Estimating Forest Ground Vegetation Cover From Nadir Photographs Using Deep Convolutional Neural Networks

Pranoy Panda
Department of Computer Science and Engineering
IIT Hyderabad
Sangareddy, Telangana, 502285, India
cs20mtech12002@iith.ac.in

Martin Barczyk
Department of Mechanical Engineering
University of Alberta
Edmonton, AB, T6G 1H9, Canada
mbarczyk@ualberta.ca

Jen Beverly
Department of Renewable Resources
University of Alberta
Edmonton, AB, T6G 1H9, Canada
jbeverly@ualberta.ca

Abstract

Forest fires, such as those on the US west coast in September 2020, are an important factor in climate change. Wildfire modeling and mitigation require mapping vegetation ground cover over large plots of land. The current forestry practice is to send out human ground crews to collect photos of the forest floor at precisely determined locations, then manually calculate the percent cover of ground fuel types. In this work, we propose automating this process using a supervised learning-based deep convolutional neural network to perform image segmentation. Experimental results on a real dataset show this approach delivers very promising performance.

1 Introduction

Vegetation data is essential for a wide range of research studies and management needs. In forested areas, detailed information about vegetation is collected by human surveying crews. Vegetation measurements are also heavily used in ecological research projects to establish baseline conditions and track changes in vegetation over time in response to management actions or disturbances such as industrial activity and for monitoring climate change impacts on ecosystems.

Documentation and measurement of live and dead vegetation are also extremely important for research on wildfires, an important factor in climate change [9, 10]. The amount and type of live and dead biomass in a location will influence how a wildfire behaves and how fast it spreads. In the province of Alberta, Canada, an inventory program was established to measure and document vegetation characteristics relevant to forest fires.

Forest fuels are assessed at different vertical layers. We focus on the classification and measurement of fuels located on the ground such as grass, moss, and dead needles from conifer trees. Ground fuels have been largely ignored in forest fire research because it is prohibitively costly to measure them across large areas. In contrast, aerial fuels (tree crowns) have long been inventoried across large areas using aerial photos to document forest tree species types.

In this work, we explore the use of machine learning and image processing techniques to extract wildfire fuel data from nadir (downward-looking) photographs taken in forested ecosystems by

Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

either drones or field personnel. We frame the problem of ground vegetation classification as a multi-class semantic segmentation task and propose to use a transfer learning methodology using a deep convolutional neural network. Our approach results in pixel masks representing regions of pre-specified ground cover types. The resulting data allows quantifying variables used in wildfire propagation modeling, such as the primary ground cover (e.g. grass, moss, needles, water, rock), the ratios between dominant fuel types (grass, moss, and needles), the presence and ground cover of shrubs, and the existence, coverage, and dimensions of deadwood segments.

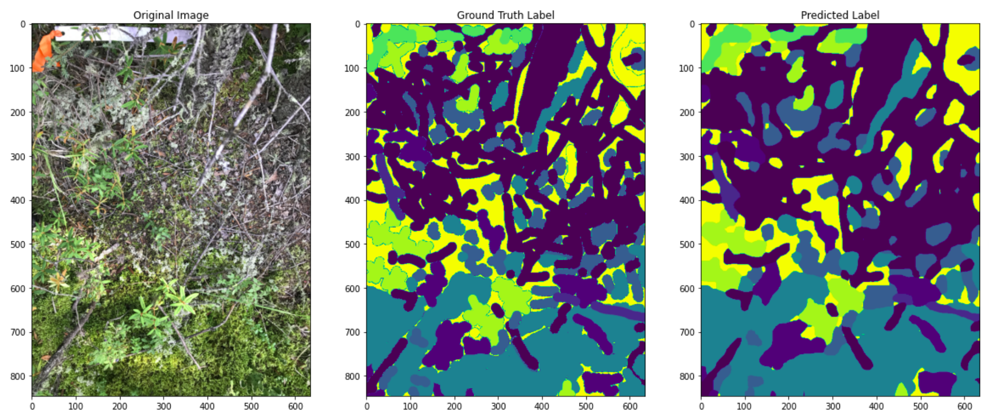


Figure 1: A sample image-label pair from our dataset along with our method's classification result

1.1 Related work

A number of groups have published works on automated vegetation classification in agriculture, for instance [1] (plants), [8] (grass and forb), and [2] (crops and weeds). Other methods based on LiDAR data [5] and remote sensing [11] have been used for forest cover estimation. However, none of these or similar publications specifically tackle forest ground cover classification using nadir photographs, which involves complex and multi-class images as seen in Figure 1. Due to the novel application area, we needed to perform the image segmentation labeling in-house, relying on a full-time employee with expertise in forest ground cover recognition to generate the training data labels.

