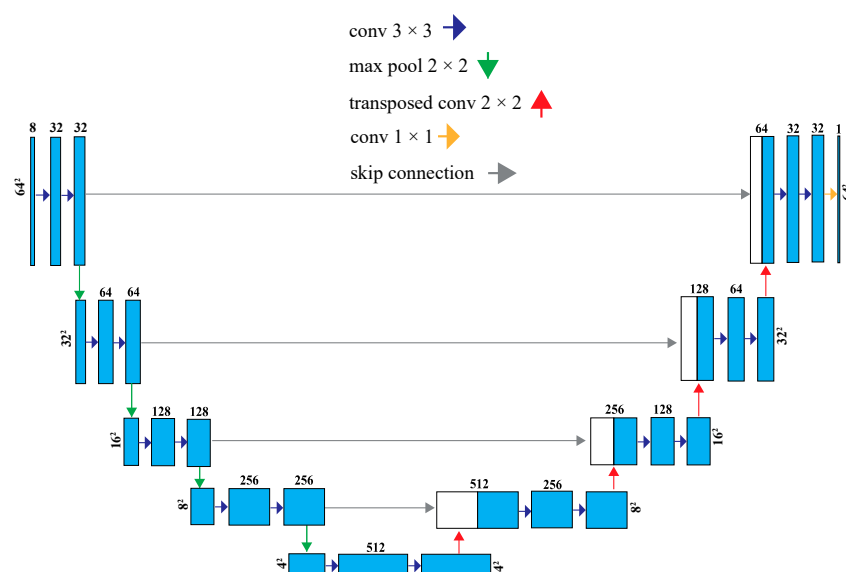


Figure 3. Architecture of the zero-padded U-Net model used as the initial Teacher network in the *Naïve-Student* in this study.

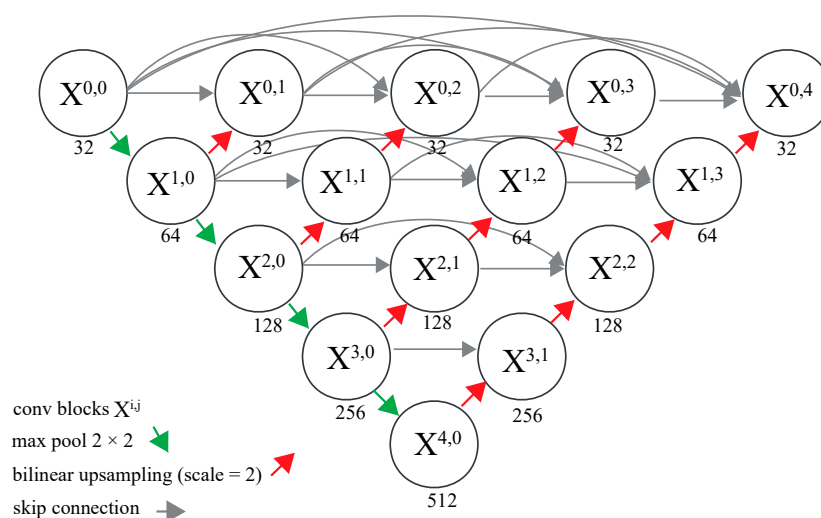


Figure 4. High-level architecture of the U-Net++ model used in the *Naïve-Student* framework in this study. Note: Values under each circle (i.e., convolutional blocks) indicate the depths of outputs from the respective layers.

The training started with the initial *Teacher* network, namely U-Net, using a learning rate of 0.001. To improve generalization, we used dropout layers (with a probability of 0.25) in both U-Net and U-Net++ networks during training, while dropout layers were disabled during inference (i.e., during pseudo-labeling and test labeling). The first *Teacher* model was trained on the labeled training data whose labels were obtained from the resampled UAV-derived map. All the training for each *Teacher* network terminated once no boost on the loss (i.e., binary cross entropy loss) of the labeled validation set was observed after a few consecutive epochs. Afterwards, the trained model was applied to the large unlabeled data set to generate the respective pseudo-labels. As with the original *Naïve-Student* approach, we applied a TTA. We used flipping functions (i.e., horizontal and vertical) with no multi-scale transformation, as it did not lead to a tangible improvement because of the partially

scale-independent nature of the lichen mapping task in this study. The pseudo-labeled data were then used as a training set for training the *Student* network (i.e., U-Net++). We used 80% of the pseudo-labeled data for training and the remaining 20% for validating the model. The trained *Student* model was then fine-tuned on the labeled training data with a smaller learning rate (i.e., 0.0001) to prevent weight explosions in the pre-trained model. For training all the networks, we used a batch size of 16, which was selected experimentally based on the directions provided by Masters and Luschi [36]. The fine-tuned *Student* was utilized as a new *Teacher* to repeat the whole process in the second iteration. This process was repeated until the smallest loss on the labeled validation set was achieved.

There were also clouds over some parts of the WorldView scene. Before training, we did not mask out clouds, cloud shadows, or water from the training process on the WorldView scene. In fact, we also aimed to analyze the robustness of SSL against cloud pixels that were spectrally similar to some lichens. For training and validation set selection, the resampled UAV-derived label map was split into 52 non-overlapping valid image patches. If an image patch had more than 256 invalid pixels (i.e., pixels not labeled), we discarded it from our analysis. We used ~70% of the image patches for training (i.e., 37 patches) and the rest for validation (i.e., 15). Then, as described earlier, 75%-overlapping image patches were extracted from neighboring training image patches.

2.2.2. Noisy Labels

Noise is an inseparable component of any data. Our data suffered from standard noise sources such as sampling noise and sensor noise. However, we also considered noise as any type of data disagreement that resulted in mislabeling. For example, assuming that the UAV and WorldView images were perfectly co-registered, there could be a lichen patch that was easily visible in the UAV image but occluded by tree shadows in the WorldView image. Another example of a data disagreement is orthorectification errors that can cause underlying land cover to appear differently in the two images. To analyze the role of noise in this research, we first need to clarify what we exactly mean by “noisy labels”. With respect to the scaled-up UAV-based lichen map, the term “noisy labels” in the context of this research had three different forms:

- *Pixels wrongly labeled as lichen*: This included pixels that were devoid of any lichen patch in the WorldView image but wrongly labeled as lichen (i.e., false positive error).
- *Pixels wrongly labeled as background*: This included pixels that were actually lichen in the WorldView image but were labeled as background (i.e., false negative error).
- *Pixels correctly labeled as lichen but shadow contaminated*: This included pixels that corresponded to a lichen patch in the WorldView image but were largely occluded by shadows (i.e., false positive error).

Past research has shown that DNNs to some degree are robust to noise compared to conventional models [15]. In fact, because of the memorization effects, DNNs first fit to clean data, and after some epochs, start overfitting to noisy data [37]. Therefore, early stopping based on a clean validation set may help prevent the model from memorizing mislabeled samples. However, providing a clean validation set in real-world applications is very cumbersome. The *Naïve-Student* approach itself is also partly robust to noisy labels but is not immune. In this study, rather than attempting to refine our training set by filtering out noisy labels, we kept them to analyze how such label noise would affect the performance of the SSL framework for lichen detection. In addition, to test the effect of noisy data on model performance, we developed a cleaner training dataset by delineating training data in the WorldView image directly. This experiment would show the effect of a noisy labeled set compared to a possibly clearer labeled set.

To expand the assessment of the SSL framework, we also deployed a pixel-based RF approach for lichen detection. The reason for using a pixel-based RF approach for comparison purposes in this study was due to the popularity of this algorithm for lichen mapping [13,38] as it is an accurate conventional model and is more computationally efficient than Support Vector Machine (SVM), which is another powerful conventional

algorithm. To select the most appropriate hyperparameters, a grid-search procedure was employed. Grid search for hyperparameter tuning resulted in a value of 500 as the number of trees, a value of 2 as the minimum samples per leaf, and the square root of the total number of randomly chosen features at each split. Along with the 8 spectral bands of the WorldView image, we calculated NDVI and gray level co-occurrence matrix (GLCM) features (mean, dissimilarity, homogeneity, variance, entropy, correlation, contrast, and second moment) to improve the performance of the model in detecting lichen cover.

