2017 IEEE International Conference on Data Mining Workshops

Convolutional Neural Network Approach for Mapping Arctic Vegetation using Multi-Sensor Remote Sensing Fusion

Zachary L. Langford*, Jitendra Kumar† and Forrest M. Hoffman†
*University of Tennessee, Knoxville, TN, USA
Email: zlangfor@vols.utk.edu
† Oak Ridge National Laboratory, Oak Ridge, TN, USA

***Abstract*—Accurate and high-resolution maps of vegetation are critical for projects seeking to understand the terrestrial ecosystem processes and land-atmosphere interactions in Arctic ecosystems, such as U.S. Department of Energy's Next Generation Ecosystem Experiment (NGEE) Arctic. However, most existing Arctic vegetation maps are at a coarse resolution and with a varying degree of detail and accuracy. Remote sensing-based approaches for mapping vegetation, while promising, are challenging in high latitude environments due to frequent cloud cover, polar darkness, and limited availability of high-resolution remote sensing datasets (e.g., $\sim$ 5 m). This study proposes a new remote sensing based multi-sensor data fusion approach for developing high-resolution maps of vegetation in the Seward Peninsula, Alaska. We focus detailed analysis and validation study around the Kougarok river, located in the central Seward Peninsula of Alaska.**

We seek to evaluate the integration of hyper-spectral, multi-spectral, radar, and terrain datasets using unsupervised and supervised classification techniques over a $\sim$343.72 km^2 area for generating vegetation classifications at a variety of resolutions (5 m and 12.5 m). We fist applied a quantitative goodness-of-fit method, called *Mapcurves*, that shows the degree of spatial concordance between the public coarse resolution maps and k-means clustering values and relabels the k values based on the best overlap. We develop a convolutional neural network (CNN) approach for developing high resolution vegetation maps for our study region in Arctic. We compare two CNN approaches: (1) breaking up the images into small patches (e.g., 6 x 6) and predict the vegetation class for entire patch and (2) semantic segmentation and predict the vegetation class for every pixel. We also perform accuracy assessments of the developed data products and evaluate varying CNN architectures. The fusion of hyperspectral and optical datasets performed the best, with accuracy values increased from 0.64 to 0.96-0.97 when using a training map produced by unsupervised clustering and *Mapcurves* labeling for both CNN models.

I. Introduction

The Arctic has emerged as an important focal point for the study of climate change due to the greatest regional warming on Earth in recent decades, which was twice the rate of the global mean warming [1]. Increased warming in the Arctic will have potential climate feedbacks [2]–[4] and consequently influence ecosystems [5], [6]. Current climate models in the Arctic project an ongoing temperature increase for the next decades [3]. Many studies suggest that environmental changes associated with a warmer climate could have considerable consequences for terrestrial ecosystems, such as Arctic plant communities [7]. Continued greening over the next century will produce multiple climate feedbacks. Model experiments in the Arctic demonstrate that the associated feedbacks with vegetation are positive, and failing to account for them in future climate studies may result in inaccurate projections [8].

Model evaluations for future climate studies also show a poleward advance of the forest–tundra boundary, an expansion of tall shrub tundra, and a dominance shift from deciduous to evergreen boreal conifer forest [7], [9]. Therefore, an accurate accounting of Arctic vegetation is important for parameterization of models that reflect accurate feedbacks to climate change. For example, [10] used WorldView-2 and LiDAR datasets to create $\sim$0.25 m spatial resolution plant functional type (PFT) datasets for driving land-surface models in Barrow, AK. Kumar et. al, 2016 [11] utilizing similar remote sensing datasets to characterize the microtopography of polygonal tundra employed the open-source PFLOTRAN (massively parallel subsurface flow and reactive transport) code to simulate variability in subsurface thermal regimes of ice-wedge polygons at the field-scale at Barrow, AK.

High-resolution vegetation datasets are needed for current modeling projects in the Arctic, such as the Next-Generation Ecosystem Experiments (NGEE Arctic). NGEE Arctic is a Department of Energy (DOE) project that seeks to improve our confidence in global climate projections through a coordinated series of model-inspired investigations (http://ngee-arctic.ornl.gov/). NGEE Arctic is working to couple 3-Dimensional surface-subsurface hydrology model PFLOTRAN with Accelerated Climate Model for Energy (ACME) Land Model (ALM) for modeling Arctic climate change at a variety of scales. High-resolution spatial datasets are needed to drive these models (e.g., vegetation, topographic, and hydrologic data). It's important to evaluate remote sensing imagery that can provide datasets to drive high-resolution models and guide field sampling campaigns. NGEE Arctic currently is conducting extensive field research at a series of sites across the Seward Peninsula, Alaska. The selection of the Seward Peninsula is based on an analysis indicating that western Alaska is a proxy for the future ecological and climatic regime of the North Slope of Alaska toward the end of the century [12].

Approaches based on hyper-spectral imagery, also named

IEEE computer society

imaging spectroscopy, have shown to be effective for identification a wide range of vegetation types [13], including Arctic vegetation types [14], [15]. However, this has mostly been shown through hand held or airborne hyper-spectral sensors [15]. Multi-spectral datasets (e.g., IKONOS, SPOT, WorldView) are also becoming more available in Arctic and Boreal ecosystems. Synthetic aperture radar (SAR) is also a promising remote sensing dataset for vegetation classifications, enabling imaging in all weather conditions at any time of day or night. These different remote sensing platforms vary in their sensor characteristics, spatial and temporal resolution and have their strengths and weaknesses. While traditionally due to their characteristics these remote sensing data sets have been used independently, while multi-sensor data fusion holds promise to harvest their complementary strengths to characterize vegetation properties. Machine learning algorithms that are based on multiple datasets (i.e., multi/hyper-spectral images) are needed to develop and evaluate high-resolution vegetation maps using data from multiple remote sensing platforms.

The resolution and accuracy of current publicly available vegetation maps are often low and are too inaccurate to be used for field-scale studies. New methods and approaches are needed to increase the accuracies at a higher resolution. Objective of this paper is to 1) explore the predictability of multi-sensor data fusion, 2) develop Covolutional Neural Networks (CNN) for predicting vegetation type on the landscape using two approaches, (2A) by breaking up the images into small patches (e.g., 6 x 6) and predicting the vegetation class for entire patch; and (2B) evaluate localization using semantic segmentation. These approaches help identify the optimal CNN network and multi-sensor data fusion for Arctic vegetation mapping.

II. Materials

A. Study Area

The study area selected for this research is located on the Seward Peninsula on the western coast of Alaska. We focused our analysis near a watershed in the Seward Peninsula (Figure 1), due to the data collected for the NGEE Arctic project. The Seward Peninsula experiences a semi-maritime climate that is controlled by the Bering Sea and sea surface conditions (especially sea–ice extent) [16]. The Seward Peninsula lies at the important zone of vegetation transitions from boreal forest to tundra [17], making this part of Alaska an important region for studying vegetation characteristics. Most of the Seward Peninsula is treeless, with some smaller isolated mountain ranges and varied geologic substrates, including limestone, deep unglaciated loess, and lava flows of various ages [18].

The bounding region of our study area was determined by the overlapping raster boundaries from the SPOT-5 and EO-1 Hyperion imagery. The study area is located near the Kougarok River, located in the center of Seward Peninsula, Alaska, and consists of low Arctic tundra dominated by open low mixed shrub-sedge tussock tundra communities [19]. The mean annual temperature for this area is -2.4°C and with the mean annual temperature in July is 11.0°C; and the mean annual precipitation is 102.1 mm [20] . This region is a zone of nearly continuous permafrost with an active layer that has a mean thickness of 56 cm [21].

