ArcticNet: A Deep Learning Solution to Classify Arctic Wetlands

Ziyu Jiang[1], Kate Von Ness[2], Julie Loisel[2] and Zhangyang Wang[1]

{*jiangziyu, k_vonn, julieloisel, atlaswang*}@*tamu.edu*

[1]Department of Computer Science and Engineering, Texas A&M University

[2]Department of Geography, Texas A&M University

https://github.com/geekJZY/arcticnet

Abstract

Arctic environments are rapidly changing under the warming climate. Of particular interest are wetlands, a type of ecosystem that constitutes the most effective terrestrial long-term carbon store. As permafrost thaws, the carbon that was locked in these wetland soils for millennia becomes available for aerobic and anaerobic decomposition, which releases carbon dioxide (CO_2) and methane (CH_4), respectively, back to the atmosphere. As CO_2 and CH_4 are potent greenhouse gases, this transfer of carbon from the land to the atmosphere further contributes to global warming, thereby increasing the rate of permafrost degradation in a positive feedback loop. Therefore, monitoring Arctic wetland health and dynamics is a key scientific task that is also of importance for policy. However, the identification and delineation of these important wetland ecosystems, remain incomplete and often inaccurate.

*Mapping the extent of Arctic wetlands remains a challenge for the scientific community. Conventional, coarser remote sensing methods are inadequate at distinguishing the diverse and micro-topographically complex nonvascular vegetation that characterize Arctic wetlands, presenting the need for better identification methods. To tackle this challenging problem, we constructed and annotated the first-of-its-kind Arctic Wetland Dataset (**AWD**). Based on that, we present **ArcticNet**, a deep neural network that exploits the multi-spectral, high-resolution imagery captured from nanosatellites (Planet Dove CubeSats) with additional Digital Elevation Model (DEM) from the Arctic-DEM project, to semantically label a Arctic study area into six types, in which three Arctic wetland functional types are included. We present multi-fold efforts to handle the arising challenges, including class imbalance, and the choice of fusion strategies. Preliminary results endorse the high promise of ArcticNet, achieving 93.12% in labelling a holdout set of regions in our Arctic study area.*

1. Background and Motivation

Over the past few decades, high-latitude environments have been undergoing fundamental structural and functional changes rapidly, as a result of rising temperatures. On land, these changes include intensified permafrost degradation, soil subsidence caused by melting of ground ice, formation of thermokarst terrain, longer and warmer growing seasons, changing fire regimes, and more [1][2][3]. Wetlands cover over 60% of the Arctic biome, of which a sizeable fraction consists of peatlands [4][5]. Peat soil accumulates over centennial to millennial timescales as a result of a net positive balance between plant production and peat decay. With growth surpassing decomposition, peat-accumulating wetlands (called peatlands) naturally sequester atmospheric carbon via photosynthesis. Under warming conditions, the fate of these soil carbon reservoirs has been questioned. On one hand, peatlands may benefit from this increased warmth by sequestering carbon at faster rates and migrating northward due to increasing plant production [6][7]. On the other hand, increasing peat decay from overall warmer and drier conditions, which would lower water table and provoke greater oxidation may occur [8].

Another important process worth considering is the transient response of wetlands to permafrost thaw. It has been shown that, in lowlands, water from permafrost thaw and melting of ground ice can form ponds, lakes, and recharge wetlands, making them even wetter. Under saturated conditions, anaerobic formation of CH_4 would increase, making wetlands strong carbon sources to the atmosphere [9][10]. Another possibility is that peat-forming plants colonize shallow ponds that develop following thaw and ground subsidence; under these conditions, peat accumulation has been shown to be extremely fast [11], which makes these young peatlands highly effective carbon sinks. Overall, the response of Arctic wetlands to warming and the resulting impacts on the global carbon cycle is still ambiguous: wetlands may either continue to act as carbon stores, with pos-

Figure 1. Overview of our proposed ArcticNet. The RGB and NIR/DEM/NDVI branches takes their corresponding modality from a same patch. Then, the feature map output by two networks are fused for classification. The semantic label map of the entire area is composed by sliding over the full image.

sible diminished levels, or may become carbon sources in the decades to come [12].

About 1700 billion tons of organic carbon is estimated to be stored in Arctic soils, double the amount presently in the atmosphere [13]. Wetland maps that rely on remote sensing and in-situ data are beginning to surface [14], but wetland location, spatial extent, and carbon reserve distribution need better characterization. To better understand wetland response to climate change, and make accurate estimates of carbon stocks and fluxes in wetlands, reliable and spatially-explicit representations of these ecosystems are needed.

Recently, Big Data coupled with new data analytics are already engendering paradigm shifts across disciplines and disrupting how research is conducted. There have been recent breakthroughs in satellite technology that make it possible to obtain daily to sub-weekly high-resolution imagery of the entire planet. Our pilot project explores the use of deep learning [15][16][17] to analyze this large volume of high spatial and high temporal resolution satellite-based images from the Arctic. Our end goal is to generate the first reliable Holarctic map of permafrost-affected ecosystems and address fundamental questions pertaining to the Arctic research. As current preliminary work, we have performed multi-fold efforts:

- We constructed the first-of-its-kind Arctic Wetland Dataset (AWD). The sample site of AWD is a 50km^2 area around the Scotty Creek Research Station, Northwest Territories, Canada, which has representative geospatial characteristics of Arctic wetlands from the discontinuous and sporadic permafrost regions. AWD includes high-resolution (3m), multispectral imagery (RGB + Near Infrared (NIR)) and 2m Digital Elevation Model (DEM). We then carefully annotated 500 $30m \times 30m$ regions or 10 pixel by 10 pixel squares in 3m imagery (15 pixel by 15 pixel in 2m imagery) into six major classes: *water*, *peat bog*, *channel fen*, *dense forest*, *sparse forest*, and *wetland*.

- We designed a patch-level multi-modal deep network to adaptively fuse the RGB and NIR/DEM/NDVI[1] modalities. In addition, we design an *Augmentation Balancing* strategy to address the class-imbalance roadblock, which is caused by non uniform distribution of each class in the real world. Being a high performance deep learning solution to generate semantic maps for an Arctic study area, we named it *ArcticNet*.