2.2.3. Programming Environment

All the neural network models were implemented using the *PyTorch* framework, and the RF model using the *scikit-learn* library. Other Python libraries including *NumPy* and *Rasterio* were also used for (geo)data handling. The framework was run on Compute Canada servers (i.e., Graham GPU clusters) using NVIDIA V100 GPUs and 100 GBs of RAM. We also used Google Colab for prototyping and initial testing of our models.

2.2.4. Accuracy Assessment

For accuracy assessment purposes, we collected 3000 pixels randomly and manually to equalize the number of samples in each class as much as possible. Samples that fell on clouds or were difficult to interpret were discarded, resulting in a total of 2649 samples. Two main measures of accuracy assessments were used: Overall accuracy (OA) (Equation (5)) and F1-score (Equation (6)); and its components Precision (Equation (7)) and Recall (Equation (8)).

$$OA = \frac{\text{True Positive} + \text{True Negative}}{\text{Positive} + \text{Negative}} \quad (5)$$

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

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (7)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (8)$$

3. Results

According to the experiments, the OA and F1-score of the classified UAV image were 97.26% and 96.84%, respectively. Figure 5a illustrates the resampled (50-cm) classified UAV map superimposed onto the corresponding area of the WorldView scene. In this figure, it is also evident that there are some non-lichen pixels wrongly labeled as lichen or vice versa. The results of our assessment on the resampled map at the WorldView level showed an OA and F1-score of 83.61% and 90.74%, respectively, which indicates a tangible accuracy degradation compared to the original, non-scaled-up UAV map. No specific preprocessing or sampling was performed before training the SSL networks at the WorldView level.

The best result of training the *Teacher-Student* framework was achieved after three iterations (where the smallest labeled validation loss was achieved) namely that the *Student* network was used as a *Teacher* in two iterations. Figure 6 shows a small portion of the WorldView image (also covering the UAV extent) that was used as an input to each of the three networks obtained after each iteration to predict lichen pixels. The *Student* network improved gradually after each iteration. The most obvious performance improvement can be seen over the road pixels and some tree types that were differentiated more accurately from lichen patches as the network evolved. Essentially, the Type I error was reduced after each iteration without adding new training samples to the process. This confirms that an iterative *Teacher-Student* SSL approach could help improve pixel-wise classification accuracy when only a small training set is available.

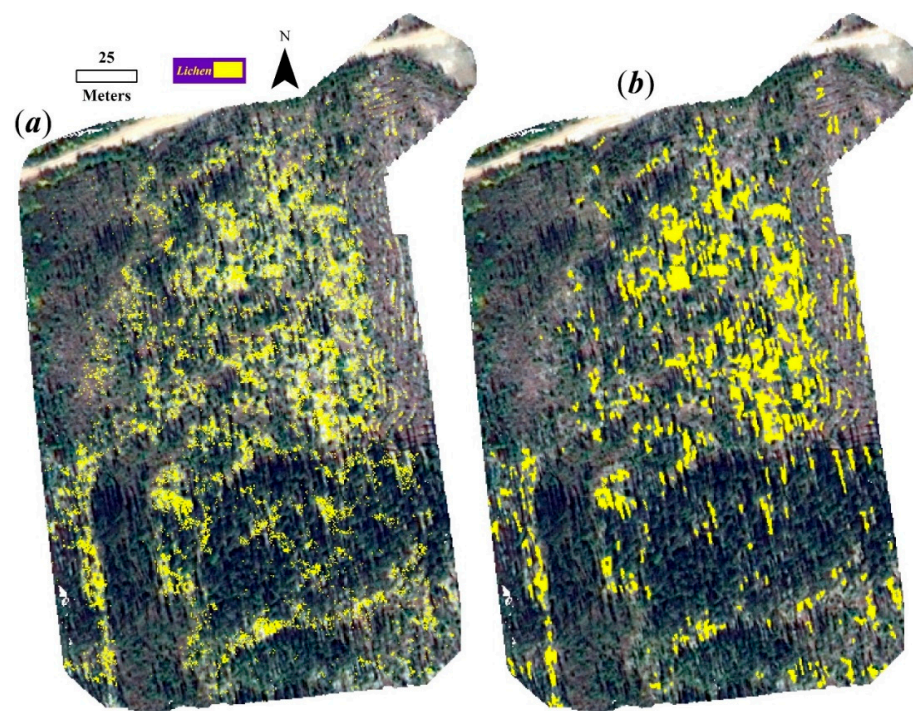


Figure 5. Lichen map (labeled data) over the UAV extent: (a) resampled UAV-derived map and (b) WorldView-derived map.

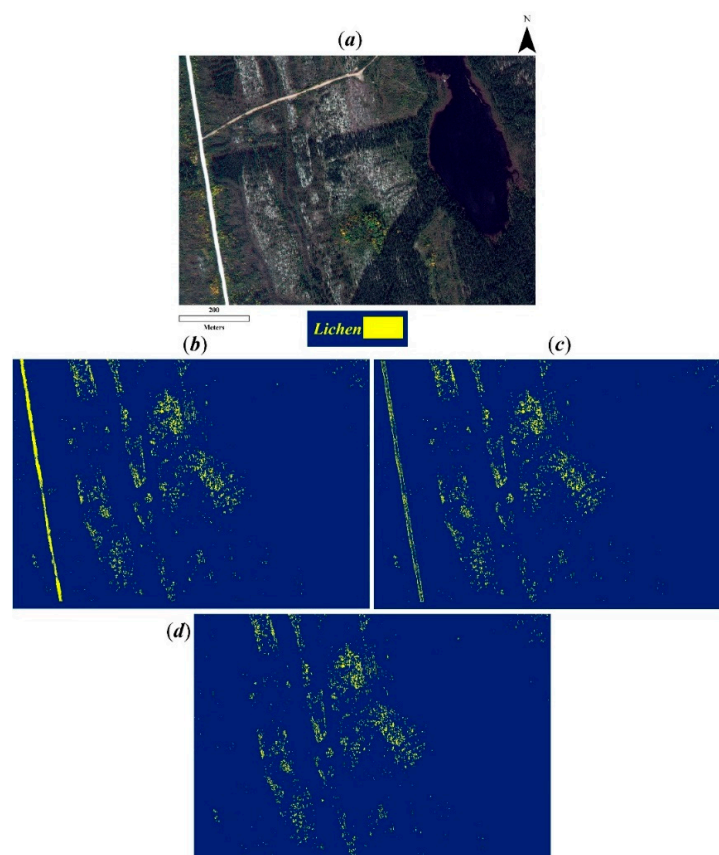


Figure 6. Progress of improvement in the Naive-Student for lichen mapping over (a) a portion of the WorldView image (which also covers the UAV extent): (b) prediction map after the first iteration, (c) prediction map after the second iteration, and (d) prediction map after the third iteration.

Figure 7 displays the map generated for the Northern part (i.e., test image) of the WorldView scene using the SSL approach. Unsurprisingly, some parts of the thick clouds were erroneously classified as lichen. However, most of the thin clouds and haze were correctly differentiated from lichen. We also observed that some rippling-water pixels were misclassified as lichen, although the number of misclassified water pixels was much less than that of misclassified cloud pixels. It should be noted that neither of these two classes were present in the labeled samples used for constructing the SSL models. Misclassification of some trees as lichen was caused by mislabeled pixels in the scaled-up UAV-derived lichen map. To handle misclassified pixels over cloud areas, we applied a simple cloud masking using the Fire Discrimination Index (FDI), which was reported to be useful for cloud-masking in WorldView imagery [39]. Although applying this mask resulted in the removal of most of the lichen misclassifications over cloudy areas, there were still few cloud pixels that required manual removal as employing a larger threshold would remove some lichen pixels as well.