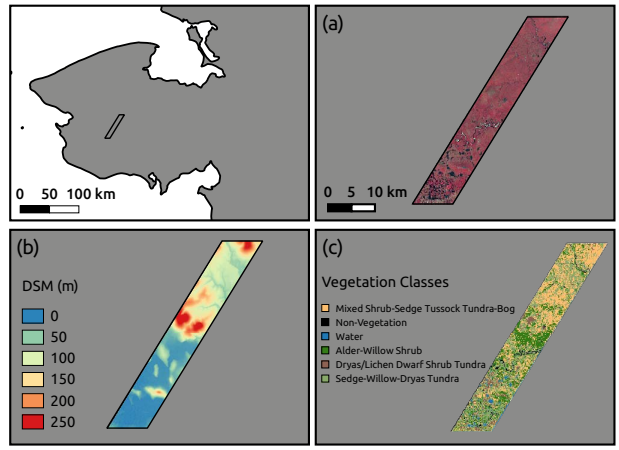


Fig. 1. Study area based on the EO-1 Hyperion footprint, showing (a) SPOT-5 false color image, (b) digital elevation model, and (c) vegetation classes.

B. Remote Sensing Datasets

All datasets were processed in Alaska Albers Equal Area Conic (EPSG:3338), NAD83 horizontal datum and NAVD88 vertical datum. Table I lists the remote sensing products in this study. One ALOS PALSAR L-band SAR image with dual polarization (FBD) of HH (horizontal transmitting, horizontal receiving) and HV (horizontal transmitting, vertical receiving) that was radiometric and terrain corrected was collected, more information can be found at https://www.asf.alaska.edu/sar-data/palsar/. The EO-1 Hyperion and Landsat 8 OLI images were collected from USGS https://earthexplorer.usgs.gov/, with EO-1 Hyperion consisting of 198 calibrated bands (0.4 - 2.5 μm) and Landsat 8 OLI consisting of 11 bands, with 9 bands (0.4 - 2.29 μm) used in this study. The terrain corrected versions for EO-1 Hyperion (L1T) and Landsat 8 OLI (Level-1) were processed. The Digital Surface Model (DSM) IfSAR and SPOT-5 data were collected from the Geographic Information Network of Alaska (http://gina.alaska.edu/). The SPOT-5 satellite images used in this study were gathered by the "Alaska Statewide Digital Mapping Initiative", which produced a new statewide orthomosaic that provides complete multispectral coverage of the state at 2.5 m spatial resolution. We used the CIR SPOT-5 orthoimage over our study area, which consists of the green (0.50 - 0.59 μm), red (0.61 - 0.68 μm), and near infrared (0.78 - 0.89 μm) bands.

C. Vegetation Data

We gathered the 30 m Alaska Existing Vegetation Type (AKEVT) (http://akevt.gina.alaska.edu/) vegetation map for our study region for training and validating our models. Some vegetation classes were very small and merged into other classes of similar strata. This was also done based on the

TABLE I
LIST OF REMOTE SENSING DATASETS USED IN THIS STUDY

Dataset	Resolution	Acquisition Date
EO-1 Hyperion	30 m	24 June 2015
Landsat 8 OLI	30 m	17 August 2016
ALOS-1 PALSAR	12.5 m	29 August 2007
IfSAR DSM	5 m	Summer 2012
SPOT-5	2.5 m	Summer 2009 – 2012

TABLE II
AREA OF THE AKEVT VEGETATION CLASSES FOR THE STUDY REGION (343.72 KM^2).

Class	Area (km^2)	Percentage
Mixed Shrub-Sedge Tussock Tundra-Bog	120.75	35.13%
Alder-Willow Shrub	74.33	21.63%
Dryas/Lichen Dwarf Shrub Tundra	20.52	5.97%
Sedge-Willow-Dryas Tundra	116.41	33.87%
Water	6.13	1.78%
Non-Vegetation	5.58	1.62%

Alaska vegetation classification system [19]. For example, the willow shrubs covered only 0.38 % of the study region and was merged into the alder-willow shrub class. We cropped the AKEVT map to our study region and condensed the classes down to six vegetation types that exist in study region: mixed shrub-sedge tussock tundra-bog, alder-willow shrub, dryas/lichen dwarf shrub tundra, sedge-willow-dryas tundra, water, and non-vegetation. Table II shows the area of the six chosen vegetation classes across the study region.

III. METHODOLOGY

Figure 2 shows a flowchart of methodology developed in this study and we describe the steps in detail in the following subsections.

A. Data Processing

The terrain corrected versions for EO-1 Hyperion (L1T) and Landsat 8 OLI (Level-1) were processed. The EO-1 Hyperion and Landsat 8 OLI datasets were converted to Top Of Atmosphere (TOA) reflectance according to the USGS documentation. The SPOT-5 satellite images used in this study were gathered by the "Alaska Statewide Digital Mapping Initiative", which have three statewide mosaics available at 2.5 m spatial resolution: color infrared (CIR), psuedo-natural color, and panchromatic (grayscale). We used the CIR SPOT-5 orthoimage over our study area, which consists of the green (0.50 - 0.59 μm), red (0.61 - 0.68 μm), and near infrared (0.78 - 0.89 μm) bands. Quantum Spatial and Fugro Geospatial, Inc. performed the image processing, orthorectification, and mosaicing of the datasets. The SPOT-5 orthoimage was radiometrically corrected for tone, balance, and geometry quality control along tile edges for terrain and linear features. The SPOT-5, PALSAR, and IfSAR datasets were normalized to between 0 and 1 for consistency among the TOA datasets. The Lee filtering algorithm (7×7 window) was applied to ALOS-1 PALSAR image to reduce the speckle [22]. All of the remote sensing datasets were re-sampled by pixel aggregation to 5 m and 12.5 m using the nearest-neighbor re-sampling.

TABLE III
MULTI-SENSOR VEGETATION CLASSIFICATION CASES

Case	Fusion Combinations of Remote Sensing Datasets
D1	EO-1, IfSAR
D2	EO-1, Landsat 8 OLI, IfSAR
D3	EO-1, ALOS-1 PALSAR, IfSAR
D4	EO-1, SPOT-5, IfSAR
D5	EO-1, ALOS-1 PALSAR, SPOT-5, IfSAR
D6	EO-1, ALOS-1 PALSAR, Landsat 8 OLI, IfSAR
D7	EO-1, SPOT-5, Landsat 8 OLI, IfSAR
D8	Landsat 8 OLI, IfSAR
D9	Landsat 8 OLI, ALOS-1 PALSAR, IfSAR
D10	Landsat 8 OLI, SPOT-5, IfSAR
D11	ALOS-1 PALSAR, IfSAR
D12	ALOS-1 PALSAR, SPOT-5, IfSAR
D13	SPOT-5, IfSAR
D14	EO-1, Landsat 8 OLI, ALOS-1 PALSAR, SPOT-5, IfSAR

To explore the predictability of various remote sensing data sets in multi-sensor data fusion approach, we designed fourteen sets of data to conduct our case studies, which are describe in Table III. We will refer to each specific dataset as D by number (e.g., D14 refers to merging all the datasets together).

B. Training Dataset

High quality training dataset is critical for building robust and accurate CNN models. We employed two different data sets to train our CNN models.

