

Project Specification

Jonathan Thomann - Student ID

June 2021

Project Code:
Project Name:
Project owner:
Project Supervisor:

Abstract

1 Introduction

Tracking soil cover is a crucial task for sustaining cultivable land. Maintaining cover over winter is one way to reduce erosion, the loss of particulate pollutants, and nutrients. In order to see the effects of proper soil cover, it is prudent to measure and monitor the cover of plants over areas of concern, and doing this manually is time-consuming and can introduce bias if done by only a small group of people. As technology improves and high-quality imagery and sensors become more ubiquitous, machine learning has become an exciting and popular way to track such soil cover.

To further realize why tracking soil cover is such an essential topic of research, the United Nations Food and Agriculture Organization (FAO) covers many of its benefits. For example, it improves the retention of moisture, decreases water stress, increases the availability of nutrients to all plants within the area, and can be used to control weeds, and pests [1].

With the growing amount of spatial imagery and in participation with the UK Centre for Ecology and Hydrology (CEH), this project will concern itself with using deep learning techniques to create scale-aware networks for the

task of segmenting flower resources in field margins using geospatial drone imagery. Successfully segmenting vegetation cover automatically in transect or quadrat surveys will significantly reduce work hours for ecological and hydrological studies and allow better insight into future research.

2 Background

Before machine learning techniques were introduced, there were two main methods of monitoring soil cover. First, the analysis of soil cover could be done manually on site. For example, studies such as [11] looked into the effect of vegetal cover on soil erosion and found that the presence of vegetation does reduce soil runoff. Second, the image analysis used to be done by hand as well. In [6], the authors attempted to estimate soil cover from slide photographs using point sampling methods. Other onsite methods, such as transect surveys, are just as time-consuming, and their results lack any reproducibility as they can be highly subjective [12]. While they showed promising results in some cases, the methods mentioned above are either too labor-intensive or too time-consuming to repeat regularly; it is, therefore, pertinent to create systems that allow for the automatic tracking and monitoring of soil cover over long periods.

Current state-of-the-art soil cover modeling systems now rely on machine learning or statistical methods to track and monitor vegetal cover over soil. [9] surveyed plant-based image processing and segmentation techniques. According to their findings, index, and threshold-based segmentation methods, while easy to understand and implement, do not generalize well and can be hindered by bad lighting conditions or complex backgrounds resulting in mis-segmented plants. Furthermore, these methods do not hold up to the state-of-the-art machine learning methods that are now being developed, as thresholds need constant optimization to keep segmentation accuracy high.

Popular machine learning methods such as support vector machines (SVM), random forest classifiers, and artificial neural networks (ANN) all lead to better generalization and accuracy results when compared to manual onsite or image processing methods. In [5], they proposed a segmentation method based on Particle Swarm Optimization and k-means clustering and were able to achieve accuracy rates of up to 91.7% when attempting to segment cotton from images. Guerrero et al. [8] used SVMs to segment weeds in maize fields with an accuracy rate of 93.1%. An issue with many of the aforementioned

techniques is that many of them (especially the thresholding and index-based methods) require that one spectral component (in the studies case, green) is stronger than the red spectral component. However, if plants are of the same spectral band as their background, accuracy results will drop significantly.

While these past results and studies show promising results, there is still the problem with scale and integrating spatial imagery taken at different scales. Atkinson and Tate [3] conducted a review of spatial scale at its various issues and applications within geospatial statistics. Spatial scale is essentially a link between an object distributed in space and its representation such as a map or image and therefore provides important features for geospatial modeling [3]. Recently convolution neural networks have performed exceptionally well in image segmentation tasks and can be used to segment objects at different scales. To bridge the gap from earlier machine learning methods such as SVMs or k-means clustering, image pyramids are now favored to detect objects at different scales. This is made all the more easy as researchers have found that convolutional neural network (CNN) architectures can be exploited to create multi-scale feature pyramids as they already follow a similar structure using convolutions and deconvolutions. According to [10], feature pyramids are a component in recognition systems to detect objects at different scales. However, due to their computational inefficiencies, many researchers have avoided using them. [10] shows that by exploiting the hierarchical structure of CNNs, it is possible to create scale-invariant feature sets. This allows recognition systems to segment any object at any scale by using different levels in the feature pyramid. The idea of image pyramids has been around for a long time. Adelson et al. [2] published one of the first ideas of image pyramids in 1984, proving that with just simple scaled convolutions, one can obtain scale-invariant image representations.