Based on our quantitative accuracy assessment, the SSL-based WorldView lichen map had an OA of 85.28% and F1-score of 84.38%. From the confusion matrix (Table 2), it can be concluded that the rate of misclassification of background pixels as lichen was less than the rate of misclassification of lichen pixels as background (i.e., more lichen underestimation than lichen overestimation). The activation maps (Grad-CAM [40]) of the SSL U-Net++ for two image titles containing lichen patches can be seen in Figure 8. We can see that the network learned low-level features (like edges, and some basic relationships between lichens and neighboring areas) in the first few layers (first convolutional block). In the last layers, the network combined abstract and high-level features to perform final predictions, resulting in CAMs looking very similar to the prediction maps. Comparing the SSL U-Net++ with the supervised one showed a significant accuracy improvement (more than 15% and 22% improvement in OA and F1-score, respectively). The more important part of accuracy improvement was due to the higher true positive rate of the SSL U-Net++ (Table 2), which is mainly because of the increase in the size of training data provided through the iterative pseudo-labeling and fine-tuning.

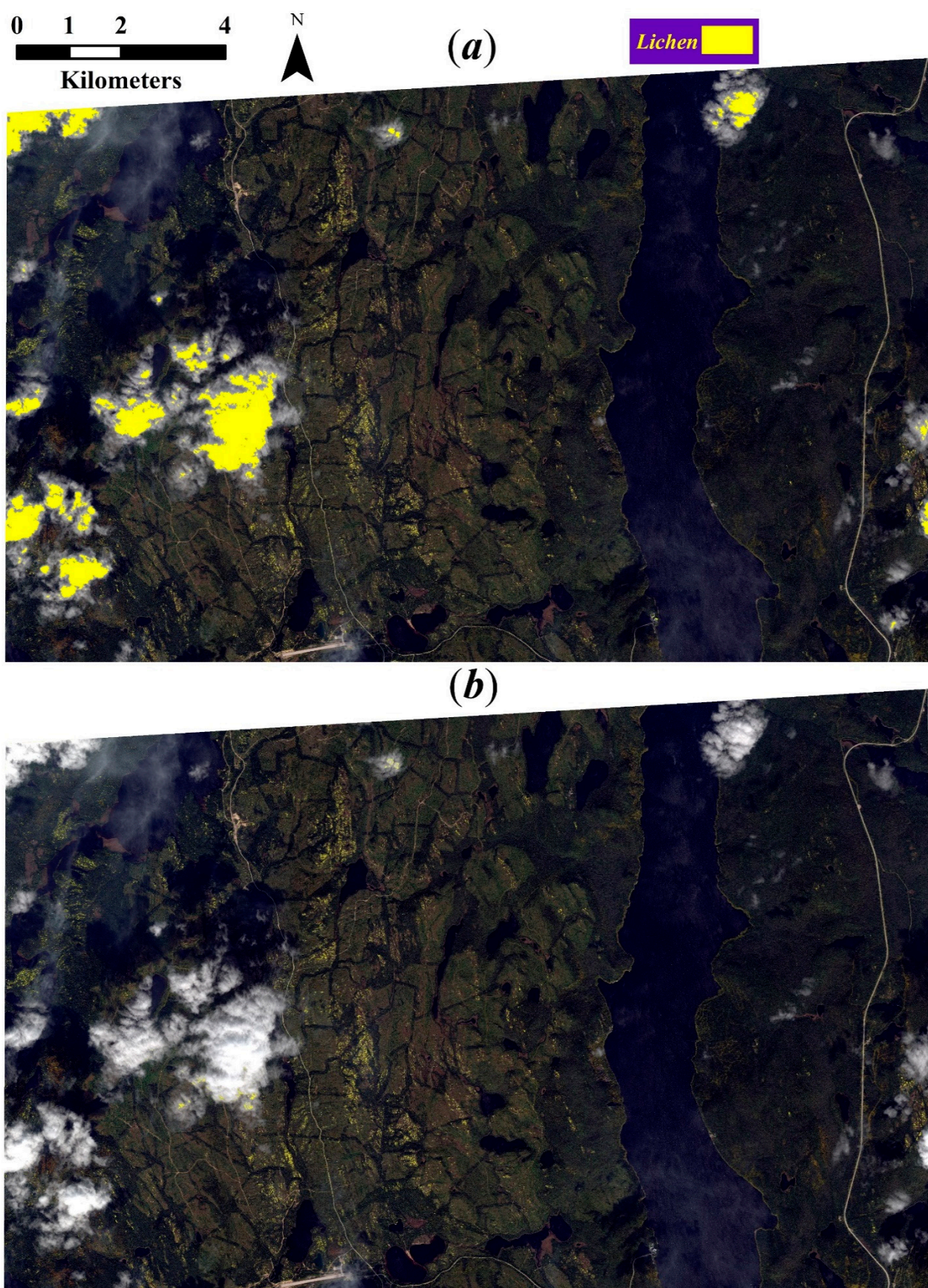
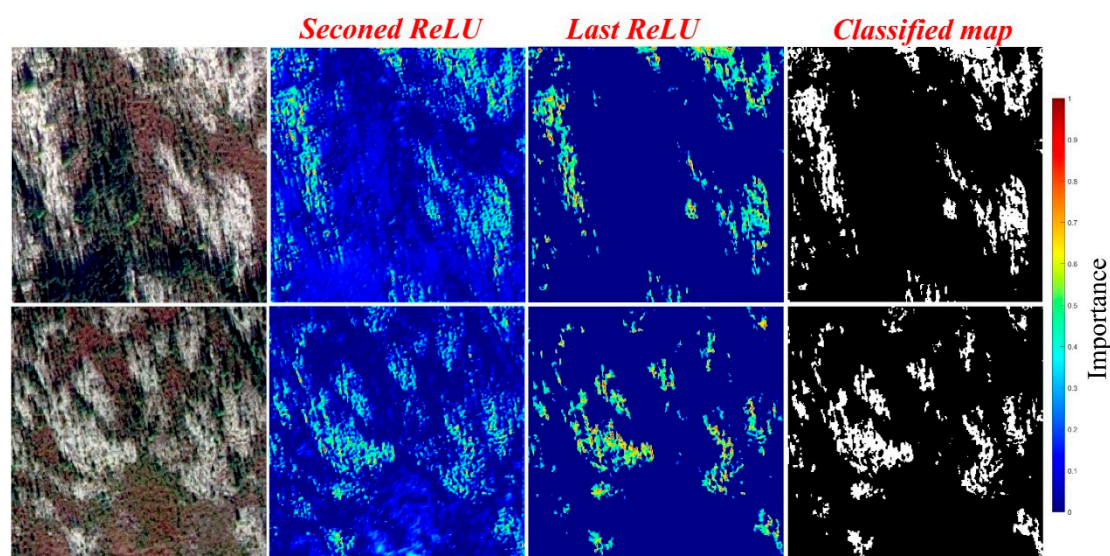


Figure 7. Lichen cover maps generated on the test WorldView image: (a) original map and (b) the map whose clouds were masked out using the FDI after classification.

Table 2. Confusion matrices calculated based on the classification maps generated by the models in this study. Note: The SSL models are based on the final U-Net++ Student networks.

Main model	OA = 85.28% F1-score = 84.38%	True label	
		<i>Lichen</i>	<i>Background</i>
	<i>Lichen</i>	1053	229
	<i>Background</i>	161	1206
Cleaner model	OA = 85.05% F1-score = 85.27%	True label	
		<i>Lichen</i>	<i>Background</i>
		1146	328
		68	1107
Overfitted model	OA = 64.82% F1-score = 48.68%	True label	
		<i>Lichen</i>	<i>Background</i>
		442	160
		772	1275
Supervised U-Net++	OA = 70.93% F1-score = 62.07%	True label	
		<i>Lichen</i>	<i>Background</i>
		630	186
		584	1249
RF model	OA = 60.89% F1-score = 27.86%	True label	
		<i>Lichen</i>	<i>Background</i>
		200	22
		1014	1413

**Figure 8.** Lichen activation maps (Grad-CAM) of the second ReLU output (the output of the first convolutional block) and the last ReLU output.

