In the Name of God

Ph.D. Thesis

_____

Aerial Image Analysis Applications in Plant Discrimination and Land Cover Mapping for Agro-Environmental Studies

_____

Author:

Mohammad Mardanikorani

Supervisor:

Professor Takafumi Saito

Graduate School of Bio-Applications and

Systems Engineering

Tokyo University of Agriculture and Technology

March 2020

Hidehiro Kamiya, Ph.D.

Professor of Department of Food and Energy Systems Science

Tokyo University of Agriculture and Technology

Atsushi Akisawa, Ph.D.

Professor of Department of Food and Energy Systems Science

Tokyo University of Agriculture and Technology

Takafumi Saito, Ph.D. (Supervisor)

Professor of Department of Computer and Information Sciences

Tokyo University of Agriculture and Technology

Kaori Fujinami, Ph.D.

Professor of Department of Computer and Information Sciences

Tokyo University of Agriculture and Technology

Ikuko Shimizu, Ph.D.

Associate Professor of Department of Computer and Information Sciences

Tokyo University of Agriculture and Technology

Aerial Image Analysis Applications in Plant Discrimination and Land Cover Mapping for Agro-

Abstract

In this thesis we introduce and evaluate a methodology to analyze, discriminate, and classify the earth surface by utilizing aerial images from different source and sensors in agro-environmental applications. In such applications, often accessing the area of interest is practically impossible or it is time consuming and requires high cost. Therefore, utilizing aerial images for such applications is the most common option for the users.

In first part, we focus on aerial images from unmanned aerial vehicles (UAVs) with Red, Green, and Blue (RGB) channel sensors. The overall objective in this part is to provide a methodology for analyzing the aerial images from a small commercial UAV in order to be used in agro-environmental applications. First, we select the hairy vetch (Vicia villosa) as target plant to be detect and discriminated in UAV images. We characterized the features and collected the data for further analyze. Then, the challenges in image mosaicking in such applications were investigated. In the next step, we introduced a methodology for plant detection in UAV images by using machine learning techniques. Finally, the proposed methodology was test and evaluated.

In the second part, we modify the methodology for landcover mapping by utilizing satellite imagery. The overall objective in this part is to provide a tool for landcover classification in national level for the country of Lesotho. The study was carried out in collaboration with the United Nations Food and Agriculture Organization. the proposed methodology utilizes free and open access satellite data, cloud based geospatial data processing platforms, and FAO datasets. The methodology was successfully test and evaluated in national level for Lesotho.

Finally, the thesis will discuss the potential, limitations and future directions of the study.

ACKNOWLEDGEMENT

First, I would like to thank you professor Takafumi Saito whom that support my study and my research with his invaluable and endless kindness. His knowledge and guide where the most valuable source for my study.

Also, I would like to appreciate professor Yoshiharu Fujii for letting me to join his lab and ideas that were formed part of the research. I thank professor Ikuko Shimizu for her kind support during the research. Also, I appreciate all the supports and guide that I received from Dr. De Simone, Dr. Samuel Varas, and Dr. Boliko at Food and Agriculture Organization of the United Nations.

Lastly, I would like to thank my friends and family, does who their unconditional love was the main reason that I could continue my way. Specially, I would like to express my deepest gratitude toward my amazing mother for her love and care. And, I am sincerely grateful to my father for his special support.

I greatly appreciate the moral and financial support provided by the Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT) and Department of Food and Energy Systems Science in Tokyo University of Agriculture and Technollogy.

Table of Contents

List of Tables

List of Figures

Chapter 1: INTRODUCTION

In the stories, religions, and novel books, looking from the above and watching the world from the sky is an advantage and superpower that belongs to the gods or supernatural creatures. In ancient civilizations, building towers, pyramids, or even kites for observing the ground from the upper levels were considered as superior technology. Looking at the ground from higher levels let the human to see the whole instead of parts. It helped the rollers of human the history to understand the overall facts on the earth. The fact that we, as humans, always try to go higher and higher was the reason for us to make Babylonia's towers, invent the balloons, kites, airplanes etc. In fact, we were so eager to see the earth from sky and have the gods power that, in 19th century, we got so high that we left the atmosphere and entered the space. Access to the sky led human being to think about how to observe the earth surface and take advantage of those new horizons. And of course, the best way to think about it and share the new knowledge with others was to capture the image of it.

The aerial imagery's history starts from the French photographer Gaspard F. Tournachon also known as Nadar. He first practiced the aerial photography from balloons in 1850's. Also, some of the first aerial photographs that exist today are from Boston, USA by James Wallace Black in1860 (Graham and Read 1986). In addition, early unmanned aerial imagery was utilizing string kites and Pigeons (known as Bavarian Pigeons). From then the aerial images were used in vast variety of application such as military, urban monitoring, forest monitoring, ocean observations, agriculture, geology, fire and earthquake monitoring etc.(Banzhaf and Hofer 2008; Lomax, Corso, and Etro 2005; Martínez-de Dios et al. 2011; Merino et al. 2010).

In general, there is four different types of aerial imagery methods for collecting the data: (1) the tower aerial imagery (2) airborne aerial imagery (3) satellite aerial imagery, and (4) drone aerial imagery. The tower imagery is cost effective and suitable for acquiring data from a fixed location during a long period of time, but its disadvantage is the angle of the sensor and field of the view that effects the quality and methodology of processing data. The airborne imagery is the imagery using sensors that are mounted on aircrafts, helicopter, or balloons. Airborne sensors are capable of capturing high spectral and spatial data in larger areas. Also, it is flexible toward the frequency of imagery and is suitable for making high resolution time series for larger areas. However, airborne imagery is expensive and requires experts to fly the equipment's.

The satellite imagery has the ability of covering very large areas on the earth. The frequency of imaging (revisiting period or temporal resolution) is different for each satellite mission and location of the area on planet. Meanwhile, satellite sensors usually provide the higher spectral

resolution (several band) data which allows further information acquisition. Disadvantage of satellite imagery is that mostly they are not open access and it need the experts to acquire the data and analyze it (Gilvear and Bryant 2003). The drone imagery uses the unmanned aerial vehicles (UAV) for data acquisition. The advantages of this type of imagery is that: first, it is very cost effective. Due to the recent advances in UAV technologies the commercial UAVs that are getting cheaper and they are accessible for everyone. Second, UAVs can fly almost everywhere and anytime. The ease of use, light wait, and cost effectivity of the UAVs allows the users to be able to fly them every time that it is necessary (Aasen et al. 2018).

The aerial image analysis is extracting the meaningful information from aerial sensors data by utilizing image processing techniques. The aerial image analysis goal is to provide effective and efficient tools for remote sensing applications. The image processing techniques that are used in different application (such as medical image processing) are able to offer new tools for remote sensing applications. However, generally remote sensing method are only depending on spectral analysis of images. It can be due to the early days of remote sensing that aerial images had very coarse spatial resolution and the most important part of data was available in spectral data. Here we should mention that the scale of images has little importance to the method that used to analyze them. A microscope image with micrometer resolution and a satellite image with several meter resolution require many of the same algorithms. The measurement of images is the principal method for acquiring scientific data, and generally requires that objects or structure be well defined, either by edges or by unique brightness, color, texture, or some combination of these factors (Russ 2016).

1.1 Current studies and methods

The use of satellite imagery in agriculture and environment applications is still the most dominant data. Satellite products make up 52% of all studies, followed by airborne (34%), field measurements (e.g. through field spectrometers; 11%) and UAV (2%) (Vaz et al. 2018). At first here, we will have a survey on UAV applications in agriculture and environment field. Then we will look at studies that use the satellite data for landcover products.
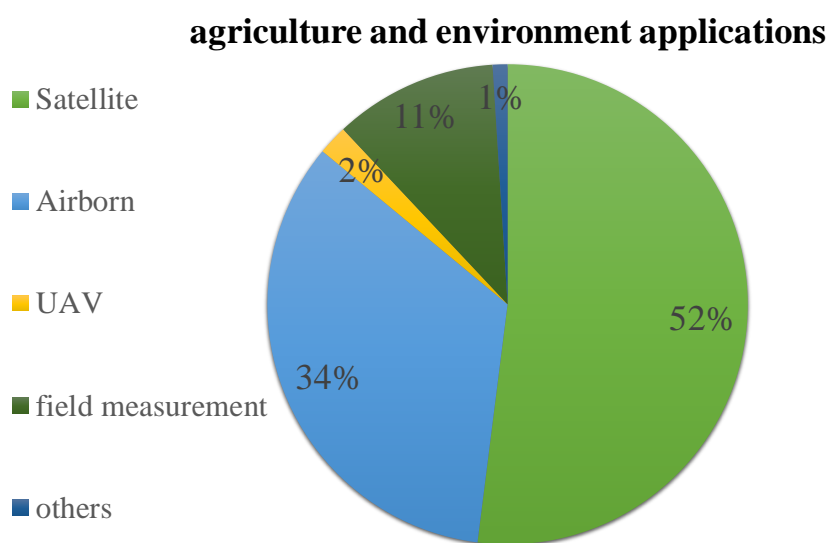
Figure 1-1 The share of each type of aerial imagery in agriculture and environmental applications

Satellite image processing is applied in a wide range of applications. the satellite data has several unique characteristics such as ability to collect the long-term time series due to the temporal resolution of satellite data (revisiting time), hyperspectral imaging (sensors with several visible and invisible bandwidth) and data collection coverage (they can cover vast area on the planet). Satellite time series have been used for obtaining the spectral footprint of plant or other target objects on the earth surface (Fattahi, Agram, and Simons 2016; Inglada et al. 2017; Molero et al. 2018; Pelletier et al. 2016). Also, time series are used to compare the plant or objects status over a period of time (Halabisky et al. 2016).

One of the most applied field for satellite data is landcover mapping and landcover processing. Landcover is the physical objects and materials on the surface of earth such as grass-land, barren-land, wetland, farms and agriculture etc. The landcover data provide information on the location

and type of the land surface on map (Almeida et al. 2016; Grekousis, Mountrakis, and Kavouras 2015; Yifang, Gong, and Gini 2015).

In last 10 years the advances in battery, flight control, and electric motor accelerated the progress in producing smaller, cheaper, and more capable UAVs. Today UAVs are utilized in a very wide range of applications including agriculture and environment. The agricultural UAVs carry two types of payloads: 1) material to be applied on farms (Material delivery) (Faiçal et al. 2017) 2) sensors to collect data. The data collection is necessary for first type of application (material delivery). There are several types of sensors for UAVs based on their type of application. The most common type is the red, green, and blue (RGB) channel sensors. The other type is multi-spectral or hyperspectral sensors. These sensors can collect the data in near infrared or infrared spectrum. However, these types of sensors are more expensive and require higher level of expertise to be utilize by individuals. There are several studies that utilizing multi-spectral UAV imagery for agriculture purposes (Ampatzidis and Partel 2019; Navia et al. 2016).

There is limited amount of effort for application of UAVs in dense and complex natural habitat of plants. For agro-environmental studies it is necessary to be able to access to remote aerial that is impossible or near impossible to reach by human. These areas are usually located in remote
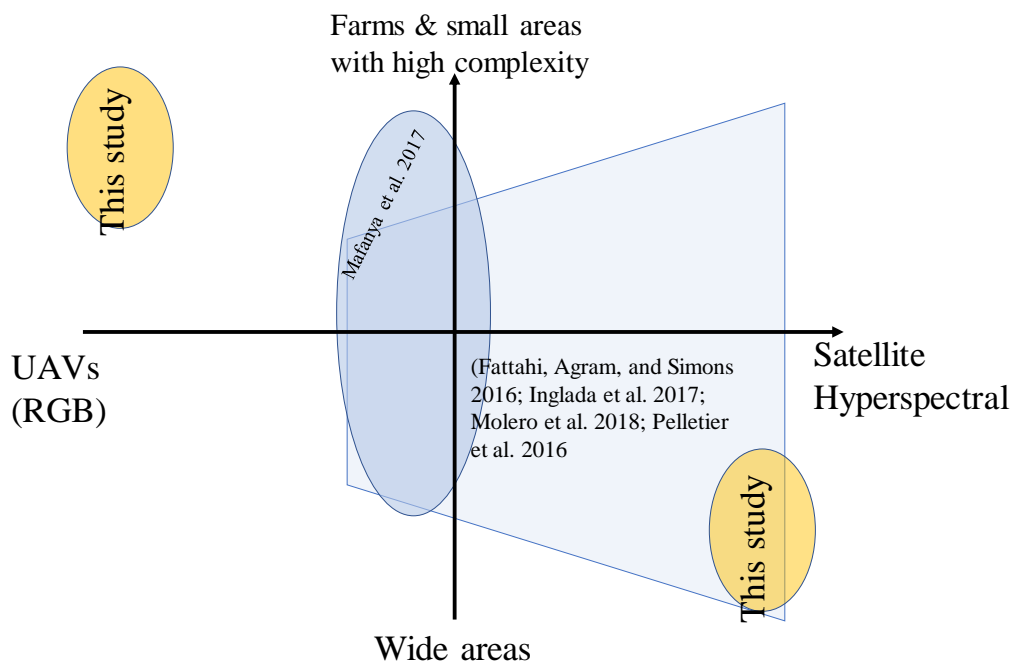
Figure 1-2 the location of this study between other similar studies in term of image complexity and type of sensor

areas and require light wait equipment's to carry. Therefore, small commercial UAVs are the most promising tools for such applications. However, such UAVs are usually equipped with RGB sensor and have moderate spatial resolution. Also, in such applications the area under the survey is usually has very complex vegetations habitat or irregular plantation order. In figure 1-2 we explain how this study is located among other studies in the term of complexity of images and the type of the sensor that has been used in other efforts. Also, figure 1-2 shows that this study covers two areas of studies in both satellite and UAV image analysis that have not been cover with other studies. These to areas are 1) utilizing low spectral and spatial sensors for plant discrimination in natural environment and 2) utilizing satellite image processing for landcover mapping of very wide areas (national level) using machine learning techniques.
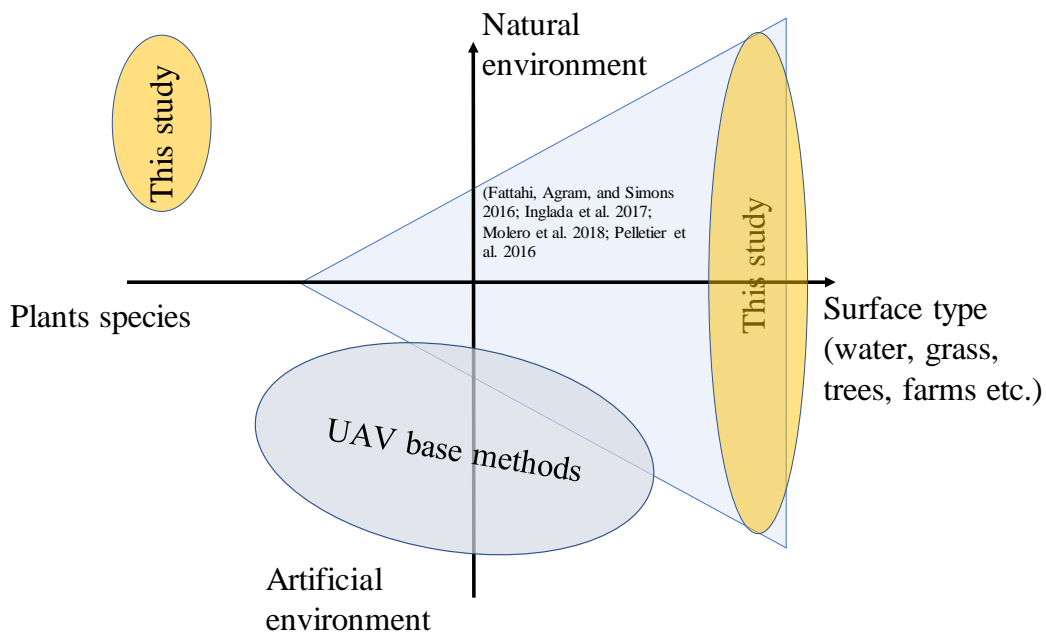
Figure 1-3 location of this study among others in term of type of the environment and the objective of studies

In this study, in order to provide a solution to this problem, we worked on the idea to recommend appropriate aerial images processing methodology that can be utilized for agro-environmental applications. By implementing this methodology, we aim to contribute to solving the problem that existing methodologies are focused on site specific or data specific application but not the target objects (figure 1-3). Here is a summary of our contributions:

- Focus on commercial UAVs with inexpensive sensors and open access satellite data

- Try to propose method that work with any environment weather it is the satellite imagery or UAV images. The goal is to provide flexible method that works with minimum modifications.
- Propose method that can classify the earth (or land) surface based on user demand. It means the method should work with minimum amount of labelled (ground truth data)

The objectives of the study are:

- To propose a land surface classification methodology based on inexpensive aerial imagery.
- To discriminate the plant species in their natural habitat and discriminate the complex land surface types using aerial imagery.
- To introduce a land surface classification method that is robust against changes in location of the area, land surface type, or etc.
- To propose a methodology that is capable of utilizing various aerial imaging sensor types.

1.2 The selected case studies

In the first part of this study we selected two plant species as our target plants to be detected in aerial images. Here we will describe the characteristics and visual specifications of each plant.

1.2.1 Hairy vetch

Hairy vetch (Vicia villosa Roth L.) is an annual or winter annual legume. It is most often grown for soil improvement, roadside, or bank stabilization. It can also be grown as a pasture or hay crop, but it can be toxic to horses, especially on high selenium soil. It grows well on a broad range of soils but is best adapted to sands or loamy sands (Singer et al, 1999).

Hairy vetch Vicia villosa is a hardy winter annual cover crop and can add 80–150 pounds/acre of nitrogen to soil (Figure 1-4). the Hairy vetch growth is small during the winter; therefore, it is recommended to grow it along with other grasses during the summer in order to achieve enough soil coverage. Hairy vetch can claim the physical support structures during the spring and adding rye is recommended to decrease the matting and leveling up the incorporation into the soil. Hairy vetch also can be planted during the spring as a follow crop; thus, it can act as a cover through late summer. It is drought tolerant it has the highest cold tolerance among other winter annual legumes. (Brust 2019).

Figure 1-4 Hairy vetch in its natural habitat near the Tone river in Japan

Hairy vetch's round stems have spreading white hairs and conspicuous ridges. The leaves clusters are up to 25 cm long and 5 cm across, and they are consisting of 8-12 pairs of leaflets and a terminal tendril that grabs adjacent plants for climbing. The central stalk of each leaf cluster has white and smooth hair spreading uniformly on it. The length of leaflets are up to 25 mm and they are 7 mm in width each. They are linear and have smooth margins with small pointed tips. long and small cluster of flowers can develop up to 18 cm long from the leaf clusters. The stalks of these flower clusters are covered with spreading hairs and usually turn purple toward their tips. The flower clusters are hanging on one side of the cluster stems and they consist of 5-20 pairs of nodding flowers. These flowers vary in color from pink to blue-violet, depending on the strain or local ecotype. This color variation which add extra challenge in discrimination tasks. Each tube shape flower is around 7-14 mm long, and they are consisting of 5 petals and a hairy calyx. Aside from its long tube-shaped base, the flower typically is similar to peas in structure. Flowers are consisting of a standard, keel, and 2 side petals. The upper petal forming the standard is usually a darker shade of color than the remaining petals (www.illinoiswildflowers.info).

1.2.2 Common reed

Common reed also known as phragmites or phragmites australis can be observed in disturbed and undisturbed wet areas including wetlands, salt and fresh-water flooding areas, river edges, lake shores and ponds, roadsides and ditches. Phragmites can tolerate fresh and salinized waters, and It prefers to grow in fully sunny places. (Hellings and Gallagher 1992). Phragmites is a tall and

Figure 1-5 Common reed also called Phragmites grow near the road and wetlands in Itakura city, Japan

warm season grass that last for a long time one area. The stems can grow up to 6 meters tall and they are smooth, rigid, and erected upward. The leaves are stiff, and between 20 to 40 cm long and 1-4 cm a cross. Flowers occur during the summer and are arranged in tawny spikelet's with tufts of silky hair. Flower's hair is purple like color but turn to brown to dark brown at its maturity. Common reed is especially common in alkaline and brackish environments. Common reed can survive in highly acidic wetlands environments. (indiananativeplants.org).

Phragmites expansion into tidal wetlands of North America is most extensive along the northern and middle Atlantic coasts, but over 80% of the US coastal wetland area (Chambers, Meyerson, and Saltonstall 1999).

1.3 References

Aasen, Helge, Eija Honkavaara, Arko Lucieer, and Pablo J Zarco-Tejada. 2018. 'Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows', *Remote Sensing*, 10: 1091.

Almeida, Cláudio Aparecido de, Alexandre Camargo Coutinho, Júlio César Dalla Mora Esquerdo, Marcos Adami, Adriano Venturieri, Cesar Guerreiro Diniz, Nadine Dessay, Laurent Durieux, and Alessandra Rodrigues Gomes. 2016. 'High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data', *Acta Amazonica*, 46: 291-302.

Ampatzidis, Yiannis, and Victor Partel. 2019. 'UAV-based high throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence', *Remote Sensing*, 11: 410.

Banzhaf, Ellen, and Rene Hofer. 2008. 'Monitoring urban structure types as spatial indicators with CIR aerial photographs for a more effective urban environmental management', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1: 129-38.

Brust, Gerald E. 2019. 'Management Strategies for Organic Vegetable Fertility.' in, *Safety and Practice for Organic Food* (Elsevier).

Chambers, Randolph M, Laura A Meyerson, and Kristin Saltonstall. 1999. 'Expansion of Phragmites australis into tidal wetlands of North America', *Aquatic botany*, 64: 261-73.

Faiçal, Bruno S, Heitor Freitas, Pedro H Gomes, Leandro Y Mano, Gustavo Pessin, André CPLF de Carvalho, Bhaskar Krishnamachari, and Jó Ueyama. 2017. 'An adaptive approach for UAV-

based pesticide spraying in dynamic environments', *Computers and Electronics in Agriculture*, 138: 210-23.

Fattahi, Heresh, Piyush Agram, and Mark Simons. 2016. 'A network-based enhanced spectral diversity approach for TOPS time-series analysis', *IEEE Transactions on Geoscience and Remote Sensing*, 55: 777-86.

Gilvear, David, and Robert Bryant. 2003. 'Analysis of aerial photography and other remotely sensed data', *Tools in fluvial geomorphology*: 135-70.

Graham, Ron, and Roger E Read. 1986. *Manual of aerial photography* (Focal Press).

Grekousis, George, Giorgos Mountrakis, and Marinos Kavouras. 2015. 'An overview of 21 global and 43 regional land-cover mapping products', *International Journal of Remote Sensing*, 36: 5309-35.

Halabisky, Meghan, L Monika Moskal, Alan Gillespie, and Michael Hannam. 2016. 'Reconstructing semi-arid wetland surface water dynamics through spectral mixture analysis of a time series of Landsat satellite images (1984–2011)', *Remote Sensing of Environment*, 177: 171-83.

Hellings, Samuel E, and John L Gallagher. 1992. 'The effects of salinity and flooding on Phragmites australis', *Journal of Applied Ecology*: 41-49.

Inglada, Jordi, Arthur Vincent, Marcela Arias, Benjamin Tardy, David Morin, and Isabel Rodes. 2017. 'Operational high resolution land cover map production at the country scale using satellite image time series', *Remote Sensing*, 9: 95.

Lomax, AS, W Corso, and JF Etro. 2005. "Employing unmanned aerial vehicles (UAVs) as an element of the Integrated Ocean Observing System." In *Proceedings of OCEANS 2005 MTS/IEEE*, 184-90. IEEE.

Martínez-de Dios, José, Luis Merino, Fernando Caballero, and Anibal Ollero. 2011. 'Automatic forest-fire measuring using ground stations and unmanned aerial systems', *Sensors*, 11: 6328-53.

Merino, Luis, Fernando Caballero, J Ramiro Martinez de Dios, Ivan Maza, and Aníbal Ollero. 2010. "Automatic forest fire monitoring and measurement using unmanned aerial vehicles." In *Proceedings of the 6th International Congress on Forest Fire Research*. Citeseer.