Building on the idea of multi-scale image representations [7] compared two different kinds of image pyramids for contrast enhancement of radiography images. Their results show that the Laplacian Pyramid enhanced large structures in images much better than fast wavelet transforms (FWT). However, when dealing with tiny details, a fast wavelet transform can outperform its counterpart. More recently, these image pyramids are now being implemented within the CNN architecture itself. In [13], Wang et al. introduced a scale-aware neural network that takes advantage of its current architectures to segment images captured at multiple spatial resolutions (MSR). Using a ResNet backbone to grab features produced at specific layers, they created multi-scale feature representations of input images that increase accuracy

results of large and small objects alike. Image pyramids are also not the only way of capturing multi-scale information from images. Methods such as multi-level feature fusion and spatial pyramid pooling architectures exist and show promising results. Audebert et al. [4] introduced a feature fusion technique in which features are fused from different modalities and layers in the neural network. Another method of retrieving multi-scale information from images is spatial pyramid pooling. Not only does this technique allow for an image of any size to be used as input into CNNs, but the architecture also allows for receptive field enlargement, which in turn helps it capture the multi-scale information.

3 General Aims

The general aim of this project is to understand appropriate deep learning methods of scale-aware image segmentation, in particular, very small flowers in field margins. Even though methods exist and show promising results, such as SVM and k-means clustering, the recent boom of deep learning has given way to a whole new area of research. Due to high results and their ability to generalize to other tasks, CNNs have become the front-runner of MSR image segmentation. Furthermore, with their inherent hierarchical structure, creating efficient scale-invariant solutions has become easier than ever and would allow CEH to apply these methods to the ever-growing amount of MSR imagery it produces.

As stated above, soil cover classification is an essential issue in today's world, and investigating these methods will make up the bulk of the work for this project. Firstly, research into image enhancement of tiny objects must occur as much of the current literature attempts to segment large vegetation in fields from a lower altitude. Secondly, scale must be addressed as simple CNN architectures have trouble segmenting the same objects at different scales.

The final product of the study should ideally be a deep learning system that can take, as input, MSR drone imagery of flowers in field margins and produce accurate segmentation maps using scale-aware features fused from multiple layers in a CNN. To reach state-of-the-art accuracy, the system must create segmentation maps with accuracy rates of at or above 90%.

4 Background Work

Initial research has already started being conducted looking at deep neural network architectures and how they can be applied to the task at hand. The first step to understanding what methods will work or not is having a processed dataset of images and their respective masks. The labeled dataset of field margins has already been created and must be masked, so the UNET architecture understands which areas to segment. One issue that arose from this is that after masking the images, there were still large areas of pure black pixels (value of 255 when looking at its matrix) that were not allowing the CNN to train correctly. To correct this, background masks were applied using the Rasterio and Geo-pandas python packages.

Once the corrections were made, the next step is to start researching image/contrast enhancement techniques and scale-invariant feature sets. From this, it was found that to increase object detection rates, image pyramids can be used to create multi-scale feature sets and enhance the contrast of details the models should be segmenting. Furthermore, [7] found that fast wavelet transforms enhance small details in images better than Laplacian pyramids. Due to these findings, work regarding multi-scale feature sets and wavelet transforms should take place, and after a literature review, it can be seen that it is a relatively unexploited area.

5 Objectives

Literature in this area has shown multiple methods and techniques that have been applied to the multi-scale enhancement of images and their subsequent segmentation. This project will aim to develop a viable deep learning architecture that will take drone MSR imagery and use it to track and monitor soil coverage over field margins. Below are key questions that will hopefully be answered throughout this project:

- How well does scale-invariant deep learning architectures work in regards to tiny objects (flowers)?
- How can the images be split to make the training dataset meaningful?
- Which wavelet transformation works best in the architecture?
- Which wavelet transform enhances the small details best?

- How well does the architecture generalize to other data sets?
- Do wavelets, rather than pooling, enhance the receptive field and produce higher accuracy results?

To answer the above questions, the following objectives must be completed:

- Review existing literature for soil cover classification and multi-scale models.
- Build the CNN and test its ability with the given data.
- Test the generalizability of the system and use completely different data sets.
- Compare accuracy results with other baseline deep learning models with the new blocks.
- Ablation study of the new architecture.