3.1. Noisy Labels vs. Cleaner Labels

We hypothesized that noise, specifically false positive detections, would have a deleterious effect on our results. In fact, since lichen maps generated based on high-resolution data are commonly used to map fractional lichen cover in coarser images (e.g., Landsat imagery), the priority may typically be to generate a map with a lower false positive rate

(instead of a higher true positive rate). In this section, by rejecting the above-mentioned hypothesis, we discuss that not only these types of noise were not deleterious, but also they, to some degree, enhanced the robustness of the SSL along with the use of unlabeled data. To demonstrate this experimentally, rather than scaling-up the UAV-based lichen map to the WorldView image and then training the SSL baseline on it, we directly classified the same area in the WorldView image. For this purpose, we used the same 1825 points employed for training and testing the DL-GEOBIA model on the UAV image. However, this time we pre-screened the points to correct mislabeled samples and replace those falling inside the same segments (i.e., in the segmented image) with new randomly chosen ones. The samples were then used for training another DL-GEOBIA model capable of detecting lichen pixels over the area. According to the accuracy assessment, the OA and F1-score measures were 88.52% and 93.46%, respectively. As expected, this map had a higher accuracy than the resampled one. In other words, compared to the scaled-up UAV-based lichen map, this map visually appeared to have higher Precision (i.e., less background pixels were wrongly classified as lichen), especially over shadow and tree classes. Furthermore, a greater number of pure lichen pixels were detected correctly, which was at the expense of not detecting spectrally mixed lichen pixels that were more difficult for the classifier to identify in the WorldView image compared to the UAV image. As a result, this map provided a suitable case to analyze the effect of Type I error on the SSL training. This map was then used for training the SSL networks using the same procedure described earlier.

Figure 9a illustrates the map generated using the trained SSL framework on this cleaner training dataset (hereafter, cleaner model). We can clearly see that more cloud pixels were classified wrongly as lichen compared to the map generated by the main model in Figure 7a. Accuracy assessment on this map showed an OA and F1-score 85.05% and 85.27%, respectively. According to the confusion matrix (Table 2), the cleaner model classified less background pixels correctly compared to the main model trained on the noisy labels. However, it detected more lichen pixels correctly than the main model did.

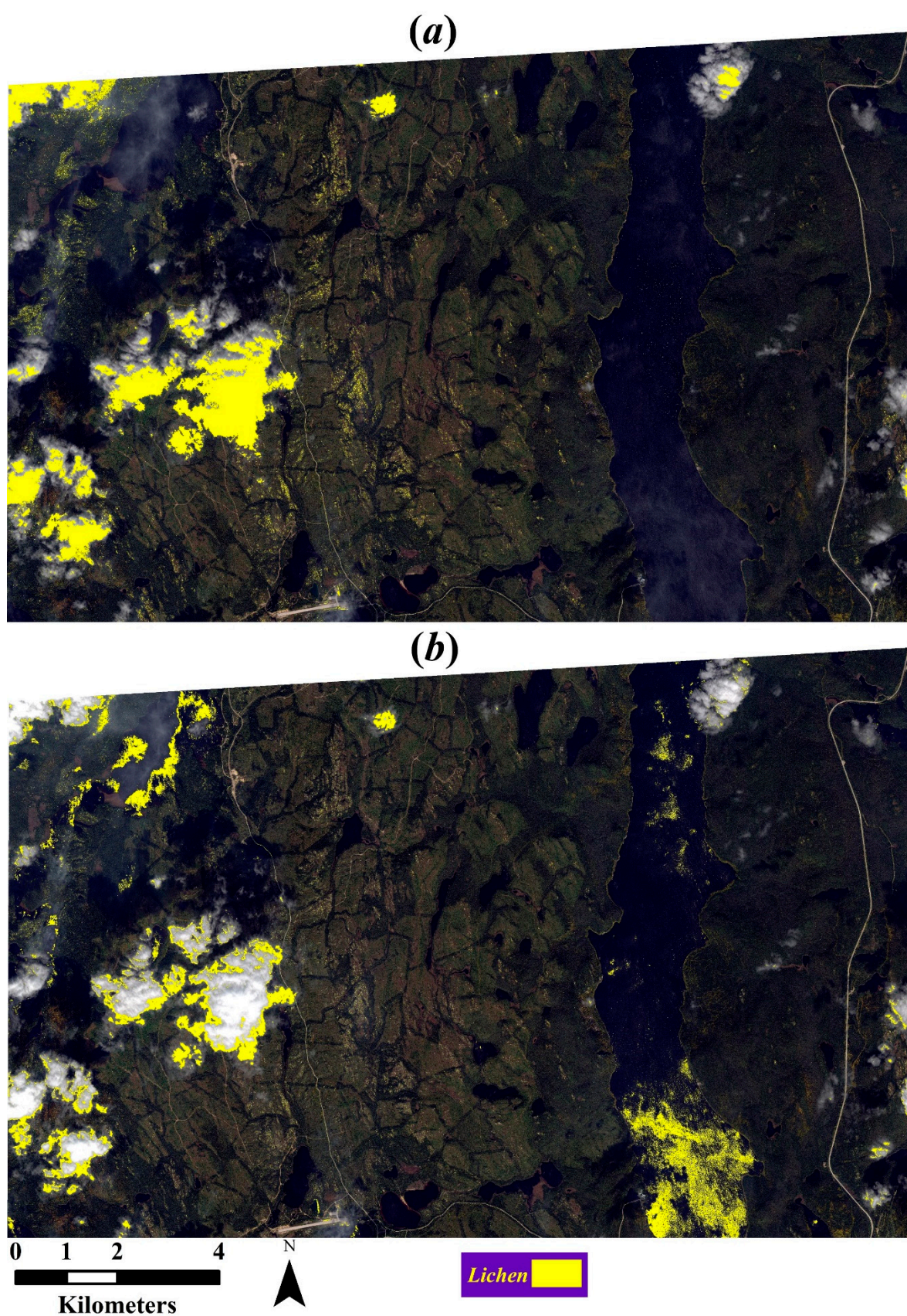


Figure 9. Lichen map generated with (a) the SSL network trained on the cleaner data set rather than the UAV-derived lichen map, and (b) the overfitted SSL network.

3.2. Overfitted SSL

As discussed earlier, since DNNs generally first fit to clean samples and then overfit to noisy samples, preventing the network from extreme overfitting can significantly improve the performance of the model. To realize the effect of overfitting to our noisy labels for lichen mapping, we experimented a scenario where all the samples were used for training the SSL approach; that is, no validation set was used to perform early stopping. In fact, we let the networks overfit to the noise in our training data to realize the negative effect of natural noise in the case of overfitting in this study. After training the overfitted SSL, we applied it to our WorldView test image resulting in the map presented in Figure 9b. Although clouds were better distinguished from lichen pixels in this map, the overfitted SSL produced a lichen map with significantly lower OA and F1-score (64.82% and 48.68%, respectively) than the main map did (Figure 7). There is also a major increase in the number of non-lichen misclassifications, especially over water bodies. We expect that such misclassifications resulted from the lack of generalization in the model. Moreover, partly discriminating clouds from lichens was to some degree indicative of learning from noise. This, however, was at the expense of considering noisy labels as true labels (i.e., water pixels and thin clouds). The corresponding confusion matrix (Table 2) shows that the network had a low performance for detecting lichen pixels correctly. This classifier had the worst accuracy compared to the other two SSL ones. The classifier detected lichen patches very sparsely as the network had a low lichen detection rate (Figure 10). In fact, of 1214 lichen pixels, only 442 of them were correctly detected. Moreover, according to Figure 9b, the unusual misclassifications of many background pixels (e.g., water bodies, thin clouds, etc.) proved the poor accuracy of the model overfitted to a noisy, limited labeled dataset.

