IMPROVING THE BINARY CLASSIFICATION OF PEAT LOCALITIES FROM MULTI-SOURCE REMOTELY-SENSED DATA USING CNN

R. Pittman, B. Hu [*]

Dept. of Earth and Space Science and Engineering, York University, 4700 Keele Street, Toronto ON M3J 1P3 Canada
(rpittman, baoxin)@yorku.ca

Commission III, WG III/10

KEY WORDS: Digital Soil Mapping, Peat, Neural Network, Convolutional Neural Network, Logistic Regression, Wetlands

ABSTRACT:

Neural networks were explored to achieve a binary classification for determining land corresponding to peat for a study area in the boreal forest of northern Ontario, Canada. Environmental covariates were employed as predictors and obtained from multiple sources, which included multispectral imagery, LiDAR, SAR, and aeromagnetic data. A dense neural network (DNN), as well as a convolutional neural network (CNN), were each implemented. Logistic regression, support vector machine (SVM) and random forest (RF) approaches were also modelled. Neighboring pixels surrounding the soil sampling sites were incorporated as input into the CNN, that permitted training on additional information that was not exploited by other methods. Preliminary results indicate that a CNN can attain improved accuracies for peat classification, when compared against other approaches.

1. INTRODUCTION

The accurate identification of physical settings comprising of peat is pertinent to digital soil mapping and a variety of associated land management activities. These applications can conform to the ascertainment of areas with higher soil carbon and organic matter contents (Vitt et al., 2009), as well as detecting wetlands (Rapinel et al., 2019; Whitcomb et al., 2009) and evaluating the extent of environmental impact for these respective regions (Sulla-Menashe et al., 2018). The recent availability and enhanced resolution of remotely-sensed imagery from multiple platforms, has facilitated digital soil mapping research for peatlands (Minasny et al., 2019). In particular, there is an impetus to identify peatlands and assess implications within the boreal transition zones, where land cover conversion is more anticipated to unfold.

For the classification of soil properties, random forest (RF) and support vector machines (SVMs) have been common (Minasny et al., 2019), as these methods can attain improved accuracies when compared to more traditional approaches such as logistic regression (Brungard et al., 2015; Heung et al., 2016). However, for many study areas there exists a dearth of soil data for modeling purposes. This situation can arise due to the difficulty of extracting soil samples, especially for remote areas within markedly forested terrain. There is the further impediment of the expense of processing soil samples, which can impart as prohibitive for obtaining an effective quorum of samples for analysis. Due to implicit connections between environmental covariates for digital soil mapping, it can be challenging to fit explicit models. There is also a need for improvement with classification accuracy.

Deep learning methods can conceivably advance accuracy, by uncovering intricate relations amongst the environmental covariates with respect to soil properties. Neural networks can be adopted, but the dense layer form of the conventional artificial neural network (ANN, or DNN) may not achieve better accuracies than other approaches (Brungard et al., 2015; Heung et al., 2016). Convolutional neural networks (CNNs), with a structure consisting of multiple hidden layers conforming to convolution and pooling that are ended with a dense layer (Buda et al., 2018; Lee and Song, 2019), can resolve features that are impractical from regarding just the dense layer structure of ANNs. Employing a filter, surrounding imagery pixels within the neighborhood of soil sample points can be convoluted, unlocking patterns within the imagery that can improve the classification. Exploiting this additional information, a CNN could attain higher accuracies than that from other approaches such as RF or SVMs.

Setting up and configuring a CNN is not necessarily a trivial matter for digital soil mapping applications. This discipline of research generally considers numerous environmental covariates obtained from a variety of sensors, that need to be aggregated for areas typically encompassing more than a couple hundred km^2 (Minasny et al., 2019). The incorporation of convolution filters, followed by max pooling and flattening layers, can complicate the modelling process. Preparation of data for usage with a CNN can be much more computationally intensive when compared to other classification methods, as image tensors about each site for each data covariate layer, need to be integrated.

A CNN was formulated to perform a binary classification for discerning localities conforming to peat for a study area in a boreal biome in northern Ontario, Canada. A brief review of the soil data and environmental covariates is stated, which is followed by presenting the implementation of the CNN. A

[*] Corresponding author

comparison of modelling results among other classification methods has been provided. This paper ends with a discussion of limitations or factors to consider when employing a CNN for digital soil mapping applications.