1) Alaska Existing Vegetation Map: AKEVT is a landcover type map produced at 30m for a portion of Western Alaska Landscape Conservation Cooperative (WALCC) region by US Fish and Wildlife Services. Map was produced using field plot data, LANDSAT 7 ETM+ spectral data and environmental variables. AKEVT is one of the best publicly available vegetation map for the data poor region of the Seward Peninsula of Alaska, however, resolution is coarser than the target resolution of our current study.

2) Unsupervised Classification-Based Vegetation Map (UCVM): Unsupervised multi-variate clustering techniques have been widely used in Earth science for delineation of ecoregions based on gridded datasets [23], [24]. We applied unsupervised classification on remote sensing data set (Table I) to develop a high resolution map and improve the accuracy of AKEVT map for use as training data set. We performed a principal component analysis (PCA) analysis of the remote sensing datasets (Table I) and selected the top three components that explain 99% variance. We used a k-means algorithm to cluster the datasets into groups containing similar spectral characteristics [25].

While our unsupervised classification technique identifies groups of similar characteristics, the output classes are defined by their spectral properties and lack translation to an understandable vegetation type. The *Mapcurves* method developed by [26] provides a map algebraic approach to compare two

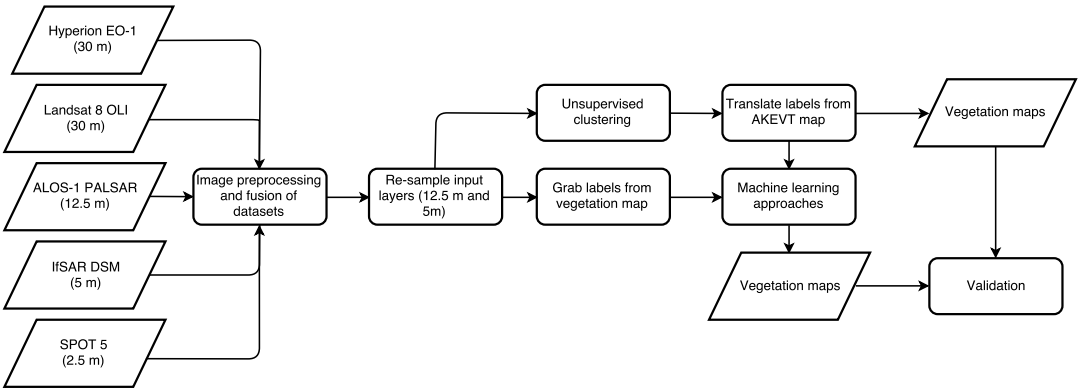


Fig. 2. Flowchart of our study for fusion of remote sensing datasets and to generate high-resolution vegetation maps.

maps. The *Mapcurves* method [26] calculates a goodness-of-fit (GOF) score that indicates the degree of spatial correspondence between polygons in two different maps, in our case the AKEVT map and the remote sensing datasets. GOF is a unitless measure of spatial overlap between map categories:

$$GOF = \sum_{polygons} \frac{C}{B+C} x \frac{C}{A+C}$$

where A is the map that is being compared, B is the reference map, and C is the proportion of the reference category (B) that shares with the tested category (A). The GOF weighs the presence of the intersecting categories, each of which might share a small or large spatial intersection with the category being tested. The GOF score is a unitless metric. In our case, a high GOF score would represent high spatial overlap with the clusters (k) from the unsupervised clustering classifications and the AKEVT map.

We used *Mapcurves* to identify the best "translation table" between dataset clusters and vegetation types defined by the AKEVT map. We take the clusters with the highest GOF score and relabel the map accordingly. The *Mapcurves* reclassifies the remote sensing based unsupervised classification maps (*UCVM*) into AKEVT vegetation classes.

When used for training, the key purpose for using unsupervised clustering and *Mapcurves* method was to evaluate if updating the AKEVT vegetation classes using this simple method could increase the accuracies of the CNN models. We finally evaluate the accuracy scores based on the *UCVM* map and the AKEVT map to determine which dataset is appropriate to train our CNN models.

C. Convolutional Neural Network

Deep learning neural networks, which learn the representative and discriminative features in a hierarchical manner from the data, are becoming more efficient at image classification in the remote sensing community [27]. Convolutional Neural Networks (CNN) has shown excellent performance in many computer vision and machine learning problems [28]. CNN take advantage of the fact that the input consists of images and they constrain the architecture in a more sensible way. In particular, unlike a regular Neural Network, the layers of a CNN have neurons arranged in 3 dimensions: width, height, depth (i.e., number of channels). CNN has one or more convolutional layers, often followed by one or more fully connected layers [29]. Conventionally, a full CNN architecture consists of Convolutional Layer, Pooling Layer, and Fully-Connected Layer [29]. CNNs have become a popular algorithm in the satellite remote sensing community [27]. Chen et al., 2014 [30] applied autoencoders to learn representations (i.e. features) of hyper-spectral image data in an unsupervised manner. Wu et al., 2017 [31] used a convolutional recurrent neural network (CRNN) to learn more discriminative features for hyper-spectral data classification. Zhang et al., 2017 [32] built a deep CNN with limited satellite image samples, using a transfer learning approach by sharing the common image features of the natural images. However, most of these approaches are built upon coarse and small datasets, such as Indian Pines and Kennedy Space Center. Additionally, these methods are solely built upon a particular dataset and don't seek multi-sensor fusion.

We use supervised CNN models throughout the paper, which maps the input image (x_i) over a series of layers to a probability vector ($\hat{y_i}$) over the different classes. The typical use of CNN is on classification tasks, where the output to an image is a single class label, such as the ImageNet competition [33]. However we would also like to evaluate pixel-wise labeling using CNNs. We implemented two CNN approaches using the TensorFlow [34] and Keras framework [35] in Python: 1) our first approach uses patches of images as inputs into a CNN architecture, where the entire patch is predicted as a single vegetation class; and 2) the use of semantic segmentation, where a class label is assigned to each pixel. We refer to the first and second approach as the patch-level and pixel-level approach, respectively.

Both CNNs training processes start with the weights of all networks randomly initialized, and the initial learning rate is set to 10^{-4}. We used a batch size of 150 in all experiments, which defines the number of samples

that is going to be propagated through the network. The `sparse_categorical_accuracy` metric was used in Keras using 10% of the data for validation while rest for training. `sparse_categorical_accuracy` metric calculates the mean accuracy rate across all predictions for our multi-class classification problem. Categorical accuracy metric was selected in Keras because of the sparse and categorical nature of the objective function. The accuracy score was computed by taking the number of correctly predicted samples and dividing by the total number of samples.

1) Patch-Level CNN Architecture: We decided to test three different patch sizes of 3×3, 6×6, and 12×12, where each patch contains 9, 36, and 144 pixels, respectively. Each layer of the CNN consists of the convolution of the previous layer output with a set of learned filters, the passing the responses through a rectified linear function ($relu(x)$ = max(x,0)), pooling over local neighborhoods, and local contrast operation that normalizes the responses across feature maps [29], [33]. The top few layers of the network are conventional fully-connected networks and the final layer is a softmax classifier [29]. Figure 3 shows our first CNN architecture used in this study.

2) Pixel-Level CNN Architecture: Semantic segmentation for images can be defined as the process of partitioning and classifying the image into meaningful parts, and classify each part at the pixel level into one of the pre-defined classes. Previous studies on segmentation break up the images into patches in order to increase the number of training images [36]. Larger patches require more max-pooling layers that reduce the localization accuracy, while small patches allow the network to see only little context [36]. More recent approaches [37] proposed a classifier output that takes into account the features from multiple layers.