A closer comparison among the maps generated by the three types of SSL models is provided in Figure 10. According to this comparison, denser lichen patches were detected by the cleaner model. In other words, within a given lichen patch, more lichen pixels were detected correctly. This was less the case for the map generated by the main model as some lichen pixels within a given lichen patch were missed. The worst case occurred in the map derived from the overfitted model. These observations are in line with the previously described confusion matrices in Table 2. As discussed earlier, the noisy model misclassified many cloudy pixels into lichen pixels. Although we were able to rule out many of those misclassifications using the FDI mask, there were still cloud pixels where the mask was not able to remove them. This was mainly the case in areas where there were thin clouds or haze. Overall, we observed that the SSL network trained on the naturally noisy labeled data was more consistent than the cleaner model was. In fact, it was found that there was a trade-off between false positive and false negative errors in this map.

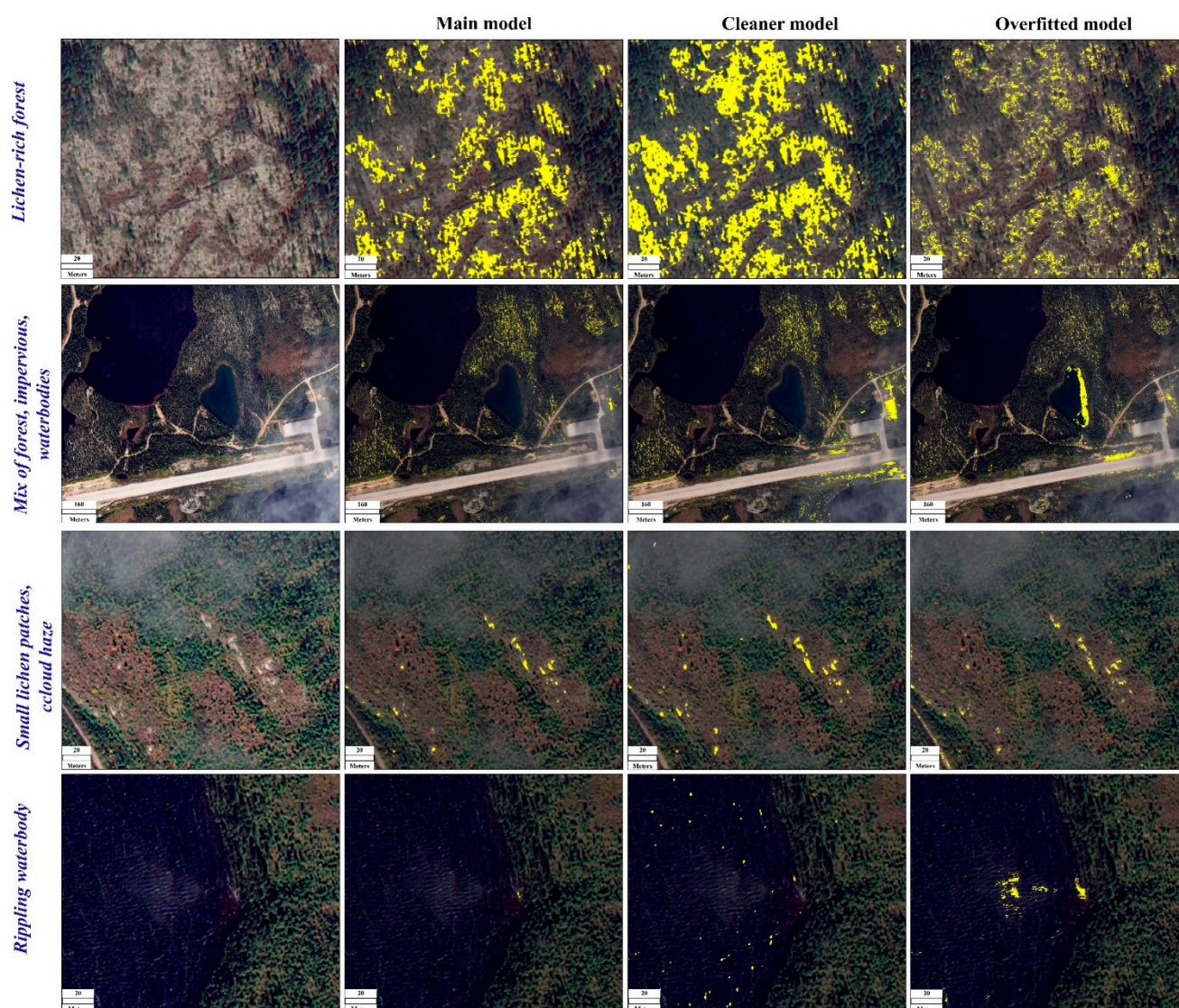


Figure 10. Comparison of four portions of the test image for which lichen pixels were detected with the main noisy model, cleaner model, and overfitted model.

3.3. Pixel-Based RF Model

In this section, we present and discuss the performance of the pixel-based RF applied for the scaling-up of the UAV-derived lichen map. As shown in Figure 11, the map generated with the RF model was able to correctly classify many cloud pixels like the overfitted model did. However, it had a low lichen detection rate, even worse than the overfitted model, which is not surprising as conventional ML models are less generalizable than CNNs. As shown in Figure 11, it is obvious that many cloud pixels were correctly classified as background, but it was at the expense of losing sufficient sensitivity for detecting many lichen pixels. According to the accuracy assessment results (Table 2 and Figure 12), the RF model had an OA of 60.89% and F1-score of 27.86%. Compared to the CNNs applied in this study, the much lower F1-score of the RF model is mainly due to its very low Recall performance. Considering the confusion matrix of the RF model (Table 2), although the model was able to correctly classify most of the background test samples, it performed very poorly in detecting true lichen pixels in the test set.

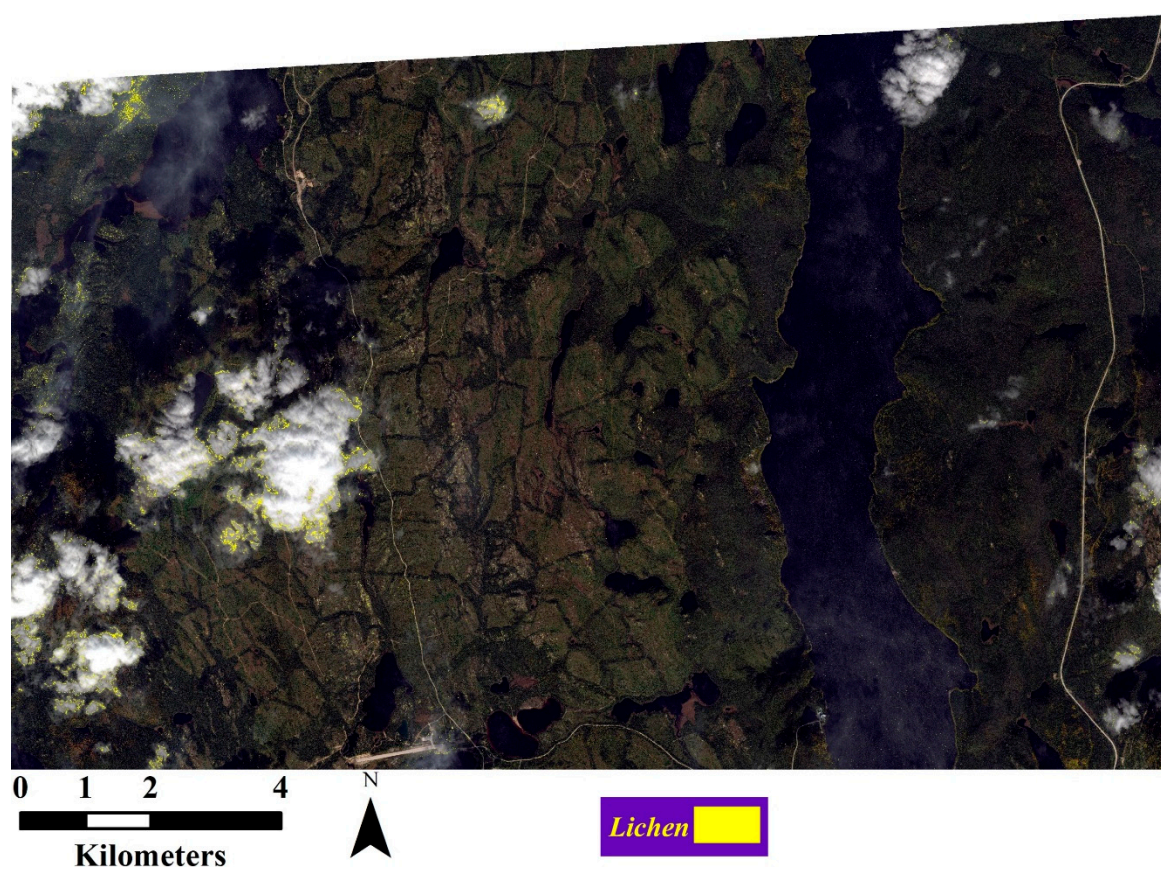


Figure 11. Lichen map generated with the RF model.

