Measuring Vegetation Density in Marsh Grass Photographs Using Deep Neural Networks (Student Abstract)

Lucas Welch, Xudong Liu
School of Computing
University of North Florida
n00797216@unf.edu, xudong.liu@unf.edu

Abstract

To protect the world's marshlands, it is of utmost importance to be able to monitor their vegetation composition and coverage. This currently is accomplished by large teams of researchers and volunteers manually looking at the marsh images and labeling randomly selected pixels by what species (or lack thereof) is present at the pixel. This task, however, is extremely labor intensive, limiting the amount of quality environmental monitoring that can be done in the field. If the task was automated, teams would be able to monitor larger swaths of land. In this paper, we propose a novel framework for such automation using deep neural networks. Then, we focus on the key component of this framework: a binary classifier to decide whether a pixel is vegetated or not. To this end, we create a dataset of labeled snippet images out of publicly available photoquadrats of the marshlands in Florida. Finally, we construct LeNet-5 and AlexNet, adjusted to our input snippets, faster training time, networks and experiment to learn them on our dataset for the binary classification task. Our results show that the AlexNet model achieves higher accuracy on the test set than the LeNet-5 model, with 92.41% for AlexNet and 91.34% for LeNet-5.

Introduction

Around the world, humans are having a large impact on the environment. These impacts come from everything from habitat loss from an expanding population to climate change. Because of these pressures, it is very important to monitor the vegetation community of a habitat to track its health. At Guana Tolomato Matanzas National Estuarine Research Reserve (GTMNERR) (https://gtmnerr.org/), researchers monitor the percent cover and species composition of several marsh sites surrounding the city of St. Augustine, Florida. To do this, they first obtain one meter-square images of the marshland with a high resolution camera (Bacopoulos, Tritinger, and Dix 2019). Determining the percent vegetation cover of the image requires a volunteer or researcher to label a set of randomly chosen points and tally the total number of each category (unvegetated and one of five vegetated categories) in order to obtain the estimate for percent cover of each image. The need for manually labeling all points leads to a bottleneck in how much area can be covered by volunteers. As such, it is our goal to automate this process. This will be done by designing a machine vision system for determining the species of vegetation (or lack thereof) contained at a randomly selected point in a ground cover image. This model will then be integrated into a framework that determines percent cover of each species in each image.

The framework for our application is as seen in Figure 1. This pipeline contains both neural network models and traditional programming aspects.

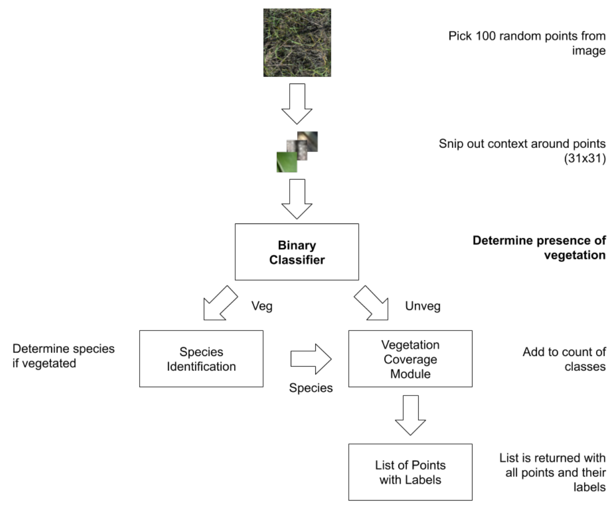

Figure 1: A flowchart representing the major components of the vegetation density pipeline. The pipeline begins with a full-size image before picking 100 random points and snipping the context around them (31x31). They are then passed into the described model, which will classify and pass them along to either the next classifier or tallier. The total number of each species for the 100 points is then returned.

Dataset

Because there was no existing dataset since GTMNERR's current labeling system does not record the pixel coordinates, we had to create our own dataset using our own labeling program. In order to obtain the points that had to be labeled, we used a script that chose 100 random points from each of the vegetation images provided by GTMNERR. We

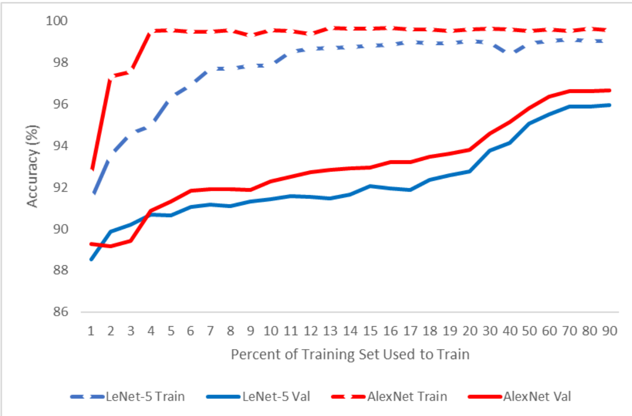

Figure 2: Learning curves for LeNet-5 and AlexNet

then provided a group of volunteers with a labeling program we developed as well as an assortment of these randomly chosen points. We have so far produced 57,410 labels.