Molero, Beatriz, DJ Leroux, P Richaume, YH Kerr, Olivier Merlin, MH Cosh, and R Bindlish. 2018. 'Multi‑Timescale Analysis of the Spatial Representativeness of In Situ Soil Moisture Data within Satellite Footprints', *Journal of Geophysical Research: Atmospheres*, 123: 3-21.

Navia, Jose, Ivan Mondragon, Diego Patino, and Julian Colorado. 2016. "Multispectral mapping in agriculture: Terrain mosaic using an autonomous quadcopter UAV." In *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, 1351-58. IEEE.

Pelletier, Charlotte, Silvia Valero, Jordi Inglada, Nicolas Champion, and Gérard Dedieu. 2016. 'Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas', *Remote Sensing of Environment*, 187: 156-68.

Russ, John C. 2016. *The image processing handbook* (CRC press).

Singer, J.W., Bobsin, N., Bamka, W.J. and Kluchinshi, D., 1999. Horse pasturemanagement. Journal of Equine Veterinary Science, 19(9), pp.540-592.Chapter 2

Vaz, Ana Sofia, Domingo Alcaraz-Segura, Joao C Campos, Joana R Vicente, and Joao P Honrado. 2018. 'Managing plant invasions through the lens of remote sensing: A review of progress and the way forward', *Science of the Total Environment*, 642: 1328-39.

Yifang, Ban, Peng Gong, and Chandra Gini. 2015. 'Global land cover mapping using Earth observation satellite data: Recent progresses and challenges', *ISPRS journal of photogrammetry and remote sensing (Print)*, 103: 1-6.

Chapter 2: Detection and mapping of hairy vetch in images obtained by UAVs

2.1 Abstract

Recent advances in unmanned aerial vehicles (UAVs) for civilian use make it possible to regularly monitor fields at spatial and temporal scales that would be difficult to achieve using conventional methods. Mapping the undercover area of a specific plant in images can contribute to compare its annual expansion ability. Hairy vetch is a moderately winter-hardy species. In this study we try to detect hairy vetch in aerial images by using color and texture features. Images obtained from Tone river shore in Gunma prefecture, Japan. Color features extract based on hairy vetch leaves and flowers separately. Grey Level Co-occurrence Matrix (GLCM) used for texture feature extraction. Results show feasibility of using purpose method in discriminates between hairy vetch and different plants in aerial imagery. The study highlights advanced techniques in sampling, processing and analyzing UAV images and identifies some research challenges and limitations in the use of UAV platforms. This project demonstrates the utility of UAV applications in agriculture and the potential for their use in many other areas of research.

2.2 Introduction

Recent advances in unmanned aerial vehicles (UAVs) for civilian use make it possible to regularly monitor fields at spatial and temporal scales that would be difficult to achieve using conventional methods. The unique characteristics of aerial images provide by UAVs like high resolution, low cost, fast, reliability and easy to use encouraged studies about plants detection and recognition systems. Another important exclusivity of UAV image refers to distance between reflection surfaces and sensors which distributes to receive more accurate information. Discrimination of the plants in their different stages of development needs images at very high spatial resolution, often in the order of millimeters or very few centimeters (Hengl, 2006 and López-Granados, 2011). Also, it is required that the image acquaintance process should be done in specific time of plant's development stage. The best suitable tools with both characteristics are unmanned Aerial Vehicles; UAVs fly at low altitude (maximum flight altitude in Japan is 150 meters) which can provide high resolution images. also, UAVs can be lunched in a very short time which allow to take images in any plants' life stage that is necessary to discrimination.

Figure 2-1 Hairy vetch flower

Detect and mapping the plants species can have a various application in agriculture and environmental field. Mapping the undercover area of a specific plant in georeferenced images can contribute to comparing the annual expansion ability of that plant species. Expansion ability determine the ability of a specific plant to be either an invasive variety or not. In this study we try to illustrate the feasibility of purposed method in discriminating plants species in real field complex aerial images. Another end for this study is to detect hairy vetch plant in its natural habitat. Detecting the hairy vetch will prepare aerial images to map the covered area for compare the annual expansion of it.

Hairy vetch also known as sand vetch is a plant species known for its moderately winter-hardy characters. organic farmers grow hairy vetch as cover crop in no-till farming since it is both winter-hardy and fixes the atmospheric nitrogen in soil. Hairy vetch can claim and scramble up to 2 meters. The whole plant has a white-woolly appearance because of the long soft hairs. self-fertilized flowers mainly purple to blue in color but sometimes white (J. Frame, FAO website). Hairy vetch was introduced to Japan as cover crop in abounded patty fields. Fujii et. al show that hairy vetch has the best results for the weed control in abandoned fields due to its ability to die off during summer season to make a thick straw-like mulch layer (Y. Fujii, 2001).

2.3 Related works

J.m. Pena et al (2015) tried to detect weeds seedling in sunflower field. The study attempt to perform early weed detection in sunflower and review the limit and efficiency of methodology with UAV images. They showed that the weed discrimination ability is significantly affected by the spectral, spatial, and temporal resolutions of the imagery and camera. The color-infrared images captured at 40 m and 50 days after sowing (date 2), when plants had 5–6 true leaves, had the highest weed detection accuracy (up to 91%). Slightly better results were observed at 40 m altitude images in compare with the images that were captured during the other flights. However, this trend changed in the visible-light images captured at 60 m and higher, which had notably better results because of the larger size of the weed plants. The study showed that sufficient spatial and spectral resolution and the right design of flight plans are required for detecting and mapping the weeds.

Their proposed algorithm consists steps as below:

- Field segmentation in sub-plots
- Sub-plots segmentation in objects
- Vegetation objects discrimination:
- Excess Green index (ExG)
- Normalized Difference Vegetation Index (NDVI)
- Crop-row classification
- Weed and crop discrimination
- Weed coverage assessment

The entire process is automatic and is composed of a sequence of routines. They implemented these steps in eCognition software.

C. Huang et al (2014) proposed learning-based approach by introducing a feature learning approach to minimize the supervised classification methods. They apply this system to the classification of invasive weed species. In this study, the feature learning was applied to generate a set of image filters that can extract image features so that the weeds and background obejcts can be discriminated. These features are pooled to summarize the image statistics and form the input to a texton-based linear classifier that classifies an image patch as weed or background. methodology was evaluated for weed classification on three weeds species in Australia namely: water hyacinth, tropical soda apple and serrated tussock. The results showed the best performace and accuracy in 5-10 flight altitude images with F1 score of 94%.

2.4 Proposed method

We propose the following new method based on color features and Gray-Level Co-occurrence Matrix (GLCM). Our methodology consist of the following steps:

a. data collection
b. image pre-processing
c. classification

A multi-copter was used to gather all the imagery in this study. Provided images was passed into a feature extractor to determine local feature vector based on color and GLCM. Also, a set of sample images provided and pass into the feature extraction step. Feature vector passed to a classifier to be classify base on similarity with sample images features. Our methodology can be described in figure (2-2).
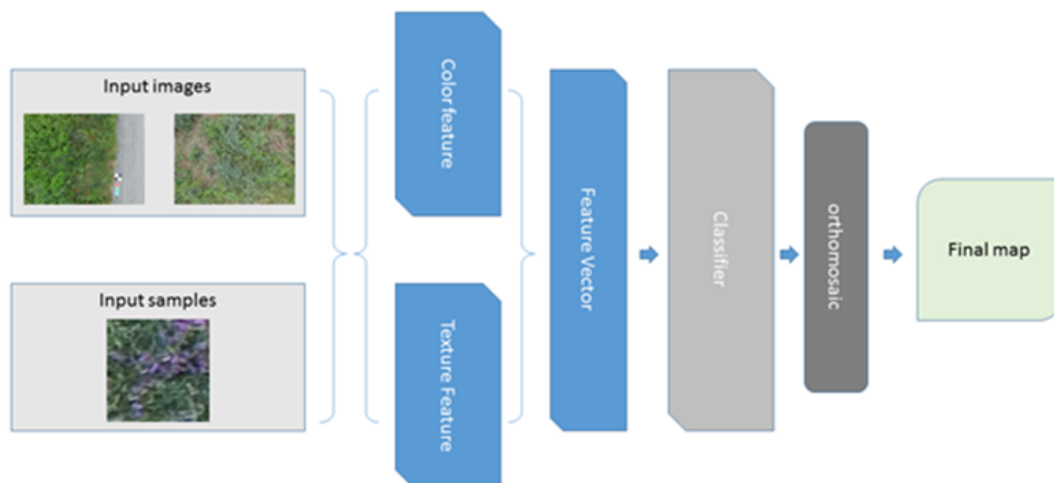
Figure 2-2 Flowchart of proposed method

a. Data collection

A multi-copter was used to gather all the imagery in this study. The UAV used in this study is DJI Phantom 3 advanced model. UAV technical characteristics are shown in table (2-1). Pix4d capture App was used for control and doing imagery in an automatic process.

Table (2-1). UAV technical characteristics.

Manufacturer, model	DJI, Phantom 3 advanced
Weight (including battery and propellers)	1280 g
Diagonal Size (Excluding Propellers)	350 mm
Max Ascent Speed	5 m/s
Max Speed	16 m/s (no wind)
Max Service Ceiling Above Sea Level	6000 m
Max Flight Time	Approx. 23 minutes

UAV camera was DJI camera with Sony EXMOR 1/2.3" sensor. Image resolution is 3000*4000 pixel. Camera technical characteristics are presented in table (2-2).

Table (2-2). Camera technical characteristics.

Manufacturer	DJI
Sensor	Sony EXMOR 1/2.3" Effective pixels: 12.4 M (total pixels: 12.76 M)
Lens	FOV 94° 20 mm (35 mm format equivalent) f/2.8, focus at ∞
ISO Range	100-1600 (photo)
Shutter Speed	8s -1/8000s
Image Max Size	4000×3000

All the imagery was done June 6th, 2016 at Itakura town, Gunma prefecture, Japan. The weather was fine with a tiny cloud covered almost all the sky. But the light condition was enough and natural. Figure 2-3 shows the location the study areas.
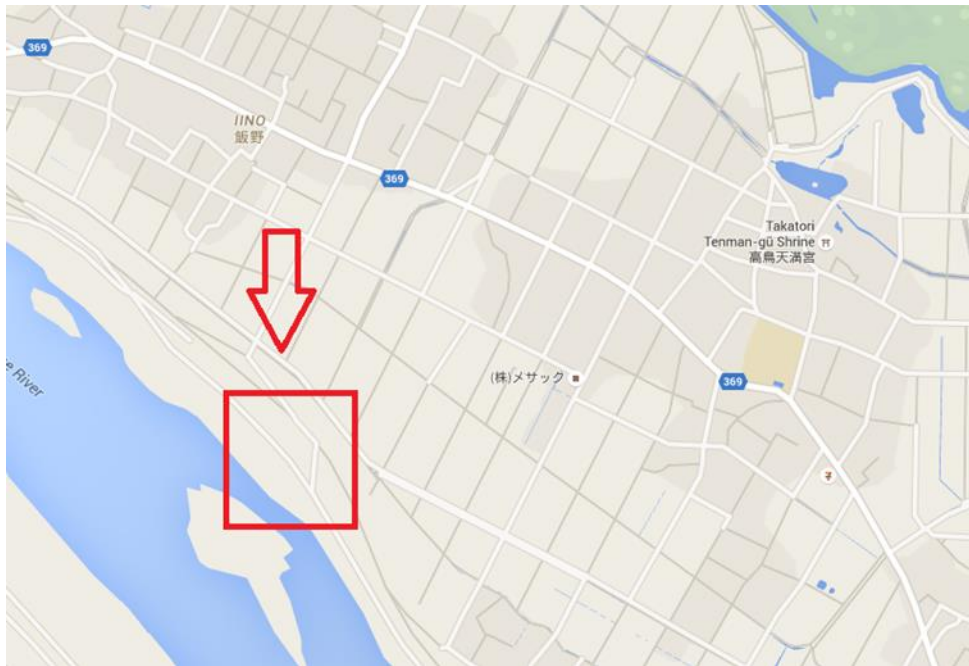
Figure 2-3 Study area: Tone river shore, Itakura city

At first two area were check by agriculture experts to select best data collection sites in real condition. C. Huang et al (2014) showed 5~10 m altitude has the highest results for weed classification in UAV imagery. Therefore, due to the time limits in imagery, altitude of 10m was selected for data collection. Sites orthomosaic images are shown in figure 2-4.
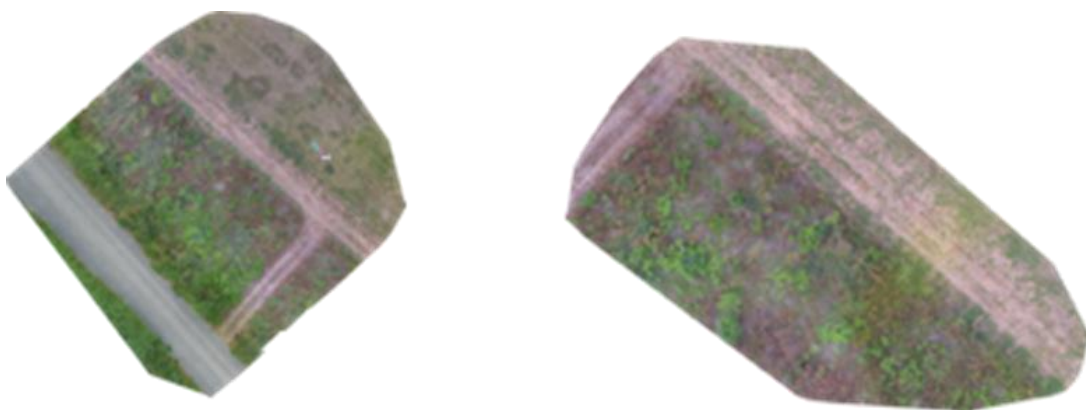
Figure 2-4 Orthomosaic images
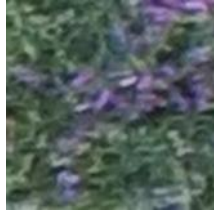
Figure 2-5 Entered sample image in algorithm

Figure 2-6 Query image and ground-trusted tag

Two 1 in 1-meter areas in one of the sites were tagged by red ribbon to select the ground-trusted samples. Sample images for extracting texture feature vectors were obtained from ground-trusted areas.

In order to extract the color features, it is necessary to obtain mean of intensity in each color channel of RGB images. Therefore, Hairy vetch cultivated in greenhouse during winter. A set of 108 images provided under natural light in greenhouse. Next, by cutting the leaves out of the images manually, a set of 287, 250 by 250 pixels sample images with black background was extracted. Figure 2-7 and 2-8 show the images taken in greenhouse and the image samples that were extracted for the next step.

Figure 2-7 Hairy Vetch grown in greenhouse

Figure 2-8 Cut leaves samples to extract color parameters

For each image of sample leaf; mean of pixel intensity calculated for Red, Green and Blue channels. Then the total maximum, minimum, mean and standard deviation of each channel determined over all 287 images. Table (2-3) illustrates the number of calculated parameters.

Table 2-3. Color parameter for leaves

	Red	Green	Blue
variance	180.2	237.8	233.0
max	113.0	121.0	108.2
min	33.2	42.2	23.7

Standard deviation	13.4	15.5	15.3
Mean	63.5	77.6	53.2

To obtain color parameters of hairy vetch flowers; samples were extracted from the original aerial images that were taken in study site (figure 2-9). The Same procedure as leaves parameters calculation applied to obtain color parameter of flowers. Table (2-4) illustrates the number of calculated parameters.
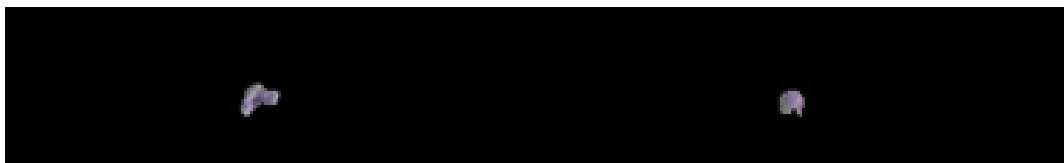
Figure 2-9 hairy vetch cut flower samples

Table 2-4. Color parameters for flowers

	Red	green	blue
variance	247.9	216.2	254.5
max	141.3	128.3	160.7
min	82.4	76.3	109.7
Standard deviation	15.7	14.7	15.9
mean	113.8	105.1	131.2

b. Image pre-processing

Color is the basic feature; color histograms are commonly used for color feature extraction. The color histogram method requires simple calculation. However, it is unsuitable for images in which there is a great color variation. But it does not include any spatial information.

To extract color feature two color mask applied on input images. First for leaves and other one for flowers. By using the color histogram and color parameters which obtained before; color masks boundaries selected as presented in table (2-5,2-6).

Table (2-5). Maximum and Minimum boundaries for leaves color.

	Red	Green	Blue
Minimum	50	62	38
Maximum	76	92	58

Table (2-6). Maximum and Minimum boundaries for flowers color.

	Red	Green	Blue
Minimum	98	91	111
Maximum	128	119	146

The result for color masking shown in figure (2-10).
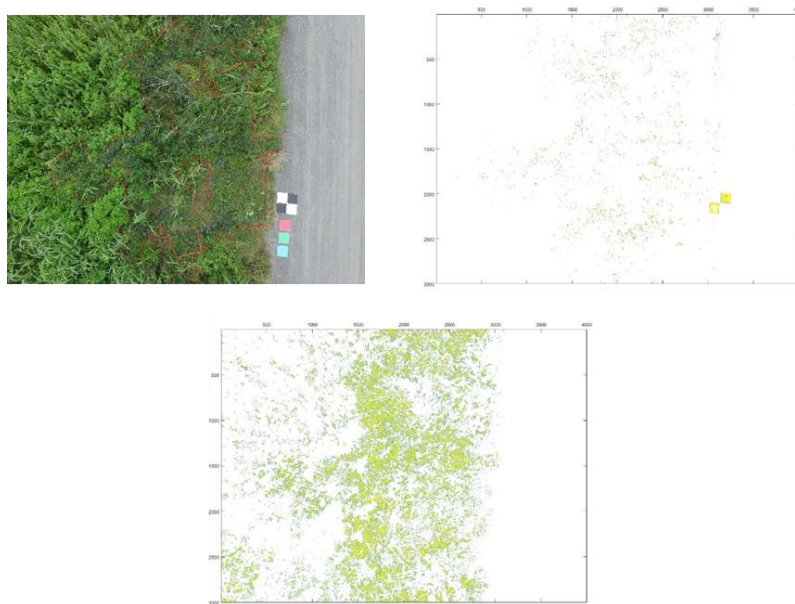
Figure 2-10 Sample results for color filtering

Next step in image pre-processing is to extract the texture features of aerial images and sample images. Texture is one of the most important features in an image. Textural features contain information about the spatial distribution of tonal variation within a band (Haralick et al 1973). A co-occurrence matrix C is a matrix that is defined over an image to be the distribution of co-

occurring values at a given offset. This matrix is defined over an $n \times m$ image I, parameterized by an offset $(\Delta x, \Delta y)$, as:

$$C_{\Delta x, \Delta y}(i,j) = \sum_{p=1}^{n} \sum_{q=1}^{m} \begin{cases} 1, & if \ I(p,q) = i \ and \ I(p + \Delta x, q + \Delta y) = j \\ 0, & otherwise \end{cases}$$

Where i and j are the image intensity values, p and q are the spatial positions in the image I and the offset $(\Delta x, \Delta y)$ depends on the direction and distance which used. The offset distance and direction are shown in figure (2-11).

```
Offsets = [0 1; 0 2; 0 3; 0 4;...

          -1 1; -2 2; -3 3; -4 4;...

          -1 0; -2 0; -3 0; -4 0;...

          -1 -1; -2 -2; -3 -3; -4 -4];
```
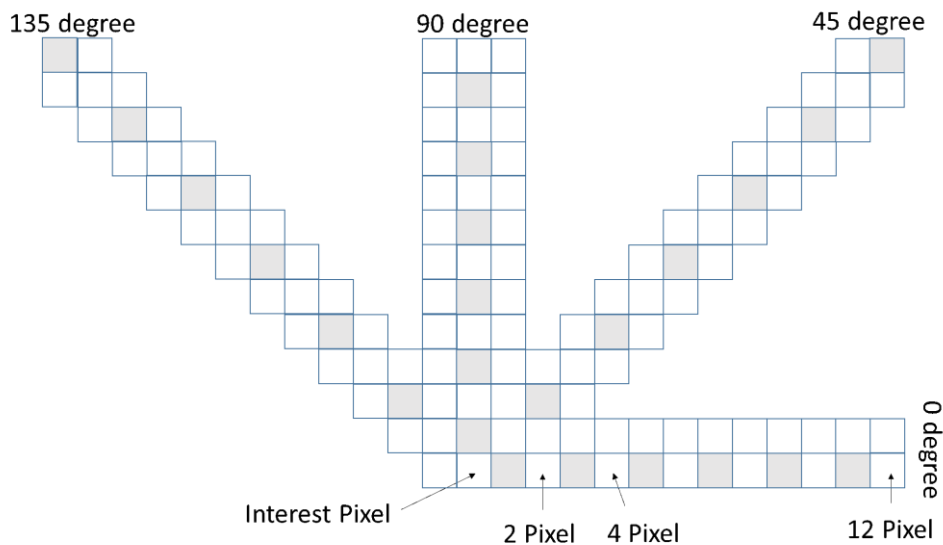

Figure 2-11 Displacement and orientation in GLCM calculation

Grey Level Co-occurrence Matrix (GLCM) used to extract texture features. For calculating the GLCM not only the displacement (distance from pixel of interest) but also the orientation of neighbor pixels must be considered. GLCM was extracted in four orientations 0, 45, 90 and 135 degree with three displacements which are 2, 4 and 12 pixels. Haralick et al (1973) suggested a set of 14 original properties for describe GLCM characteristics.

Below are described all the texture features used in this study. The meaning of each property in actual analysis case is explained in Haralick et al (1973), Gebejes et al., 2013 and Clausi et al., 2002.

Contrast

Contrast is a local grey level variation in the grey level co-occurrence matrix. It can be thought of as a linear dependency of grey levels of neighboring pixels.

$$Contrast = \sum_{i,j} |i-j|^2 p(i,j)$$

Where i and j are the horizontal and vertical cell coordinates and p is the cell value. In case of texture, the grey level variations show the variation of texture itself. High contrast values are expected for heavy textures and low for smooth, soft textures. The range of Contrast is $[0, (size\ (GLCM, 1) - 1)\char`\^2]$ where Contrast is 0 for a constant image (Gebejes et al, 2013).

Energy

Energy is a measure of local homogeneity and therefore it represents the opposite of the Entropy. Basically, this feature will tell us how uniform the texture is

$$Energy = \sum_{i,j} p(i,j)^2$$

The higher the Energy value, the bigger the homogeneity of the texture. The range of Energy is [0,1], where Energy is 1 for a constant image (Gebejes et al, 2013).

Correlation

Correlation describe the correlation between the cell pairs

$$Correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j} p(i,j)$$

Where (μ_i, μ_j) represents the mean of row i and column j; and (σ_i, σ_j) represents standard deviation of row i and column j in GLCM (Clausi et al, 2002).

Homogeneity

Homogeneity measures the uniformity of the non-zero entries in the GLCM. It weights values by the inverse of contrast weight.

$$Homongeneity = \sum_{i,j} \frac{1}{1-(i-j)^2} p(i,j)$$

The GLCM homogeneity of any texture is high if GLCM concentrates along the diagonal, meaning that there are a lot of pixels with the same or very similar grey level value. The larger the changes in grey values, the lower the GLCM homogeneity making higher the GLCM contrast. The range of homogeneity is [0,1]. If the image has little variation, then homogeneity is high and if there is no variation then homogeneity is equal to 1. Therefore, high homogeneity refers to textures that contain ideal repetitive structures, while low homogeneity refers to big variation in both, texture elements and their spatial arrangements (Gebejes et al, 2013).