2. STUDY AREA AND COVARIATE DATA

The study area encloses the vicinity of Hearst, within the District of Cochrane in northern Ontario (ON), Canada. This region consists of roughly 897 km², with latitudes extending from 49.6° N to 49.8° N and longitudes varying from 83.3° W to 84.0° W. The topography for this region is relatively flat, with steep slopes along segments of the riverbanks of the Mattawishkwia River, which traverses on a northeast trajectory through the area. Elevations range from a minimum of 209 m in the north to a maximum of 283 m in the southwest. The largest community in this region is Hearst, situated in the center of the study area, which is partially surrounded by cleared land comprised of pasture and hay fields. This region is depicted in Figure 1, which denotes the locations of 157 sites with soil data that were obtained by the Ontario Forest Resources Inventory (FRI). Also shown is true-color composite imagery visualized from median satellite reflectance from Landsat-8 for cloud-free (less than 1% clouds) scenes from June and July of 2018. Note for this composite image, the bands B2 (0.452-0.512 μm), B3 (0.533-0.590 μm) and B4 (0.636-0.673 μm) were specified for blue, green and red, respectively. The grid coordinates bordering the figure are in the NAD 1983 Lambert conformal conic projection.
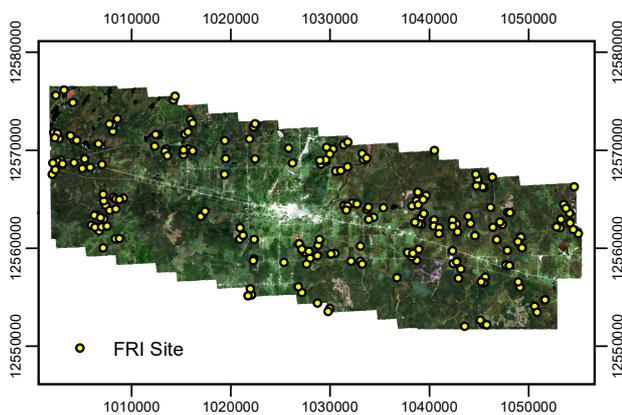
Figure 1. True-color composite of the Hearst study area, with Ontario Forest Resources Inventory (FRI) sites denoted.

Soil samples were retrieved during field campaigns in August and October of 2011 by the Ontario FRI (Paloniemi, 2018). These soil samples were extracted near the surface, primarily at 5–15 cm depths. Soil texture classification was inferred by expert assessment in the field, where samples conforming to peat were evaluated according to a van Post (vP) system of decomposition (Malterer et al., 1992; Paloniemi, 2018). For this study, the soil texture family classifications were condensed into the categories of peat and non-peat.

Environmental covariates relating to various soil formation factors were compiled from multispectral satellite imagery, LiDAR, SAR, and aeromagnetic data. These predictors included optical surface reflectance and structure for vegetation, topographic covariates, and geophysical surveys

for parent material. The complete listing of environmental covariates utilized is summarized in Table 1. In total there were 72 different covariates exploited. Each covariate layer was rasterized, reprojected and coregistered to a common coordinate system, with a cell size of 30 m. These covariates were wielded as predictors for the classification of peat versus non-peat, for pixels corresponding to the site locations of where soil samples were extracted.

Predictor	Spatial Resolution
DEM (derived from LiDAR)	
Aspect	30 m
Convergence Index	30 m
Elevation	30 m
General Curvature	30 m
Hillshading	30 m
Mid-Slope Position	30 m
Multiresolution Ridge Top Flatness (MRRTF)	30 m
Multiresolution Valley Bottom Flatness (MRVBF)	30 m
Plan Curvature	30 m
Profile Curvature	30 m
SAGA Topographic Wetness Index (SAGA TWI)	30 m
Sky View Factor	30 m
Slope	30 m
Slope Height	30 m
Slope Length	30 m
Standardized Height	30 m
Stream Power Index	30 m
Terrain Ruggedness Index (TRI)	30 m
Terrain View Factor	30 m
Topographic Wetness Index (TWI)	30 m
Total Curvature	30 m
Valley Depth	30 m
Visible Sky	30 m