- We provided extensive ablation experiments of different models and fusion strategies, and observed significant gains by our progressive model improvements. A competitive accuracy of 93.12% was obtained by our final model in a hold-out testing set.

Despite calling our model ArcticNet to indicate our focus on Arctic vegetation mapping, the methodology is indeed broadly applicable to analyzing any high-resolution geospatial imagery. **We have open-sourced AWD[2], our associated codes and pre-trained ArcticNet models**, with the hope that they can benefit similar efforts in Earth observation and remote sensing.

2. Related work

2.1. Deep learning in remote sensing

Recently, deep learning has become increasingly integrated with high-resolution remotely sensed imagery [18][19]. [20] proposed a deep semantic segmentation model by changing the fully connected layer into convolutional layers. This technique has been well received at many remote sensing tasks, including road extraction[21], building detection[22] and land cover classification[23]. Ilke et al.[15] further developed an automatic generative algorithm to create street addresses from satellite imagery based on a deep learning road extraction technique. Due to the variety of remote sensing technologies and sensors (active vs.

---

[1]Normalized Difference Vegetation Index, extracted via Eqn. 1
[2]Please refer to our github repository

passive; radar, lidar, spectroradiometer, etc.), there are often multiple sources of overlapping data available for the same geographic areas. Therefore, many studies focus on the fusion of heterogeneous datasets. For example, multispectral imagery and depth data were fused using CNN features, hand-crafted features, and Conditional Random Fields (CRFs) by [24]. [16] further investigated fusion of remote sensing imagery in an end-to-end manner. OpenStreetMap (OSM) was also fused in [25]. Remote sensing data for one area through time can also be accessed and analyzed, making it a popular and useful data source for temporal analyses. [26] used the latter to identify forest type changes over a 20-year period. It was also used to predict traffic speed with long short-term memory neural network (LSTM) [27]. Other challenges in remote sensing image segmentation, such as high resolution [28], and low resolution and visual quality [29], are also addressed.

2.2. Wetlands classification

Remote sensing is being increasingly utilized for mapping wetlands and peatlands due to its extensive spatial coverage at low costs [30]. However, these ecosystems have presented a particular challenge to conventional remote sensing techniques due to 1) the incapability of optical and radar sensors to directly estimate peat depth, and therefore tease apart organic vs. mineral soils; 2) the difficulty to distinguish between wetland types due to similar vegetation cover, and 3) the seasonal changes in soil moisture in wetlands, which make sensors with high penetration capabilities ineffective [31]. To further complicate the matter, Arctic wetlands experience additional mapping challenges when vegetation is used as an identification proxy. Traditional remote sensing methods are tailored towards vascular vegetation, but a majority of Arctic wetland vegetation is highly heterogeneous, low-lying, and non-vascular, presenting complications. Additionally, Arctic vegetation has complex micro-topographies that are harder for commercial, coarser sensors to discern [10].

Hyperspectral sensors are starting to become more relevant in wetland studies (e.g. [32][33][34]), but this data is generally less commercially-available to scientists and is more cost- and labor-intensive. Therefore, a focus needs to be placed on developing better methods for wetland identification using multispectral imagery. This study presents a unique set of high-resolution (3m), multispectral imagery (RGB + NIR) acquired from CubeSats. These nanosatellites have been mostly used for educational purposes and only recently have started to be applied to Earth Science missions [35]. While CubeSats have inevitable limitations compared to large-scale acquisition systems with respect to payload constraints, data storage, geolocational controls, power, propulsion, and thermal control, CubeSats have shown high success [35]. Cooley et al.[36] have so far been the only study to apply CubeSat imagery to Arctic research, using a time-series of imagery to track surface water area changes in lakes. Their study suggests high success, using machine learning and object-based classifications to overcome some of the data accuracy limitations, thanks to the high spatial and temporal imagery resolutions.

Several studies started utilizing the deep learning method for wetland identification and classification. Siewert et al.[37] employed random forest to map soil organic carbon content in permafrost terrain across the (sub-)Arctic regions. Räsänen et al. [38] mapped areal coverage and changes in bare peat area over a peat plateau located in north-western Russia between 2007 and 2015 with random forest and achieved an F-score of 0.57. In our work, we show that by leveraging the deep learning method with the data fusion technique, we can achieve an accuracy as high as 93.12% in wetland classification.

3. Method

For developing our solution, we first introduce a new dataset named *Arctic Wetland Dataset* (**AWD**). We then detail the ArcticNet model, which consists of two single modality patch-wise classification models for RGB and NIR/DEM/NDVI bands respectively, followed by a learnable fusion step. The model is illustrated in Figure 1.

3.1. Arctic Wetland Dataset

Our first contribution is the Arctic Wetland Dataset (AWD). The sample site of AWD is a 50km^2 area located around the Scotty Creek Research Station, Northwest Territories, Canada ($61.3°N, 121.3°W$). Scotty Creek is an intensively studied watershed when it comes to permafrost dynamics (e.g. [39][40]), hydrology (e.g. [41][42][43]), and vegetation (e.g. [44][45][46][47]). Located in Canada's boreal region, it is an environment characterized by discontinuous to sporadic permafrost, coniferous forests, and wetlands ([48][47]).

We argue that this site is reasonably representative of Arctic wetlands due to being characterized by: 1) forested permafrost peat plateaus, which are elevated areas dominated by black spruce trees with an understory of shrubs, lichen, and mosses; 2) channel (flow-through) fens, which are lowland areas that act as water drainage pathways with some scattered trees, healthy sedges, and small-stature vegetation overlaying peat, and 3) peat bogs, which are blanketed by lichen and mosses, with some of the bogs having sparse dwarfed trees. However, we **do not claim** that this site suffices to train a model that can generalize to the whole Arctic area: it is merely a starting point. We are working to add multiple different study areas and increase the dataset size. The transferability and generalizability across different sites would be further investigated.
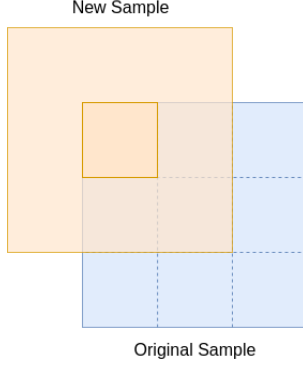
New Sample

Original Sample

Figure 2. *Augmentation Balancing*. We can pick one $10m \times 10m$ area on an original sample, to be the center to crop a new sample from the imagery, sharing the same label.