Entropy

Entropy in any system represents disorder, where in the case of texture analysis is a measure of its spatial disorder.

$$Entropy = \sum_{i,j} p(i,j)\log{(p(i,j))}$$

A completely random distribution would have very high entropy because it represents chaos. Solid tone image would have an entropy value of 0. This feature can be useful to tell us if entropy is bigger for heavy textures or for the smooth textures giving us information about which type of texture can be considered statistically more chaotic (Gebejes et al, 2013).

In our method we compute the GLCM characteristics for input sample images; also, those characteristics will compute for a window with same size as sample image so that all the pixels in each query image select as the center of window. Then the calculated characteristics for each window will store in a matrix as parameters for corresponding interest pixel with same position in query image. This procedure will make the feature vector for each query image.

classification

In this study we used the simple absolute difference of each parameter in feature vector from the corresponding parameter of sample images as the classifier. A set of boundaries were selected as interest distance for each parameter by experience. Boundaries selected by a supervised procedure. Boundaries are presented in table (2-7).

Table (2-7). Selected boundaries for each parameter in feature vector.

Parameter	Boundary value
Contrast	0.05
Energy	0.003
Correlation	0.001
Homogeneity	0.001
Entropy	0.375

results for texture base classification presented in figure (2-12~14)

Figure 2-12 Sample of results for texture analysis
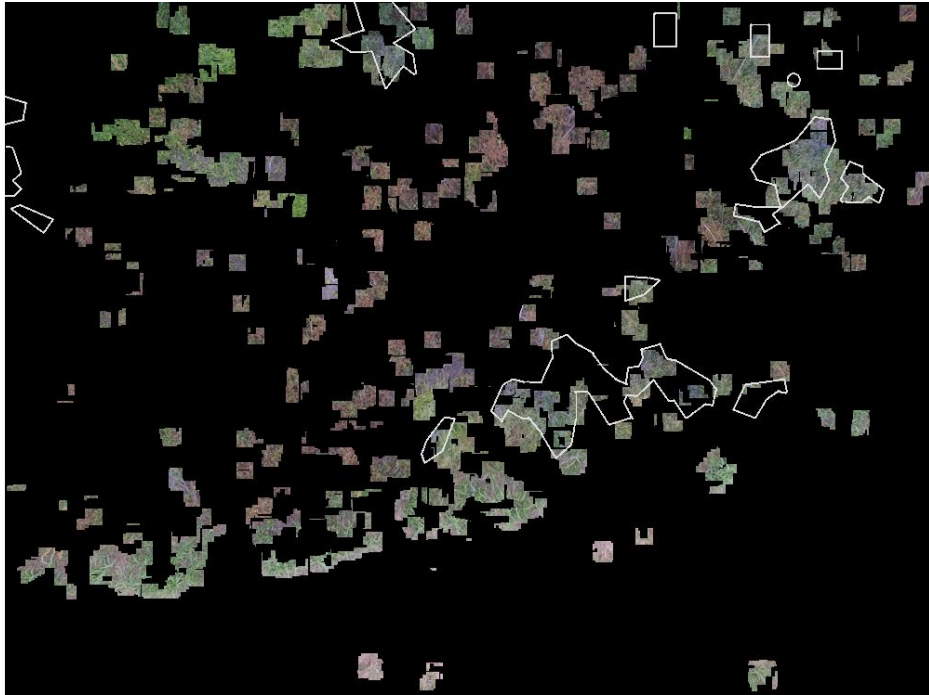
Figure 2-13 Sample of results for texture analysis
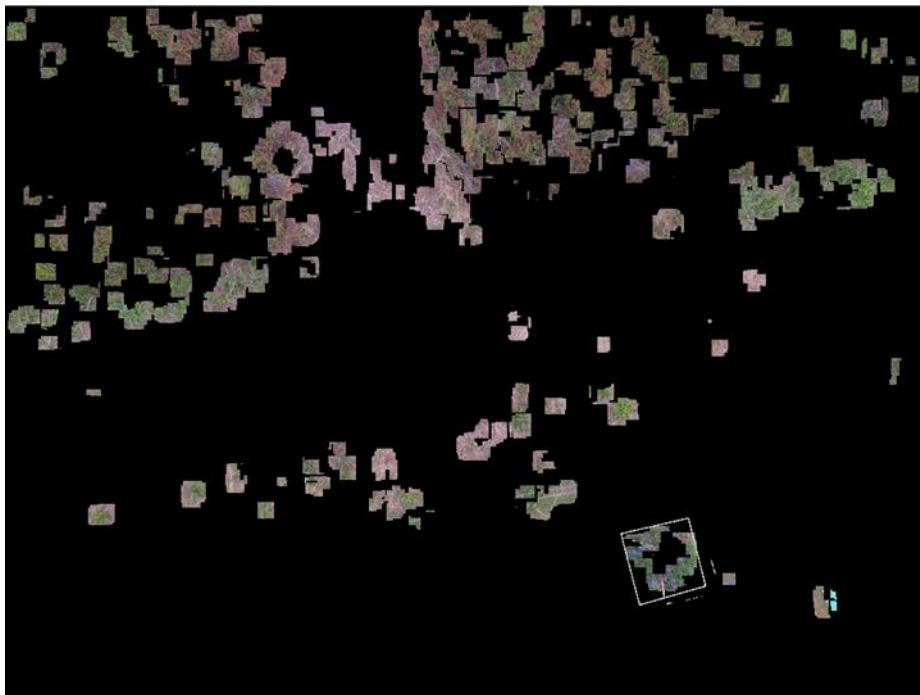
Figure 2-14 Sample of results for texture analysis

2.5 Discussion

This research was intended to explore the feasibility of the methodology that was described in the chapter 2. In the current proposed method, there are several limitations, and correspondingly, several potential improvements can be listed. First, Complexity of images in real field are very high. This complexity cause difficulty in matching images for orthomosaic them. One of the reasons for complexity is change in orientation and overlapping of the plant leaves from one image to other.

Second, color is highly light dependent which means color of same species can vary from image to other. This variation also could happen for different species. Colors of same plant are different in different life stages of plant. Using the other color space than RGB might help to solve the problem.

Third, the camera lens angel was 94°. It causes high variation in shape and color of plants such that a same plant in center of image are vary from one near the edge of image. This issue addresses the importance of scale and orientation invariant methods for feature extraction.

Four samples selected from all of the study area. This sample selected randomly and include two hairy vetch sample, one soil and one sample of other plants. Samples were used to extract the four first parameters of GLCM texture, means Contrast, correlation, energy and homogeneity. Graph (1~4) show these parameters for each image by different orientation and angels. Orientation and angels are same as what we used in our study. Graphs shows that the parameters are enough for first four orientations and for other orientations same patterns will appear. This could distribute to decrease the size of feature vector. Comparison between graphs can show s that energy have the highest variation and lowest variation belongs to homogeneity for different textures. Most similarity in distribution patterns of hairy vetch can be discriminates in can be seen in contrast and correlation. It shows that correlation and contrast have higher potentials to show the texture characteristics.
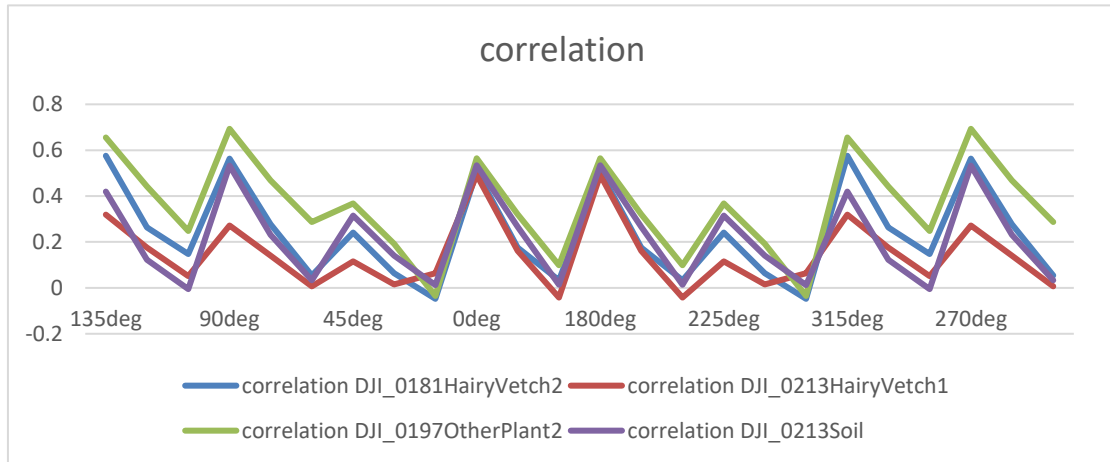
Figure 2-15 Correlation variation for different orientation and distances of Soil, Hairy vetch and other plant.
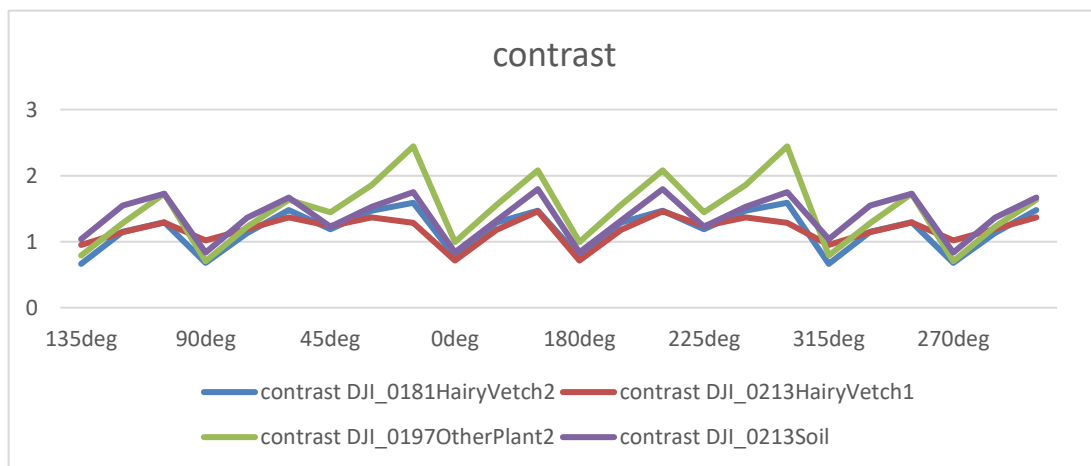
Figure 2-16 Contrast variation for different orientation and distances of Soil, Hairy vetch and other plant.
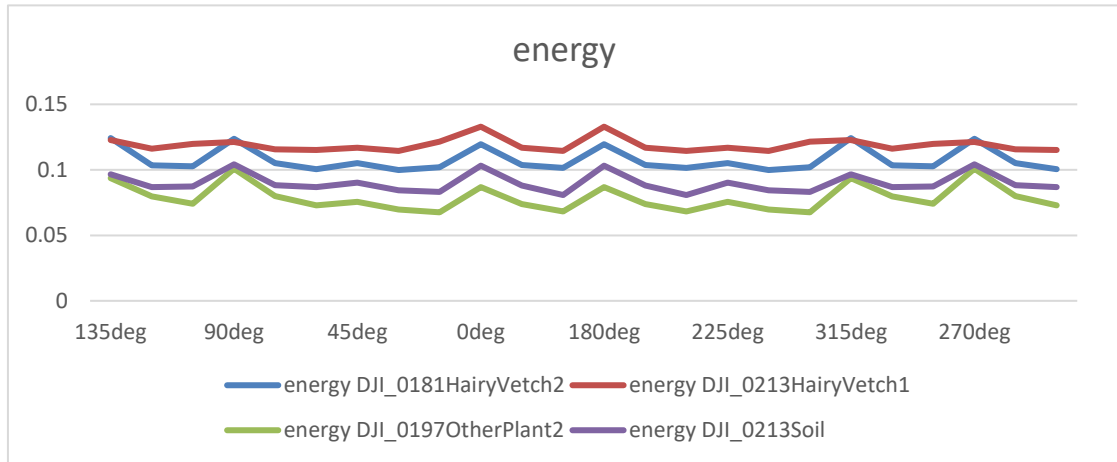
Figure 2-17 Energy variation for different orientation and distances of Soil, Hairy vetch and other plant.
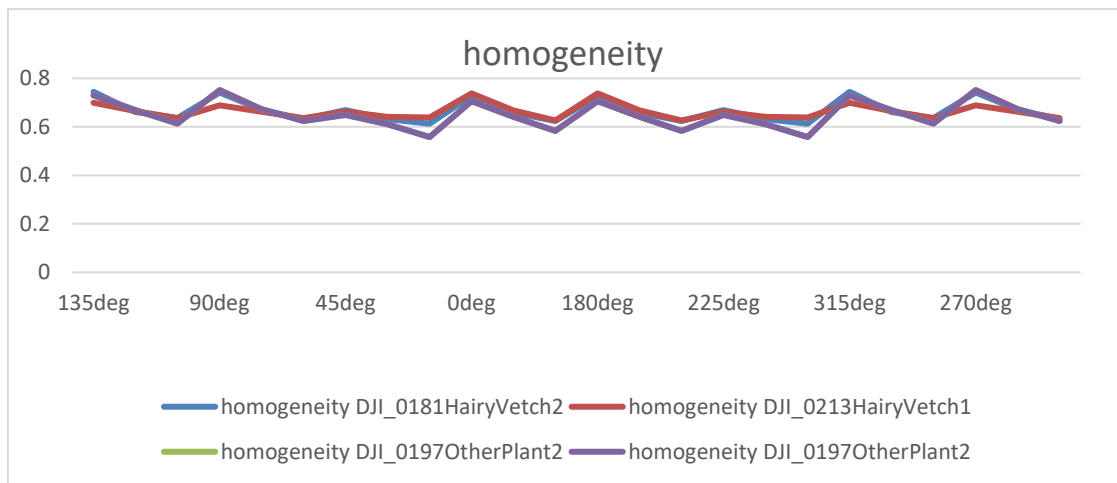
Figure 2-18 Homogeneity variation for different orientation and distances of Soil, Hairy vetch and other plant.

In addition, by comparing the results of texture filtering and color masking we can realize the potentials for combination of these filters to increase the efficiency of the method. However, miss classifications in results remind the extra efforts in improving the classifier in next steps works.

2.6 Conclusion

Results of initial classification show the feasibility of proposed method to detect the hairy vetch in UAV imageries. By comparing the results of color filtering and texture analysis we can realize

the potential of proposed method to implement an automatic algorithm to detect and map the hairy vetch (or other plants) in aerial imagery. However, the lack of ground-trust information and unproportioned hairy vetch leaf size to imagery resolution ratio prevent us from calculate the accuracy of method in this stage. In the next works we will try to emphasize on approve the accuracy of algorithm. Experiments show that classification step play a significant role in plant detection. High variation in color shows that in the next steps we should pay attention to effectiveness of color filtering or apply other color spaces to approve the algorithm.

2.7 References

Clausi, David A., and Yongping Zhao. "Rapid extraction of image texture by co-occurrence using a hybrid data structure." Computers & Geosciences28.6 (2002): 763-774.

Frame, J. (2019). Forage legumes for temperate grasslands. CRC Press.

Fujii, Yoshiharu. "Screening and future exploitation of allelopathic plants as alternative herbicides with special reference to hairy vetch." Journal of Crop Production 4.2 (2001): 257-275.

Gebejes, A., and R. Huertas. "Texture characterization based on grey-level co-occurrence matrix." Proceedings in Conference of Informatics and Management Sciences. No. 1. 2013.

Haralick, Robert M., and Karthikeyan Shanmugam. "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics6 (1973): 610-621.

Hengl, Tomislav. "Finding the right pixel size." Computers & Geosciences32.9 (2006): 1283-1298.

Hung, Calvin, Zhe Xu, and Salah Sukkarieh. "Feature learning-based approach for weed classification using high resolution aerial images from a digital camera mounted on a uav." Remote Sensing 6.12 (2014): 12037-12054.

LÓPEZ-GRANADOS, Francisca. "Weed detection for site-specific weed management: mapping and real-time approaches." Weed Research 51.1 (2011): 1-11.

Peña Barragán, José Manuel, et al. "Quantifyng Efficacy and Limits of Unmanned Aerial Vehicle (UAV) Technology for Weed Seedling Detection as Affected by Sensor Resolution." (2015).

Chapter 3: The challenges and aspects of UAV Images Mosaicking

3.1 Abstract

Advances in drone technology made them suitable tools for monitoring farm and fields. They can fly easily in anytime and anywhere. High resolution images provide by drone should be mosaic to be capable for producing maps and be used in other application like plant detection etc. an automatic image stitching method was used to mosaic the drone images obtained from Tone river shore in Gunma prefecture and Tokyo University of Agriculture and technology farm; and its advantages and weaknesses was studied and some suggestions have been provided for future works.

3.2 Introduction

Unmanned Aerial Vehicles (UAV) are becoming more and more advance and pervasive for some of their unique abilities. UAVs are cheap, they can be flown very easy and fast, and can fly almost everywhere. These characteristics makes them suitable tools for monitoring the field and farms. Provided images by UAVs have high resolution which make them suitable to produce aerial maps by combining them and making bigger mosaic image.

Image mosaicking and in this study, aerial image mosaicking, stitches a number of images into a composite image (orthomosaic image). The mosaic image must include all the images that are placed at the right position. In mosaicking the aim is to make the edges between images invisible or seamless. Thus, the quality of stitching is articulated by measuring both the correspondence between adjacent stitched images that form the orthomosaic image and the visibility of the seam between the stitched images (Levin, Anat, et al.; Rankov, Vladan, et al. 2005)

Usually, one can categorize the image alignment and stiching into two major categories— First, direct (Szeliski and Kang, 1995; Irani and Anandan, 1999) and second, feature based (Zoghlami et al., 1997). Direct methods as an advantage utilize all the available image data. Therefore, the registrations have a very high accuracy in these methods, but a close initialization required in this category. In the feature-based registration, initialization step is not required, but on the other hand , traditional feature matching methods (e.g., correlation of image patches around Harris corners

(Harris, 1992; Shi and Tomasi, 1994))   are incapable of detecting invariant features that are required for a reliable matching results.

In this study we tried to have a survey on an automatic image stitching method described in B. Matthew and G. Lowe (2003 and 2007) for UAV driven aerial image stitching. Automatic matching and blending method have been presented in method have several advantages for aerial image mosaicking, like it works with arbitrary image sequences. Thus, it doesn't have problem with side or up and down overlaps or rotations which exist in aerial images.

3.3 Proposed method

the method is including three main steps:

a.   Keypoint extraction
b.   Registration
c.   Blending

The first step in the automatic image mosaicking algorithm is to extract and match the scale invariant feature transform (SIFT) (Lowe, 2004) features between all the images. SIFT starts with calculating the difference of Gaussian function for images. The first step is to find the local maxima/minima of DoG function. SIFT features are located at scale-space maxima/minima of a DoG function. For each feature location, the scale and orientation will be calculated. This process will give a similarity invariant widow to each feature location that helps with measuring the similarities. While simply sampling intensity values in this window is similarity invariant, the invariant descriptor is computed by accumulating local gradients in orientation histograms. This allows edges to shift slightly without altering the descriptor vector, giving some robustness to affine change. Brown showed that the spatial accumulation is important for robustness against shifts because interest point locations are accurate in up to 3 pixels distance (Brown et al., 2005; Sivic and Zisserman, 2003). The algorithm gains the robustness against the light variation by utilizing the gradients and normalizing the descriptor vector which will result in eliminating bias and gain in the algorithm.

We should note that the Harris corners and ordinary correlation are note invariant to scale changes. This means that Harris corners and correlation of image patches feature matching are not able to find matches in case of rotation and changes in scale of the images. But SIFT features are invariant under rotation and scale changes, therefore, the algorithm can handle images with varying orientation and zoom.

In the case that camera rotates along its optical center the transformation that apllies to the image is a special group of homographies. We assume that each camera has a rotation vector $\theta = [\theta_1, \theta_2, \theta_3]$ and focal length f. This gives pairwise homographies $\tilde{u}_i = H_{ij}\tilde{u}_j$ where:

$$H_{ij} = K_i R_i R_j^T K_j^{-1} \qquad (1)$$

and $\tilde{u}_i$, $\tilde{u}_j$ are the homogeneous image positions ($\tilde{u}_i = s_i [u_i, 1]$, where u_i is the 2-dimensional image position).

The 4-parameter camera model is defined by:

$$K_i = \begin{bmatrix} f_i & 0 & 0 \\ 0 & f_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

and (using the exponential representation for rotations)

$$R_i = e^{[\theta_i]\times}, \qquad [\theta_i]_\times = \begin{bmatrix} 0 & -\theta_{i3} & \theta_{i2} \\ \theta_{i3} & 0 & -\theta_{i1} \\ -\theta_{i2} & \theta_{i1} & 1 \end{bmatrix} \qquad (3)$$

ideal conditions image features that are invariant under this transformations will be used. However, for small changes in image position

$$u_i = u_{i0} + \left.\frac{\partial u_i}{\partial u_j}\right|_{u_{i0}} \Delta u_j \qquad (4)$$

Or equivalently $\tilde{u}_i = A_{ij}\tilde{u}_j$, where

$$A_{ij} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

That is affine transformation and it can be calculated by linearizing the homography about u_{i0}. One can conclude that for each patch in images an affine transformation applies and it shows that utilizing the SIFT features are partially invariant under the affine transformations.

After extracting the features for all n aerial images, they must be matched. Because multiple images can be overlapped in a single point on ground, each feature vector should be matched to its k nearest neighbors in feature space (in this study k = 4). This can be done in O(n log n) time by using a k-d tree to find approximate nearest neighbors (Beis and Lowe, 1997). K-D tree is aligned binary space partitioning method that recursively partitions the feature space by using the mean in that dimension of the feature space which has the highest variance.

in this study, RANSAC was utilized to select a set of inliers that are compatible with a homography between the images. In next step a probabilistic model was used to verify the accuracy of matching features. When a geometrically consistent set of match features are obtained the bundle adjustment (Triggs et al.,1999) was applied to solve all the camera parameters once together. The following algorithm shows the overall steps for produce orthomosaic image:

Algorithm: Automatic Stitching

Input: n unordered images

I. Extract SIFT features from all n images

II. Find k nearest-neighbors for each feature using a k-d tree

III. For each image:

Select m candidate matching images that have the most feature matches to this image.

Find geometrically consistent feature matches using RANSAC to solve for the homography between pairs of images.

Verify image matches using a probabilistic model

IV. Find connected components of image matches

V. For each connected component:

Perform bundle adjustment to solve for the rotation $\theta_1, \theta_2, \theta_3$ and focal length f of all cameras

Render panorama using multi-band blending

Output: Panoramic image(s)

This algorithm concluded in license free application AutoStitch. This application is accessible from link below:

http://matthewalunbrown.com/autostitch/autostitch.html

A multi-copter was used to gather all the imagery in this study. The UAV used in this study is DJI Phantom 3 advanced model. Pix4d capture App was used for control and doing imagery in an automatic process. UAV camera was DJI camera with Sony EXMOR 1/2.3" sensor. Image resolution is 3000*4000 pixel. Camera technical characteristics are presented in table 3-1.

Table 3-1. Camera technical characteristics.