Predictor	Spatial Resolution	Spectral Resolution
PALSAR		
HH 2017	30 m	HH Polarization
Density of Water Bodies	30 m	HH Polarization
Distance to Water Bodies	30 m	HH Polarization
LiDAR		
Canopy Height Model (CHM)	10 m	1.064 μm
Gap Fraction	10 m	1.064 μm
Tree Species Maps		
Black Spruce Indicator	30 m	
Balsam Fir Indicator	30 m	
Landsat*		
B1 (Winter, May, Summer, Autumn) 2017	30 m	0.435-0.451 μm
B2 (Winter, May, Summer, Autumn) 2017	30 m	0.452-0.512 μm
B3 (Winter, May, Summer, Autumn) 2017	30 m	0.533-0.590 μm
B4 (Winter, May, Summer, Autumn) 2017	30 m	0.636-0.673 μm
B5 (Winter, May, Summer, Autumn) 2017	30 m	0.851-0.879 μm
B6 (Winter, May, Summer, Autumn) 2017	30 m	1.566-1.651 μm
B7 (Winter, May, Summer, Autumn) 2017	30 m	2.107-2.294 μm
B10 Summer 2017	100 m	10.60-11.19 μm
B11 Summer 2017	100 m	11.50-12.51 μm
NDVI (Winter, May, Summer, Autumn) 2017	30 m	0.636-0.673 μm 0.851-0.879 μm
NDWI (Winter, May, Summer, Autumn) 2017	30 m	0.851-0.879 μm 1.566-1.651 μm
NRCan**		
Gravity Anomaly	2 km	
Magnetic Residual	200 m	

* Surface reflectance bands; seasons were Winter (January, February, March), May, Summer (June, July) & Autumn (Sep. 10 - Oct. 10). For May & Autumn, median surface reflectance was obtained for same period over 5 years (2015-2019).
** First vertical derivatives for each attribute were also included

**Table 1**. Listing of environmental covariates utilized as predictors for the models.

Surface reflectance was acquired from Landsat imagery coinciding to various seasons throughout the year, with

normalized difference vegetation index (NDVI) and normalized difference water index (NDWI) each calculated. LiDAR data collected during the autumn of 2017 for the Ontario Ministry of Forestry and Natural Resources was wielded to generate rasterized imagery. A digital elevation model (DEM) was constructed from the LiDAR data, which was subsequently fielded to compute an assortment of topographic covariates using SAGA GIS (Conrad et al., 2015). A canopy height model (CHM) and gap fraction were each derived from the LiDAR data. Tree species maps for black spruce (*Picea mariana*) and balsam fir (*Abies balsamea*) were incorporated from previous research by the authors for this region. JAXA PALSAR imagery (Shimada et al., 2014) from 2017 for L-band SAR (Lucas et al., 2010) was exploited to detect surface water, which was applied to calculate distances to water bodies. Gravity anomaly and magnetic residual were obtained from aeromagnetic data for 2016 from Natural Resources Canada (NRCan).

3. METHODOLOGY

3.1 Convolutional Neural Network Image Tensor

The advantage of employing a CNN is that its conformation can decipher patterns, commonly known as features within the imagery, amongst the neighboring pixels for a site. In effect, one can consider an image tensor about each site pixel, of a window set to the initial kernel size for the first convolution layer. For the training and validation data required for modelling and evaluating a CNN, respectively, these corresponding pixels must be extracted for each site across the various covariate layers, as illustrated in Figure 2.
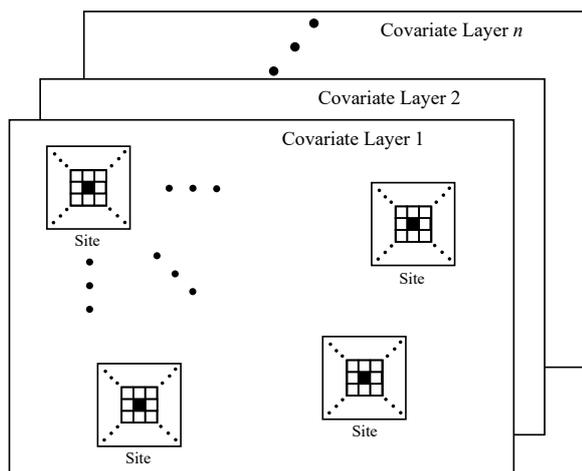

Figure 2. Extraction of data pertaining to each sampling site, from each of the total *n* different covariate layers.