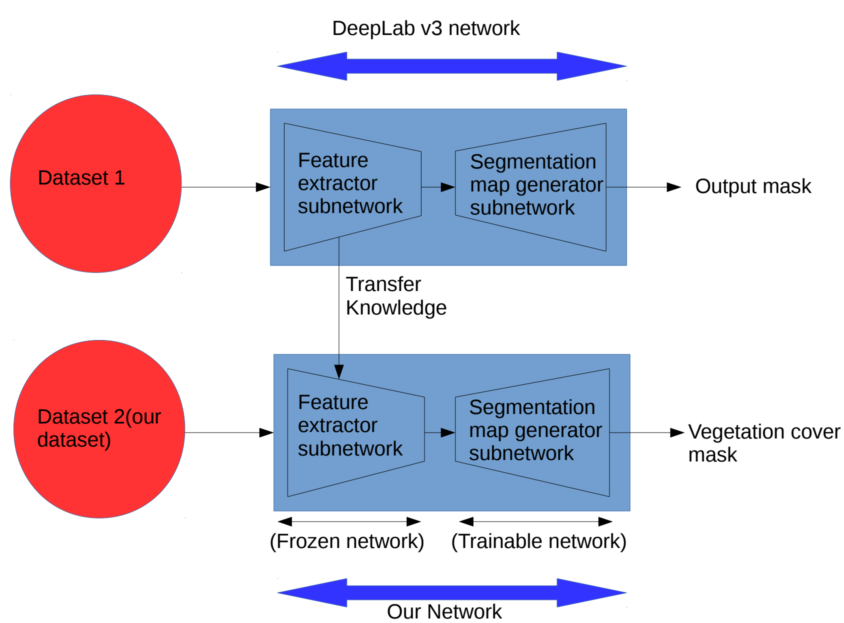


Figure 2: Transfer Learning framework

2 Dataset and Methodology

The dataset consists of photos taken by ground crews at 28 field sample plots. At each sample plot, 16 ground fuel photos were taken in a uniform 4×4 grid for a total of 448 images across all plots. Of these, 330 images were manually labeled by the aforementioned expert, using ten label types: firewood, forb, grass, Lichen, Moss-Feath, Moss-Other, Moss-Sphag, shrub, non-fuel, and void (label for pixels which do not belong to any other classes). We divided this dataset into 290 images for training and 40 images for testing. We are unfortunately not able to make the dataset publicly available due to IP issues with the agency responsible for the data collection.

The segmentation task we are facing is complex since many vegetation types have subtle differences in appearance, and exhibit irregular shapes and sizes. We used the supervised deep learning framework in our work as it has been found to work well for semantic segmentation purposes. However, this framework is notorious for requiring a large number of sample points for the training process, which we did not have available. Therefore, we opted to use *Transfer Learning* when dealing with our (small) dataset. Transfer learning involves transferring knowledge acquired by a machine learning model in one particular task to another, related task.

In our case, we used the pre-trained Deeplabv3-Resnet101 model [3] as our base network for deriving knowledge to solve our current task. This base network has been trained on a subset of the COCO train2017 dataset, consisting of a few thousand images, using the 20 categories present in the Pascal VOC dataset. The transfer learning methodology used in our work is shown in Figure 2. Our deep convolutional neural network uses the convolution layers of the Deeplabv3-Resnet101 network as a learned feature extractor. We then attached a classification head of 2048 neurons with sigmoid activation. Therefore, each pixel was assigned 10 scores in the range of 0 to 1, for the 10 classes, and the class with the highest score was declared as the predicted class for that pixel. We used the cross-entropy loss as the objective function for training our network, and the Adam [6] optimizer for updating the trainable parameters.