We assembled the AWD by collecting high-resolution (3m), multispectral imagery (RGB and Near Infrared (NIR) from CubeSats), as well as a 2m Digital Elevation Model (DEM) from the ArcticDEM project by the Polar Geospatial Center for complement. We then selected 500 non-overlapping local areas (or patches) of $30m \times 30m$ (10 pixels by 10 pixels for 3m imagery and 15 pixels by 15 pixels for 2m) from our study site. We categorized each patch into one of the six geospatial categories (as per the dominant type in this patch): *water*, *peat bog*, *channel fen*, *dense forest*, *sparse forest* and *wetland*. The categories of interest are *peat bog*, *channel fen* and *wetland*, since they are the three critical types of wetlands our study aims at disentangling. The patch classification was performed by a geoscientist with expertise in Arctic vegetation. In addition, a drone flight video footage covering part of the study area was used to distinguish between the six main classes due to its high-resolution coverage over the varying ecosystems in this watershed. Furthermore, many published classifications and vegetation surveys related to this watershed have been used as a cautious reference. Moreover, a combination of false color imagery and a normalized difference vegetation index (NDVI) were used to complement the visual identification of the patches. Lastly, we i.i.d. split the dataset to create a training set of 300 patches, a validation set of 100, and a hold-out testing set of 100. All data are aligned by latitude and longitude.

3.2. Single-Modality Backbone

To start with, we chose two single band modality mod we employed ResNet-50 [49] as the backbone patch-wise classification model for each single band modality (RGB, or NIR/DEM/NDVI band). For the RGB input, we stacked all three channels together. For the NIR/DEM/NDVI input, the Normalized Difference Vegetation Index (NDVI) band

was first calculated as

$$NDVI = \frac{NIR - R}{NIR + R} \qquad (1)$$

Then NIR and DEM are extracted from CubeSat and ArcticDEM project, respectively, and scaled to $(0, 1)$ in order to stay within "comparable" value ranges with the NDVI data. Afterward, they were stacked and fed to the network. The choice of these two single modalities follows the successful practice in [16], which explore the fusion of RGB with NDSM[3]/DSM[4]/NDVI. We replace the DSM with DEM, which is similar with it. However, NDSM band cannot be extracted from DEM. So we replace it with NIR for utilizing pretrained model which requires 3 channels' input.

Addressing the Class Imbalance   AWD closely follows and reflects the distribution of the six classes. However, from the perspective of model training, the class distribution is too unbalanced to ensure robust performance, especially when it comes to classifying sparse classes (such as *wetland*) that are actually critical for analyzing wetlands. The class distribution in one of the training sets (since we used 3-fold cross-validation, we have multiple training sets) could be found in Table 1 as *Original*. We propose an *Augmentation Balancing* to calibrate the class balance and to enable a more robust training.

Considering that the labeled patches are cropped from the holistic high-resolution image, a possible augmentation method could be first divide each original $30m \times 30m$ sample into nine $10m \times 10m$ sub-patches. Each sub-patch is then considered as a "center". One of nine "centers" is randomly chosen, to which a new $30m \times 30m$ sample is cropped. Finally, left-right and up-down flipping are applied to the newly-cropped $30m \times 30m$ sample in a probability of 0.5. This augmentation strategy is illustrated in Figure 2. This augmentation strategy is employed for generating new samples from old ones to make all the classes the same size in terms of their number of patches (same with the largest class from the original samples). The class distribution after *Augmentation Balancing* is also listed in Table 1 as *After*.

This approach is based on two reasons: 1) We observe that the geospatial characteristic changes smoothly in those selected areas; therefore, sliding the window for a little typically will not change the label. 2) Our goal is to classify the dominant class within each patch. Therefore, introducing a small part of other class will increase the robustness instead of hurting the model.

---

[3]Normalized Digital Surface Map
[4]Digital Surface Map

Table 1. The distribution of the training dataset before and after *Augmentation Balancing*

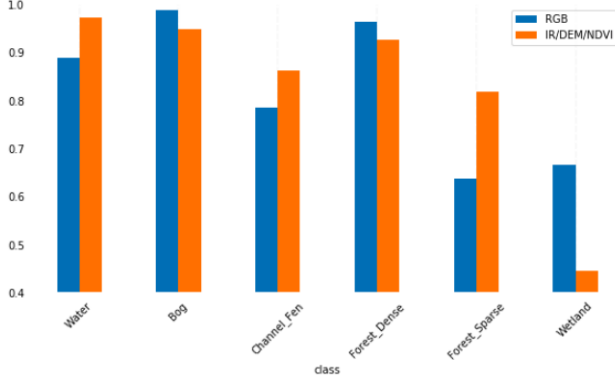| Phase	Water	Bog	Channel Fen	Forest Dense	Forest Sparse	Wetland
Original	29	94	44	52	60	15
After	94	94	94	94	94	94

Figure 3. Comparison of two single modality backbones with the accuracy on six classes.

3.3. Multi-Modality Fusion

We compared the performance of each single modality network as shown in Figure 3. Apparently, while both can classify the AWD reasonably well, there is a large discrepancy between their class-wise discriminative abilities. For example, NIR/DEM/NDVI performs better in discriminating *water*, *channel fen* and *sparse forests*, while RGB seems to be particularly more reliable at classifying *Bog*, *dense forest* and *wetlands*. A closer look shows that the two modalities too often yield very different classification confidences, or misaligned classification results, on individual patches. This result naturally motivated a fusion to explore their complementary powers.