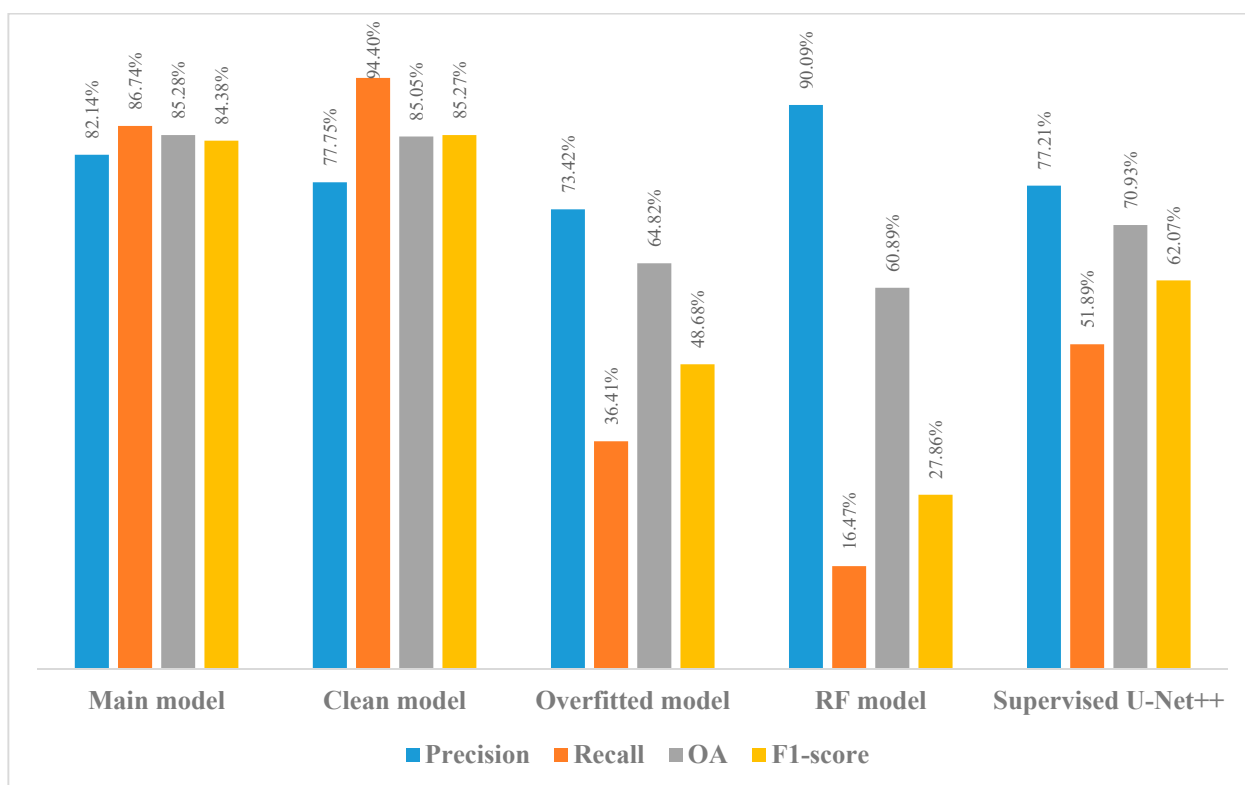


Figure 12. Accuracy assessment results of the four classifiers used in this study.

4. Discussion

Scaling-up a small, very-high-resolution lichen map to a coarser, larger scale is a challenging task. In fact, given limited prior information, we expect an ML model to produce quality maps, which is a very hard problem for any model or training approach. Results presented in this study indicated that even using a small base map, we were able to carry out the scale-up process to obtain a reasonably accurate distribution map (i.e., OA of ~85% and F1-score of ~84%) for bright caribou lichens at a much larger scale. In addition to the small extent of the UAV image, another limiting factor was the lack of sufficient land-cover diversity over the UAV extent. The area surveyed with the UAV contained four primary land-cover types (i.e., soil, tree, bright caribou lichen, and grass). Beyond the extent of the area surveyed with UAV, we encountered several new land-cover types, some of which could mislead the classifier. Another potential issue for the scaling-up process was the presence of label noise in the training set prepared for the *Naïve-Student* training. Although the behavior of DNNs against noise is not yet fully understood, past research has shown that DNNs have a tendency to first fit to clean samples and then overfit to noise after some epochs [15,41]. However, in this study, we had natural, not synthetic, noise in our label data. The resampled UAV-derived map had several mislabeled pixels caused by spatial and temporal data disagreements between the UAV and WorldView images. All these sources of label noise are different from data noise and classification-driven noise. Such noise can also affect the quality of the scaled-up maps that are generally difficult to avoid. The fourth problem was the presence of clouds and haze in the WorldView image. We depicted that the main model occasionally misclassified thick clouds as lichen and thus overestimated lichen cover. We addressed this issue by removing most of these misclassifications with an FDI mask. Conversely, the presence of haze over some areas with lichen led to misclassifications of lichen and therefore caused an underestimation of lichen cover. Given this, we found that clouds caused both false positive and false negative errors in cases where target lichens would be bright colored, as in this study.

The better performance of the cleaner model compared to the main model in detecting more lichen patches correctly (i.e., higher Recall) can be ascribed to the fact that more pure lichen training samples (i.e., less noisy labels) were used for training the network. This, however, degraded the Precision of the model compared to the main model.

Given the limited labeled samples and their lack of sufficient diversity, it appears that noisy labels helped prevent the model from memorizing homogenous patterns in the data. These findings are in line with those reported by Xie, Luong, Hovy, and Le [21] where they found that injecting noise to the *Student* network was one of the main reasons for improving the generalization power of their SSL approach. However, in our study, we had natural noise resulting from the scaling-up of the UAV-derived lichen map. This type of data noise affected both the *Teacher* and *Student* networks, as opposed to the *Noisy-Student* approach in which data noise (i.e., image augmentation) is only injected to inputs of the *Student* network, and as opposed to *Naïve-Student* in which no explicit data noise is applied. This natural noise ultimately improved robustness related to Type I error in the *Naïve-Student* approach applied in this research. Although the natural noise in this study partially improved the robustness of the model, overfitting to noise would be a critical problem if models are not trained appropriately. Considering this issue, we showed that overfitting to noise caused very poor accuracies, resulting in detecting many non-lichen pixels erroneously as lichen.

We also presented a mapping result generated with an RF model and a high-dimensional feature set. Although RF is a popular method in lichen mapping, accuracy assessment results showed that it did not lead to an accurate map. This was, however, to some degree expected as pixel-based models do not generalize well. Processing times of SSL and RF are presented in Table 3. Training the SSL was more time consuming than RF and supervised U-Net++. However, the inference time of the RF was higher, although this is without considering the time spent on computing the GLCM features in the test image.

Table 3. Training time of SSL and RF models.

Models	Training Time
SSL (average)	~2 h (3 GPUs)
RF	~15 min (10-core CPU)
Supervised U-Net++	~5 min (1 GPU)

Given the above limitations, the use of unlabeled samples in a full-fledged SSL model was advantageous in scaling-up the UAV-derived lichen map to the WorldView scale. However, it is obvious that if lichens of interest are of different colors, and not all of them are present within the extent of the very-high-resolution maps, SSL will fail to detect those lichens in larger-scale images. The results also showed that although the UAV site was very homogenous, it was possible to obtain a map with a reasonable rate of false positive, which is important for lichen mapping tasks. The use of unlabeled data and noisy labels was found to improve the robustness of the network against samples that could have caused Type I error.