The outcomes of all objectives listed above will be:

- Fully documented convolutional neural network architecture that can segment flowers from field margins.
- A detailed comparison of current state-of-the-art methods along with the novel architecture created as a result of this project.

This project aims to create a novel CNN architecture for using MSR imagery to segment very small objects. The second goal is to understand the role that image/contrast enhancement plays in the results of these scale-invariant deep learning models. The research done in this project will help further CEH's knowledge into how well deep learning architectures can do given the task of extremely high-resolution image segmentation of small objects and hopefully give insight into soil cover distributions across the UK.

6 Methodology

This section will detail the methodological approach taken to complete this project given current time constraints. First, an initial literature review of current state-of-the-art methods will be done to find any existing gaps and begin understanding what works or does not given the geospatial data. Once that has concluded, the training set must be created. For this task, the python libraries Rasterio and Geo-pandas will be used to break down the data into masks and smaller subsets for the training process. No labeling has to be done as the data has been provided, fully labeled, from CEH.

Next, it will be time to start developing the novel architecture. Given the information acquired during the literature review, a CNN architecture will be created. After the new model has been created, an ablation study will be done on the model to stress-test it, and it will also be compared to other current state-of-the-art methods given the dataset.

7 Implementation

To complete this project, and given that I do not own a laptop that will reliably train such a network, I have been graciously given access to the High-End Computing Cluster at Lancaster University. This includes access to 40 cores, 192Gb of memory, and 6 Tesla V100 GPUs. Regarding the tech stack; Python will be used as the primary programming language, PyTorch as the neural network library given its ease of use when changing deep learning architectures, QGIS to allow easy access and edits to the geospatial data (GeoTIFF files), and all results will be visualized in python using various libraries.

8 Gantt Chart

Having issues formatting this, took it out to upload to GitHub.

References

- [1] Fao.org - the main principles of conservation.

- [2] Edward H Adelson, Charles H Anderson, James R Bergen, Peter J Burt, and Joan M Ogden. Pyramid methods in image processing. *RCA engineer*, 29(6):33–41, 1984.
- [3] Peter M. Atkinson and Nicholas J. Tate. Spatial scale problems and geo-statistical solutions: A review. *The Professional Geographer*, 52(4):607–623, 2000.
- [4] Nicolas Audebert, Bertrand Le Saux, and Sébastien Lefèvre. Beyond rgb: Very high resolution urban remote sensing with multimodal deep networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 140:20–32, 2018.
- [5] Xiaodong Bai, Zhiguo Cao, Yu Wang, Zhenghong Yu, Zhu Hu, Xuefen Zhang, and Cuina Li. Vegetation segmentation robust to illumination variations based on clustering and morphology modelling. *Biosystems Engineering*, 125:80–97, 2014.
- [6] SJ Corak, TC Kaspar, and DW Meek. Evaluating methods for measuring residue cover. *Journal of soil and water conservation*, 48(1):70–74, 1993.
- [7] Sabine Dippel, Martin Stahl, Rafael Wiemker, and Thomas Blaffert. Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform. *IEEE Transactions on medical imaging*, 21(4):343–353, 2002.
- [8] J.M. Guerrero, G. Pajares, M. Montalvo, J. Romeo, and M. Guijarro. Support vector machines for crop/weeds identification in maize fields. *Expert Systems with Applications*, 39(12):11149–11155, 2012.
- [9] Esmael Hamuda, Martin Glavin, and Edward Jones. A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture*, 125:184–199, 2016.
- [10] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.

- [11] María José Marques, Ramón Bienes, Luis Jiménez, and Raquel Pérez-Rodríguez. Effect of vegetal cover on runoff and soil erosion under light intensity events. rainfall simulation over usle plots. *Science of The Total Environment*, 378:161–165, 5 2007.
- [12] Peter Riegler-Nurscher, Johann Prankl, Thomas Bauer, Peter Strauss, and Heinrich Prankl. A machine learning approach for pixel wise classification of residue and vegetation cover under field conditions. *Biosystems Engineering*, 169:188–198, 5 2018.
- [13] Libo Wang, Shenghui Fang, Ce Zhang, Rui Li, Chenxi Duan, Xiaoliang Meng, and Peter M. Atkinson. Sanet: Scale-aware neural network for semantic labelling of multiple spatial resolution aerial images, 2021.