Since the multi-class (here 10 classes) semantic segmentation problem is quite complex, the set of 290 training images was found to be insufficient to obtain satisfactory results, even with the use of the transfer learning framework. Therefore we employed data augmentation strategies to increase the training set size fourfold: horizontal flip, Gaussian noise addition, and contrast reduction. The first two augmentation strategies are fairly common. The contrast reduction strategy was specifically motivated by our application, namely given that we expect the images to contain shadows occluding the vegetation types, contrast reduction helps to simulate areas of low lighting which can be expected in the dataset.

Also, in general, datasets generated from natural scenes for semantic segmentation have a significant variation in the occurrence frequencies of different classes. To deal with this, we used median frequency balancing to weigh the loss based on the correct label/class, where the weight given to each class in the loss function is the ratio of the median of class frequencies over the entire dataset divided by the class frequency [4].

3 Results and Discussion

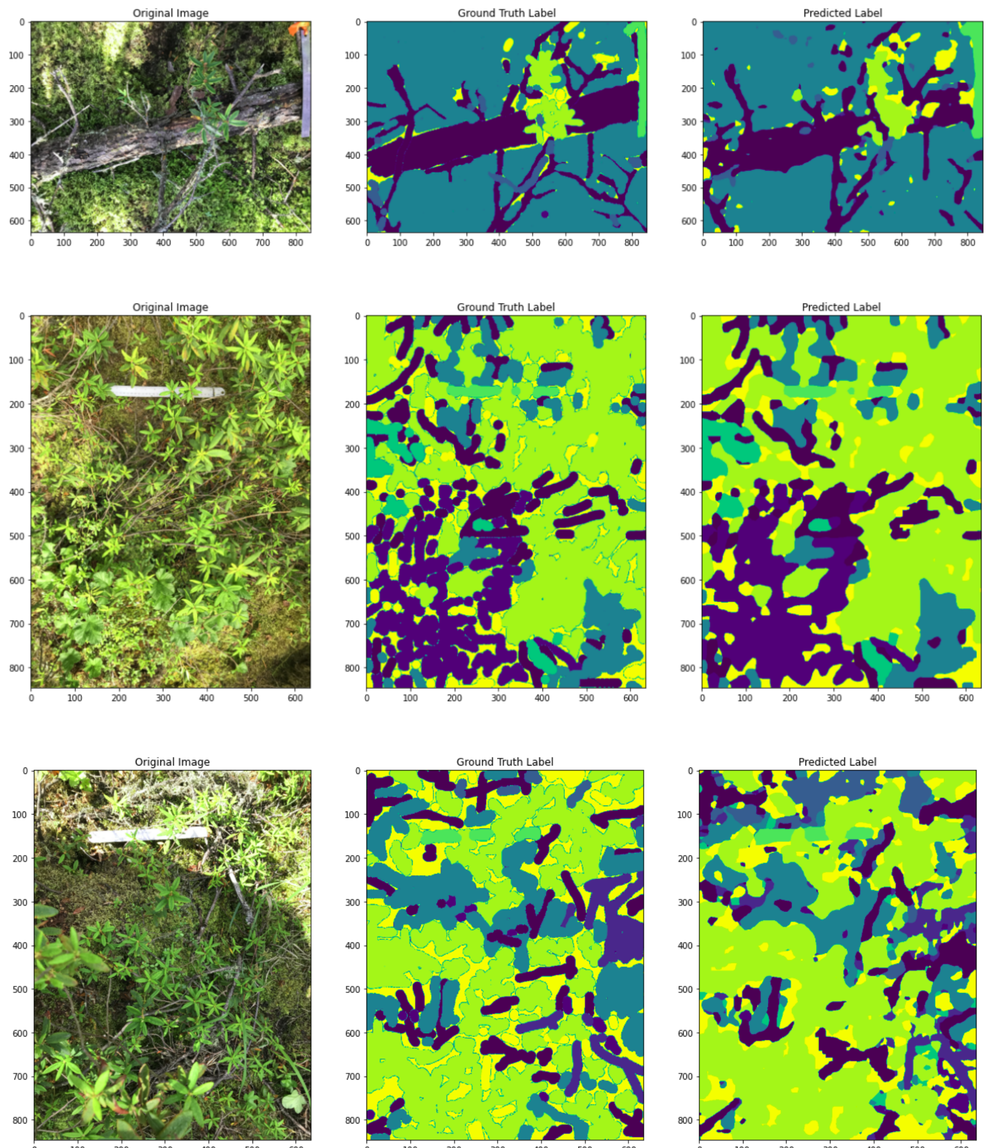
Table 1: Ablation study results

Model	mIoU	Accuracy
base:Deeplabv3-ResNet101+no class balancing	0.285	0.936
base:FCN-ResNet101+no class balancing	0.321	0.942
base:Deeplabv3-ResNet101+class balancing (primary model)	**0.352**	**0.950**
base:FCN-ResNet101+class balancing	0.286	0.940