Our semantic segmentation CNN model was inspired by the family of u-net architectures, where low-level feature maps are combined with higher-level ones, which enables precise localization [36]. This type of network architecture was especially designed to effectively solve image segmentation problems. This architecture allowed for more detail in the segmentation by using shortcut connections from each layer [36]. The contracting path consists of the repeated application of two 3×3 convolutions, each followed by a Rectified Linear Unit (ReLU) and a 2×2 max pooling operation with stride 2 for downsampling [36]. Every step in the expansive path consists of an upsampling of the feature map followed by a 2×2 convolution (up-convolution), a concatenation with the correspondingly feature map from the contracting path, and two 3×3 convolutions, each followed by a ReLU [36]. The number of feature channels during the downsampling and upsampling stayed the same at 64. At the final layer a 1×1 convolution is used to map each 64-component feature vector to the vegetation classes using the softmax function. Figure 4 shows the network architecture for the pixel-level CNN approach.

IV. Results and Discussion

1) Unsupervised Classification based Vegetation Map (UCVM): Unsupervised k-means clustering algorithm was used to identify patterns of vegetation using remote sensing data set. Top three dominant principal components from remote sensing data set (Table I) were classified at varying levels of division (k = 10, 25, 50). While classifying data among fewer number of clusters (k) tends to lump vegetation classes together, using larger number of cluster may define clusters with small differences in remote sensing signatures across vegetation types. For that reason, we conducted our unsupervised classification at varying levels of division from small to large. *Mapcurves* however is capable of identifying small difference between categories and eliminating them by reclassifying the similar categories as one.

We applied *Mapcurves* to add labels to the unsupervised classifications, identify unique vegetation classes while collapsing similar ones. While the *Mapcurves* classification method was able to identify unique vegetation types, it also identified where outliers or noisy data could exist within the AKEVT map. Figure 5 shows an example of using the *Mapcurves* method to generate vegetation map using unsupervised k-means clustering. Figure 5(c) shows the re-labeled vegetation map (UCVM) generated from k=25 clustering map (Figure 5(a))using data set D14 (EO-1, Landsat 8 OLI, ALOS-1 PALSAR, SPOT-5, IfSAR), showing large differences from the original AKEVT map (i.e., accuracy of 0.51) (Figure 5(b)). Table IV shows the accuracy of unsupervised clustering based vegetation map using *Mapcurves*.

We performed the analysis at both 5 m and 12.5 m resolutions. Increasing the spatial resolution from 5 m to 12.5 m did not increase the accuracy substantially. This could be due to the fact that the AKEVT map was only available at 30 m resolution. Cases using the hyper-spectral datasets displayed the best accuracies (D1–D7 and D14), with unsupervised classification at k = 25 and k = 50 levels of division performing the best. After assessing the accuracy scores, we used the D4 and D14 for 5 m and 12.5 m datasets, respectively. These datasets were used in addition to AKEVT maps to train our CNN models.

A. Patch-Level CNN Approach

We developed patch-level CNN models for patch sizes of 6×6, 12×12, and 16×16. Models were trained using AKEVT and UCVM data sets. Table V shows the accuracy values for our patch-level CNN models corresponding to the image patch (width and height), resolution, and training data set used (i.e., AKEVT and UCVM). The CNN models trained using AKEVT had accuracies ranging from 52% to 68%. The hyper-spectral datasets performed the best with the AKEVT labels, ranging from 59% to 68% accurate, with patch size of 6×6 and 12×12 achieving the highest scores. Increasing the accuracy from 12.5 m to 5 m seems to increase the accuracy for certain datasets and patch size. For example, for D14 we saw an increase of 10% when increasing the spatial resolution to 5 m. D11 (ALOS-1 PALSAR, IfSAR) performed poorly

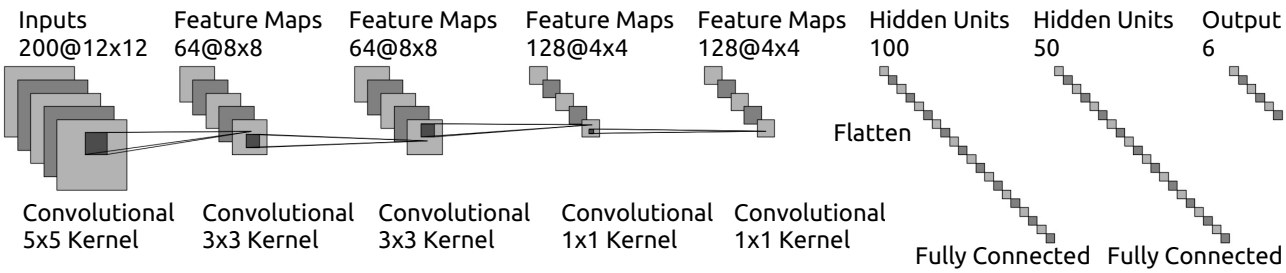


Fig. 3. Example architecture of a seven layer patch-based CNN model. A 12 by 12 crop of an image (with 200 bands) is presented as the input. This is convolved with 64 different 1st layer filters, with a kernel size of 5×5, using pooling with stride 2. Similar operations are repeated in the following layers. The last two layers are fully connected, taking features from the top convolutional layer as inputs. The final layer is a softmax function, corresponding to the number of classes.

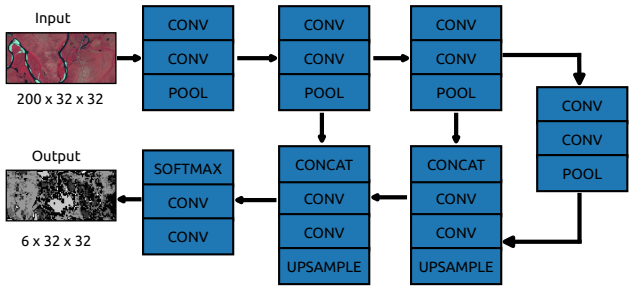


Fig. 4. An example of the u-net pixel-based architecture from taking a 32×32 pixel patch with 200 bands. The architecture has a series of convolutional and max pooling layers when downsampling. We then perform upsampling and feature concatenating, which means this process uses information from the previous layers. There is a constant number of 64 filters throughout the network. The final output feature maps are fed to a softmax classifier for pixel-wise classification.

TABLE IV
ACCURACY VALUES FOR THE VEGETATION MAPS (UCVM) PRODUCED BY THE *Mapcurves* METHOD USING UNSUPERVISED CLASSIFICATION AT DIFFERENT LEVELS (k) AT 5 M AND 12.5 M RESOLUTION, WITH CASES WITH BEST SCORES INDICATED BY BOLD FONT.

	k=10		k=25		k=50	
Case	5 m	12.5 m	5 m	12.5 m	5 m	12.5 m
D1	0.47	0.46	0.50	0.51	0.54	0.53
D2	0.47	0.46	0.50	0.51	0.54	0.54
D3	0.47	0.46	0.50	0.51	0.54	0.53
D4	0.47	0.46	0.52	0.52	**0.55**	0.54
D5	0.46	0.46	0.50	0.51	0.54	0.54
D6	0.47	0.45	0.49	0.52	0.53	0.54
D7	0.47	0.46	0.51	0.51	0.54	0.54
D8	0.44	0.45	0.47	0.47	0.48	0.48
D9	0.44	0.44	0.46	0.46	0.47	0.46
D10	0.45	0.45	0.47	0.47	0.48	0.50
D11	0.44	0.44	0.44	0.44	0.45	0.45
D12	0.45	0.44	0.48	0.46	0.50	0.47
D13	0.44	0.44	0.48	0.48	0.49	0.50
D14	0.47	0.46	0.50	0.51	0.54	**0.55**

the when using AKEVT for training, with accuracy values around 52%. This could be due to SAR datasets generally have high variations due to speckle or to different types of scattering mechanisms [38]. However, several multi-sensor sets that included SAR didn't see drop in performance. For

TABLE V
ACCURACY OF THE PATCH-LEVEL CNNS USING VARYING PATCH SIZES (I.E., 6×6, 12×12, AND 16×16) AT 5 M AND 12.5 M RESOLUTION USING AKEVT AND UCVM MAPS FOR TRAINING.