Concepts from deep learning-based video classification were used to explore three different fusion structures [50]: early fusion, middle fusion and late fusion. The three fusion strategies represents different levels of flexibility control over fusing multi-modality information. Their structures are shown in Figure 4.

- **Early Fusion.** RGB and NIR/DEM/NDVI modalities are directly concatenated as the network input. The resulting model will be re-trained from scratch. We observed that in early fusion, the NIR/DEM/NDVI input needs to be re-normalized to the same scale as the RGB input, to ensure stable training.

- **Middle Fusion.** The feature maps produced by layer2 or layer3 of either single-modality backbone are concatenated together for the next stage joint processing (see the Middle Fusion Layer2 and Middle Fusion Layer3 in Figure 4, as two different progressive fusion
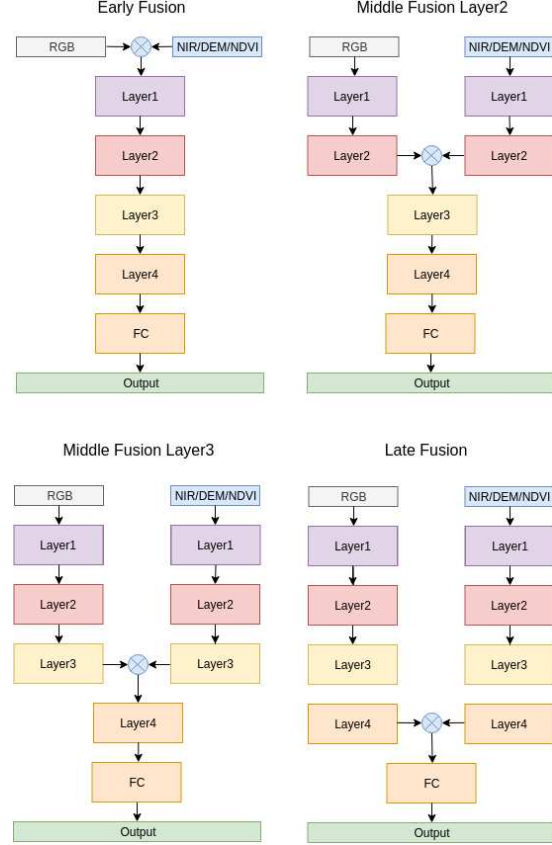

Figure 4. Different fusion structures explored in our work. Middle Fusion Layer2 and Middle Fusion Layer3 are two variations of Middle Fusion. $\otimes$ stands for concatenation operation.

ways that we tried). The resulting model will inherit all single modality model weights before the fusion, and re-train the layers after the concatenation step.

- **Late Fusion.** The activation vectors by layer4 of either single-modality backbone are concatenated as the input for the last fully connected layer for final classification. The resulting model will inherit all convolutional weights of both single modality models, and re-train only the last fully-connected layer.
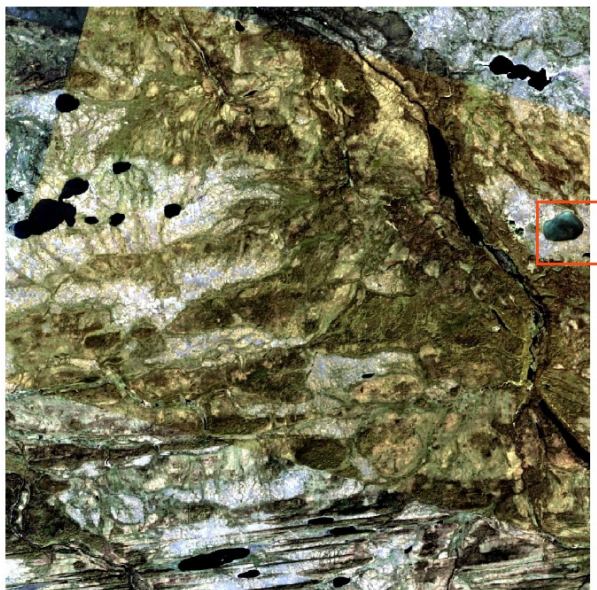
4. Experiment
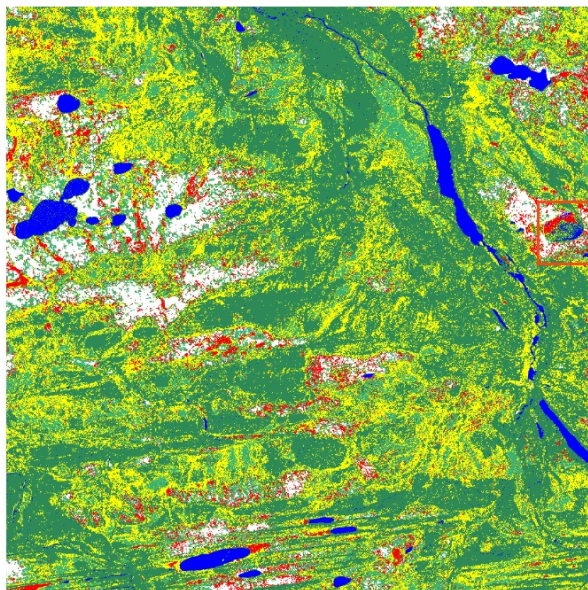
4.1. Training setting

We trained our models in AWD dataset in an end-to-end fashion. We used the ImageNet pre-trained ResNet-50 model as an initialization, and fine-tuned it on the AWD training set using Stochastic Gradient Descent (SGD). We employed the "poly" learning rate policy described in [51]:

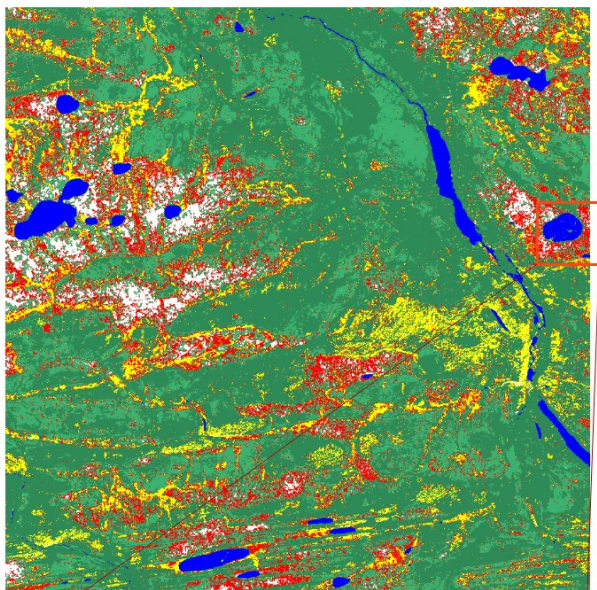$$r = r_i \times (1 - \frac{epoch}{max\ epoch})^{power}, \qquad (2)$$

(a) Remote Sensing Image

(b) Semantic Labeling of RGB Track Model

(c) Semantic Labeling of Late Fusion Model
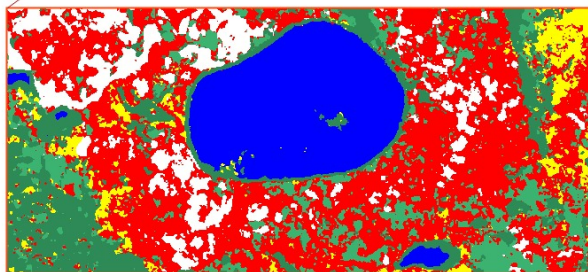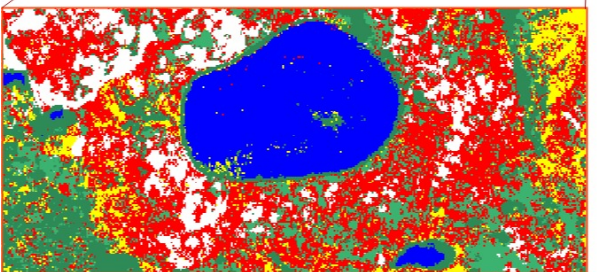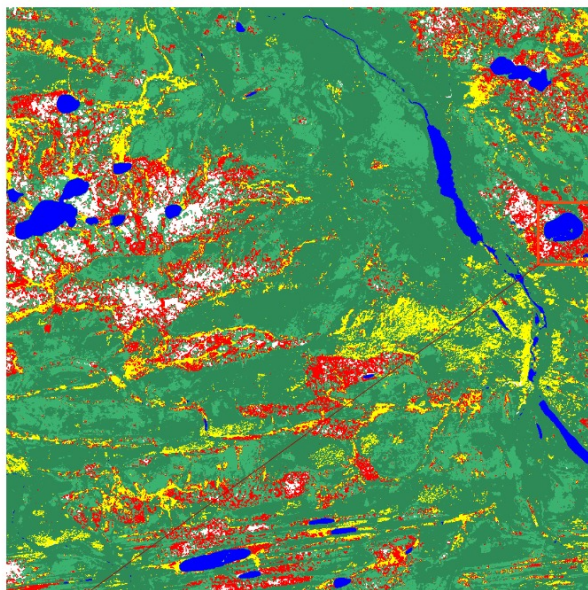
(d) Semantic Labeling of Late Fusion Model with CRF

Figure 5. The semantic labeling results of several variats of ArcticNet. The colors of blue, white, yellow, dark green, light green and red denote the classes of *water*, *bog*, *channel fen*, *dense forest*, *sparse forest* and *wetland*, respectively.

Table 2. The ablation experiment results of single modality backbones and fusion models. The first column indicates the name of each model. The first row lists the name of each class; and *Overall* stands for the averaged accuracy of all classes.

Model	Water	Bog	Channel Fen	Forest Dense	Forest Sparse	Wetland	Overall
RGB Track no balance	0.9167	**1.0**	0.8824	0.9383	0.6667	0.1111	0.7525
RGB Track with re-weighted loss	0.8889	**1.0**	0.8824	0.9630	0.6364	0.3333	0.7840
RGB Track	0.8889	0.9872	0.7843	0.9630	0.6364	0.6667	0.8211
NIR/DEM/NDVI Track	0.9722	0.9487	0.8627	0.9259	0.8182	0.4444	0.8287
Early Fusion	**1.0**	0.9744	0.8627	**1.0**	0.8333	0.5556	0.8710
Middle Fusion Layer2	**1.0**	0.9744	0.8235	0.9877	**0.9091**	0.7778	0.9121
Middle Fusion Layer3	0.9722	0.9744	**0.9020**	0.9877	0.8636	0.5556	0.8758
Late Fusion	0.9167	0.9615	0.8824	0.9630	0.8636	**1.0**	**0.9312**

where r_i is set as 1×10^{-4}, *power* is set as 4. The weight decay is set as 2×10^{-5} and the momentum is 0.5. The classification accuracy on the AWD testing set was used as the evaluation metric. A three-fold cross-validation was employed and the final accuracy is reported as the averaged result. Our model is trained on *Nvidia 1080 Ti*, and it takes around 2 hours and 2 GB GPU memory to train.

4.2. Performance evaluation and analysis

A single modality model on RGB is a strong baseline. We train the single modality backbone with only RGB images . The result is shown in Table 2, referred to as as *RGB Track no balance*. The overall accuracy achieves 75.25%, with some classes showing very high accuracy, e.g., *Bog*. We consider this result to be promising, in reference to performance levels reported in similar studies [38]. However, the accuracy of the *Wetland* class is very low (only 11.11%), as a result of suffering from class imbalance.

Augmentation Balancing yields significant improvement. When it comes to re-balancing classes, a common off-the-shelf option is to replace the standard cross entropy loss with a re-weighted loss [52]. The loss has a weight vector $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_n]$, which is calculated by the inverse class frequencies:

$$\alpha_i = \frac{\frac{1}{f_i}}{\sum_{q=0}^{n} \frac{1}{f_q}} \quad (3)$$

f_i represents the frequency of corresponding class. The resulting model, referred to as *RGB Track with re-weighted loss* in Table 2, shows overall accuracy improvement as well as on the *wetland* class accuracy. We then apply the proposed *Augmentation Balancing* to re-training the original model (without using re-weighted loss), and find it (called *RGB Track* in Table 2) to boost the performance even more than the competitive alternative of *RGB Track with re-weighted loss*. That endorses the remarkable effectiveness of *Augmentation Balancing* in conquering the class imbalance challenge. We therefore adopt *Augmentation Balancing* by default hereinafter.