Experimental Results and Analysis

Two convolutional neuron networks we employ in this paper are Le-Net5 (LeCun et al. 1998) and AlexNet (Krizhevsky, Sutskever, and Hinton 2012), both adjusted to our input image. To speed up training, we also modified AlexNet to have only 100 units in its two fully connected layers at the end of the network. For our experiments, we create a training set of 80% of our data and a test set of 20% using a stratified split. We further split the training set into 100 nearly equally sized buckets for constructing learning curves. Finally, we perform a 10-fold cross validation to produce a model with the best validation accuracy to test on the test set.

In order to gauge how well the models fit to the data, we start off with a learning curve for both Le-Net5 and AlexNet, which is pictured in Figure 2, where we present accuracies in relation to percentiles of the training set, and at each percentile the experiment instance is repeated 10 times to report the average. Both AlexNet and LeNet-5 seem to underfit. This underfitting points toward the potential for increasing accuracy as the dataset continues to expand. However, with the relative simplicity of LeNet-5 compared to AlexNet, it is likely that adding data will do more to improve AlexNet than it would LeNet-5.

In addition to generating a learning curve, our other main goal for the binary classifier was to develop a usable model for classifying a point as vegetated and unvegetated. To do this, we perform a 10-fold cross validation and keep the version of each model that had the highest validation accuracy. For LeNet-5, the best trained model has a training accuracy of 99.05% and a validation accuracy of 98.41%—it has a test accuracy of 91.34%. AlexNet, similarly, has a model with a training accuracy of 99.48% and a validation accuracy of 98.54% with a test accuracy of 92.41%. It should be noted, however, that there is a large discrepancy between the winning validation accuracy and that model's test accuracy, which could very well be caused by the fact that many classes are underrepresented in the dataset. Expanding this dataset, along with several goals to be discussed, will be our next step in continuing this application.

Conclusion and Future Work

In this work, we proposed a framework to automate the process of identifying whether an image of marshlands is vegetated, built a new collection of such images, and experimented with CNNs to show their effectiveness. Our results show that AlexNet achieves 92.41% accuracy on the test set, whereas LeNet-5 with 91.34%.

In order to create a more robust dataset, one of the primary goals for the future is to expand the size and increase the diversity of our dataset. This will be done by identifying images that include the underrepresented classes and then taking more samples from them for identification. The overall goal of this process is to even out the distribution of the vegetation classes in order to make them more represented and allow the models to more easily identify them. While expanding the number of entries in the dataset, we need to begin working on the species identifier. This model will complete the machine learning portion of our data pipeline, which will allow us to deploy a preliminary application for thorough testing. Alongside working on a new model, we hope to focus on fine tuning our models and pipeline for the task as much as possible. A major aspect of this is testing some sort of ensemble model that can more accurately classify points. This will consist of several different models (Le-Net5, AlexNet, and others) that each have a vote on what a point is; the class with the most votes is then selected. A similar direction is to condense our two different models in the pipeline down to one model, which will be a 6 class classifier. Depending on what is discovered to work, we can combine these ideas gradually and produce a working model that can accurately and reliably classify these points. While the machine learning models are necessary for identifying the points, there is no way to truly use this data without some sort of wrapper that formats the data and records the results. This wrapper will first be written in Django and hosted on a server in order to produce a preliminary application that can load in images and record the percent cover for each class. Later on, a more permanent offline version of the app will be produced that will be used by GTMNERR to predict vegetation cover for their vegetation monitoring program.

References

Bacopoulos, P.; Tritinger, A. S.; and Dix, N. G. 2019. Sea-Level Rise Impact on Salt Marsh Sustainability and Migration for a Subtropical Estuary: GTMNERR (Guana Tolomato Matanzas National Estuarine Research Reserve). *Environmental Modeling & Assessment* 24(2): 163–184.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.

LeCun, Y.; Bottou, L.; Bengio, Y.; and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11): 2278–2324.