		AKEVT			UCVM		
		6×6	12×12	16×16	6×6	12×12	16×16
D1	5 m	0.64	0.65	0.63	0.94	0.93	0.91
	12.5 m	0.63	0.63	0.59	0.95	0.93	0.87
D2	5 m	0.66	0.66	0.66	0.94	0.92	0.91
	12.5 m	0.65	0.64	0.59	0.95	0.92	0.85
D3	5 m	0.64	0.64	0.64	0.94	0.93	0.90
	12.5 m	0.64	0.62	0.61	**0.96**	0.92	0.88
D4	5 m	0.66	0.67	0.66	0.95	0.93	0.91
	12.5 m	0.66	0.64	0.61	0.95	0.92	0.87
D5	5 m	0.67	0.67	0.66	0.95	0.93	0.91
	12.5 m	0.67	0.64	0.63	0.95	0.92	0.88
D6	5 m	0.66	0.67	0.66	0.94	0.92	0.90
	12.5 m	0.66	0.64	0.61	0.95	0.92	0.87
D7	5 m	0.67	0.68	0.67	0.95	0.93	0.91
	12.5 m	0.67	0.66	0.59	0.95	0.93	0.86
D8	5 m	0.63	0.64	0.65	0.78	0.77	0.78
	12.5 m	0.64	0.62	0.59	0.82	0.80	0.80
D9	5 m	0.63	0.64	0.63	0.78	0.77	0.77
	12.5 m	0.65	0.63	0.59	0.83	0.82	0.80
D10	5 m	0.66	0.67	0.67	0.80	0.79	0.78
	12.5 m	0.67	0.67	0.62	0.83	0.81	0.81
D11	5 m	0.52	0.53	0.53	0.67	0.66	0.67
	12.5 m	0.53	0.52	0.52	0.72	0.74	0.73
D12	5 m	0.62	0.63	0.63	0.76	0.76	0.76
	12.5 m	0.64	0.64	0.56	0.81	0.80	0.78
D13	5 m	0.62	0.63	0.64	0.76	0.75	0.76
	12.5 m	0.64	0.64	0.56	0.81	0.80	0.78
D14	5 m	0.67	0.68	0.67	**0.96**	0.93	0.91
	12.5 m	0.67	0.65	0.57	0.95	0.93	0.86

example, D12 (ALOS-1 PALSAR, SPOT-5, IfSAR) and D13 (SPOT-5, IfSAR) performed the same using the AKEVT or UCVM data sets for training.

The UCVM vegetation map increased the accuracy for all datasets by a large margin (Table V) when used for training. The highest scores were achieved for the hyper-spectral datasets, having 96% accuracy for D3 (EO-1, ALOS-1 PALSAR, IfSAR) for the 12.5 m dataset. The highest score from the 5 m datasets was achieved using D14 dataset with a 96% accuracy. The Landsat 8, SPOT-5, and SAR datasets (D8, D9, D10, D11, D12, and D13) perform better when trained using the AKEVT map with an accuracy range of 66% to

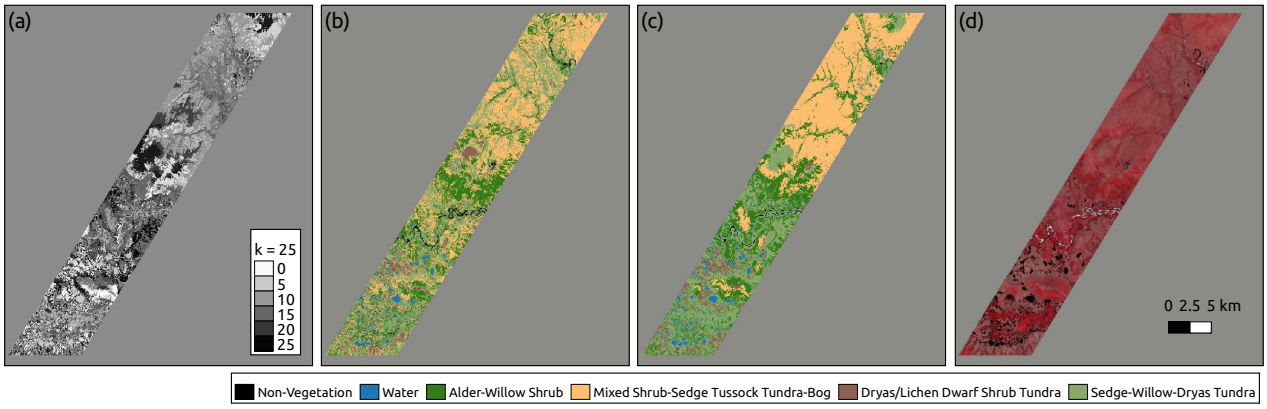


Fig. 5. (a) Unsupervised classification map (k=25) from using D14 (EO-1, Landsat 8 OLI, ALOS-1 PALSAR, SPOT-5, IfSAR); (b) The original AKEVT vegetation maps; (c) UCVM vegetation map generated using *Mapcurves* method; (d) and EO-1 Hyperion false color image.

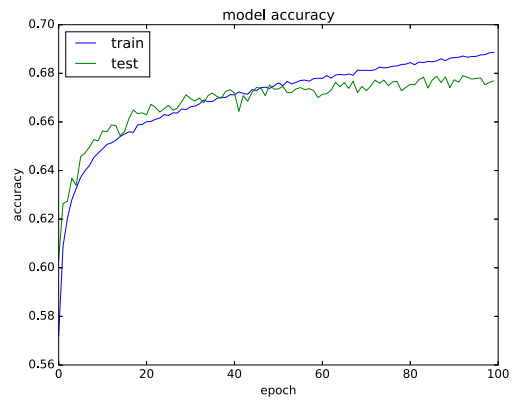
83%, however this is 10% lower than models developed using the hyper-spectral datasets.

Figure 6 shows the accuracy scores for D14 at 5 m resolution when increasing the epoch to 100 when using AKEVT and UCVM vegetation maps for training. When trained using AKEVT, each case showed overfitting after certain epochs. Figure 6(a) shows the model overfitting after epoch 45, indicating that the training dataset was getting substantially higher accuracies than the test dataset. However, when trained using UCVM (Figure 6(b)) the accuracy of the test dataset (i.e., 10% of the data) was always higher than the training dataset, meaning that model overfitting was not happening and the CNN architecture was performing well.