For each site, an image tensor window (here 7 pixels by 7 pixels) is extracted; the pixel matching to the sampling site is depicted in the center (colored black), with 3 pixel widths on each of the four sides. The cell size of each pixel corresponds to 30 m resolution, which is the spatial resolution of the environmental covariate layers. The size of the image tensor indicates the spatial scale of imagery considered for the convolution for each sampling location, which here amounts to an area of 44,100 m^2 for each site. By incorporating data of neighboring pixels within an image tensor window, there is the potential for a CNN to more effectively classify a site.

Examples of scenarios where there can be this type of information gain is conveyed in Figures 3 and 4.

A site bordering a wetland environment is exhibited in Figure 3, with an image tensor window (here of 7 pixels by 7 pixels) considering pixels coinciding to the attributes of the nearby wetland. The darkest pixel intensities depict the lowest canopy heights, which here would correspond to cut hayfields or the surface of water bodies. For the boreal biome in northern Canada, there is a notable association between wetlands and peat (Haynes et al., 2021), where canopy heights are generally stunted for trees growing within or along the periphery of wetlands. Even if a site has a taller CHM, the site can still conform to peat if it is bordering a wetland, populated with the upland variety of black spruce, which can grow nearby scrubby lowland black spruce within the wetland (El Abidine et al., 1994).
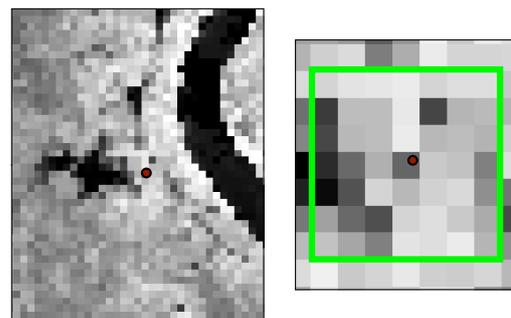
Figure 3. CHM pixels for a site adjacent a wetland, with variation of heights shown in the image tensor window.

Comparatively, classifications can be challenging for secondary forest sites, or for localities which were previously cleared but where trees have regrown. The soil from these sites tend to be compacted, and subsequently do not normally consist of peat. There can exist a great deal of variability with covariate attribute intensities amongst pixels for this sort of site. An image tensor window can capture a mix of pixel intensities, which can pertain to patterns with how the land was logged or managed, such as field ridge boundaries from a previous agriculture usage. An example of pixels in the neighborhood of a site from a secondary forest plot, again with CHM depicted with darker intensities corresponding to lower canopy heights, is displayed in Figure 4. Here the canopy is of short to intermediate height, where the image tensor window can resolve a site location that was previous cleared, which may not be possible to ascertain by just considering one pixel.
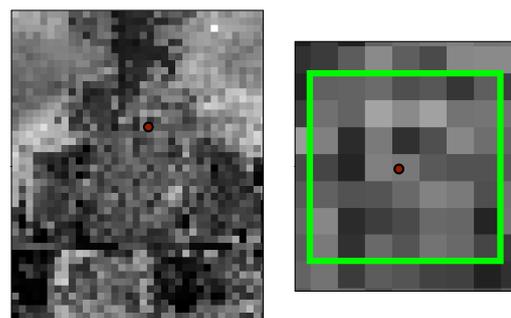
Figure 4. CHM pixels for a site within secondary forest, with corresponding image tensor window.

The information extracted by the image tensor window for each site can be incorporated into a CNN consisting of 2-dimensional convolution layers. However, a CNN comprised of 1-dimensional convolution layers can be devised by considering the pixel selections matching directly for each site (across the covariate layers). The 1-dimensional convolution layer would require fewer parameters and computational resources. Rather than integrating neighboring pixels about a site, a CNN of this form could unravel data patterns amongst the covariates.

3.2 Modelling

Keras (Chollet, 2015; Lee and Song, 2019) harnessing TensorFlow (Abadi et al., 2015) in RStudio (RStudio Team, 2020) were utilized for fitting the neural networks. Torch (Collobert et al., 2011; Paszke et al., 2019) was applied for preparing data input for the CNN. The caret package in R (Kuhn, 2008) was employed for the other classification methods. The data was split randomly 70:30 for training and validation, respectively. For each model, a total of 72 predictors were wielded. With the CNN, an image tensor window of 7 pixels by 7 pixels was adopted for each site for each of the covariates. For the other approaches, data was extracted for just the pixels directly corresponding to soil sampling sites; that is, one pixel per site per covariate layer.