Manufacturer	DJI
Sensor	Sony EXMOR 1/2.3" Effective pixels: 12.4 M (total pixels: 12.76 M)
Lens	FOV 94° 20 mm (35 mm format equivalent) f/2.8, focus at ∞
ISO Range	100-1600 (photo)
Shutter Speed	8s -1/8000s
Image Max Size	4000×3000

Tree different sets of aerial images have been prepared. For the first set (set 1) the altitude of imagery flight is 25 meters. A 30cm by 30cm plat was inserted in imagery site 1 as ground trust scale. Set 1 is include 45 images. Based on camera characteristics and the ground trust scale the spatial resolution of aerial images is 1 cm. images have been taken in Tokyo University of Agriculture and technology farm on May 25th, 2015. Sample images for the site 1 are show in figure 3-1.
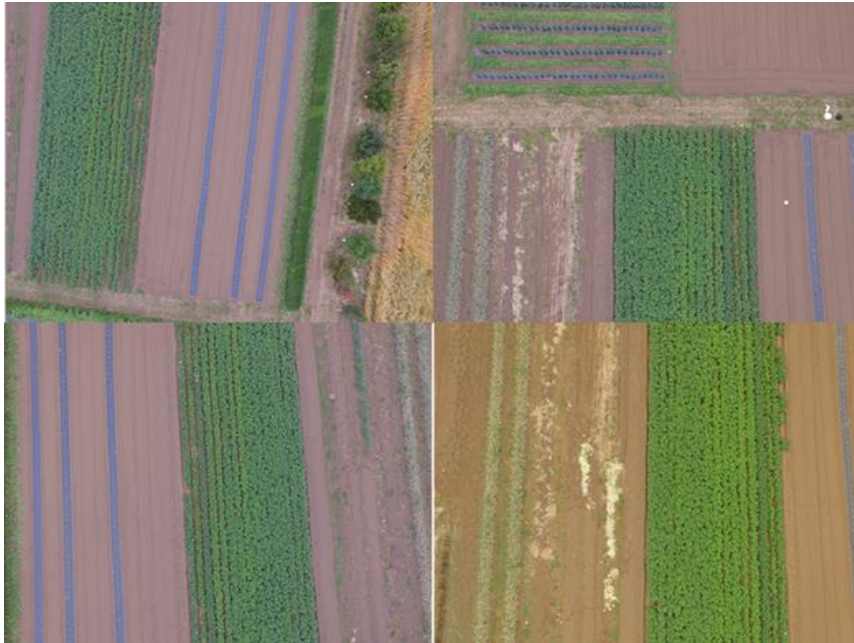
Figure 3-1 Sample images of set 1

The second set (set 2) is include 93 images. The imagery altitude is 10 meters and imagery haven done on June 4th, 2016 in Itakura city, Gunma, Japan. Sample images are shown in figure 3-2.

Figure 3-2 Sample images of set 2

Set 3 is include 63 images and same characteristic with set 2. Sample images for set 2 are shown in figure 3-3.

Figure 3-3 Sample images of set 3

Figure 3-4 Study Areas 2 and 3 and their difference in height of plants

3.4 Result and discussion

This research was intended to explore the feasibility of the methodology described in this report to make mosaic images from UAV's aerial images. In the current system, there are several limitations, and correspondingly, several potential advantages that can be listed. First, Complexity

of images in real field are very high. This complexity cause difficulty in matching images for orthomosaic them. One of the reasons for complexity is similarity of target objects. In our work the goal is to stitch aerial images of plants. Typically, plants' prospect is very similar between one species. Other reason for high complexity can be change in orientation and overlapping of the plant leaves from one image to other which can affect the DoG in SIFT.

Results of images stitching for image sets 1 and 3 are shown in figures 3-5 and 3-6. Presented method was successful in stitching aerial images for these two study areas. However, as shown in figure 3-7 image stitching for image set 2 is defective. This defectiveness caused by inability of SIFT in extracting accurate feature points to match and calculating the homography of images. Lack of accuracy of feature points can be caused by several factors.
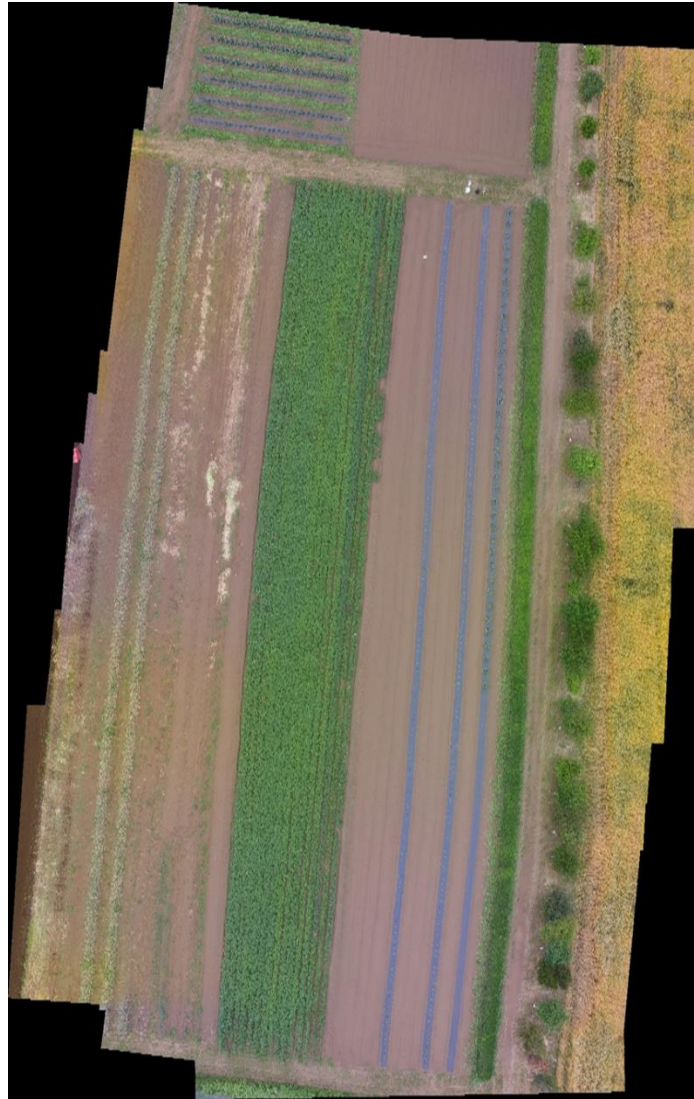
Figure 3-5 Image stitching results for set 1
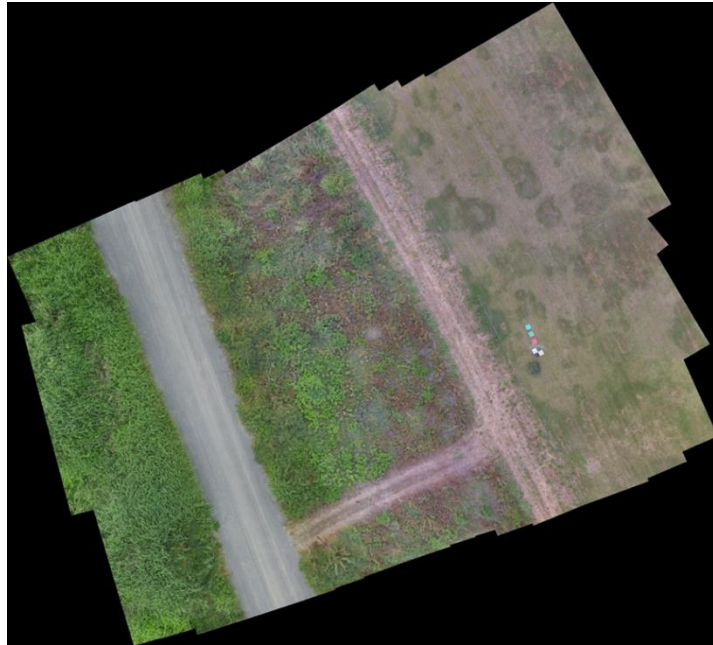
Figure 3-6 Image stitching results for set 3
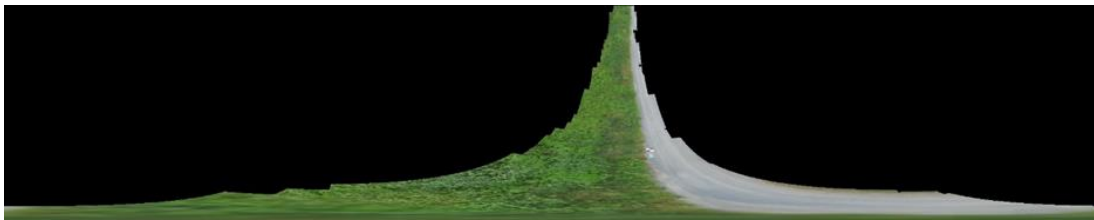
Figure 3-7 Image stitching results for set 2

First, it could be caused by height of plants in study area set 2. Since study area of set 2 and 3 were close to each other and the flight altitude is same for both areas, one of differences between two areas is the kind of plants with variation in height. This difference can be seen in figure 3-4 which plants in set are taller. It can also result in instability of plant stem and leave in different images. Also objects of set 3 are include road or dead plants but in set 2 number of significant objects are few.

Another reason that can be consider for SIFT problem is the difference in adjacent images caused by camera lens distortion. Lens distortion is capable to change the images homography. Therefore, camera lens parameters were calculated in camera calibration experiments. Results of calibration showed that UAV camera has slight distortion with radial distortion coefficient vector:

$$[k_1, k_2] = [-0.04224, 0.02602] \tag{6}$$

The distorted point in images $(x_{distorted}, y_{distorted})$ can be denote as:

$$x_{distorted} = x(1 + k_1 * r^2 + k_2 * r^4)$$
$$y_{distorted} = y(1 + k_1 * r^2 + k_2 * r^4) \tag{7}$$

Which:

$$(x, y) = undistorted\ pixel\ location$$

$$r^2 = x^2 + y^2 \tag{8}$$

Results of images undistorting show in figure 3-8. Since the amount of distortion is small, it cannot be considered as a reason for SIFT problem.

Figure 3-8 Original image (left) and undistorted image (right).

The camera which used for obtaining the images equipped with 94° diagonal angle lens. Considering the flight altitude $h = 10\ meters$, the shape and characteristics of objects in images will change from image to image so that in one image top side of plants appear in the middle of image and in other adjacent images side of the plants will appear in images (figure 3-9) which can affect the accuracy of SIFT in extracting the feature points. This effect shown in results of non-similarity registration and simple blending of images figure 3-10.
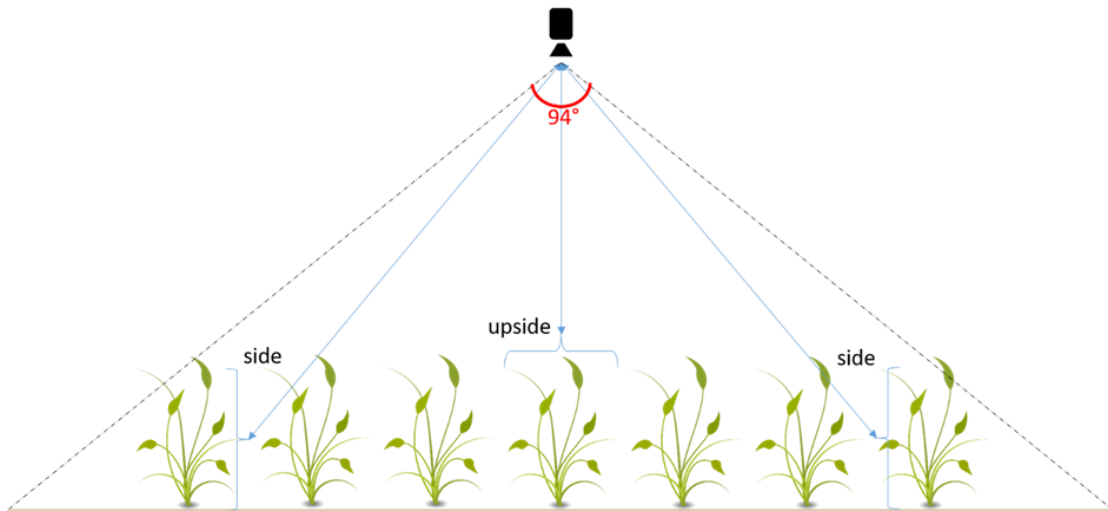
Figure 3-9 Camera angle and plants in different positions
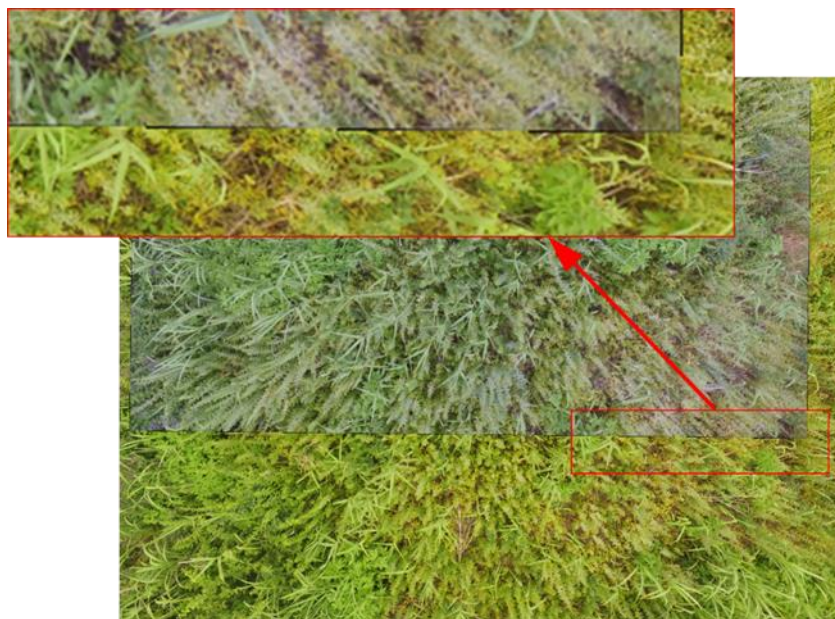
Figure 3-10 Result for blending two adjacent images without adjustments

3.5 Conclusion

The results of experiments address important points that should be focused on in future efforts. The applied method for mosaicking the aerial images initially developed for stitching images and making panorama photos. This initial aim changes the developers to consider different kind of problems. This issue occurs when movement of camera supposed to be around one axis and the camera would not twist horizontally. The effect of this assumption can be the reason for distorted mosaic image in figure 3-5. This problem should be considered in future works. Also, inability of

SIFT in extracting accurate feature points addresses the future efforts to increase the number of feature point by installing ground tags in imagery area or increase the flight altitude and applying different lenses.

3.6 References

Beis, Jeffrey S., and David G. Lowe. "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces." Computer Vision and Pattern Recognition, 1997. Proceedings, 1997 IEEE Computer Society Conference on. IEEE, 1997.

Brown, Matthew, and David G. Lowe. "Automatic panoramic image stitching using invariant features." International journal of computer vision 74.1 (2007): 59-73.

Brown, Matthew, and David G. Lowe. "Recognising panoramas." ICCV. Vol. 3. 2003.

Brown, Matthew, Richard Szeliski, and Simon Winder. "Multi-image matching using multi-scale oriented patches." 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Vol. 1. IEEE, 2005.

Harris, Chris. "Geometry from visual motion." Active vision. MIT Press, 1993.

Irani, Michal, and P. Anandan. "About direct methods." International Workshop on Vision Algorithms. Springer Berlin Heidelberg, 1999.

Levin, Anat, et al. "Seamless image stitching in the gradient domain." European Conference on Computer Vision. Springer Berlin Heidelberg, 2004.

Lowe, David G. "Distinctive image features from scale-invariant keypoints." International journal of computer vision 60.2 (2004): 91-110.

Rankov, Vladan, et al. "An algorithm for image stitching and blending." Biomedical Optics 2005. International Society for Optics and Photonics, 2005.

Shi, Jianbo, and Carlo Tomasi. "Good features to track." Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on. IEEE, 1994.

Sivic, Josef, and Andrew Zisserman. "Video Google: A text retrieval approach to object matching in videos." Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. IEEE, 2003.

Szeliski, Richard, and Sing Bing Kang. "Direct methods for visual scene reconstruction." Representation of Visual Scenes, 1995. (In Conjuction with ICCV'95), Proceedings IEEE Workshop on. IEEE, 1995.

Triggs, Bill, et al. "Bundle adjustment—a modern synthesis." International workshop on vision algorithms. Springer Berlin Heidelberg, 1999.

Zoghlami, Imad, Olivier Faugeras, and Rachid Deriche. "Using geometric corners to build a 2D mosaic from a set of images." Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on. IEEE, 1997.

Chapter 4: Application of UAVs in Agro-environmental Studies: Discrimination of Natural Vegetation in High Complex Aerial Images

4.1 Abstract

In agro-environmental studies, it is necessary to have access to a vast study area in high temporal resolution. Detect and discrimination of vegetation is an important part of the efforts to produce land cover map of remote study areas. However, having access to all the area is not practical in most cases by humans. Land cover maps of natural and agriculture land uses are essential to monitor and understand the land use changes, invasive plants discrimination, biomass estimation and herbicide/pesticide application. In this study, we utilized a machine learning algorithm based on spectral and spatial features to analyze the high complex natural vegetation in aerial images driven from a commercial drone. The algorithm was trained using an implemented training set of 7 different land covers classes. This study discusses the ability of different spatial and spectral feature extractors and classifiers in detecting the high complex and low-resolution natural vegetation patterns. trained models yielded on average 80% accuracy in 5-fold cross validation. The model was used to predict the land cover types of natural vegetation at Tone river shore, Gunma prefecture, Japan. The results showed the potential ability of the algorithm in turning a low spectral resolution commercial drone into a powerful tool for agro-environmental studies and land cover monitoring.

4.2 Introduction

In general, textures are complex visual patterns that have characteristics like brightness, color, slope, size, etc. Thus, texture can be regarded as a similarity grouping in an image (Kak et al. 1982). There are several approaches for texture analysis and texture feature extraction like structural, statistical, model-based transform, etc. some of the most popular statistical texture features are derive from Grey Level Co-occurrence Matrices (GLCM) (Materka et al. 1998). Haralick introduced 14 textural features derived from GLCMs (Haralick et al. 1973) that was explained in previous chapter in details. Segmentation-based texture analysis (SFTA) describes the texture of an image by decomposing the input image into a set of binary images from which the fractal dimensions of the resulting regions are computed in order to describe segmented texture patterns (Costa et al. 2012). the SFTA extraction algorithm has two parts:

A. Two-Threshold Binary Decomposition

The Two-Threshold Binary Decomposition (TTBD) takes as input a grayscale image $I(x, y)$ and returns a set of binary images by applying Otsu algorithm (Liao et al 2001).

B. SFTA extraction algorithm

After applying the TTBD to the input gray level image, the SFTA feature vector is constructed as the resulting binary images' size, mean gray level and boundaries' fractal dimension. The fractal measurements are employed to describe the boundary complexity of objects and structures in the input image. Figure 4-1 shows the structure of the SFTA feature extraction algorithm. The GLCM and SFTA were used in this study to describe the textural properties of images.

Spectral property of image is the basic feature. Color explains the spectral feature of the objects in RGB sensors. The basic spectral features are the statistical features such as mean and standard deviation of each R, G, or B channel. Each image also can be represented in HSI color space. In
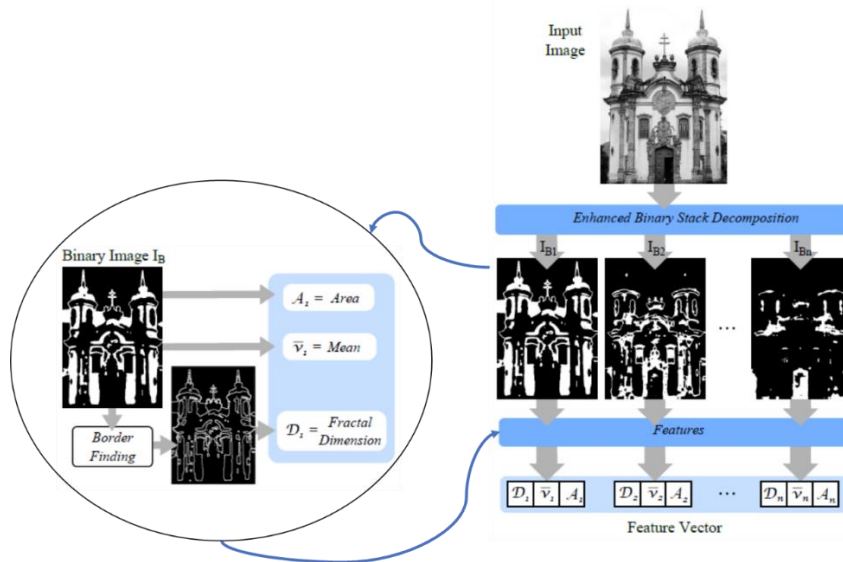
Figure 4-1 The structure of SFTA algorithm

HSI color space each pixel has three property: Hue, Saturation, and Intensity of each pixel. The mean and standard deviation on each channel in HSI color space also can represent a basic feature to the describe the spectral feature.

The other types of spectral features for RGB images are the different indices that can be obtain by band-wise calculations. The Synthetic Normalized Difference Vegetation Index (or Synthetic NDVI or Green Red Vegetation Index (GRVI)) is a measure that try to predict the NDVI values only using red and green channels. NDVI is the most common vegetation index that is used in remote sensing applications. We will describe and use the NDVI in this Thesis latter on and we will explain it in detail. The Synthetic NDVI is calculate as:

$$Synthetic\ NDVI = \frac{(Green - Red)}{(Green + Red)}$$

Synthetic NDVI values are between -1 and 1, And the greater the value, the healthier the plant is. It tries to predict the chlorophyll contents of plant leaves. Each plant species has a unique chlorophyll content in its leaves. This can help to describe the spectral features of the plant (Motohka et al., 2010).

The next vegetation index that can be calculated for RGB sensors is The Excess Green Index (ExGI). The ExGI proved to outperform the other indices that can be calculated by RGB channels (Sonnentag et al., 2012, Larrinaga et al. 2019). It is defined as:

$$\text{Excess Green Vegetation Index} = 2 \times \text{Green} - (\text{Red} + \text{Blue})$$

It contrasts the green against the other visible channels to distinguish between soil and or residue background and green plants. It is also able to minimize the effect of changes in illumination. Since, it can recognize how green an object in the image is, it is a suitable feature descriptor for plant spectral properties.

Machine learning (ML) techniques have become one of the most powerful tools for different range of applications. ML is consisting of supervised, semi-supervised, and unsupervised methods. Also, ML is a branch of artificial Intelligence (AI). In each problem, the ML offers a different solution. Support vector machine (SVM) and bagged decision trees (BT) are two of the best-known ML supervised classifiers. Unlike the neural network models that have superiority in accuracy, SVM and BT models can be trained on small datasets while still have high accuracy. Besides, the SVM and BT models are fast in training in compare to neural networks which make them suitable for classification tasks that require customizing and modifications frequently.

A support vector machine (SVM) is a supervised learning algorithm which learns a given independent and identically distributed set of training instances $\{(x_1, y_1), \ldots, (x_N, y_N)\}$, where $y \in \{-1, 1\}$ are binary classes to which data points belong (Wittek, 2014).

A hyperplane in $\mathbb{R}^d$ has the generic form

$$w^T x - b = 0. \tag{4.1}$$

where w is the normal vector to the hyperplane, and the bias parameter b helps determine the offset of the hyperplane from the origin.

We assume that the data instances are linearly separable—that is, there exists a hyperplane that completely separates the data instances belonging to the two classes. In this case, we look for two

hyperplanes such that there are no points in between and we maximize their distance. The area between the hyperplanes is the margin.