Due to the novel approach of our work, there does not exist any prior method which we can compare against our results. We thus performed an ablation study with another standard base model, FCN (Fully Convolutional Network, [7]). We also studied the impact of class balancing on the results of these two different models.

We used two standard metrics to evaluate our segmentation networks: mean intersection over union (mIoU), i.e. the percentage of overlap between the true mask and our prediction output, and accuracy, i.e. the percentage of pixels in the image which were correctly classified.

The results for our primary model show that the mIoU is low whereas the accuracy or percentage of pixels classified correctly is high. This means that our model accurately predicts the prominent vegetation classes in individual images. The low mIoU values are misleading since they give equal weight to every class, irrespective of its proportion in an image. We, therefore, urge the reader to focus on the latter metric, since the majority of the complex vegetation cover is classified reasonably well. Below, we provide a sample of results obtained in our testing dataset images.



4 Conclusion

We have employed a deep convolutional neural network to perform semantic segmentation of ground forest vegetation cover, used for modeling and mitigation of wildfires. A proprietary dataset of pictures collected by ground crews was labeled by a human expert and used to train the classification algorithm. Using an ablation study, the proposed approach was shown to provide very promising results, specifically in terms of the percentage of testing image pixels correctly classified by the network. Future work will test the method on other datasets, compare the classification results against traditional human-calculated ground cover estimates, and investigate deploying autonomous drones to collect and analyze ground cover data in real-time.

The authors do not believe the work presented involves ethical aspects or future societal consequences.

References

[1] Alwaseela Abdalla, Haiyan Cen, Liang Wan, Reem Rashid, Haiyong Weng, Weijun Zhou, and Yong He. Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. *Computers and Electronics in Agriculture*, 167, December 2019.

[2] Petra Bosilj, Erchan Aptoula, Tom Duckett, and Grzegorz Cielniak. Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. *Journal of Field Robotics*, 37(1):7–19, January 2020.

[3] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

[4] David Eigen and Rob Fergus. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. In *Proceedings of the 2015 IEEE international conference on computer vision*, pages 2650–2658, Santiago, Chile, December 2015.

[5] Víctor González-Jaramillo, Andreas Fries, Jörg Zeilinger, Jürgen Homeier, Jhoana Paladines-Benitez, and Jörg Bendix. Estimation of above ground biomass in a tropical mountain forest in southern ecuador using airborne LiDAR data. *Remote Sensing*, 10(5):660, 2018.

[6] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[7] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, Boston, MA, June 2015.

[8] Chris McCool, James Beattie, Michael Milford, Jonathan D. Bakker, Joslin L. Moore, and Jennifer Firn. Automating analysis of vegetation with computer vision: Cover estimates and classification. *Ecology and Evolution*, 8(12):6005–6015, June 2018.

[9] Melania Michetti and Mehmet Pinar. Forest fires across italian regions and implications for climate change: a panel data analysis. *Environmental and Resource Economics*, 72(1):207–246, January 2019.

[10] Christine Ribeiro-Kumara, Jukka Pumpanen, Jussi Heinonsalo, Marek Metslaid, Argo Orumaa, Kalev Jogiste, Frank Berninger, and Kajar Koster. Long-term effects of forest fires on soil greenhouse gas emissions and extracellular enzyme activities in a hemiboreal forest. *Science of the Total Environment*, 718, May 2020.

[11] Hui Yang, Philippe Ciais, Maurizio Santoro, Yuanyuan Huang, Wei Li, Yilong Wang, Ana Bastos, Daniel Goll, Almut Arneth, Peter Anthoni, et al. Comparison of forest above-ground biomass from dynamic global vegetation models with spatially explicit remotely sensed observation-based estimates. *Global Change Biology*, 26(7):3997–4012, July 2020.