A single modality model on NIR/DEM/NDVI performs reasonably well and provides complementary
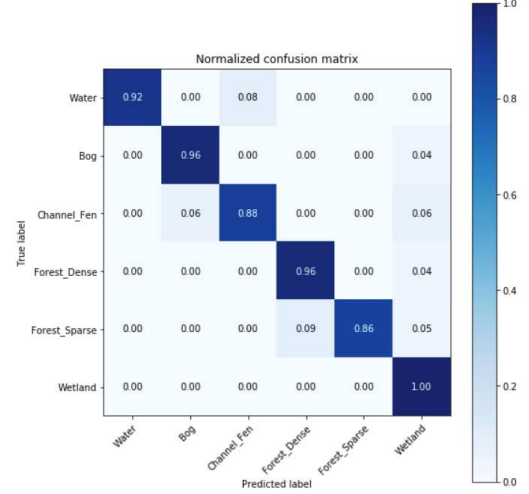

Figure 6. The confusion matrix of late fusion.

power. We trained single modality backbone on NIR/DEM/NDVI with *Augmentation Balancing*. The result was displayed in Table 2 as *NIR/DEM/NDVI Track*. As can be observed, its class-wise accuracy distribution notably differs from that of the RGB track. While its accuracy is higher in *water*, *channel fen* and *sparse forest*, the performance on other classes is less competitive than in RGB, especially on *wetland*. Meanwhile, the overall accuracy is comparable between these two tracks.

Fusion always helps, and Late Fusion helps the most. We conducted experiments based on the fusion options described in Section 3.3, including two variants of middle fusion. According to Table 2, all fusion methods seem to improve overall classification accuracy over either single modality backone, and usually lead to more "balanced" class-wise accuracies, e.g., remarkably improving the *wetland* class accuracy. The best model, using Late Fusion, achieves a high overall accuracy of 93.12% and achieves perfect class-wise accuracies on *wetland*. The confusion matrix of late fusion are given in Figure 6.

Different performance shows among different fusion strategies. Although we cannot explain why it perform differently well, this is consistent with the results show in the
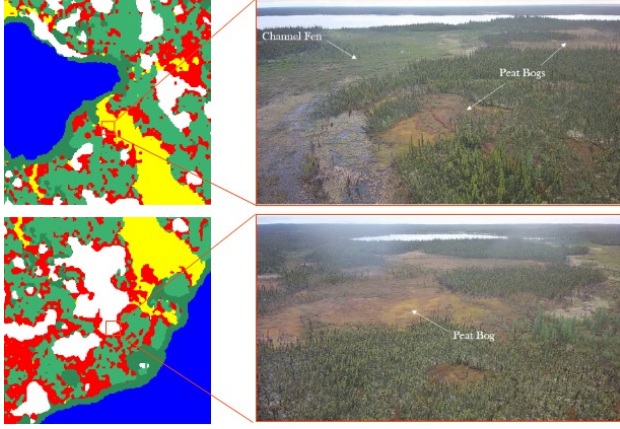
Figure 7. The semantic segmentation with the drone footage in the corresponding spot.

fusion experiment in [50].

Qualitative comparison. Figure 5 visualizes the semantic labeling maps. Figure 5.(a) shows the visualization of raw data for target area, which contains Mosaicking caused by software visualization algorithm (Notably, we used the raw data instead of this mosaicking visualization). From Figure 5.(b), which is the predicted map by the RGB Track model, we can clearly see that the lake region depicted in the red bounding box suffers from substantial misclassifications. Also, a large number of areas are mistakenly classified as *(Channel Fen)*, which were verified to be incorrect by geospatial experts. By applying late fusion, those problems seem to be well alleviated. Further, we apply conditional random field (CRF) [24] as a common post-processing tool in semantic segmentation to further enhance spatial consistency. Comparing the zoom in part of Figure 5.(c) and 5.(d), CRF alleviates the occasional image aliasing and suppresses isolated outliers. For example, the wrongly classified wetland pixels in the lake are eliminated after CRF post-processing. Figure 7 shows some of our semantic segmentation results with the drones photos in depicted areas. From the high resolution of the drone imagery, we can identify the key vegetation differences of wetland functional types, proving the classifying results' correctness. Channel or flow-through fens can be identified by their lowland nature and general behavior as hydrological pathways. These ecosystems are additionally uniquely characterized by healthy sedges, short-story vegetation, and sparse trees which overlay peat. Peat bogs, on the other hand, are characterized by peat-covering lichen and moss, with some having scattered, small trees.

Results make sense for geospatial experts. With geospatial experts on the project, we have confirmed that the results are reasonable, primarily in the Late Fusion model. The RGB Track model highly overestimates *channel fens*, when these ecosystems should only comprise about 20% of

the area. This is likely due to the fact that at 3m, it becomes difficult to distinguish *channel fens* based solely on visible wavelengths that cannot capture the high amounts of healthy biomass or soil moisture that are the most defining features of these ecosystems. The Late Fusion model appears to still overestimate *channel fens* in certain areas but to a much lower degree. Overall, the Late Fusion model succeeds where the RGB Track model fails by incorporating NIR information which extends the models ability to distinguish environments based on biophysical characteristics that the visible wavelengths cannot capture. A limitation with both models is the ability to classify other features that may be present (for example, there is a road throughout the area which is being grouped into the *water* and *bog* classes in the Late Fusion model). Despite this issue, all the model results are interpretable, with the Late Fusion model results the most reasonable.