In its simplest, linear form, a support vector machine is a hyperplane that separates a set of positive examples from a set of negative examples with maximum margin. The distance between the two planes is $\left\|\frac{2}{w}\right\|$ hence, minimizing w will lead to a maximal margin, which in turn leads to good generalization performance. The formula for the output of a linear support vector machine is

$$\hat{y}_i = sign\ (w^T x_i + \ b) \tag{4.2}$$

Where the x_i is the $i\ th$ training example. With this, the conditions for data instances for not falling into the margin are as follows:

$$
\begin{aligned}
w^T \mathrm{x} - b &\geq 1 \quad for\ y_i = 1, \\
w^T \mathrm{x} - b &\leq -1\ \ for\ y_i = -1.
\end{aligned}
\tag{4.3}
$$

These conditions can be written briefly as

$$y_i\bigl(w^T \mathrm{x} - b\bigr) \geq 1, \qquad i = 1, \dots, N \tag{4.4}$$

The optimization is subject to these constraints, and it seeks the optimal decision hyperplane with

$$arg\ \min_{w,b}\frac{1}{2}\,\|w\|^2. \tag{4.5}$$

The margin is also equal to the distance of the decision hyperplane to the nearest of the positive and negative examples. Support vectors are the training data that lie on the margin (Figure 4-2).
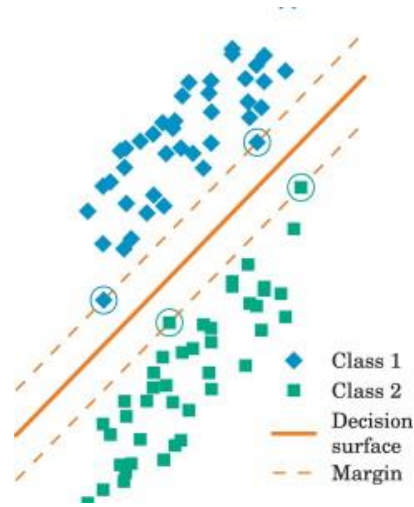
Figure 4-2 Support vectors are the training data that lie on the margin

3. Bagging decision tree classifiers

In this section, we review bootstrap aggregated tree classifiers as well as estimators of the misclassification error.

3.1. Bagging

Let $L = \{(y_1, x_1), \dots, (y_N, x_N)\}$ denote a learning sample of N independent observations consisting of p-dimensional vectors of predictors $x_i = (x_{i1}, \dots, x_{ip}) \in R^p$ and class labels $y_i \in \{1,2\}$. We assume that the observations in the learning sample are independent identical distributed (iid) random variables from some distribution function F_L

$$(y_1, x_1), \dots, (y_N, x_N) \sim^{iid} F_L.$$

A classifier $C(\tilde{x}, L)$ predicts future $\tilde{y}$-values for a vector of predictors $\tilde{x}$ based on a learning sample L. In the following we denote future observations by $(\tilde{y}, \tilde{x})$ distributed according to $F_{\tilde{y}, \tilde{x}}$.

Classifiers C can be stabilized by averaging over multiple learning samples (Breiman, 1996). An aggregated classifier C_A for a new observation $\tilde{x}$ is given by

$$C_A(\tilde{x}) = E_{F_L} C(\tilde{x}, L),$$

where the expectation is over learning samples L distributed according to F_L. Bagging as introduced by (Breiman, 1996, Breiman, 1996) estimates the aggregated rule $C_A(\tilde{x})$ using the bootstrap by

$$\hat{C}_A(\tilde{x}) = E_{\hat{F}_L} C(\tilde{x}, L^*),$$

where L^* is a random sample from the empirical distribution function $\hat{F}_L$

$$(y_1{}^*,x_1{}^*),\ldots,(y_N{}^*,x_N{}^*) \sim^{iid} \hat{F}_L.$$

The bagged classifier $\hat{C}_A{}^B$ based on B bootstrap samples is computed as follows.

1. Draw B random samples $L^{*\,(1)},\ldots,L^{*\,(B)}$ of size N with replacement from L.
2. Construct the classifier C using the bootstrap sample $L^{*\,(b)}$.
3. Iterate step 2 for all $b = 1,\ldots,B$ bootstrap samples.
4. A new observation $\tilde{x}$ is classified by majority voting:

$$\hat{C}_A{}^B(\tilde{x}) = \underset{0 \le j \in \{1,2\}}{\operatorname{argmax}} \sum_{b=1}^{B} \chi_{\{j\}} \left(C(\tilde{x},L^{*\,(b)}) \right)$$

Where the χ is the indicator function

$$\chi_Z(x) = \begin{cases} 1 & x \in Z \\ 0 & else \end{cases}.$$

4.3 Dataset

We collected the data from the study site that has been described in previous chapter. The aerial imagery was repeated for the same site in 2016 and 2017. The UAV that was used in previous research was utilized in second year. The area was checked by agriculture experts and then the obtained imagery was classified manually in order to provide an estimation of the location for two target plants. The nature of study site did not allow the experts to go into the site and have complete survey. Therefore, the grand truth data for next step were collected from the aerial images later at laboratory. AgiSoft Photoscan software (Agisoft LLC) was used to prepare orthomosaic images with up to 6.5 mm spatial resolution. The result of orthomosaic and manual classification for both years are shown in figure 4-3.
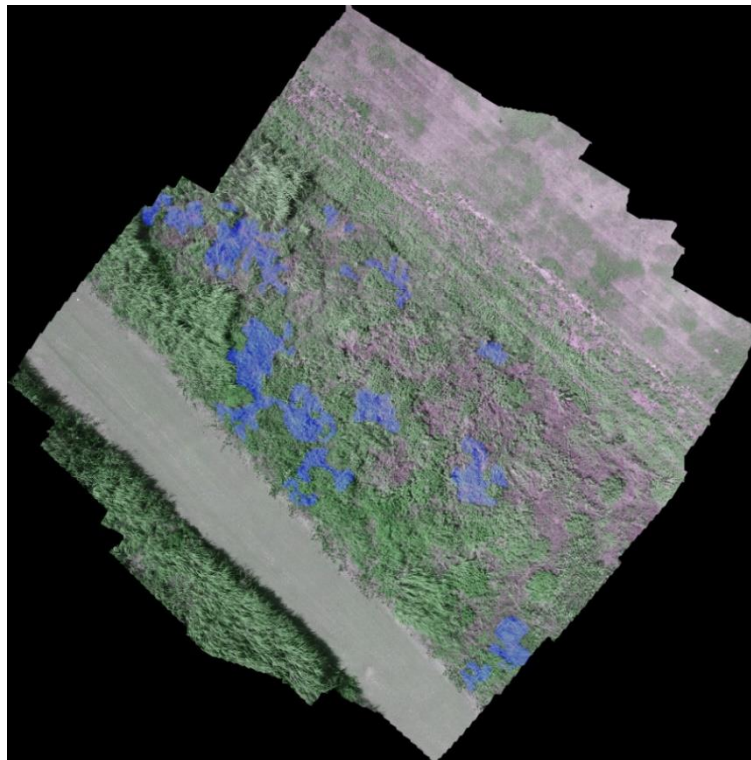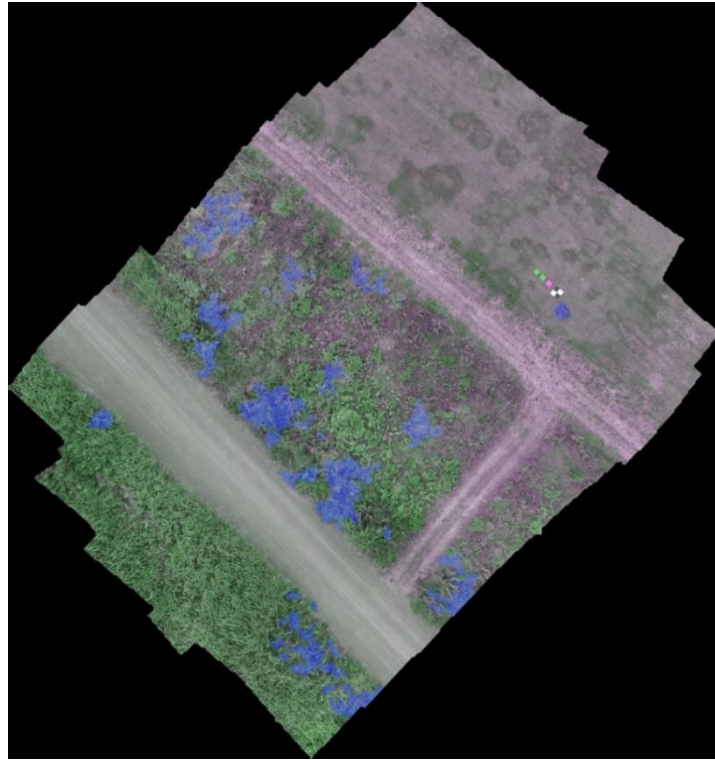
Figure 4-3 result of orthomosaicing and manual classification for hairy vetch 2016 (up) and 2017 (down)

In the next step, a data set of 374 sample images were prepared for training the ML models. The sample images were selected manually by our colleagues in agriculture department. The image

size for each sample image is 90 by 90 pixel. This image size was selected by considering the image resolution and the size of the leaves and flowers for each plant. The dataset consists of 7 category of land surface including common reed and hairy vetch samples. Figure 4-4 shows the image samples for each category in dataset.
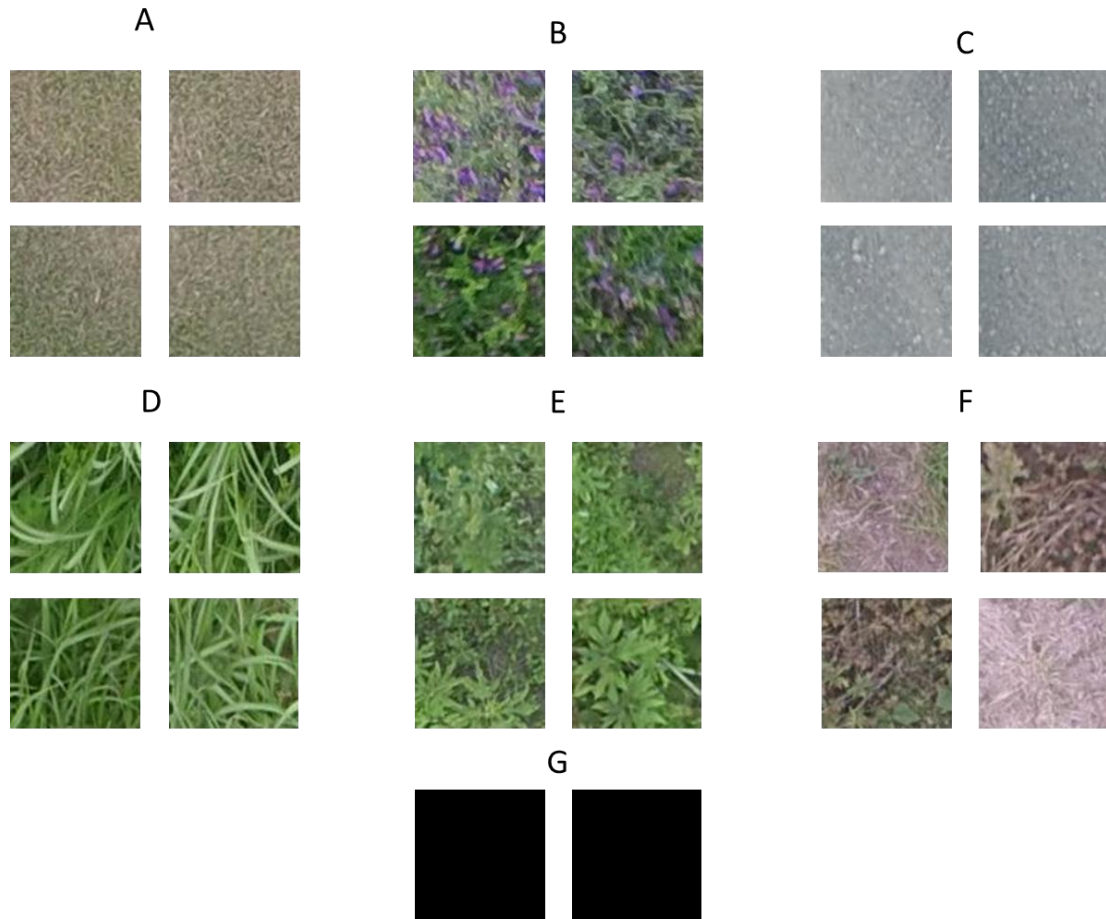
Figure 4-4 The sample images in each category in implemented dataset. A) Natural grass. B) Hairy vetch. C) Gravel read. D) common reed. E) Other types of plants. F) Dead plants or bare soil. G) black image.

4.4 Methodology

A new method based on color and texture features was proposed. Our methodology can be decomposed in following steps:

3.1. Pre-processing and feature extraction

Figure 4-5. shows the overall methodology's workflow. First the aerial images were preprocessed to obtain the orthomosaic image and the sample data in dataset as it was explained in the previous

section. After the preprocessing the data, we extract all the features of each sample image in the dataset and obtain the feature vectors for all the samples in the dataset. For the feature extraction, first we extracted the basic color features (mean and standard deviation of each band in RGB and HSI) and the ExGI. The GLCM second order statistics and the SFTA features were extracted as texture features of each sample image in the dataset.

In this study the GLCM was extracted for 2, 4, and 12 pixels displacement with 0, 45, 90, 135 degrees rotation with the following offset parameter:

```
Offsets = [0 2; 0 4; 0 12; ...
           -2 2; -4 4; -12 12; ...
           -2 0; -4 0; -12 0; ...
           -2 -2; -4 -4; -12 -12];
```

The distances were selected based on the size and shape the plants leaves and flowers.
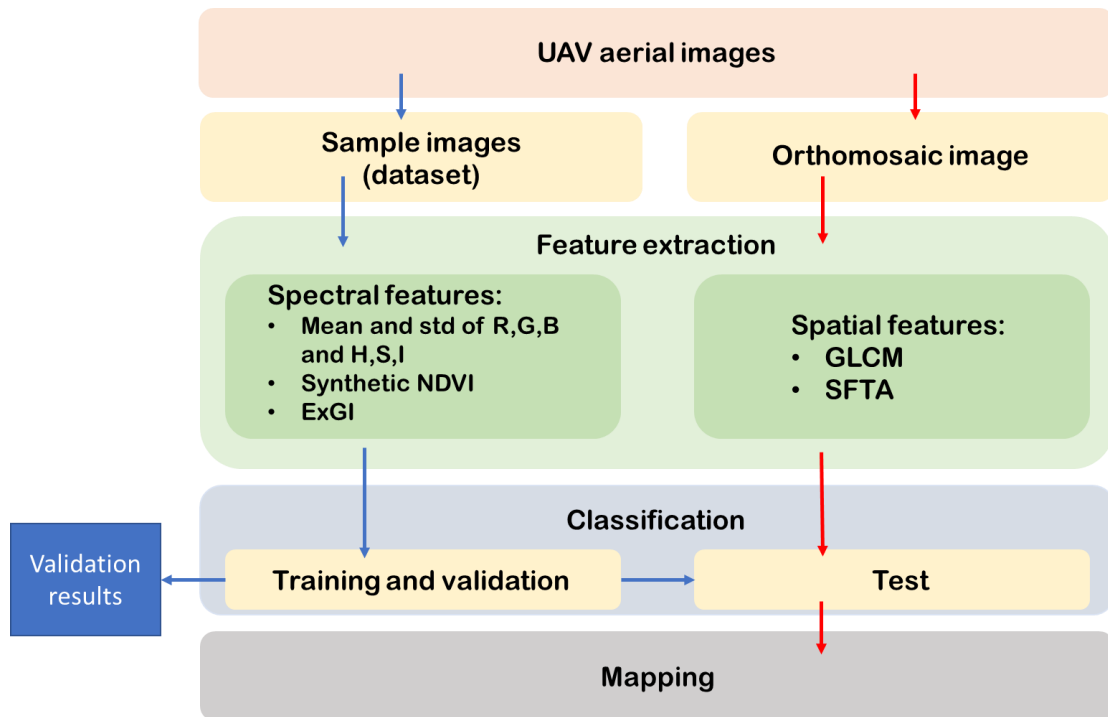
Figure 4-5 Methodology's workflow.

3.2. Training the models

The next step in methodology is to train ML models for classification. We trained SVM with the cubic kernel and bagged trees classifiers. We trained 2 models for each set of features (in total 6

models). Each classifier was trained on each of feature sets separately (e.g. one for basic color features + ExGI, one for GLCM, one for SFTA). The models were validated in a 5-fold cross validation process. The results of validation and receiver operator curves for each model were calculated.

3.3. Classification

Provided orthomosaic images were passed into a feature extractor as input images to detect local color and texture features based on tile-based image objects (Alhichri et al., 2018). Regarding to the hairy vetch and Phragmites leaves' size and images resolution, 90 by 90 pixels tiles were used. The algorithm processes all orthomosaic image with a sliding window and extract the features for each processed block. Then processed block's features appended to a data table and form the feature vectors for classification step. The feature vector was passed in each trained model for prediction. The result of prediction step then returned to the block processing step and all the pixels in the correspondent block of orthomosaic image was marked the same as the predicted class.

4.5 Results and discussion

4.1 validation

This study was intended to explore the feasibility of the described methodology. In the current system, there are several limitations, and correspondingly, several potential improvements can be listed. Results showed high accuracy for SVM model in classifying color properties of images. This trained model was used to classified orthomosaic images of two years from the same area. Figure 4-9 and 4-10 show the results of SVM and BT classification of hairy vetch base on color and textural properties in 2016.

The study shows (table 4-1.) that the color has the highest overall (OA) accuracy among other feature sets. However, the SFTA feature set has the lowest performance accuracy. This trend can be observed for in class accuracy of the models too. The trained model on SFTA features set has moderate accuracy in discriminating hairy vetch although its performance on common reed class is acceptable. This could reflect the fact that the hairy vetch has significantly smaller leaves in compare with common reed and the spatial resolution of the orthomosaic images.

Table 4-1. The overall accuracy for trained models

Feature	SVM OA accuracy (%)	BT OA accuracy (%)
Color feature set Mean and std of RGB&HSI +ExGI	97.7	91.4
SFTA feature set nt = 8 → 48 features	69.3	68.4
GLCM feature set 3 distance* 4 angles * 4 statistics + entropy →49 features	84.0	78.1

Fan et al. 2018 applied the SVM, random forest (a type of bagged trees), and neural network (CNN) for detecting the Tobacco plant in the farm. The study report that SVM and random forest achieved 93% accuracy. By comparing the results of two studies confirms the ability of our method in detecting the hairy vetch and common reed in an environment with significantly higher complexity.
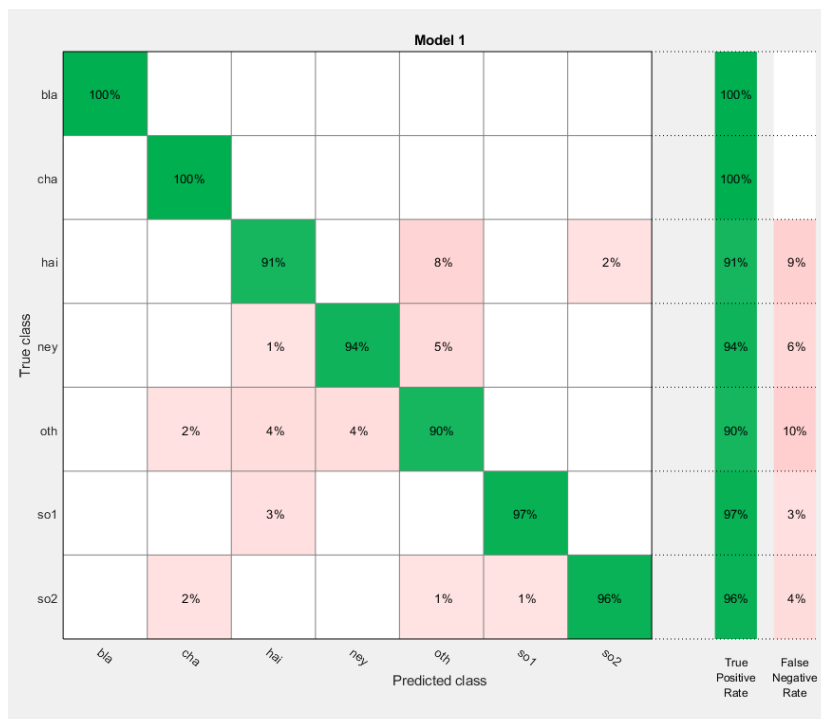
Figure 4-6 The confusion matrix for training models on color features. SVM model 94% (up) and ensembled trees 91.7% (down).

Figure 4-6 and 4-8 show the confusion matrix of trained models. Comparing the in-class accuracies shows that the hairy vetch has the highest confusion rate with the untagged plants class and the soil and dead plants class. The untagged plants class includes the arbitrary plant species that were exist in study sites during the data collections. These plants have very high variation in their appearance and color. Confusion matrix of models also show that the same confusion happens for all the classes and the dead plants and soil class. This situation appears less for the color feature. Therefore, we can conclude that the texture of so2 class is local similar with other classes and the color is a critical feature for eliminate the confusions.

Müllerová et al. 2017 applied similar set of texture features and ML models to detect the two type of plants in natural environment. The study achieved moderate accuracy in majority of their experiments. The comparison between our method and Müllerová work show that our methodology has higher performance despite the larger number of classes.
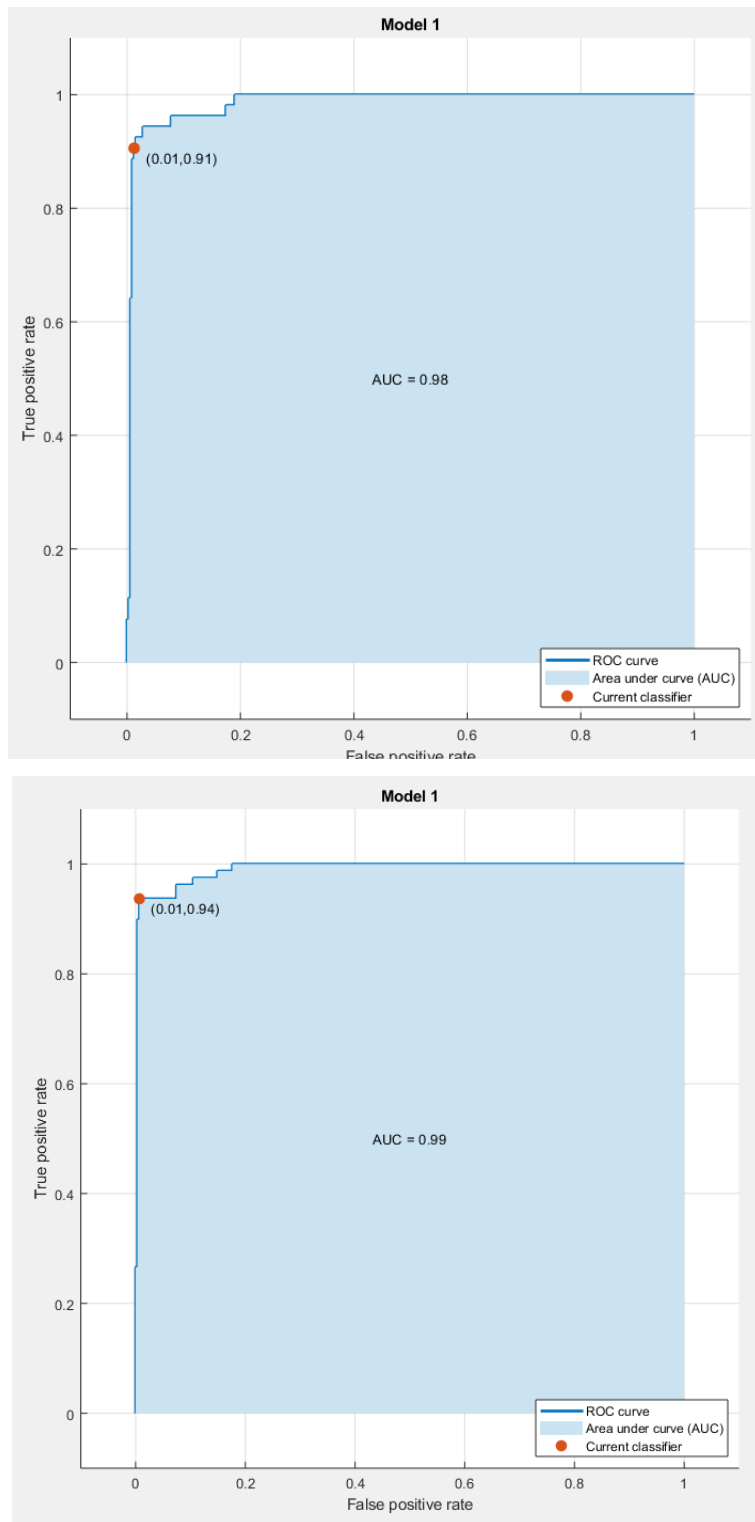
Figure 4-7 The receiver operator curve and area under the curve for SVM models trained on color features