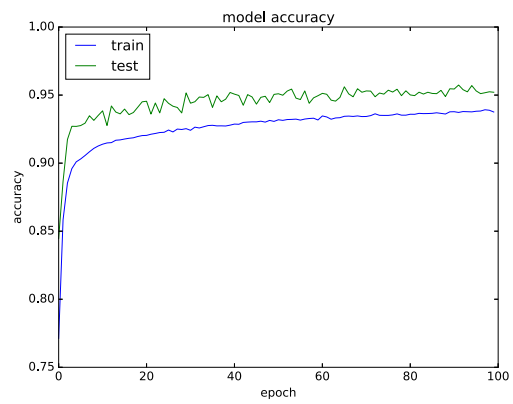
Patch size 6×6 performed the best for the UCVM training data set, however the drop in performance was lower when increasing the resolution to 5 m. We believe that small patch sizes work better since a 6×6 patch may contain many vegetation classes, but the number of samples used in the model training is higher. However, in the patch-level CNN approach, each image patch is input to the model independently, which means that only the "intra-patch" context information is considered [39]. The correlations between patches are not taken into account, which might lead to gaps between patches [39]. However, this may only apply to objects with strong continuity, such as urban features, which is not present in our imagery. Additionally, other patch-level CNN architectures may perform better with large patch sizes (i.e., 32×32) [40].

B. Pixel-Level CNN Approach

Table VI shows the accuracy values for our pixel-level CNN models corresponding to the image patch width and height, resolution, and training vegetation map used (i.e., AKEVT and UCVM) . The CNN models developed using the AKEVT labels had accuracies ranging from 54% to 68%. D11 (ALOS-1 PALSAR, IfSAR) performed the worst with the AKEVT labels, similar to the patch-level CNN architecture. D7 (EO-1, SPOT-5, Landsat 8 OLI, IfSAR), D10 (Landsat 8 OLI, SPOT-5, IfSAR), and D14 (EO-1, Landsat 8 OLI, ALOS-1 PALSAR,



(a) The model accuracy for D14 by varying the epoch to 100 using the AKEVT for training and patch size of 6×6.



(b) The model accuracy for D14 using the UCVM for training and varying the epoch to 100.

Fig. 6. CNN Model accuracies for D14. The green line represents the validation set (10% of data) and the blue line represents the training dataset (90% of data).

TABLE VI
ACCURACY VALUES FOR THE AKEVT AND MAPCURVES VEGETATION MAPS PRODUCED BY U-NET VARYING BY PATCH SIZE (I.E., 16, 32, 64) AT 5 M AND 12.5 M RESOLUTION.

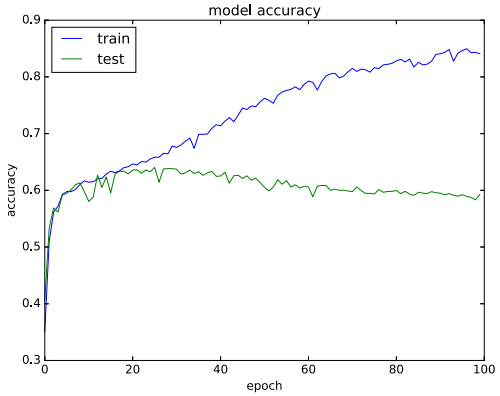
		AKEVT			Mapcurves		
		16	32	64	16	32	64
D1	5 m	0.64	0.62	0.61	0.95	0.92	0.89
	12.5 m	0.61	0.59	0.61	0.94	0.90	0.86
D2	5 m	0.66	0.64	0.61	0.94	0.92	0.91
	12.5 m	0.62	0.60	0.61	0.94	0.92	0.87
D3	5 m	0.64	0.63	0.62	0.94	0.93	0.89
	12.5 m	0.62	0.60	0.60	0.94	0.90	0.86
D4	5 m	0.67	0.65	0.63	**0.97**	0.93	0.89
	12.5 m	0.64	0.60	0.62	0.94	0.89	0.86
D5	5 m	0.67	0.65	0.62	**0.97**	0.94	0.89
	12.5 m	0.64	0.62	0.62	**0.95**	0.90	0.86
D6	5 m	0.66	0.65	0.63	0.94	0.92	0.90
	12.5 m	0.64	0.60	0.63	0.94	0.91	0.87
D7	5 m	0.68	0.65	0.62	0.95	0.93	0.91
	12.5 m	0.64	0.62	0.62	0.94	0.91	0.86
D8	5 m	0.64	0.62	0.61	0.76	0.74	0.74
	12.5 m	0.61	0.60	0.56	0.80	0.79	0.80
D9	5 m	0.65	0.65	0.65	0.76	0.75	0.75
	12.5 m	0.65	0.62	0.63	0.82	0.79	0.80
D10	5 m	0.68	0.66	0.66	0.78	0.76	0.76
	12.5 m	0.65	0.64	0.65	0.83	0.80	0.82
D11	5 m	0.55	0.55	0.56	0.67	0.66	0.68
	12.5 m	0.55	0.54	0.55	0.74	0.71	0.74
D12	5 m	0.66	0.65	0.63	0.76	0.75	0.74
	12.5 m	0.64	0.63	0.63	0.81	0.79	0.81
D13	5 m	0.65	0.65	0.64	0.76	0.73	0.75
	12.5 m	0.65	0.62	0.63	0.81	0.78	0.81
D14	5 m	0.68	0.66	0.64	0.96	0.93	0.91
	12.5 m	0.64	0.62	0.63	0.94	0.90	0.87

SPOT-5, IfSAR) performed the best with the AKEVT at 5 m and 16×16 patch size. Increasing the resolution and having a patch size of 16×16 worked best for our segmentation architecture. The smaller patch size worked best, due to the increase number of samples used for training the network. However, larger patch sizes may work best for features with strong continuity [39].

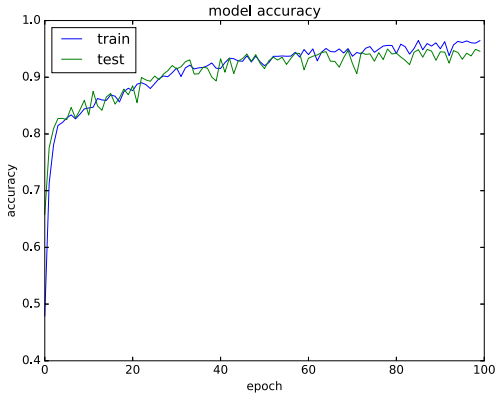
The UCVM vegetation map increased the accuracy for all datasets by a large margin (Table VI), similar to the patch-level CNN model. The highest scores were achieved from the hyper-spectral datasets, having 97% accuracy for D4 (EO-1, SPOT-5, IfSAR) and D5 (EO-1, ALOS-1 PALSAR, SPOT-5, IfSAR) for the 5 m dataset. The highest score from the 12.5 m datasets were the hyper-spectral datasets (D1 – D7 and D14) with a 94%-95% accuracy. The SAR dataset (D11) performed the worst with 66%-67% accuracy. The Landsat, SAR, and SPOT datasets (D8, D9, D10, D12, and D13) had accuracies from 0.77% to 0.83%.

C. CNN Architectures

Increasing the accuracy from 12.5 m to 5 m didn't have a significant impact on the accuracies. This is probably due to the nearest-neighbor re-sampling method used in this study. Sophisticated re-sampling methods could be performed for a more rigorous analysis. For example, deep learning methods can perform image super-resolution [41]. Additionally, other image enhancement methods could be performed that is specific to hyper-spectral imagery [42].



(a) The model accuracy for D5 by varying the epoch to 100 using the AKEVT for training and patch size of 16×16.



(b) The model accuracy for D5 using the *Mapcurves* for training and varying the epoch to 100.

Fig. 7. Model accuracies for D5. The green line represents the validation set (10% of data) and the blue line represents the training dataset (90% of data).