Although the main and cleaner SSL models led to better accuracies compared to the other approaches used in this study, one of the challenges with training these models was their computational complexity compared to supervised approaches. Not only were multiple models iteratively trained, but also the number of data used for training was also increased compared to the supervised models. This caused a significant increase in the execution time of the training process. In general, one of the main computational bottlenecks in the SSL approach was the dense prediction of pseudo-labels and then using them for training the *Student* models, which, as mentioned above, were then trained on a much larger data set. Another relevant factor affecting the computational time of the SSL approach was the TTA that was performed during pseudo-labeling. These two factors were the major computationally intensive components of the SSL in this study. Given the accuracy improvement provided by the main SSL model, and the expectations of the lichen mapping task under consideration, this inefficiency compared to the supervised models and to the very expensive labeling tasks is justifiable provided that sufficient computational resources are available.

5. Conclusions

In this study, we assessed the possibility of scaling-up a very fine lichen map acquired over a very small, homogenous area to a much larger area without collecting any new labeled samples. We used a *Teacher-Student* SSL approach (i.e., *Naïve-Student*) that trains *Teacher* and *Student* networks iteratively on both labeled and unlabeled data. This approach produced a reasonably accurate (OA of ~85% and F1-score of ~84%) lichen map at the WorldView scale. The main findings in this research are as follows:

- A powerful SSL approach capable of taking advantage of abundant unlabeled data is beneficial for scaling-up small lichen maps to large scales using high-resolution satellite imagery, provided that there are samples for lichens of interest within the small lichen maps under consideration.
- Different types of image misalignments can introduce noisy labels in the scaled-up training set. However, we found that if this noise is not massive, it may ultimately improve the model robustness without conducting any intensive data augmentation on training data.
- The map generated by the cleaner model indicated a higher Recall but a lower Precision than the map generated by the main model.
- Overfitting in the presence of noise significantly degraded the performance of the trained network when applied to the test image. That network failed to detect many lichen patches and resulted in low accuracy. In some cases, we observed that due to overfitting to noisy labels (especially dark areas), many lichen misdetections occurred over water bodies resembling noisy labels in the training set.

There is a need for more comprehensive studies on the effect of label noise on final classification results. This is an important gap in the context of lichen mapping although when using multi-source/sensor data, there will be several types of misalignments that cause label noise. In this regard, a potential future direction for RS-based lichen mapping could be to conduct more in-depth studies on the nature and amount of noise that can improve the robustness of a model. This can be performed based on a systematic study introducing synthetic and natural noise to the classification procedure. Similarly, it could be investigated whether an ensemble of fine-tuned *Student* networks is able to produce more accurate results than a single network. In fact, different networks have different prediction capabilities. Therefore, if they are ensembled together (e.g., based on an averaging/majority-voting output or a more complex aggregation strategy), they may improve lichen detection. This may be also achieved by distilling several interconnected *Student* networks to reach a single, lighter final *Student* network.

If sufficient labeled data are available, it is generally recommended to use the maximum input size fitting to the GPU memory when splitting a given image. However, due to the use of a small label map, one of the most important limitations in this study was the use of small image patches (i.e., 64 by 64 pixels) for training the *Naïve-Student* networks. Such an input size caused both of the *Teacher* and *Student* networks to be unable to extract more useful, representative contextual features in deeper encoding layers of the networks. We also found a shallower U-Net(++) less accurate than the one used in this study, although this improvement could have been more significant if larger image patches were used. If the likelihood of having more untraceable noise and uncertainty as well as training new, complex networks can be justified, one way of mitigating this problem could be to use Generative Adversarial Networks (GANs) to either synthesize more image patches from larger image patches or to improve the resolution of small image patches through a super-resolution setup.

Author Contributions: Conceptualization, S.J.; methodology S.J., D.C., W.C.; software, S.J., D.C.; validation, S.J.; data curation, all authors.; writing—original draft preparation, S.J.; writing—review and editing, all authors; visualization, S.J.; supervision, D.C. and W.C. All authors have read and agreed to the published version of the manuscript.