Figure 4-7 indicates the receiver operator curve (ROC) for SVM model and color features. The ROC indicates the reliability of model in selecting the true positive samples (TP). The area under

curve measure indicates the current model's ability in separate the TP over all the accepted error rates.
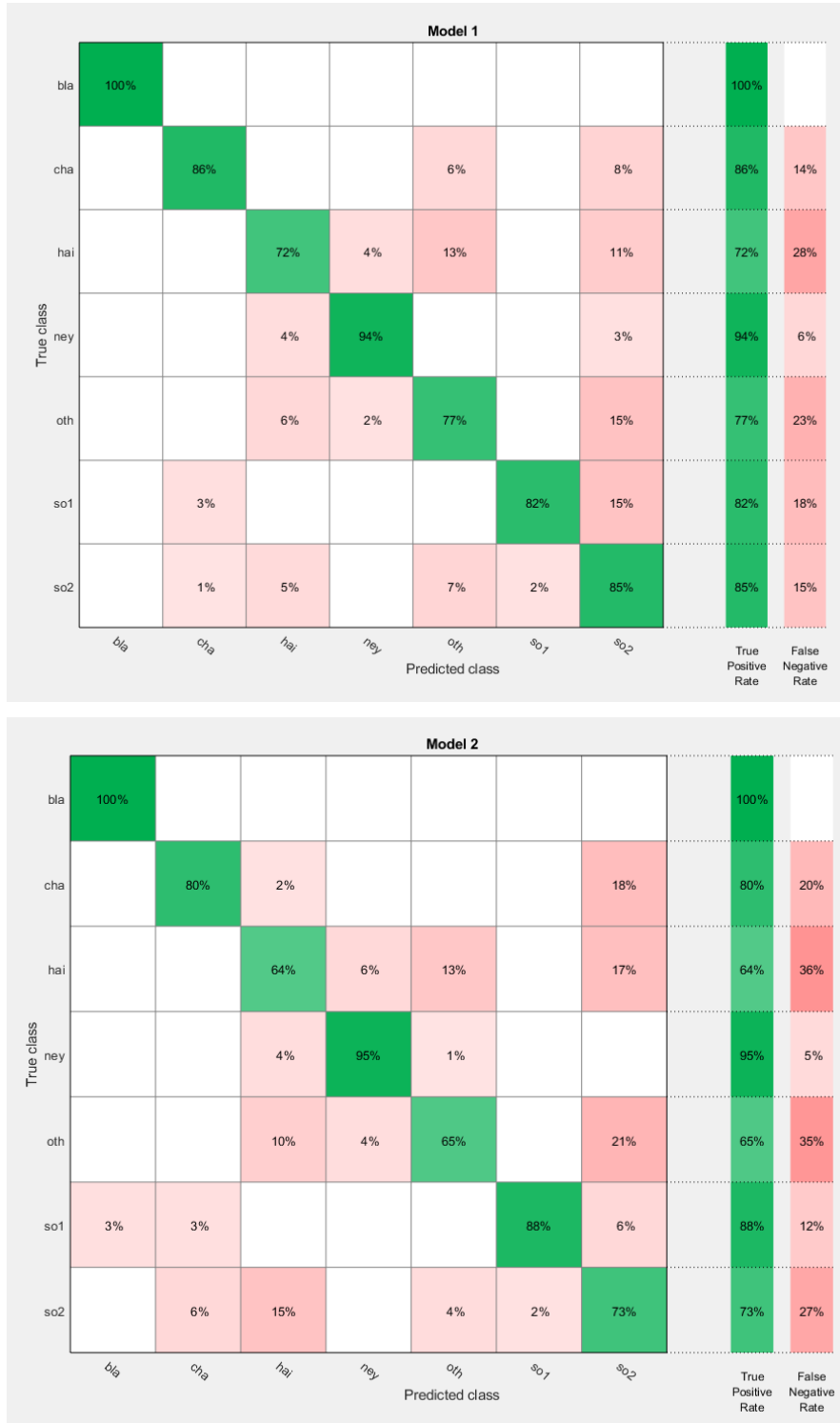
Figure 4-8 The confusion matrix for trained models on GLCM features. SVM model 84% (up) and ensembled trees 78% (down).
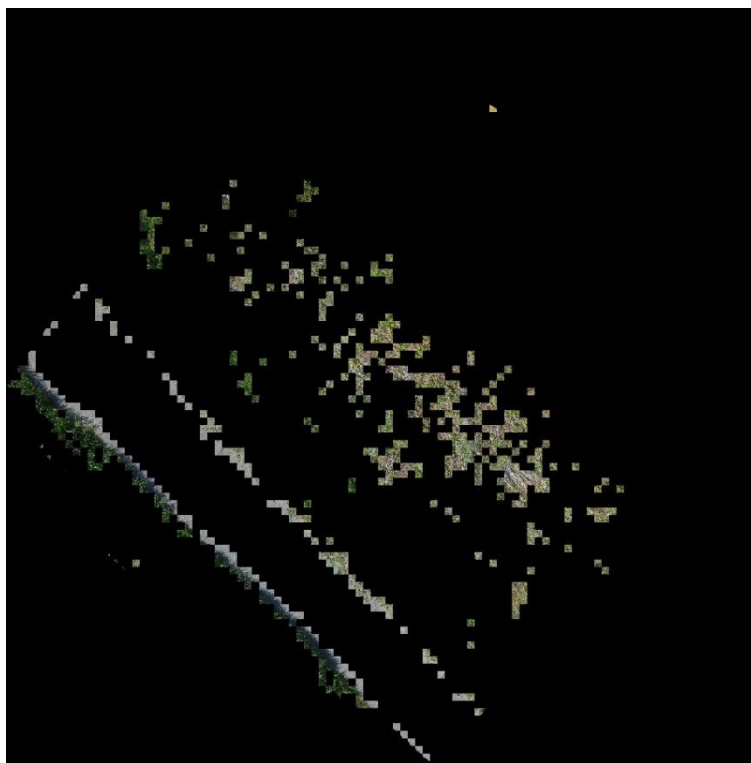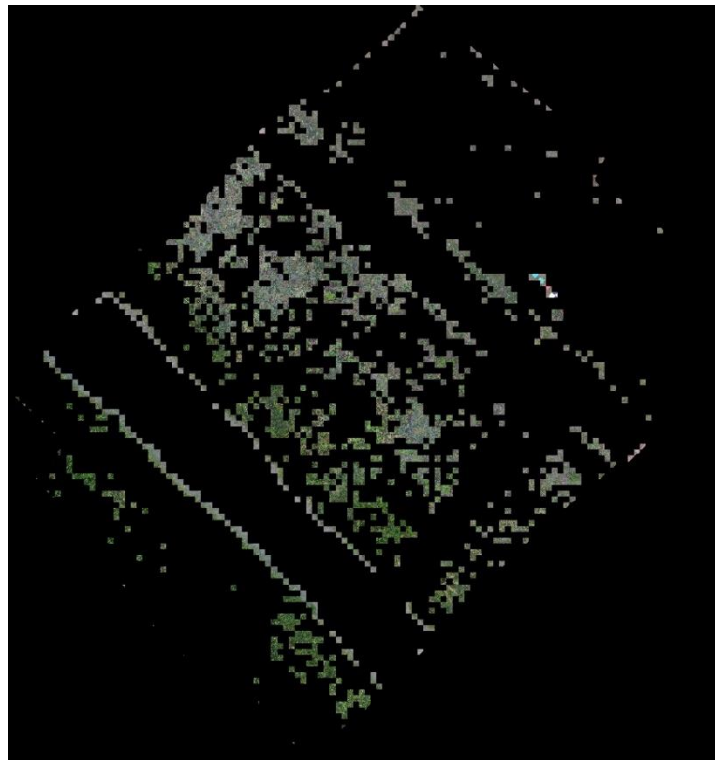
Figure 4-9 The test results of hairy vetch discrimination for SVM model and color features.
The 2016 orthomosaic (up) and 2017 orthomosaic (down).
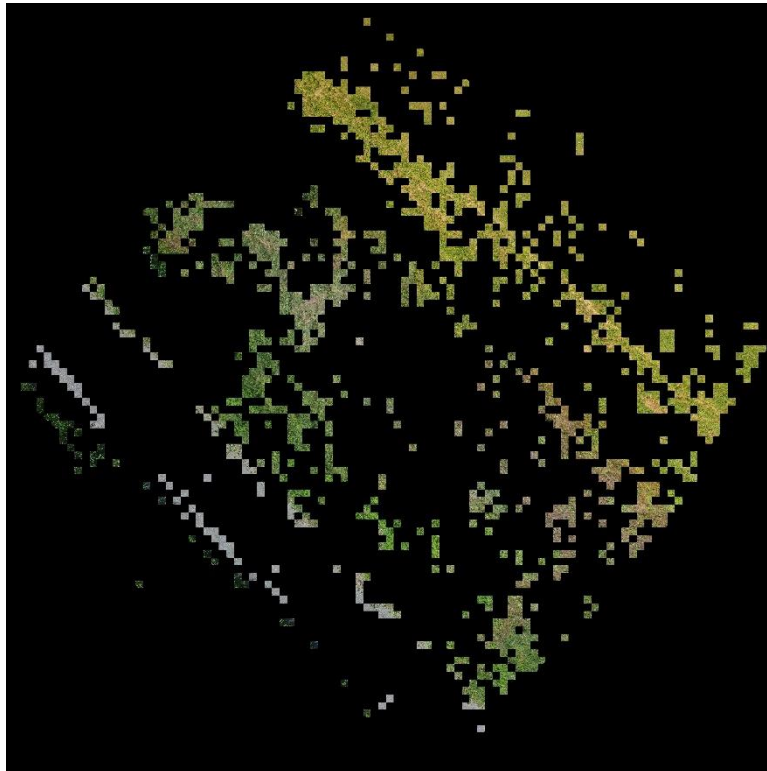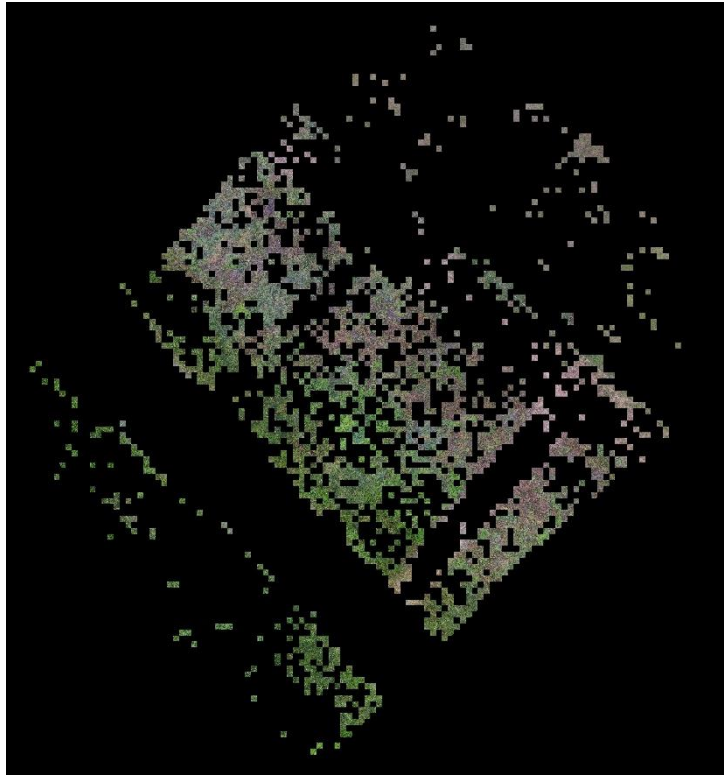
Figure 4-10 The test results of hairy vetch discrimination for SVM model and GLCM features.

The 2016 orthomosaic (up) and 2017 orthomosaic (down).

4.2. Testing and classification results

The result of classification for the trained SVM on color and GLCM features models are shown in figure 4-9 and 4-10. The comparison between the result of manual classification of hairy vetch (figure 4-3) and the models results indicates the ability of trained models in discriminating the hairy vetch in UAV image. Also, it should be noted that the results for color features set is more accurate for the 2016 image. Furthermore, SVM results for GLCM features is more accurate in detecting hairy vetch in both orthomosaic images although the number misclassified blocks are higher that the color feature set model.

The other important outcome of the classification results is that all the models are light dependent and a slight difference in color temperature of two mosaic images introduces higher misclassification to the result of all models. This color difference effects the color features models more.

4.6 Conclusion

Results showed the highest accuracy of 94% for SVM classifier on color properties. Phragmites and Hairy vetch were classified using SVM model with the accuracy of 94 and 91%. Although, the accuracy for color properties are the highest classification of orthomosaic images showed that the models are more successful in discriminating different color and textural patterns in 2016 images. GLCM texture and SVM classifier showed the highest potential in detecting hairy vetch and Phragmites plants in both orthomosaic images.

4.7 References

Kak, Avinash C., and Azriel Rosenfeld. "Digital picture processing." New York (1982).

Materka, Andrzej, and Michal Strzelecki. "Texture analysis methods–a review." Technical university of lodz, institute of electronics, COST B11 report, Brussels (1998): 9-11.

Haralick, Robert M., and Karthikeyan Shanmugam. "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics6 (1973): 610-621.

Costa, Alceu Ferraz, Gabriel Humpire-Mamani, and Agma Juci Machado Traina. "An efficient algorithm for fractal analysis of textures." In Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on, pp. 39-46. IEEE, 2012.

P. Liao, T. Chen, and P. Chung, "A Fast Algorithm for Multilevel Thresholding," Journal of Information Science and Engineering, vol. 17, no. 5, pp. 713–727, 2001.

Motohka, Takeshi, Kenlo Nishida Nasahara, Hiroyuki Oguma, and Satoshi Tsuchida. "Applicability of green-red vegetation index for remote sensing of vegetation phenology." Remote Sensing 2, no. 10 (2010): 2369-2387.

Larrinaga, Asier, and Lluis Brotons. "Greenness Indices from a Low-Cost UAV Imagery as Tools for Monitoring Post-Fire Forest Recovery." Drones 3, no. 1 (2019): 6.

Sonnentag, Oliver, Koen Hufkens, Cory Teshera-Sterne, Adam M. Young, Mark Friedl, Bobby H. Braswell, Thomas Milliman, John O'Keefe, and Andrew D. Richardson. "Digital repeat photography for phenological research in forest ecosystems." Agricultural and Forest Meteorology 152 (2012): 159-177.

Wittek, Peter. Quantum machine learning: what quantum computing means to data mining. Academic Press, 2014.

Breiman, Leo. "Bagging predictors." Machine learning 24, no. 2 (1996): 123-140.

Breiman, Leo. Bias, variance, and arcing classifiers. Tech. Rep. 460, Statistics Department, University of California, Berkeley, CA, USA, 1996.

Alhichri, Haikel, Essam Othman, Mansour Zuair, Nassim Ammour, and Yakoub Bazi. "Tile-Based Semisupervised Classification of Large-Scale VHR Remote Sensing Images." Journal of Sensors 2018 (2018).

Fan, Zhun, Jiewei Lu, Maoguo Gong, Honghui Xie, and Erik D. Goodman. "Automatic tobacco plant detection in UAV images via deep neural networks." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, no. 3 (2018): 876-887.

Müllerová, Jana, Josef Brůna, Tomáš Bartaloš, Petr Dvořák, Michaela Vítková, and Petr Pyšek. "Timing is important: unmanned aircraft vs. satellite imagery in plant invasion monitoring." Frontiers in plant science 8 (2017): 887.

Chapter 5: Integration of Machine Learning and Open Access Geospatial Data for Land Cover Mapping

5.1 Abstract:

In-time and accurate monitoring of land cover and land use are essential tools for countries to achieve sustainable food production. However, many developing countries are struggling to efficiently monitor land resources due to the lack of financial support and limited access to adequate technology. This study aims at offering a solution to fill in such a gap in developing countries, by developing a land cover solution that is free of costs. A fully automated framework for land cover mapping was developed using 10-meter resolution open access satellite images and machine learning (ML) techniques for the African country of Lesotho. Sentinel-2 satellite images were accessed through Google Earth Engine (GEE) for initial processing and feature extraction at a national level. Also, FAO land cover (LC) data were used to train a support vector machine (SVM) and bagged trees (BT) classifiers. SVM successfully classified urban and agricultural lands with 62 and 67 % accuracy, respectively. Also, BT could classify the two categories by 81 and 65% accuracy, correspondingly. Trained models could provide precise LC maps in minutes or hours. It can also be utilized as a viable solution for developing countries as an alternative to traditional geographic information system (GIS) methods, which are often labor intensive, require acquisition of very high-resolution commercial satellite imagery, time consuming and call for high budgets.

5.2 Introduction

The United Nations (UN) predicts that the world population will increase to 8.7 billion by 2030 and 9.7 billion by 2050 (UN, 2015). This population growth impacts natural resources utilization and causes land use changes. Hence, the UN 2030 agenda for sustainable developments goals (SDG) aims to, achieve food security, improve nutrition, and promote sustainable agriculture so as to end hunger (Kamwi et. al, 2015). This aim has been highlighted under target 2.4 of the SDGs by promoting political and technological efforts to ensure sustainable food production systems, implementing resilient agricultural practices that increase productivity and production, helping maintain ecosystems, strengthening capacity for adaptation to climate change, extreme weather, drought, flooding, and progressively improving land and soil quality (UN, 2015). In more detail, indicator 2.4.1 of the SDGs is defined as the proportions of the agricultural area within productive

and sustainable agriculture and was specifically designed to focus on agricultural lands which are used to grow crops and raise livestock (UN, 2019). Therefore, characterizing and mapping land cover are essential for planning and managing natural resources including agricultural lands (Gomez et al, 2016). For this purpose, implementation of efficient operational land cover requires advanced remote sensing methodologies with the ability to provide inexpensive, on-demand and accurate land cover products using availably free and open access data and free cloud-based data processing platforms.

Regarding land cover mapping resources, Global Land Cover – SHARE database (GLC-SHARE) developed by the UN's Food and Agriculture Organization (FAO) is a database with 30 arc-second spatial resolution based on available national and regional land cover databases (Latham et. al, 2014). Inglada et al, provided the land cover of France using Sentinel-2 products for 17 land cover classes (Inglada et.al, 2017). Also, a cloud-based platform and dense stack satellite time series were utilized to provide artic land cover (Nyland et. al, 2018). Belgiu et al, investigated the ability of a Machine Learning (ML) methodology in land cover mapping in different agro-ecological regions of the planet (Belgiu et. al, 2018). Cardille et al, Introduced Bayesian Updating of Land Cover (BULC) for the ongoing updating of land cover classification (Cardille et. al, 2016). Xiaong et al. developed an automated cropland scheme for the continent of Africa by utilizing Moderate Resolution Imaging Spectroradiometer (MODIS) data (250-meter resolution) and Google Earth Engine (GEE) (Xiong et. al, 2017). Furthermore, the Climate Change Initiative (CCI) team of the European Space Agency (ESA) released a land cover map at 20-meter resolution over Africa for the year 2016 (Myroslava et. al, 2017).

The above researches were performed using supervised or unsupervised methods to map and classify land cover. However, such methods are time-consuming and require substantial labor and funds (Inglada et.al, 2017). Moreover, the mentioned methodologies have been applied to sets of data collected at a certain period of time (Hachigonta et. al, 2013). This leads to the inability to monitor changes in land cover or to conduct further necessary post-analysis. Furthermore, implementation of such methods on cloud-based platforms are very limited. Thus, many attempts with remote sensing have been carried out in order to overcome the challenges of producing less costly and more time efficient land cover mapping (Mumby et. al, 1999).

In this study, we are proposing a combination of an automated land cover mapping methodology and machine learning technique using sets of data obtained from the UN's FAO land cover maps and free of cost open access Sentinel-2 high-resolution imagery adapted through open access and a cloud-based platform to produce high precision and on-demand land cover maps. The FAO

provides land cover maps at the national and global level for a diverse range of purposes such as forest management, global land cover, water management, etc (Latham et. al, 2014; Ridder et. al, 2007; UN 2019). These land cover maps are generated from different methods including field, airborne or space-borne data. Provided data are valuable sources of information for various applications (Stibig et. al, 2007). GEE is a cloud-based platform providing access to free satellite and airborne image services and offering computational power (Gorelick et. al, 2017), as well as access to a wide range of satellite data through its Application Program Interfaces (API) including the ESA's Copernicus Programme, (ESA, 2019) NASA and the U.S. Geological Survey (Woodcock et. al, 2008; Loveland et. al, 2012). Furthermore, it allows users to ingest and utilize other geospatial and/or in-situ data as fusion tables or images. GEE has been utilized in several studies covering topics such as global forest change (Hansen et. al, 2013), global surface water change (Pekel et. al, 2016), crop yield estimation (Lobell et. al, 2015), rice paddy mapping (Dong et. al, 2016), urban mapping (Patel et. al, 2015; Zhang et. al, 2015), flood mapping (Coltin et. al, 2016), and land cover mapping (Huang et. al, 2017; Sidhu et. al, 2018). The computational power of GEE along with its comprehensive data access make GEE a capable option for implementation of land cover mapping platforms which are timely, accessible from remote areas and free of cost for researchers and countries with economical and technical difficulties.

The objectives of this study are summarized as following:

- To investigate the potential, limitations, and utilization of GEE for feature extraction.
- To study the advantages of adding spatial feature to classify land cover and the feasibility of high dimensional feature space in similar applications.
- To evaluate the performance of machine learning models to classify the land surface by using high dimensional feature space.
- To evaluate the methodology on three different areas in Lesotho to ensure that it is independent from climatic variables and agro-ecological zones.

This study is structured as follows: Section 2 describes the study areas and the data; Section 3 presents the proposed methods; Section 4 is dedicated to the results and highlights the main findings and the implications of this study and is followed by our conclusion.

5.3 Study Area and Data

2.1 Study Area.

This study has been carried out over territory of the Kingdom of Lesotho (Figure 5-1). The Kingdom of Lesotho is a landlocked country completely surrounded by South Africa. The total

area of the country is 30,450 square kilometers, making it the 137th largest country on the planet. Lesotho's population was estimated at 2,174,645 in 2015 by the UN and expected to reach 2,607,957 by 2030 and 3,203,470 by 2050 (UN, 2015). The poverty level in Lesotho is high with 62 percent of the population living on less than 2 USD per day. It is estimated that 36.4 percent are living on 1 USD or less per day. Agriculture is the main economic sector in Lesotho such that 60-70 % of the country's laborers obtain supplemental income from agriculture.
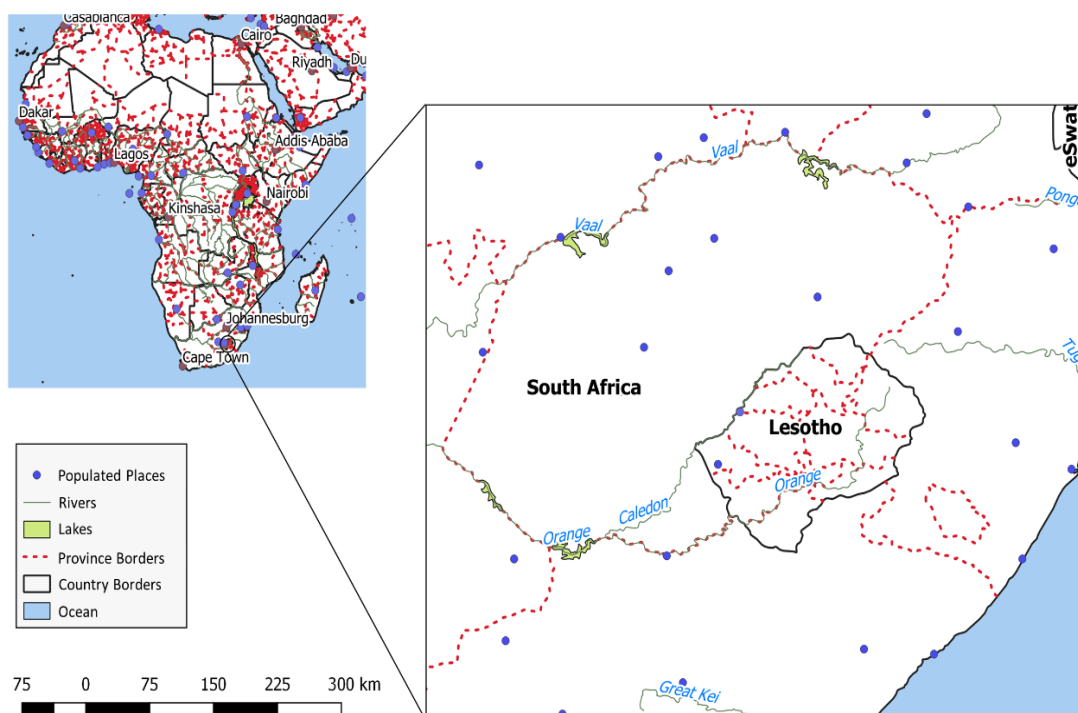
Figure 5-1 The Kingdom of Lesotho (study area) within the continent of Africa.