The pixel-level segmentation architecture consists of a contracting path and a symmetric expanding path that capture context and localization. This architecture performs well with very little training data available, by applying elastic deformations to the available training images [36]. This allows the network to learn invariance to such deformations, without the need to see these transformations in the annotated image corpus [36]. However, other segmentation algorithms could be looked into for comparisons. For example, a segmentation framework specifically tailored to remote sensing images could be applied [43]. Looking into other activation functions, such as the exponential linear units (ELUs), could provide performance increase [43]. Additionally, looking more closely into the hyperparameters, (e.g., learning rate, loss function, weight initialization) could help optimize our CNN approaches.

Smaller patch sizes helped for both CNN models, most likely due to the structure and shape of the vegetation classes.

We believe that the number of training examples didn't significantly impact our model accuracies, since we didn't see a high performance gain when increasing the resolution to 5 m. Better image enhancement methods may be necessary (e.g., super-resolution) in order to verify that small patch sizes work better for satellite remote sensing, especially in the case of Arctic vegetation mapping. Patch aggregation approaches could also be looked into that better utilize the interactions of patch features, such as constructing multiple, shared columns in the neural network and feeding multiple patches to each of the columns [44].

V. Conclusion

We implemented an CNN approach to develop Arctic vegetation map using remote sensing data sets for a range of sensors and platforms. We evaluated two CNN architectures, in one approach we break up the images into small patches (e.g., 6×6) and estimate the vegetation class for entire patch and the other approach uses semantic segmentation and pixel-wise labeling. Both model approaches had similar accuracy metrics for all cases. Unsupervised clustering and the *Mapcurves* method based vegetation map (UCVM) helped improve the accuracy scores compared to using the AKEVT map alone. We believe the unsupervised clustering and the *Mapcurves* method help initially reduce the noise associated with the AKEVT map, since all the *Mapcurves* scores were generally low (45-50%).

We explored and demonstrated the potential of multi-sensor remote sensing fusion for mapping vegetation in data sparse region of Arctic. Among data sets considered in the study, hyper-spectral remote sensing datasets provided highest data content to our CNN models, indicating that the spectral signatures of vegetation with high bandwidths are important for predictability of vegetation. SAR datasets performed the worst, however further evaluation of SAR datasets is needed. The Landsat and SPOT-5 datasets demonstrated improved predictability when using the UCVM for training and were important data set for our analysis. We believe the CNN architectures performed in this study provides an accurate method for Arctic vegetation mapping. However, we identified several different approaches that could improve our vegetation maps. Methods developed in this study helps enhance and reduce noise in publicly available vegetation maps, utilizing high-resolution datasets and open-source toolkits (e.g., Keras and Tensorflow). Approaches like these are needed for to provide accurate and high-resolution datasets needed for current modeling projects in the Arctic.

Acknowledgment

The Next Generation Ecosystem Experiments (NGEE) Arctic project (http://ngee-arctic.ornl.gov/) is supported by the Office of Biological and Environmental Research in Department of Energy Office of Science. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

References

[1] T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013.

[2] I. Mahlstein and R. Knutti, "September Arctic sea ice predicted to disappear near 2°C global warming above present," *Journal of Geophysical Research: Atmospheres*, vol. 117, no. D6, pp. n/a–n/a, 2012, d06104. [Online]. Available: http://dx.doi.org/10.1029/2011JD016709

[3] J. E. Overland, M. Wang, J. E. Walsh, and J. C. Stroeve, "Future Arctic climate changes: Adaptation and mitigation time scales," *Earth's Future*, vol. 2, no. 2, pp. 68–74, 2014. [Online]. Available: http://dx.doi.org/10.1002/2013EF000162

[4] E. A. G. Schuur, A. D. McGuire, C. Schädel, G. Grosse, J. W. Harden, D. J. Hayes, G. Hugelius, C. D. Koven, P. Kuhry, D. M. Lawrence, S. M. Natali, D. Olefeldt, V. E. Romanovsky, K. Schaefer, M. R. Turetsky, C. C. Treat, and J. E. Vonk, "Climate change and the permafrost carbon feedback," *Nature*, vol. 520, no. 8, pp. 171–179, 2015.

[5] L. Hinzman, N. Bettez, W. Bolton, F. Chapin, M. Dyurgerov, C. Fastie, B. Griffith, R. Hollister, A. Hope, H. Huntington, A. Jensen, G. Jia, T. Jorgenson, D. Kane, D. Klein, G. Kofinas, A. Lynch, A. Lloyd, A. McGuire, F. Nelson, W. Oechel, T. Osterkamp, C. Racine, V. Romanovsky, R. Stone, D. Stow, M. Sturm, C. Tweedie, G. Vourlitis, M. Walker, D. Walker, P. Webber, J. Welker, K. Winker, and j. Yoshikawa, "Evidence and implications of recent climate change in Northern Alaska and other Arctic regions," *Climatic Change*, vol. 72, no. 3, pp. 251–298, 2005.

[6] L. E. Ogden, "Plants Duke It Out in a Warming Arctic," *BioScience*, 2015.

[7] W. Zhang, P. A. Miller, B. Smith, R. Wania, T. Koenigk, and R. Dscher, "Tundra shrubification and tree-line advance amplify arctic climate warming: results from an individual-based dynamic vegetation model," *Environmental Research Letters*, vol. 8, no. 3, p. 034023, 2013. [Online]. Available: http://stacks.iop.org/1748-9326/8/i=3/a=034023

[8] P. A. Miller and B. Smith, "Modelling tundra vegetation response to recent Arctic warming," *Ambio*, vol. 41, no. 3, pp. 281–291, 2012. [Online]. Available: http://dx.doi.org/10.1007/s13280-012-0306-1

[9] L. Stewart, I. G. Alsos, C. Bay, A. L. Breen, C. Brochmann, N. Boulanger-Lapointe, O. Broennimann, H. Bltmann, P. K. Bcher, C. Damgaard, F. J. A. Danils, D. Ehrich, P. B. Eidesen, A. Guisan, I. S. Jnsdttir, J. Lenoir, P. C. le Roux, E. Lvesque, M. Luoto, J. Nabe-Nielsen, P. Schnswetter, A. Tribsch, L. U. Tveraabak, R. Virtanen, D. A. Walker, K. B. Westergaard, N. G. Yoccoz, J.-C. Svenning, M. Wisz, N. M. Schmidt, and L. Pellissier, "The regional species richness and genetic diversity of Arctic vegetation reflect both past glaciations and current climate," *Global Ecology and Biogeography*, vol. 25, no. 4, pp. 430–442, 2016. [Online]. Available: http://dx.doi.org/10.1111/geb.12424

[10] Z. Langford, J. Kumar, F. M. Hoffman, R. J. Norby, S. D. Wullschleger, V. L. Sloan, and C. M. Iversen, "Mapping Arctic plant functional type distributions in the Barrow Environmental Observatory using WorldView-2 and LiDAR datasets," *Remote Sensing*, vol. 8, no. 9, p. 733, Sep. 2016.

[11] J. Kumar, N. Collier, G. Bisht, R. T. Mills, P. E. Thornton, C. M. Iversen, and V. Romanovsky, "Modeling the spatiotemporal variability in subsurface thermal regimes across a low-relief polygonal tundra landscape," *The Cryosphere*, vol. 10, no. 5, pp. 2241–2274, 2016. [Online]. Available: http://www.the-cryosphere.net/10/2241/2016/

[12] F. M. Hoffman, J. Kumar, R. T. Mills, and W. W. Hargrove, "Representativeness-based sampling network design for the State of Alaska," *Landscape Ecology*, vol. 28, no. 8, pp. 1567–1586, Oct. 2013.