5. Conclusion

In this work, we first introduced Arctic Wetland Dataset (AWD), a first-of-its-kind dataset sampled from a representative Arctic wetlands, with six important classes annotated. Based on this dataset, the ArcticNet was proposed for classifying imagery patches. With techniques from dataset class balancing and multi-modality fusion, our model achieves a high accuracy of 93.12%, as well as qualitatively promising semantic label maps of the whole region. However, this work is still preliminary for mapping the whole Arctic area. In the future work, we would extend the study area and increase the dataset size. The transferability and generalizability of the model in different sites would be further investigated. Moreover, we plan explore the semi-supervised training on the AWD, to fully unleash the power of our collected (albeit unannotated) data.

Future work will also involve further analysis on different band composites (i.e. false color composite) and even different backbone neural networks that may distinguish *channel fens* better. By using different variations of band combinations and backbones, we should be able to capture different types of ecosystem information throughout the study area which should help resolve model overestimation/underestimation. We will also apply these models to even higher-resolution CubeSat imagery (1m) that we have for our study area. With the higher-resolution imagery captured over the same area for a duration of 3 months, we will test how the models classification performs when there are distinct phenological or ecosystem changes present throughout time.

References

[1] Gensuo J Jia, Howard E Epstein, and Donald A Walker. Greening of arctic alaska, 1981–2001. *Geophysical Research Letters*, 30(20), 2003.

[2] David G Vaughan, Josefino C Comiso, Ian Allison, Jorge Carrasco, Georg Kaser, Ronald Kwok, Philip Mote, Tavi Murray, Frank Paul, Jiawen Ren, et al. Observations: cryosphere. *Climate change*, 2103:317–382, 2013.

[3] F Stuart Chapin, M Sturm, Mark C Serreze, JP McFadden, JR Key, AH Lloyd, AD McGuire, TS Rupp, AH Lynch, Joshua P Schimel, et al. Role of land-surface changes in arctic summer warming. *science*, 310(5748):657–660, 2005.

[4] Katrin Kohnert, Bennet Juhls, Sina Muster, Sofia Antonova, Andrei Serafimovich, Stefan Metzger, Jörg Hartmann, and Torsten Sachs. Toward understanding the contribution of waterbodies to the methane emissions of a permafrost landscape on a regional scalea case study from the mackenzie delta, canada. *Global change biology*, 24(9):3976–3989, 2018.

[5] Tatiana Minayeva, Andrey Sirin, Peter Kershaw, and Olivia Bragg. Arctic peatlands. *The Wetland Book: II: Distribution, Description, and Conservation*, pages 275–288, 2018.

[6] Miriam C Jones and Zicheng Yu. Rapid deglacial and early holocene expansion of peatlands in alaska. *Proceedings of the National Academy of Sciences*, 107(16):7347–7352, 2010.

[7] Julie Loisel and Zicheng Yu. Recent acceleration of carbon accumulation in a boreal peatland, south central alaska. *Journal of Geophysical Research: Biogeosciences*, 118(1):41–53, 2013.

[8] Takeshi Ise, Allison L Dunn, Steven C Wofsy, and Paul R Moorcroft. High sensitivity of peat decomposition to climate change through water-table feedback. *Nature Geoscience*, 1(11):763, 2008.

[9] A Harris, RG Bryant, and AJ Baird. Detecting near-surface moisture stress in sphagnum spp. *Remote Sensing of Environment*, 97(3):371–381, 2005.

[10] M Kalacska, M Lalonde, and TR Moore. Estimation of foliar chlorophyll and nitrogen content in an ombrotrophic bog from hyperspectral data: Scaling from leaf to image. *Remote Sensing of Environment*, 169:270–279, 2015.

[11] Miriam C Jones, Robert K Booth, Zicheng Yu, and Paul Ferry. A 2200-year record of permafrost dynamics and carbon cycling in a collapse-scar bog, interior alaska. *Ecosystems*, 16(1):1–19, 2013.

[12] Jianghua Wu and Nigel T Roulet. Climate change reduces the capacity of northern peatlands to absorb the atmospheric carbon dioxide: The different responses of bogs and fens. *Global Biogeochemical Cycles*, 28(10):1005–1024, 2014.

[13] Edward AG Schuur and Benjamin Abbott. Climate change: High risk of permafrost thaw. *Nature*, 480(7375):32, 2011.

[14] Jiren Xu, Paul J Morris, Junguo Liu, and Joseph Holden. Peatmap: Refining estimates of global peatland distribution based on a meta-analysis. *Catena*, 160:134–140, 2018.

[15] Ilke Demir, Forest Hughes, Aman Raj, Kleovoulos Tsourides, Divyaa Ravichandran, Suryanarayana Murthy, Kaunil Dhruv, Sanyam Garg, Jatin Malhotra, Barrett Doo, et al. Robocodes: towards generative street addresses from satellite imagery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–10, 2017.

[16] Nicolas Audebert, Bertrand Le Saux, and Sébastien Lefèvre. Beyond rgb: Very high resolution urban remote sensing with multimodal deep networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 140:20–32, 2018.

[17] Chao Tian, Cong Li, and Jianping Shi. Dense fusion classmate network for land cover classification. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 262–2624. IEEE, 2018.

[18] Andong Ma, Anthony M Filippi, Zhangyang Wang, and Zhengcong Yin. Hyperspectral image classification using similarity measurements-based deep recurrent neural networks. *Remote Sensing*, 11(2):194, 2019.

[19] Zhangyang Wang, Nasser M Nasrabadi, and Thomas S Huang. Semisupervised hyperspectral classification using task-driven dictionary learning with laplacian regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3):1161–1173, 2014.

[20] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[21] Jun Wang, Jingwei Song, Mingquan Chen, and Zhi Yang. Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine. *International Journal of Remote Sensing*, 36(12):3144–3169, 2015.

[22] Maria Vakalopoulou, Konstantinos Karantzalos, Nikos Komodakis, and Nikos Paragios. Building detection in very high resolution multispectral data with deep learning features. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1873–1876. IEEE, 2015.

[23] Nataliia Kussul, Mykola Lavreniuk, Sergii Skakun, and Andrii Shelestov. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5):778–782, 2017.

[24] Sakrapee Paisitkriangkrai, Jamie Sherrah, Pranam Janney, Van-Den Hengel, et al. Effective semantic pixel labelling with convolutional networks and conditional random fields. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 36–43, 2015.