Firstly, a DNN of five dense layers was employed, with categorical cross entropy implemented for the loss function. Rectified linear unit (ReLU) activation functions were adopted for the first four dense layers, with units of 64 followed by 32, 16 and 8, respectively. A softmax activation was applied for the last layer, as output comprised of two classes (non-peat, peat). The structure of the DNN is presented in Table 2. The corresponding formulation for the CNN is depicted in Table 3.

Layer	Specifications
Input	
Dense	Units = 64, Activation = ReLU
Dense	Units = 32, Activation = ReLU
Dense	Units = 16, Activation = ReLU
Dense	Units = 8, Activation = ReLU
Dense	Units = 2, Activation = Softmax

Table 2. Structure of DNN.

Layer	Specifications
Input	
Convolution 2D	Filters = 32, Kernal Size = (7,7), Activation = ReLU
Max Pooling 2D	Pool Size = (2,2)
Convolution 2D	Filters = 32, Kernal Size = (3,3), Activation = ReLU
Max Pooling 2D	Pool Size = (2,2)
Dropout	Rate = 0.25
Flatten	
Dense	Units = 16, Activation = ReLU
Dropout	Rate = 0.5
Dense	Units = 2, Activation = Softmax

Table 3. Structure of CNN.

As expected, the structure for the CNN is more complex than that for the DNN. The CNN consisted of two convolution layers, each succeeded by a max pooling layer. Dropout was enacted between the last max pooling and flattening layers, as well as before the final dense layer. The dropout was executed so as to reduce the over-fitting of the neural network. As with the DNN, for the CNN the loss was set to optimize the categorical cross entropy, with ReLU activation functions imposed for the respective interior layers, and the final activation being softmax (Lee and Song, 2019). Both the DNN and CNN were each trained for 100 epochs.

Other modelling approaches were considered, so as to compare the accuracy results. A logistic regression, a linear SVM, and a RF were each implemented. Repeated cross-validation was set with 10-folds and 10 repeats. For the RF, the mtry parameter was fixed as the square root of the number of predictors, and 1000 trees were utilized for training. The accuracy in the form of percent correct classification (here denoted as accuracy) as well as Cohen's kappa score, were calculated for each model from the corresponding confusion matrix on the validation data.

4. RESULTS

The validation accuracies for the binary classification of peat from the logistic regression, linear SVM, RF, DNN and CNN approaches are reported in Table 4. The poorest accuracy was attained from the logistic regression model, which had an accuracy of 0.43, and a negative kappa score indicating less agreement than expected by chance. There were convergence issues with the training of the logistic regression model on the full set of predictors. The highest accuracies were realized by the DNN, RF and CNN approaches. The CNN achieved an accuracy of 0.70 with a corresponding kappa of 0.39, indicating fair to moderate agreement. A prediction map of the binary classification from the CNN for the study area is rendered in Figure 5.

Model Accuracies	Peat Classification [Peat / Non-Peat]	
	Accuracy	Kappa
Logistic Regression	0.43	-0.14
Support Vector Machine Linear-basis (SVM Linear)	0.57	0.12
Random Forest (RF)	0.67	0.25
Dense Neural Network (DNN)	0.65	0.30
Convolutional Neural Network (CNN)	0.70	0.39

Table 4. Modeling accuracies attained on validation data.
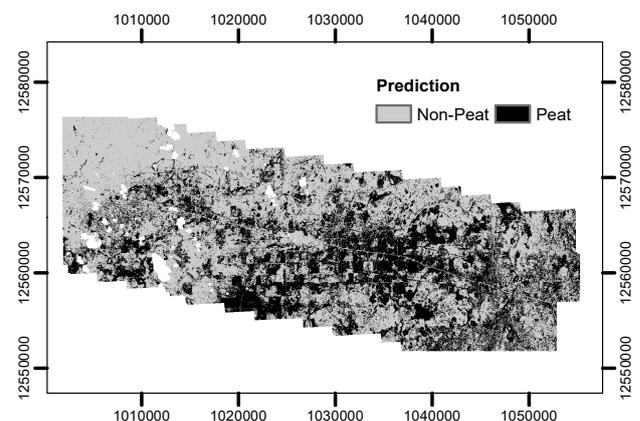
Figure 5. Prediction map for the binary classification of peat for the Hearst study area.

The regions with the greatest prevalence of peat correspond to wetland regions to the east and southeast of the community of Hearst. The cleared land in the study area, which would conform to the pastureland and settlements, is depicted as non-peat. This representation is reasonable, as the soil in these areas should be compacted, whereas peatlands are generally not suitable for agricultural activities or uses.