Poverty is the worst in rainfed farming areas. The country's territory is divided into four agro-ecological zones based on altitude and landscape, namely: lowlands, the foothills, the Senqu River Valley (SRV), and the mountains (Hachigonta et. al, 2013). The FAO Global Agro-ecological Zones (Fischer et. al, 2012) divide the country into four major agro-ecological zones which are characterized by four different lengths of growing period (LGP) based on analysis of climate, soil, and terrain data (Figure 5-2).
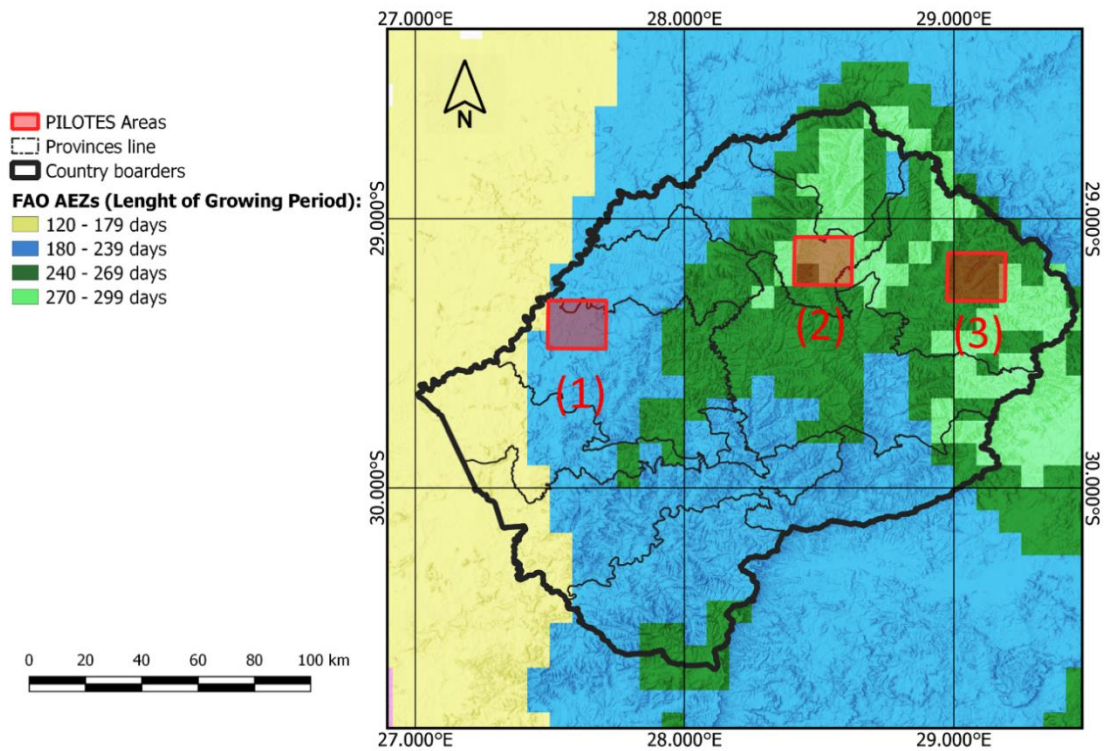
Figure 5-2 Agro-ecological zones of Lesotho based on length of growing period (LGP) and foot-print of pilot areas. 1) Lejone, 2) Maseru, and 3) Mokhotlong

The climate of Lesotho is characterized by two main seasons: a rainy season from October to the end of March and a dry season from April to the end of September. Normally the majority of precipitation occurs during the southern hemisphere summer thunderstorms (Figure 5-3).
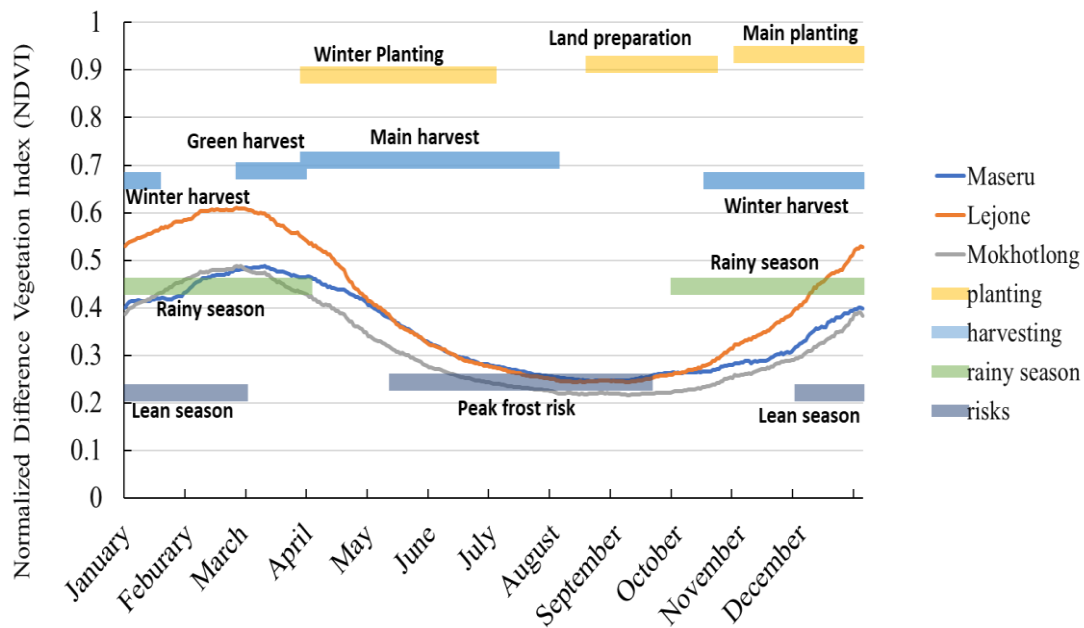
Figure 5-3 Comparison between mean of normalized difference vegetation index (NDVI) of three studied pilot areas and agricultural seasons in Lesotho.

In this study three different pilot areas labelled as 1) Lejone, 2) Maseru, and 3) Mokhotlong with surface areas of 482.65, 484.45, and 492.39 square kilometers, respectively, were selected from the territories of Lesotho in a way that each area would fall within a distinct agro-ecological zone to represent Lesotho's variability of different geological features, land cover and land use.

Footprints of the selected areas are shown in Figure 5-2. This will ensure the efficiency and the independency of the proposed algorithm for providing comprehensive training data. It is well-known that different agro-ecological zones form unique ecosystems influencing plant growth characteristics including photosynthesis (plant chlorophyll content) and therefore spectral reflection (Mokarram et. al, 2015).

2.2 FAO Land Cover Lesotho Classes

FAO Land Cover of Lesotho (FAO LCL) was developed in 2016 using commercial satellite images and conventional GIS methods, along with supervised image classification techniques with very high-resolution satellite and airborne images with spatial resolution of 1.5 meter. It utilizes an object-based classification on pan-sharpened images. The FAO LCL consists of 32 land cover categories based on FAO ISO 19144 Land Cover Meta Language, which provides a common reference structure for the comparison and integration of data for any generic land cover

classification system (UNFAO, 2012). A list of land cover classes and the specifications of satellite images that were used in FAO LCL data are shown in Tables 5-1 and 5-2, respectively (UNFAO, 2017). In this study, the FAO LCL resolution was resampled in order to change the spatial resolution from 1.5 m to 10m in order to match Sentinel 2 spatial resolution. The pixel values of the original dataset represented the dominant land cover for each land cover class within each 10 m X 10 m cell. Such values were aggregated across the new pixel size using a majority filter of 100m X 100m cell size. The majority filter algorithm determined the new value of the cell based on the most popular land cover value within the filter window. Such a method is mainly used with discrete data. Such method has been used by FAO in 2014 to harmonize national land cover datasets into the GLC-SHARE land cover and by other studies (Latham et. al, 2014; Campbell et. al, 1996; Vancutsem et. al, 2013).

Table 5-1. FAO LCL classes and classes definitions

Class code	LC Type	LC Name	LC description
1	BUILT-UP (4 classes)	Urban Areas	Relatively larger urban built-up areas, commonly with presence of trees
		Urban Commercial and/or Industrial areas	Commercial and/or industrial built-up areas
		Rural Settlements, Plain Areas	Rural houses in flat lying plain areas + small cultivated herbaceous crops + closed herbaceous natural vegetation, often together with trees and/or shrubs employed for demarcation
		Rural Settlements,	Rural houses in sloping and mountainous areas + herbaceous natural vegetation,

		Slopping and Mountain Areas	occasionally with shrubs employed for demarcation, usually treeless
2	AGRICULTURE (5 Classes)	Rainfed Agriculture, Plain Areas	Rainfed herbaceous crops cultivated in flat-lying plains, relatively larger sized fields
		Rainfed Agriculture, Sloping & Mountainous regions	Rainfed herbaceous crops in sloping land and mountains with terracing and/or contour ploughing, small and medium sized fields, sometimes with lines of shrubs demarcating fields
		Rainfed Agriculture, Sheet Erosion	Rainfed herbaceous crops with visible water sheet erosion, commonly with associated gully erosion
		Irrigated Agriculture	Small size irrigated herbaceous crops near water courses
		Rainfed Agriculture + Rainfed Orchards	Small rainfed herbaceous crops + regular rainfed orchard plantations (usually as rows of fruit trees separating elongated fields)
3	TREES (7 Classes)	Trees, Needle leaved, (Closed)	Closed evergreen needle-leaved trees, sometimes occurring as plantations
		Trees, Needle leaved, (Open)	Open evergreen needle-leaved trees + herbaceous natural vegetation
		Trees, Broadleaved, (Closed)	Closed deciduous broadleaved trees, commonly along river beds

		Trees, Broadleaved, (Open)	Open deciduous broadleaved trees + herbaceous natural vegetation
		Trees, Undifferentiated (Closed)	Closed undifferentiated trees
		Trees, Undifferentiated, (Open)	Open undifferentiated trees + herbaceous natural vegetation
		Trees, (Sparse)	Sparse trees + herbaceous natural vegetation (closed - open)
4	HYDROLOGY (4 Classes)	Large Waterbody	Large perennial fresh water lake or dam reservoir
		Small Waterbody	Small fresh water seasonal and/or perennial reservoir, Pool, Waterhole, etc.
		Wetland (Perennial and/or seasonal)	Natural perennial and/or seasonal fresh waterbody + Perennial closed-open natural vegetation
		River Bank	River Bank (soil/sand deposits) + perennial or periodic flowing fresh water (river)
5	SHRUBLAND (2 Classes)	Shrub-land – (Closed)	Natural Shrubs (H=0.5 to 1.5m), Closed
		Shrub-land - Open	Natural Shrubs (H=0.5 to 1.5m), Open + Natural herbaceous vegetation (Open Closed)
6	GRASSLAND	Grassland	Grassland - Natural vegetation

	(1 Class)		
7	BARREN LAND (5 Classes)	Bare Rock	Rock outcrops
		Bare Area	Bare areas - undifferentiated areas not used for cultivation and usually devoid of grass or shrub cover
		Boulders & Loose Rocks	Areas with large scattered boulders and/or unconsolidated loose rocks, commonly sloping, usually together with patchy natural vegetation and/or shrubs and/or natural trees
		Gullies	Gully erosion, occasionally with trees and/or tall shrubs
		Mines & Quarries	Major mines and quarries as well as temporary building material extraction sites

Table 5-2. FAO LCL specifications of utilized images

Image source	Spatial Resolution (meter)	Spectral Resolution
Rapid Eye	5	5 bands (440 to 850 nm)
Spot 5	2.5	5 Bands (480 to 1750 nm)
Aerial orthophotos	0.5	3 Bands (visible light)

2.3 Test and Training Data Set Generation

Machine learning (ML) models are often used to classify satellite images. In order to train the ML models, a set of ground truth labeled data points are required. In this study, we have randomly selected over 12,000 points from three pilot areas. To ensure that the accuracy measurements for

all classifications are in the same order, the training and validation points were equally distributed between all land cover classes. FAO LCL was used for labelling the training points.

5.4 Methods

Figure 5-4 shows the workflow of the proposed methodology. First, the FAO LCL data was prepared and ingested into the GEE as data assets. Data preprocessing and feature extraction was done using GEE on cloud machine by utilizing GEE JavaScript API. All the available Sentinel-2 imagery between the first of October 2015 until the end of December 2017 were used as open access data. Images were preprocessed to remove cloudy pixels and formed 1) an image stack divided into four seasons and 2) a single image for all the territory of Lesotho. The image stack was used to extract the spectral features as explained in Section 3.3.1. The single image was sharpened using Principal Component Analysis (PCA)coefficients and then spatial features were extracted as explained in Section 3.3.2. Finally, the feature vectors of three pilot areas were exported for training, validation, and classification steps on local or cloud machines.

3.1 Google Earth Engine Data

In this regard, GEE was used to collect and extract data at the national level for Lesotho. First, all the available Sentinel-2 image tiles were selected for Lesotho and clipped to the country's borders. Clouds were filtered based on a pixel-wise filtering technique by the band QA60 of Sentinel-2 products with a 5% cloudy pixel rate. This guaranteed that all the image tiles used in the next steps were only removing cloudy pixels of each image without removing the whole tile. Then image collection was divided into four groups based on the imagery season. Lesotho possesses four distinct seasons: spring-rainy season (October first to the end of December), summer-rainy season (January first to end of March), Autumn-dry season (April to end of June) and Winter-dry season (July to end of September). The season time periods are based on the data available on the Famine Early Warning Systems (FEWS) (www.fews.net) and the MODIS NDVI band (Figure 4-3).
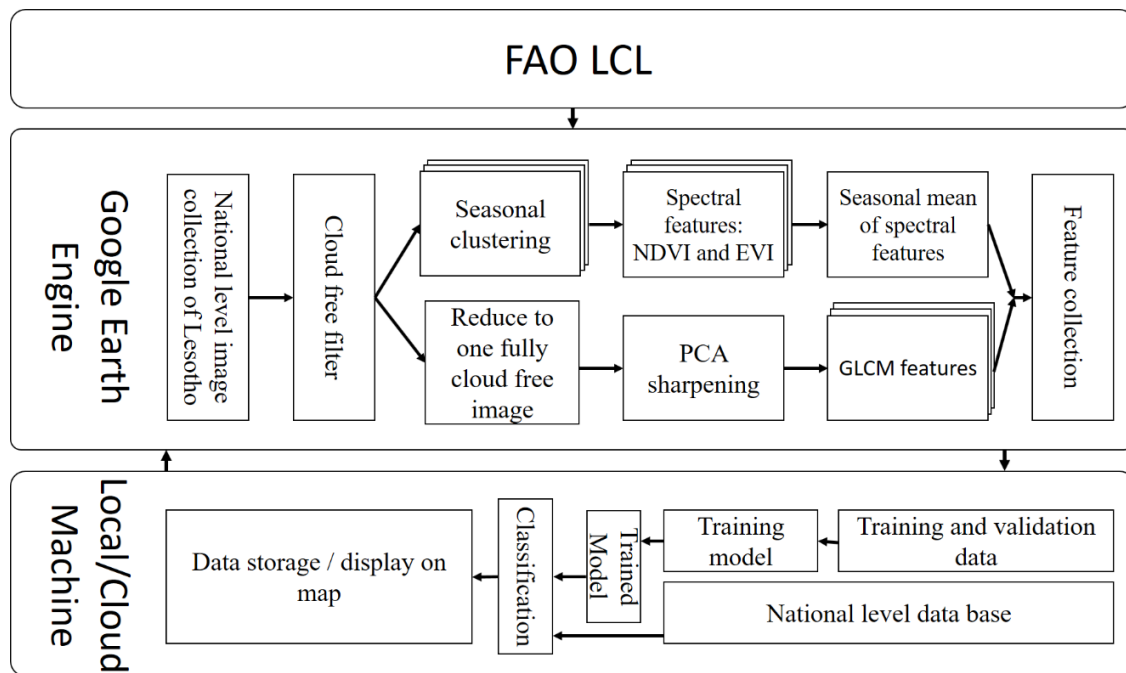
Figure 5-4 Algorithm workflow of proposed machine learning classification and accuracy assessment methodology.

3.2. Data Preparation

3.2.1. Spectral Features

High spectral resolution satellite sensors produce images in a higher number of bands in both visible and invisible spectrums with narrower bandwidth (UNFAO, 2017). Higher spectral sensitivity of sensors could help in obtaining more information from earth surfaces including vegetation. One of the benefits of high spectral images is the possibility of calculating different vegetation indices (VI). VIs, which are obtained from satellite sensors, have been intensively studied and applied in many environmental and remote sensing studies. Xue et al. listed 118 different VIs and reviewed the most popular ones (Xue et. al, 2017). In this regard, the Normalized Vegetation Index (NDVI) is the most utilized VI, which was introduced by Rouse Jr. et al. (Rouse et. al, 1974) as follows:

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

NDVI is sensitive to vegetation even in conditions of scarce vegetation. However, soil brightness, soil color, atmosphere, clouds, and shadows affect the NDVI. Since NDVI has negative

correlation with soil brightness and atmospheric effect. Enhanced Vegetation Index (EVI) was later introduced which simultaneously correlates soil and atmospheric effects as:

$$EVI = G \times \frac{NIR-Red}{NIR+C_1 \times Red-C_2 \times Blue+L},$$ (2)

Where G is the gain factor equal to 2.5, L = 1 is the soil-adjustment factor, C1 = 6 and C2 = 7.5 are the coefficients of the aerosol resistance term, which uses the blue band to correct for aerosol influences in the red band (Jiang et. al, 2008; A database for remote sensing indices, 2019).

In this study, the seasonal mean of NDVI and EVI were calculated as spectral features based on Lesotho climatic seasons on all existing data. The combined application of NDVI and EVI decreases the effect of atmospheric transmission and insures the reliability of the spectral features.

3.2.2. Spatial Features

In addition to the spectral properties of satellite images, the spatial properties (i.e., image texture) are important characteristics that can be used to describe land cover. Texture is defined as feel, appearance, or consistency of a surface or a substance. In the digital image processing field, the texture is the spatial arrangement of colors or pixel intensities. Different sets of metrics are designed in image processing to quantify the perceived texture of an image. Hereby, spatial properties were extracted from the Sentinel-2 images into two steps.

3.2.2.1. Image Pre-Processing with PCA

To extract the textural properties of images, it is necessary to pre-process the images in order to remove the noise or increasing the amount of details in each images tile. Therefore, to increase the differences between different textures in an image, principal component analysis (PCA) was applied.

PCA uses orthogonal transformation to convert possibly correlated data into linearly uncorrelated values so that the first principal component has the largest possible variance (Wold et. al, 1987). In general, almost any data matrix can be simplified by PCA. PCA estimates the correlation structure of variables. Therefore, principal component analysis of a data matrix extracts the dominant patterns in the matrix.

PCA was applied on visible channels (B2, B3, and B4) and NIR channel (B8) of Sentiel-2 images. Then the first component was combined with the original image with a coefficient. Pan-sharpening is the injection of high-resolution panchromatic image into lower resolution multispectral images to get higher resolution multispectral images (Shah et. al, 2008). However,

this study takes advantage of PCA pan-sharpening method on Sentinel-2 images by applying PCA coefficients that were obtained from visible bands and bands with higher atmospheric transmission (i.e. band 8 and 12) on all the bands. Figure 5-5 compares the first and second principal components with Google Earth images, Sentinel-2 images and the final sharpening results using first and second components.

3.2.2.2. Texture Features: Grey Level Co-occurrence Matrix

A co-occurrence matrix is a matrix that is defined over an image I in which it is the distribution of co-occurring pixel values (grayscale values or colors) at a given offset $(\Delta x, \Delta y)$ that is a position operator to be applied on any pixel in an image. The $(i,j)^{th}$ value of the co-occurrence matrix gives the number of times in the image that the ith, and j^{th} pixel values occur in the relation given by the offset. For an image with p different pixel values, the $p \times p$ co-occurrence matrix C is defined over an $n \times m$ image I, as:

$$C_{\Delta x, \Delta y}(i,j) = \sum_{x=1}^{n} \sum_{y=1}^{m} \begin{cases} 1, & if\ I(x,y) = i\ and\ I(x + \Delta x, y + \Delta y) = j \\ 0, & otherwise \end{cases}, \qquad (3)$$

where, $I(x,y)$ indicates the pixel value at the pixel. Here, the co-occurrence matrices are calculated over 3 visible and 2 NIR bands for an offset of 8 pixels distance in all directions.

Second order statistics could be driven from co-occurrence matrices to quantify the properties of an image. In this study, we have calculated 18 second-order statistics introduced by Haralick et al. and Conners et al., e.g. correlation, sum of variance, contrast, etc. (Conners et. al, 1984; Haralick and Shanmugam, 1973).
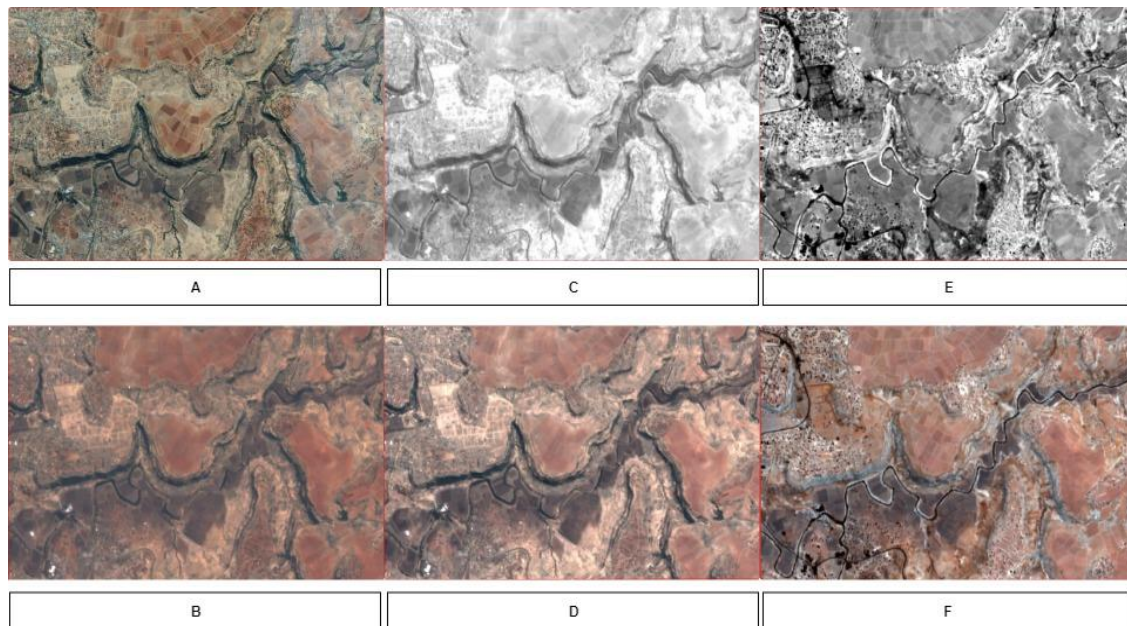
Figure 5-5 A) Google Earth high resolution image. B) Sentinel-2 10-meter resolution image. C) First principal component of 4 selected bands. D) PCA sharpened image with first principal component. E) Second component of PCA. F) Sharpened image with second component of PCA.