[13] P. S. Thenkabail, J. G. Lyon, and A. Huete, *Hyperspectral Remote Sensing of Vegetation*. CRC Press, 2011.

[14] S. N. Bratsch, H. E. Epstein, M. Buchhorn, and D. A. Walker, "Differentiating among Four Arctic Tundra Plant Communities at Ivotuk, Alaska Using Field Spectroscopy," *Remote Sensing*, vol. 8, no. 51, 2016. [Online]. Available: http://www.mdpi.com/2072-4292/8/1/51

[15] S. J. Davidson, M. J. Santos, V. L. Sloan, J. D. Watts, G. K. Phoenix, W. C. Oechel, and D. Zona, "Mapping Arctic Tundra Vegetation Communities Using Field Spectroscopy and Multispectral Satellite Data in North Alaska, USA," *Remote Sensing*, vol. 8, no. 12, 2016. [Online]. Available: http://www.mdpi.com/2072-4292/8/12/978

[16] S. Hunt, Z. Yu, and M. Jones, "Lateglacial and Holocene climate, disturbance and permafrost peatland dynamics on the Seward Peninsula, western Alaska," *Quaternary Science Reviews*, vol. 63, pp. 42 – 58, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0277379112004908

[17] C. Silapaswan, D. Verbyla, and A. McGuire, "Land Cover Change on the Seward Peninsula: The Use of Remote Sensing to Evaluate the Potential Influences of Climate Warming on Historical Vegetation Dynamics," *Canadian Journal of Remote Sensing*, vol. 27, no. 5, pp. 542–554, 2001.

[18] C. Racine, R. Jandt, C. Meyers, and J. Dennis, "Tundra Fire and Vegetation Change along a Hillslope on the Seward Peninsula, Alaska, U.S.A." *Arctic, Antarctic, and Alpine Research*, vol. 36, no. 1, pp. 1–10, 2004. [Online]. Available: http://www.jstor.org/stable/1552423

[19] L. A. Viereck, *The Alaska Vegetation Classification*, ser. General technical report. USDA Pacific Northwest Research Station, 1992. [Online]. Available: https://books.google.com/books?id=mSU2jwEACAAJ

[20] K. Narita, K. Harada, K. Saito, Y. Sawada, M. Fukuda, and S. Tsuyuzaki, "Vegetation and Permafrost Thaw Depth 10 Years after a Tundra Fire in 2002, Seward Peninsula, Alaska," *Arctic, Antarctic, and Alpine Research*, vol. 47, no. 3, pp. 547–559, 2015.

[21] L. D. Hinzman, D. L. Kane, K. Yoshikawa, A. Carr, W. R. Bolton, and M. Fraver, "Hydrological variations among watersheds with varying degrees of permafrost," in *Proceedings of the 8th International Conference on Permafrost*. A. A. Balkema, July 2003, pp. 407–411.

[22] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. CRC Press, Taylor and Francis, 2009.

[23] F. M. Hoffman and W. W. Hargrove, "Multivariate geographic clustering using a Beowulf-style parallel computer," in *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA '99)*, H. R. Arabnia, Ed., vol. III. CSREA Press, Jun. 1999, pp. 1292–1298.

[24] J. Kumar, J. Weiner, W. W. Hargrove, S. P. Norman, F. M. Hoffman, and D. Newcomb, "Characterization and classification of vegetation canopy structure and distribution within the Great Smoky Mountains National Park using LiDAR," in *Proceedings of the 15th IEEE International Conference on Data Mining Workshops (ICDMW 2015)*, P. Cui, J. Dy, C. Aggarwal, Z.-H. Zhou, A. Tuzhilin, H. Xiong, and X. Wu, Eds., Institute of Electrical and Electronics Engineers (IEEE). Conference Publishing Services (CPS), Nov. 2015, pp. 1478–1485.

[25] J. A. Hartigan, *Clustering Algorithms*. John Wiley & Sons, 1975.

[26] W. W. Hargrove, F. M. Hoffman, and P. F. Hessburg, "Mapcurves: A quantitative method for comparing categorical maps," *Journal of Geographical Systems*, vol. 8, no. 2, pp. 187–208, Jul. 2006.

[27] L. Zhang, L. Zhang, and B. Du, "Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, June 2016.

[28] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *CoRR*, vol. abs/1512.00567, 2015. [Online]. Available: http://arxiv.org/abs/1512.00567

[29] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," *CoRR*, vol. abs/1311.2901, 2013. [Online]. Available: http://arxiv.org/abs/1311.2901

[30] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep Learning-Based Classification of Hyperspectral Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2094–2107, June 2014.

[31] H. Wu and S. Prasad, "Convolutional Recurrent Neural Networks for Hyperspectral Data Classification," *Remote Sensing*, vol. 9, no. 3, 2017. [Online]. Available: http://www.mdpi.com/2072-4292/9/3/298

[32] P. Zhang, X. Niu, Y. Dou, and F. Xia, "Airport Detection on Optical Satellite Images Using Deep Convolutional Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. PP, no. 99, pp. 1–5, 2017.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[34] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "Tensorflow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 265–283. [Online]. Available: https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf

[35] F. Chollet *et al.*, "Keras," https://github.com/fchollet/keras, 2015.

[36] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015. [Online]. Available: http://arxiv.org/abs/1505.04597

[37] B. Hariharan, P. A. Arbeláez, R. B. Girshick, and J. Malik, "Hypercolumns for Object Segmentation and Fine-grained Localization," *CoRR*, vol. abs/1411.5752, 2014. [Online]. Available: http://arxiv.org/abs/1411.5752

[38] Y. Duguay, M. Bernier, E. Lévesque, and F. Domine, "Land Cover Classification in SubArctic Regions Using Fully Polarimetric RADARSAT-2 Data," *Remote Sensing*, vol. 8, no. 9, 2016. [Online]. Available: http://www.mdpi.com/2072-4292/8/9/697

[39] G. Fu, C. Liu, R. Zhou, T. Sun, and Q. Zhang, "Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional Network," *Remote Sensing*, vol. 9, no. 5, 2017. [Online]. Available: http://www.mdpi.com/2072-4292/9/5/498

[40] S. Bianco, C. Cusano, and R. Schettini, "Single and Multiple Illuminant Estimation Using Convolutional Neural Networks," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4347–4362, Sept 2017.

[41] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, Feb 2016.

[42] E. K. Ghasrodashti, A. Karami, R. Heylen, and P. Scheunders, "Spatial Resolution Enhancement of Hyperspectral Images Using Spectral Unmixing and Bayesian Sparse Representation," *Remote Sensing*, vol. 9, no. 6, 2017. [Online]. Available: http://www.mdpi.com/2072-4292/9/6/541

[43] T. Panboonyuen, K. Jitkajornwanich, S. Lawawirojwong, P. Srestasathiern, and P. Vateekul, "Road Segmentation of Remotely-Sensed Images Using Deep Convolutional Neural Networks with Landscape Metrics and Conditional Random Fields," *Remote Sensing*, vol. 9, no. 7, 2017. [Online]. Available: http://www.mdpi.com/2072-4292/9/7/680

[44] X. Lu, Z. Lin, X. Shen, R. Mech, and J. Z. Wang, "Deep Multi-patch Aggregation Network for Image Style, Aesthetics, and Quality Estimation," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec 2015, pp. 990–998.