[25] Nicolas Audebert, Bertrand Le Saux, and Sébastien Lefèvre. Joint learning from earth observation and openstreetmap data to get faster better semantic maps. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 67–75, 2017.

[26] Manqi Li, Jungho Im, and Colin Beier. Machine learning approaches for forest classification and change analysis using multi-temporal landsat tm images over huntington wildlife forest. *GIScience & Remote Sensing*, 50(4):361–384, 2013.

[27] Xiaolei Ma, Zhimin Tao, Yinhai Wang, Haiyang Yu, and Yunpeng Wang. Long short-term memory neural network

for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54:187–197, 2015.

[28] Wuyang Chen, Ziyu Jiang, Zhangyang Wang, Kexin Cui, and Xiaoning Qian. Collaborative global-local networks for memory-efficient segmentation of ultra-high resolution images. *arXiv preprint arXiv:1905.06368*, 2019.

[29] Ding Liu, Bihan Wen, Xianming Liu, Zhangyang Wang, and Thomas S Huang. When image denoising meets high-level vision tasks: A deep learning approach. *arXiv preprint arXiv:1706.04284*, 2017.

[30] KJ Lees, T Quaife, RRE Artz, M Khomik, and JM Clark. Potential for using remote sensing to estimate carbon fluxes across northern peatlands–a review. *Science of the Total Environment*, 615:857–874, 2018.

[31] ON Krankina, D Pflugmacher, M Friedl, WB Cohen, P Nelson, and A Baccini. Meeting the challenge of mapping peatlands with remotely sensed data. *Biogeosciences*, 5(6):1809–1820, 2008.

[32] Thierry Erudel, Sophie Fabre, Thomas Houet, Florence Mazier, and Xavier Briottet. Criteria comparison for classifying peatland vegetation types using in situ hyperspectral measurements. *Remote Sensing*, 9(7):748, 2017.

[33] Nanfeng Liu, Paul Budkewitsch, and Paul Treitz. Examining spectral reflectance features related to arctic percent vegetation cover: Implications for hyperspectral remote sensing of arctic tundra. *Remote sensing of environment*, 192:58–72, 2017.

[34] Zachary Langford, Jitendra Kumar, Forrest Hoffman, Amy Breen, and Colleen Iversen. Arctic vegetation mapping using unsupervised training datasets and convolutional neural networks. *Remote Sensing*, 11(1):69, 2019.

[35] Daniel Selva and David Krejci. A survey and assessment of the capabilities of cubesats for earth observation. *Acta Astronautica*, 74:50–68, 2012.

[36] Sarah W Cooley, Laurence C Smith, Jonathan C Ryan, Lincoln H Pitcher, and Tamlin M Pavelsky. Arctic-boreal lake dynamics revealed using cubesat imagery. *Geophysical Research Letters*, 2019.

[37] Matthias B Siewert. High-resolution digital mapping of soil organic carbon in permafrost terrain using machine learning: a case study in a sub-arctic peatland environment. *Biogeosciences*, 15(6):1663–1682, 2018.

[38] Aleksi Räsänen, Vladimir Elsakov, and Tarmo Virtanen. Usability of one-class classification in mapping and detecting changes in bare peat surfaces in the tundra. *International Journal of Remote Sensing*, pages 1–21, 2019.

[39] Alastair F McClymont, Masaki Hayashi, Laurence R Bentley, and Brendan S Christensen. Geophysical imaging and thermal modeling of subsurface morphology and thaw evolution of discontinuous permafrost. *Journal of Geophysical Research: Earth Surface*, 118(3):1826–1837, 2013.

[40] KM Haynes, RF Connon, and WL Quinton. Permafrost thaw induced drying of wetlands at scotty creek, nwt, canada. *Environmental Research Letters*, 13(11):114001, 2018.

[41] Masaki Hayashi, William L Quinton, Alain Pietroniro, and John J Gibson. Hydrologic functions of wetlands in a discontinuous permafrost basin indicated by isotopic and chemical signatures. *Journal of Hydrology*, 296(1-4):81–97, 2004.

[42] W L Quinton and JL Baltzer. The active-layer hydrology of a peat plateau with thawing permafrost (scotty creek, canada). *Hydrogeology Journal*, 21(1):201–220, 2013.

[43] Katheryn Burd, Suzanne E Tank, Nicole Dion, William L Quinton, Christopher Spence, Andrew J Tanentzap, and David Olefeldt. Seasonal shifts in export of doc and nutrients from burned and unburned peatland-rich catchments, northwest territories, canada. *Hydrology & Earth System Sciences*, 22(8), 2018.

[44] L Chasmer, W Quinton, C Hopkinson, R Petrone, and P Whittington. Vegetation canopy and radiation controls on permafrost plateau evolution within the discontinuous permafrost zone, northwest territories, canada. *Permafrost and Periglacial Processes*, 22(3):199–213, 2011.

[45] L Chasmer, C Hopkinson, T Veness, W Quinton, and J Baltzer. A decision-tree classification for low-lying complex land cover types within the zone of discontinuous permafrost. *Remote Sensing of Environment*, 143:73–84, 2014.

[46] Marie-Ève Garon-Labrecque, Étienne Léveillé-Bourret, Kellina Higgins, and Oliver Sonnentag. Additions to the boreal flora of the northwest territories with a preliminary vascular flora of scotty creek. *The Canadian Field-Naturalist*, 129(4):349–367, 2016.

[47] Michael A Merchant, Justin R Adams, Aaron A Berg, Jennifer L Baltzer, William L Quinton, and Laura E Chasmer. Contributions of c-band sar data and polarimetric decompositions to subarctic boreal peatland mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(4):1467–1482, 2017.

[48] Jerry Brown, OJ Ferrians Jr, JA Heginbottom, and ES Melnikov. *Circum-Arctic map of permafrost and ground-ice conditions*. US Geological Survey Reston, VA, 1997.

[49] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[50] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 1725–1732, 2014.

[51] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2018.

[52] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.