Funding: This study is funded as a part of the Government of Canada’s initiative for monitoring and assessing regional cumulative effects, a recently added requirement to the new Impact Assessment Act (2019). It was also financially supported by National Science and Engineering Research Council (NSERC) Discovery grant and Queen’s University Graduate Award.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We are grateful to Robert H. Fraser of CCRS for contributing to constructive discussions on this research. Many others participated in the field data collection, including André Arsenault of the Atlantic Forestry Centre-NRCan, Isabelle Schmelzer of the Department of Fisheries and Land Resources—Government of Newfoundland and Labrador, Darren Pouliot of Landscape Science and Technology Division-ECCC, Jurjen van der Sluijs of the Department of Lands, Government of Northwest Territories, as well as co-op students, Monika Orwin and Yiqing Wang from University of Waterloo and Alex Degrace from the University of Ottawa. Shahab Jozdani conducted this research in collaboration with CCRS through a Research Affiliate Program (RAP). The authors would like to acknowledge Compute Canada and Google Colab for providing cloud computing platforms used for different analyses in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Gunn, A.; Tarandus, R. The Iucn Red List of Threatened Species. 2016. Available online: <https://www.iucnredlist.org/> (accessed on 10 September 2020).
- Wittmer, H.U.; Bruce, N.; McLellan, R.S.; Clayton, D.A. Changes in Landscape Composition Influence the Decline of a Threatened Woodland Caribou Population. *J. Anim. Ecol.* **2007**, *76*, 568–579. [[CrossRef](#)] [[PubMed](#)]
- Festa-Bianchet, M.; Ray, J.C.; Boutin, S.; Côté, S.D.; Gunn, A. Conservation of Caribou (*Rangifer tarandus*) in Canada: An Uncertain Future. This Review is Part of the Virtual Symposium “Flagship Species—Flagship Problems” That Deals with Ecology, Biodiversity and Management Issues, and Climate Impacts on Species at Risk and of Canadian Importance, Including the Polar Bear (*Ursus maritimus*), Atlantic Cod (*Gadus morhua*), Piping Plover (*Charadrius melodus*), and Caribou (*Rangifer tarandus*). *Can. J. Zool.* **2011**, *89*, 419–434.
- Barber, Q.E.; Parisien, M.-A.; Whitman, E.; Stralberg, D.; Johnson, C.J.; St-Laurent, M.-H.; DeLancey, E.R.; Price, D.T.; Arsenault, D.; Wang, X.; et al. Potential Impacts of Climate Change on the Habitat of Boreal Woodland Caribou. *Ecosphere* **2018**, *9*, e02472. [[CrossRef](#)]
- Sharma, S.; Serge, C.; Steeve, D.C. Impacts of Climate Change on the Seasonal Distribution of Migratory Caribou. *Glob. Chang. Biol.* **2009**, *15*, 2549–2562. [[CrossRef](#)]
- Joly, K.; Cole, M.J.; Jandt, R.R. Diets of Overwintering Caribou, Rangifer Tarandus, Track Decadal Changes in Arctic Tundra Vegetation. *Can. Field Nat.* **2007**, *121*, 379–383. [[CrossRef](#)]
- Kyle, J.; Cameron, M.D. Early Fall and Late Winter Diets of Migratory Caribou in Northwest Alaska. *Rangifer* **2018**, *38*, 27–38.
- Petzold, D.E.; Goward, S.N. Reflectance Spectra of Subarctic Lichens. *Remote Sens. Environ.* **1988**, *24*, 481–492. [[CrossRef](#)]
- Nordberg, M.L. Vegetation and Biomass Changes in Mountainous Areas in Sweden Using Satellite and Airborne Imaging Scanner Data. In Proceedings of the 27th International Symposium on Remote Sensing of Environment, Tromsø, Norway, 8–12 June 1998.
- Théau, J.; Duguay, C.R. Lichen Mapping in the Summer Range of the George River Caribou Herd using Landsat TM Imagery. *Can. J. Remote Sens.* **2004**, *30*, 867–881. [[CrossRef](#)]
- Gilchinsky, M.; Sandström, P.; Reese, H.; Kivinen, S.; Moen, J.; Nilsson, M. Mapping Ground Lichens Using Forest Inventory and Optical Satellite Data. *Int. J. Remote Sens.* **2011**, *32*, 455–472. [[CrossRef](#)]
- Falldorf, T.; Olav, S.; Manuela, P.; Hans, T. Estimating Lichen Volume and Reindeer Winter Pasture Quality from Landsat Imagery. *Remote Sens. Environ.* **2014**, *140*, 573–579. [[CrossRef](#)]
- Macander, M.J.; Palm, E.C.; Frost, G.V.; Herriges, J.D.; Nelson, P.R.; Roland, C.; Russell, K.L.; Suitor, M.J.; Bentzen, T.W.; Joly, K.; et al. Lichen Cover Mapping for Caribou Ranges in Interior Alaska and Yukon. *Environ. Res. Lett.* **2020**, *15*, 055001. [[CrossRef](#)]
- Kennedy, B.; Pouliot, D.; Manseau, M.; Fraser, R.; Duffe, J.; Pasher, J.; Chen, W.; Olthof, I. Assessment of Landsat-Based Terricolous Macrolichen Cover Retrieval and Change Analysis over Caribou Ranges in Northern Canada and Alaska. *Remote Sens. Environ.* **2020**, *240*, 111694. [[CrossRef](#)]
- Zhang, C.; Bengio, S.; Hardt, M.; Recht, B.; Vinyals, O. Understanding Deep Learning Requires Rethinking Generalization. *arXiv* **2016**, arXiv:1611.03530.
- Jozdani, E.; Shahab, A.; Brian, J.; Dongmei, C. Comparing Deep Neural Networks, Ensemble Classifiers, and Support Vector Machine Algorithms for Object-Based Urban Land Use/Land Cover Classification. *Remote Sens.* **2019**, *11*, 1713. [[CrossRef](#)]
- Hendrycks, D.; Kimin, L.; Mantas, M. Using Pre-Training Can Improve Model Robustness and Uncertainty. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 2712–2721.
- Nogueira, K.; Penatti, O.A.; dos Santos, J.A. Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification. *Pattern Recognit.* **2017**, *61*, 539–556. [[CrossRef](#)]
- Van Engelen, J.E.; Holger, H.H. A Survey on Semi-Supervised Learning. *Mach. Learn.* **2020**, *109*, 373–440. [[CrossRef](#)]
- Lee, D.-H. Pseudo-Label: The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks. *Workshop Chall. Represent. Learn. ICML* **2013**, *3*, 896.
- Xie, Q.; Luong, M.-T.; Hovy, E.; Le, Q.V. Self-Training With Noisy Student Improves Image Net Classification. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Virtual, 14–19 June 2020; pp. 10684–10695.
- Chen, L.-C.; Lopes, R.G.; Cheng, B.; Collins, M.D.; Cubuk, E.D.; Zoph, B.; Adam, H.; Shlens, J. Naive-Student: Leveraging Semi-Supervised Learning in Video Sequences for Urban Scene Segmentation. In *Transactions on Petri Nets and Other Models of Concurrency XV*; Springer, Cham: Basel, Switzerland, 2020; pp. 695–714.
- Johnson, B.; Tateishi, R.; Kobayashi, T. Remote Sensing of Fractional Green Vegetation Cover Using Spatially-Interpolated Endmembers. *Remote Sens.* **2012**, *4*, 2619–2634. [[CrossRef](#)]
- Johnson, B.; Tateishi, R.; Xie, Z. Using Geographically Weighted Variables for Image Classification. *Remote Sens. Lett.* **2012**, *3*, 491–499. [[CrossRef](#)]
- Alvarez-Taboada, F.; Paredes, C.; Julián-Pelaz, J. Mapping of the Invasive Species *Hakea sericea* Using Unmanned Aerial Vehicle (UAV) and WorldView-2 Imagery and an Object-Oriented Approach. *Remote Sens.* **2017**, *9*, 913. [[CrossRef](#)]
- Shao, Z.; Hu, Y.; Deren, L.; Orhan, A.; Cheng, T. Remote Sensing Monitoring of Multi-Scale Water-sheds Impermeability for Urban Hydrological Evaluation. *Remote Sens. Environ.* **2019**, *232*, 111338. [[CrossRef](#)]
- Campbell, M.J.; Dennison, P.E.; Tune, J.W.; Kannenberg, S.A.; Kerr, K.L.; Coddling, B.F.; Anderegg, W.R. A Multi-Sensor, Multi-scale Approach to Mapping Tree Mortality in Woodland Ecosystems. *Remote Sens. Environ.* **2020**, *245*, 111853. [[CrossRef](#)]

28. Zhang, X.; Chen, G.; Wang, W.; Wang, Q.; Dai, F. Object-Based Land-Cover Supervised Classification for Very-High-Resolution UAV Images Using Stacked Denoising Autoencoders. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3373–3385. [[CrossRef](#)]
29. Baatz, M.; Schäpe, A. Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. In *Beiträge Zum AGIT-Symposium Salzburg 2000*; Wichmann: Heidelberg, Germany, 2000.
30. Benz, U.C.; Peter, H.; Gregor, W.; Iris, L.; Markus, H. Multi-Resolution, Object-Oriented Fuzzy Analysis of Remote Sensing Data for Gis-Ready Information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258. [[CrossRef](#)]
31. Hinton, G.; Oriol, V.; Jeff, D. Distilling the Knowledge in a Neural Network. *arXiv* **2015**, arXiv:1503.02531.
32. Bernal, J.; Kushibar, K.; Cabezas, M.; Valverde, S.; Oliver, A.; Llado, X. Quantitative Analysis of Patch-Based Fully Convolutional Neural Networks for Tissue Segmentation on Brain Magnetic Resonance Imaging. *IEEE Access* **2019**, *7*, 89986–90002. [[CrossRef](#)]
33. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Strasbourg, France, 27 September–2 October 2021; pp. 234–241.
34. Zhou, Z.; Siddiquee, M.M.; Nima, T.; Liang, J. Unet++: A Nested U-Net Architecture for Medical Image Segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, 14 September 2017*; Springer: Berlin, Germany, 2018; pp. 3–11.
35. Liu, Y.; Ren, Q.; Geng, J.; Ding, M.; Li, J. Efficient Patch-Wise Semantic Segmentation for Large-Scale Remote Sensing Images. *Sensors* **2018**, *18*, 3232. [[CrossRef](#)]
36. Dominic, M.; Luschi, C. Revisiting Small Batch Training for Deep Neural Networks. *arXiv* **2018**, arXiv:1804.07612.
37. Yu, X.; Han, B.; Yao, J.; Niu, G.; Tsang, I.; Sugiyama, M. How Does Disagreement Help Generalization against Label Corruption? In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; Kamalika, C., Ruslan, S., Eds.; Proceedings of Machine Learning Research: Long Beach, CA, USA, 2019; pp. 7164–7173.
38. Fraser, R.H.; Pouliot, D.; van der Sluijs, J. UAV and High Resolution Satellite Mapping of Forage Lichen (*Cladonia* spp.) in a Rocky Canadian Shield Landscape. *Can. J. Remote Sens.* **2021**, 1–14. [[CrossRef](#)]
39. Whiteside, T.; Bartolo, R. *Vegetation Map for Magela Creek Floodplain Using Worldview-2 Multispectral Image Data*; Internal Report 628; Supervising Scientist: Darwin, NT, Australia, 2014.
40. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2020**, *128*, 336–359. [[CrossRef](#)]
41. Han, B.; Yao, Q.; Yu, X.; Niu, G.; Xu, M.; Hu, W.; Tsang, I.W.; Sugiyama, M. Co-Teaching: Robust Training of Deep Neural Networks with Extremely Noisy Labels. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, QC, Canada, 2–8 December 2018; pp. 8536–8546.