5. DISCUSSION

Generating a prediction map for a CNN, or related maps for the corresponding study area (such as entropy maps for determining regions of modelling uncertainty) are a greater endeavor than producing the same mapping for other methods. This complication is due to having to consider an image tensor window about essentially each pixel of the targeted study area, for each environmental covariate layer utilized a predictor. Due to the size of the image tensor window adopted, one may end up with a nominally reduced plot for the targeted study area. In terms of pixels, the study area here comprised of 1,785 columns by 827 rows. These dimensions were reduced to 1,779 columns by 821 rows, as 3 pixels depth along the edge of each side were omitted due to the input image tensor window entailing 7 pixels by 7 pixels.

The size of the image tensor window and convolution filter adopted should depend upon what is suitable for each study area or application. For this analysis, a max filter size was set as 7 pixels by 7 pixels, in part to mitigate data constraints; this amounted to an image tensor window of spatial extent of 210 m per side, that was centered upon each site. A variogram analysis might be warranted to determine appropriate input image window dimensions. It would be sensible to only consider neighboring pixels within an image tensor window that would conform to the soil properties of the site; if an image tensor window is too large, then some of those pixels would not correspond to the same soil attributes.

Patterns of the features within an image tensor might be specific to a particular application or study area. There is the possibility that a CNN could attain worse results than a DNN or other methods, as trends within these features may not persist among the sites. This analysis can be sensitive to the number of convolution layers utilized, and the corresponding sizes of the filters. Future work for this research will focus on experimentation with the CNN structure, as well as with adjustments to the filter windows, so as to maximize accuracy.

A main drawback of neural networks is that it is difficult to determine variable importance from amongst the predictors. Subsequently, it can be challenging to ascertain what predictors are important, and which features are redundant (or where their inclusion negatively impacts modelling). It is imperative to implement variable importance for a CNN that can take into account the interaction of covariates. There is a possibility that two predictors can attain great significance independently, but when acted together can actually diminish the accuracy of a model.

CNNs can necessitate the optimization of vast amounts of parameters when training. Due to the involvement of extracting soil samples, sample sizes can be very sparse, which may not justify the utilization of a CNN. To alleviate issues of over-fitting, alternatively one can implement 1-dimensional convolution layers, as these would entail fewer parameters and resources. Image tensors are not required for a 1-dimensional convolution layer, and instead site pixel data as used for less complex methods can be inputted. Another possibility is to evaluate a CNN trained with just a few covariate layers, but then one would foremost have to determine what covariates have the greatest significance. Taking into account these considerations, a CNN might not be a feasible method of choice. If viable models can be achieved from less complicated approaches, specifically SVM or RF, then employing those methods might be more appropriate.

Peat modelling can be intrinsic to a location, as what environmental covariates are important for one biome may not be veritable for another. Even within the boreal biome, there can exist dissimilarities between wetlands in the boreal shield of eastern Canada in northern Ontario, versus the boreal plain of western Canada in Saskatchewan. In northern Ontario, peat is commonly encountered in wetland environments, coinciding with black spruce. In part because of a lack of economic incentives, peatlands are generally avoided and directly spared from human activities. However, peatlands are vulnerable to environmental change (Minasny et al., 2019), and hence should be researched to monitor these potential future impacts.

6. CONCLUSION

A CNN was able to achieve gains in modelling accuracy when evaluated against conventional classification methods, for the prediction of localities with peat for a boreal biome. The highest accuracy was 0.70 and attained from a CNN, compared to 0.65 and 0.67 from a DNN and RF, respectively.

The incorporation of neighboring pixels via image tensor windows about a site pixel point, which was subsequently convoluted and pooled, added information for use in the classification of peat. However, more analysis might have to be performed in order to determine optimal sizes of input image tensor windows and filters. A CNN approach necessitates greater computational resources, and might not be suitable for low site counts due to over-fitting. With regards to application, the features unlocked from the image tensors would be specific in context to each study area.

For peat prediction within the Hearst study area, the localities conforming to peat are primarily the wetlands, situated to the east and southeast of the community of Hearst. As expected, the agricultural and settled locations correspond to non-peat.

extracting tensors for the CNN. ArcMap 10.6 software from Esri was utilized for producing figures of the maps.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Schuster, M., Monga, R., Moore, S., Murray, D., Olah, C., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems.

Brungard, C.W., Boettinger, J.L., Duniway, M.C., Wills, S.A., Edwards, T.C., 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*, 239, 68–83. doi.org/10.1016/j.geoderma.2014.09.019.

Buda, M., Maki, A., Mazurowski, M.A., 2018. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, 106, 249–259.

Chollet, F., 2015. Keras.

Collobert, R., Kavukcuoglu, K., Farabet, C., 2011. Torch7: A Matlab-like Environment for Machine Learning. *BigLearn*. NIPS Workshop.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.*, 8, 1991–2007. doi.org/10.5194/gmd-8-1991-2015.

El Abidine, A.Z., Bernier, P.Y., Stewart, J.D., Plamondon, A.P., 1994. Water stress preconditioning of black spruce seedlings from lowland and upland sites. *Can. J. Bot.*, 72, 1511–1518.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* doi.org/10.1016/j.rse.2017.06.031.

Haynes, K.M., Smart, J., Disher, B., Carpino, O., Quinton, W.L., 2021. The role of hummocks in re-establishing black spruce forest following permafrost thaw. *Ecohydrology*, 14. doi.org/10.1002/eco.2273.

Heung, B., Ho, H.C., Zhang, J., Knudby, A., Bulmer, C.E., Schmidt, M.G., 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. *Geoderma*, 265, 62–77.

Kuhn, M., 2008. Building Predictive Models in R using the caret Package. *J. Stat. Softw.*, 28.

Lee, H., Song, J., 2019. Introduction to convolutional neural network using Keras; An understanding from a statistician. *Commun. Stat. Appl. Methods*, 26, 591–610.

Lucas, R., Armston, J., Fairfax, R., Fensham, R., Accad, A., Carreiras, J., Kelley, J., Bunting, P., Clewley, D., Bray, S., Metcalfe, D., Dwyer, J., Bowen, M., Eyre, T., Laidlaw, M., Shimada, M., 2010. An Evaluation of the ALOS PALSAR L-Band Backscatter—Above Ground Biomass Relationship Queensland, Australia: Impacts of Surface Moisture Condition and Vegetation Structure. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 3, 576–593.

Malterer, T.J., Verry, E.S., Erjavec, J., 1992. Fiber Content and Degree of Decomposition in Peats: Review of National Methods. *Soil Sci. Soc. Am. J.*, 56, 1200–1211.

Minasny, B., Berglund, Ö., Connolly, J., Hedley, C., de Vries, F., Gimona, A., Kempen, B., Kidd, D., Lilja, H., Malone, B., McBratney, A., Roudier, P., O'Rourke, S., Rudiyanto, Padarian, J., Poggio, L., ten Caten, A., Thompson, D., Tuve, C., Widyatmanti, W., 2019. Digital mapping of peatlands – A critical review. *Earth-Science Rev.*, 196, 102870. doi.org/10.1016/j.earscirev.2019.05.014.

Paloniemi, J., 2018. FRI Field Guide. Forest Resources Inventory, Ontario Ministry of Natural Resources and Forestry, Ontario.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*, 32, 8024–8035.

Rapinel, S., Fabre, E., Dufour, S., Arvor, D., Mony, C., Hubert-Moy, L., 2019. Mapping potential, existing and efficient wetlands using free remote sensing data. *J. Environ. Manage.*, 247, 829–839.

RStudio Team, 2020. RStudio: Integrated Development Environment for R.

SAGA Development Team, 2020. System for Automated Geoscientific Analysis (SAGA).

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data (2007-2010). *Remote Sens. Environ.*, 155, 13–31. doi.org/10.1016/j.rse.2014.04.014.

Sulla-Menashe, D., Woodcock, C.E., Friedl, M.A., 2018. Canadian boreal forest greening and browning trends: An analysis of biogeographic patterns and the relative roles of disturbance versus climate drivers. *Environ. Res. Lett.*, 13.

Vitt, D.H., Wieder, R.K., Scott, K.D., Faller, S., 2009. Decomposition and peat accumulation in rich fens of boreal Alberta, Canada. *Ecosystems*, 12, 360–373.

Whitcomb, J., Moghaddam, M., McDonald, K., Kellndorfer, J., Podest, E., 2009. Mapping Wetlands of Alaska Using L-Band Satellite Radar Imagery. *Can. J. Remote Sens.*, 35, 54–72.