3.3. Trained Machine Learning Models

For machine learning models, we have trained bagged trees (BT) and support vector machine (SVM). BT and SVM have been used in various studies for land surface classification (Hao et. al, 2015; Mountrakis et. al, 2011; Mardani et. al, 2017). The BT that was used in this study is an ensemble of decision trees model with 30 learner trees (Breiman et. al, 1994). The BT model in this study only accepts one parameter (number of trees). It should be noted that, typically, in remote sensing applications, random forest classifiers are used which are a specific type of bagged trees in which it needs two parameters. First, the number of trees, and second is the number of features (predictors) to train each decision tree. However, in this study, the BT model was trained on all features. The SVM model was trained based on a degree three polynomial kernel (Boser et. al, 1992) and one verses one (ovo) decision function.

5.5 Results

4.1. Trained Models' Performance

Table 5-3 shows the overall accuracy and training time for both models. The training time for the SVM model was significantly higher than the BT model. The BT and SVM were trained and

validated based on a randomly generated dataset (12,000 point). BT yielded a higher overall accuracy in a 5-fold cross validation process.

Table 5-3. Results of the overall accuracy.

Classifier	Training time (seconds)	Over-all accuracy (%)
Bagged Trees	76	62.6
Support Vector Machine	1197	60.4

Both models were trained and validated on a laptop PC with 1.8 GHz dual-core Intel Core i5 CPU and 8 GB of memory to demonstrate the efficiency and practicality of the proposed method in real world scenarios. Training and segmentation steps in remote sensing applications require high computational power as well as memory resources (Inglada et. al, 2017; Gorelick et. al, 2017). In this study, the training time for both models showed very high efficiency despite the size of the training and validation datasets. Furthermore, the achieved time efficiency would allow the users to modify or add ground truth data samples and to retrain the models in the future. This will guarantee the operational ability of the algorithm.

Tables 5-4 and 5-5 show the confusion matrices for the BT and SVM models. The BT achieved highest accuracy, in the built-up class, 81%, while the SVM gained the highest accuracy in the hydrology class, 76%. The differences of models' performance are not significant in agriculture, tree and barren-land classes. The lowest accuracy was observed in shrub-land and grass-land classes for both models. However, in the case of in grass-land class, the BT showed a lower score than the SVM. The ESA CCI 20 m land cover of Africa reported the overall accuracy of 65%, however, the overall accuracy for Lesotho was reported between 22 to 45%. The proposed methodology in this research outperformed the ESA CCI 20m land cover by 20% (Myroslava et. al, 2017).

Table 5-4. Confusion Matrix - Bagged Trees Model (%).

Class No.	Class name	Built-up	Agriculture	Trees	Hydrology	Shrub-land	Grassland	Barren-land

1	Built-up	81	6	3	1	1	5	3
2	Agriculture	9	65	2	2	6	11	5
3	Trees	10	3	66	3	11	4	3
4	Hydrology	6	7	5	73	2	4	3
5	Shrub-land	4	6	13	1	55	11	10
6	Grass-land	11	15	5	3	14	38	14
7	Barren-land	7	6	3	3	8	9	63

Table 5-5. Confusion Matrix - Cubic SVM Model (%).

Class No.	Class name	Built-up	Agriculture	Trees	Hydrology	Shrub-land	Grass-land	Barren-land
1	Built-up	62	8	5	3	2	15	5
2	Agriculture	5	67	2	3	5	13	6
3	Trees	4	2	64	4	13	8	4
4	Hydrology	2	6	4	76	2	5	5
5	Shrub-land	2	5	9	2	55	17	9
6	Grass-land	5	13	4	4	17	43	14
7	Barren-land	3	7	3	4	9	14	60

4.2 Classes Accuracy and Inter-Class Similarities

Tables 5-4 and 5-5 also shows the highest confusion rate for each class. It should be noted that both ML models have high confusion rates in similar classes. The built-up class allocated the highest confusion rate in agriculture and grass-land classes with 6% and 5% for the BT model and 8% and 15% for the SVM model, respectively. Similarly, the highest confusion rate for the agriculture class was observed in the grass-land class with 11% and 13% for the BT and SVM, respectively.

Tree and shrub-land classes showed a 13% and 9% confusion rate where the tree class was selected as shrubs or vice versa. Also, the highest confusion rate for shrub-lands was occurred in grass-land class with 17%. Grass-land class has the highest confusion rates with other classes with 17%, 14%, and 13% for shrub-land, barren-land, and agriculture areas respectively.

4.3 Discriminating Ability of The Train Models: Precision, Recall and Receiver Operator Curve

Figure 5-6 shows the recall, precision and F-score values for both trained models. As can be seen in figure 5-6, both models have similar precision and recall values. Precision is the fraction of relevant points among the retrieved points in a specific class, while recall is the fraction of relevant points that is retrieved over the total amount of relevant points in a class. Our results indicate that both models are successful in retrieving relevant classes for both in-sample and out-sample data. Precision comparing false positives to true positives. It captures the effect of the large number of negative examples on the model's performance, while higher recall means that the trained model is successful in returning most of the relevant results (Davis et. al, 2006).

Figure5-7 compares the recall, area under the curve (AUC) in a receiver operator curve (ROC), and a false positive rate (FPR) for both ML models. When comparing the AUC values of trained models, the BT model showed higher ability in retrieving relevant data than the SVM. Also, the false positive rate for both models was very low, meaning the number of incorrect points that were assigned to each class of the land cover is low.
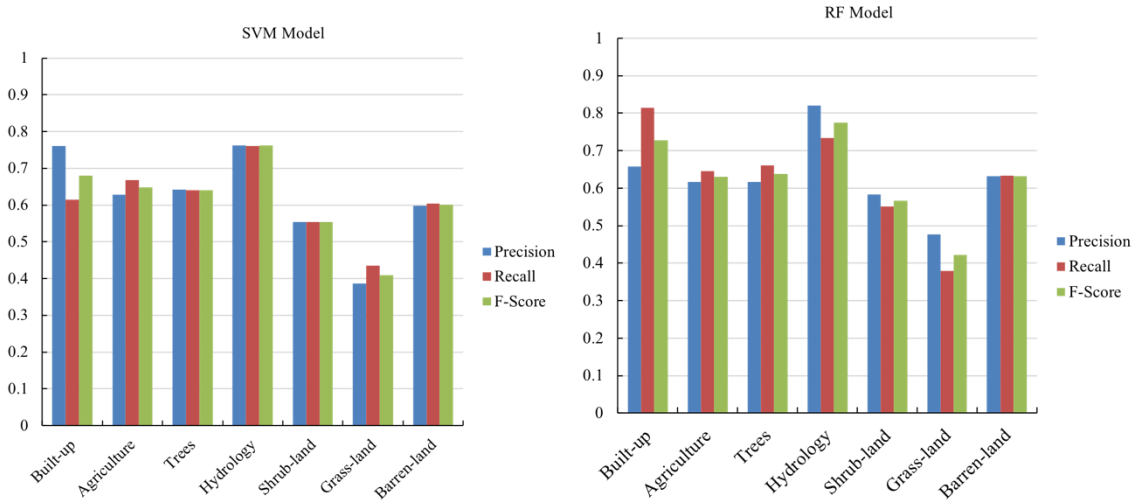
Figure 5-6 Precision recall and F-Score parameters for each class. Left: Support Vector Machine model. Right: Bagged Trees model.
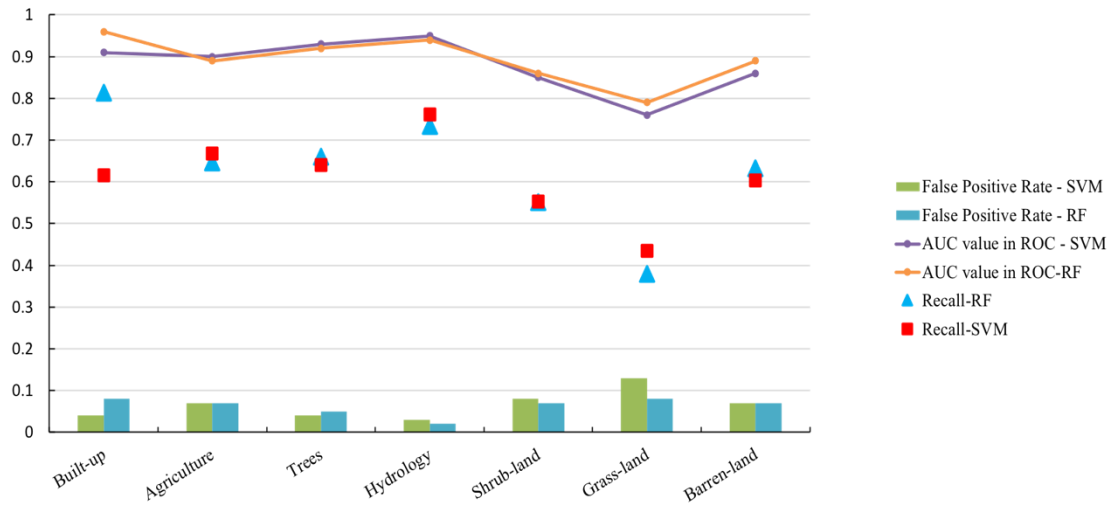
Figure 5-7 Area under curve (AUC) values in receiver operator curve (ROC), recall, and false positive rate per class for trained models.

4.4 Classification Results and Final Land Cover Product

The trained models were tested on the selected pilot areas to generate the final land cover products. The testing step was run on the same PC as training and validation step. Segmentation results for studied pilot areas are shown in Figure 5-8. The images in column A show the FAO LCL as the base map for training ML models. Images in columns B and C show the results for the SVM and BT models, respectively. The BT model classified the pilot areas in average 51 minutes for each pilot area. The BT model was 48% faster than those of SVM in classifying pilot areas.
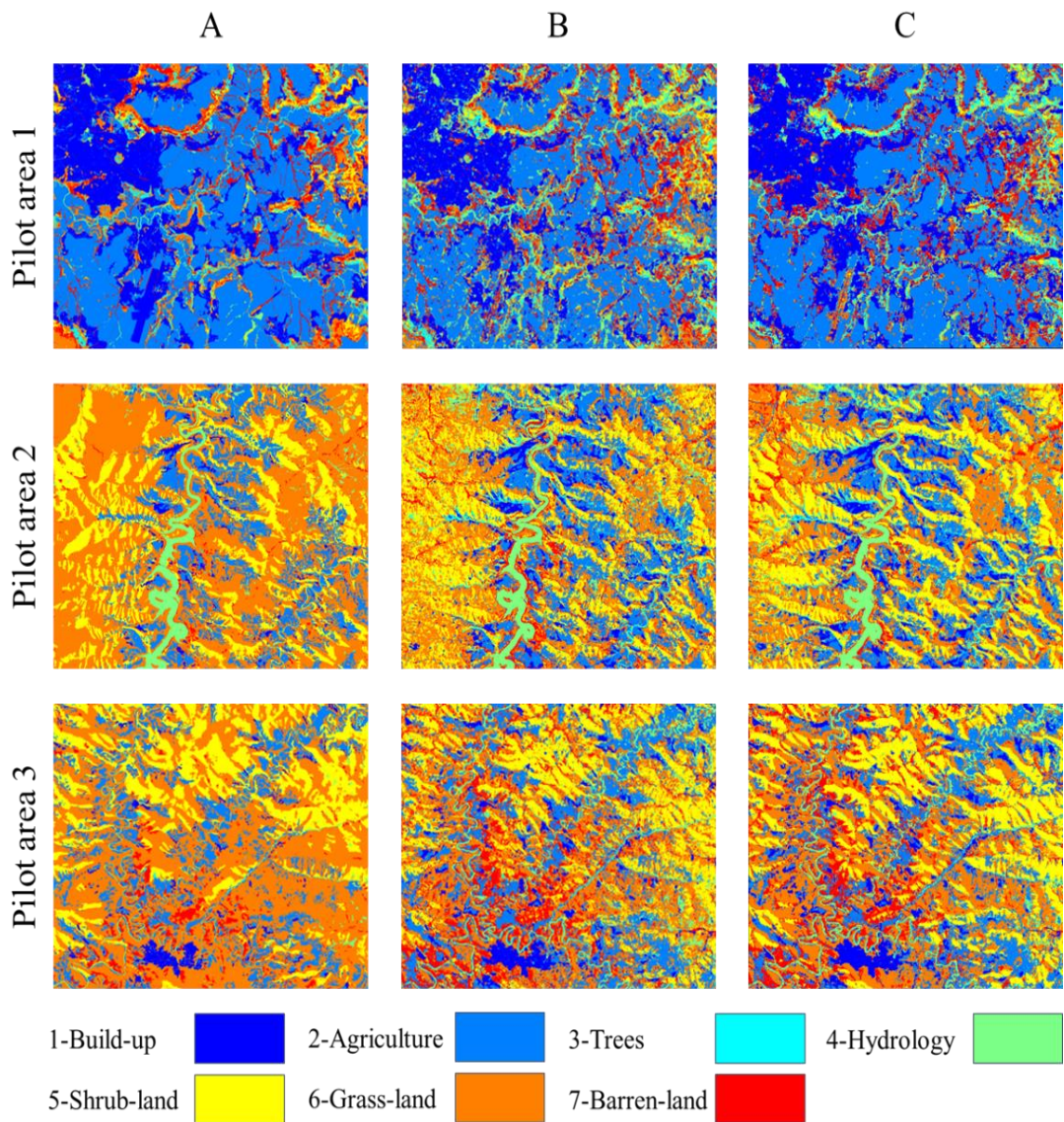
Figure 5-8 Comparison between the pilot areas 1, 2 and 3 for A: FAO land cover of Lesotho (FAO LCL), B: Results of trained SVM model and C: Results of BT model.

5.6 Discussion

5.1. Google Earth Engine as a Cloud Base Remote Sensing Platform

GEE was able to calculate the feature vector for the entire territory of Lesotho. Computational power of GEE enables feature extraction at the national level; however, it was out of computation-time in some circumstances. It could be due to high dimensional matrix operations for PCA analysis and GLCM passing the system limits on allowed processing time for each request (

Gorelick et. al, 2017). This problem can be avoided by extracting the features at a local level instead of a national level, once the initial parameters for PCA (i.e. the mean and covariance of pixel values for the selected bands) were extracted for the country's territory. GEE allows processing of data through its APIs in different ways. However, this study selected the JavaScript API and GEE code editor as the main data processing tool. The goal was to minimize the cost and increase the time efficiency of GEE in land cover mapping. Meanwhile, the processing of generated feature vectors required to be partly executed on cloud machine (GEE), then exported and partly processed on the local machine (PC), since such processing method was causing the GEE code editor to run out of computation-time. This process of exporting and processing the feature vector on a local/cloud machine acts as the main bottleneck of the algorithm's pipeline. Further efforts are required to overcome this problem by using new GEE features in the future.

5.2. The Effect of Spectral and Spatial Features on Accuracy Performance

Image segmentation has traditionally been applied to single-date satellite images (Desclée et. al, 2006). Several studies report the advantages of satellite time series segmentation, such as automated detection of agricultural fields (Davis et. al, 2006), better and faster forest change analysis (Desclée et. al, 2006), robustness against shadowing and registration errors (Desclée et. al, 2006; Mäkelä et. al, 2001), reduced salt-and-pepper effect apparent in per-pixel classifications (Matton et. al, 2015), and the segmentation of multi-temporal images for cropland mapping (Belgiu et. al, 2018). In this study, we trained the ML models based on the mean of a two years' period in order to allow the ML models to learn the spectral and temporal features of different land cover classes. Therefore, the proposed methodology successfully takes advantage of the power of multi-temporal satellite time series, but still has the flexibility to deliver accurate land cover products on an annual base period for land cover and land use change applications.

Nyland et al. and Xiong et al. used GEE for land cover mapping by using fully spectral features of available satellite time series (Nyland et. al, 2018; Xiong et. al, 2017). However, using a times series introduces a level of confusion to the final land cover products since the land cover is subject to change during the time, i.e. a landscape could vary from forest to agriculture in due time because of agricultural development (Inglada et. al, 2017). In this study, we attempted to overcome this problem by using the GEE's potentials in extracting texture features of land surfaces in addition to spectral features for the first time. Introducing spatial (texture) features as a descriptor for land cover mapping enabled our proposed method to perform on an annual base process. It means the method needs the data of one annual period to be able to predict land cover classes. The experiment showed that in the case of Lesotho the same method performs with 40%

and 51% overall accuracy by utilizing only spectral or spatial features, respectively. On the other hand, when utilizing the GLCM in a feature vector, it is important to select the spatial distance parameter in the GLCM offset relative to the resolution of the satellite images. Therefore, the effect of GLCM parameters in land cover mapping applications ask for more investigation in future studies.

5.3 The Inter-Class Confusion Rates

Inglada et al. reported high confusion between continuous and discontinuous urban fabrics. Our study outperformed their results with 81% accuracy for the BT model, even though in our study, we merged all continuous and discontinuous urban fabrics into the built-up class (Inglada et. al, 2017). The confusion between the built-up, agriculture, and grass-land classes happened in two major situations: first, in areas where the built-up class contains large abandoned land between buildings and second in areas where small agricultural activities occur in cites or rural areas.

Moreover, agricultural lands were selected as grassland in areas with higher soil moisture. They are mainly consisting of wet land like areas near springs or on slopes or at the bottom of wide valleys where vegetation growth is more continuous and in similar temporal patterns with natural vegetation (Belgiu et. al, 2018). Also, confusion between agriculture and grass-land was common in rainfed and sloping lands where natural vegetation and crops have similar growth periods. These results also can be confirmed based on the Table 5-1 definition of each class.

Shrubs (0.5~1.5 m high) and trees (1.5 m~ high) normally have very similar growth patterns and spatial distribution of these two classes are similar. This will cause almost identical spectral and spatial characteristics of trees and shrubs. In addition, shrub-lands in FAO LCL data are defined as open shrubs with natural vegetation in between, which could lead to confusion between shrub-land and grass-land classes in which both contains natural vegetation. Moreover, the two classes are containing bare rocks or areas without any vegetation and therefore were classified as barren-land.

The earth surface geometry influences the quality of segmentation (Yan et. al, 2014). Meanwhile, woody shrubs naturally grow on surfaces toward the south while natural vegetation is more adopted to areas with shorter daylight time (Daubenmire et. al, 1980). Therefore, the surface geometry typically affects the type of land cover in each area. As an example, this effect could be seen in the final map product (Figure 5-8) where the shrub-lands are located mostly on areas sloping toward the south. The comparison between results, FAO LCL, and visual inspections showed that the study results are more compatible with pilot areas' topology and geometry.

The proposed methodology was able to provide an accurate land cover of urban and agriculture lands. These land cover products are based on continuously updating satellite data with an annual cycle. The annually based product will allow the authorities in the country to monitor urban and agriculture lands in line with the indicator 2.4.1 of SDGs to measuring the proportion of agricultural area within productive and sustainable agriculture (UN, 2019).

5.7 Conclusions

In this paper, an operational land cover mapping methodology is presented based on available free access Sentinel-2 data and the GEE cloud processing platform. By utilizing a combination of GEE and the FAO LCL data, our main objective was to measure the performance of machine learning models and classify the land surface by using high dimensional feature space. The validation with the proposed method and satellite imagery revealed an accuracy of 62 and 67%, respectively, for the important land cover categories such as 'urban' and 'agricultural lands'. Moreover, the trained models provided a precise land cover tool for annual land cover and land use change comparison applications. The models were trained, validated and tested on a simple laptop PC to demonstrate the cost and time efficiency of the methodology. Such methodology opens doors for in field processing and mapping land cover for future studies.

The study showed the GEE potentials in land cover mapping by processing the feature vector at a national level for Lesotho in a short time. Further investigation is required to unveil the ability of GEE in applying new ML techniques, i.e. deep learning models, on geospatial data. The current GEE pipeline for processing the available data on GEE through the Python or JavaScript APIs requires exporting large volumes of data to cloud or local storage as well as running the ML models on a cloud machine. These processes are time consuming and require extra funds for cloud processing and cloud storage.

Knowing the scale of the country, moderate resolution of the images, and natural changes and errors in FAO LCL as the base map, the proposed methodology achieved accuracy that could be considered satisfactory. Therefore, it is hoped that the individuals and organizations involved in rural and urban development strategies, as well as forest and land conservation at regional and sub-regional scales, can utilize the presented models for developing countries as an alternative to traditional higher-cost GIS methods.

5.8 Application of Deep Learning and Open Access Geospatial Cloud Processing Platforms in Land Cover Mapping

5.8.1 Introduction

Inspired by the learning processing of human beings, artificial neural networks (ANNs) employ a general structure of connected units to learn feature representation. Deep learning models, or deep ANNs with more than two hidden layers, provide sufficient model complexity to learn feature representations from data (LeCun et al. 1995). In this section we study the abilities and potentials of Artificial Neural Networks (ANN) for the land cover classification purposes. We introduced two DNN and tried to utilize the data from previous sections in chapter 5 in order to improve the accuracy and the reliability of the models.

5.8.2 study area and data

5.8.2.1 Study Area

in continue of the previous sections this utilizes the same study area. The study has been carried out over territory of the Kingdom of Lesotho in Africa.

5.8.2.2 Data

FAO Land Cover of Lesotho (FAO LCL) was developed in 2016 using commercial satellite images and conventional GIS methods. In this study, the FAO LCL were modified and used as base line for training and validation data. First, the Sentinel-2 satellite images were processed to form a single image for whole the country's territory using GEE. The produced image consists of 115 bands including 24 spectral (Normalized Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI)) and 90 bands from Gray Level Co-occurrence Matrices texture descriptors (LeCun et al. 1995).

5.8.3 methodology

Two ANN architecture were used in this study. Both ANN were used for purpose of classification task. The main goal was to compare the potential of ANNs in classifying preprocessed signal shaped data and compare the results with other machine learning methods (i.e. support vector machine and decision trees) performance which were investigated in previous studies.

5.8.3.1 *Dense Neural Network*

The Dese network consists of input and output layers with 4 fully connected hidden layers. Hidden layers have rectified linear unit (Relu) activation function and the SoftMax activation was applied in output layer.

5.8.3.2 *Convolutional Neural Network*

The CNN in this study was consist of a reshape layer as input layer 4 one dimensional convolution layers with 100 nodes. Each pair of convolutional layers were followed by a max-pooling layer. At the end a fully connected layer was used as output layer with SoftMax activation function.

5.8.4 results and discussion

The result showed the potential of ANNs in LC mapping and visualization application. The DNN achieved high accuracy in training and testing steps with 97% accuracy. 1D CNN was not able to achieve a better accuracy result than DNN, SVM, or RF. In CNN convolutional layers are extracting features and dense layers classify them. Therefore, lower accuracy in 1D CNN might be the result of applying convolutional layers on signal like image pixel in which the pixel value by itself is a set of extracted features.

The study will apply the 1D and 2D CNN on raw satellite images in future work to improve the accuracy. Also, the DNN will be trained with a larger dataset to prevent overfitting in future work.

Supplementary Materials: Programming codes and supplementary materials are available at: https://github.com/coderdud/Supp_materials_fao

5.9 References

A database for remote sensing indices. Available online: https://www.indexdatabase.de/db/s-single.php?id=96 (accessed on 2019-05-23).

Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote sensing of environment 2018, 204, 509-523.

Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classiï¬ ers. In Proceedings of Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory; pp. 144-152.

Breiman, L. Bagging predictors. Univ. California Technical Report No. 421. 1994.

Campbell, J.B. Introduction to Remote Sensing, Virginia Polytechnic Institute and State University. The Guildford Press, NY: 1996.

Cardille, J.A.; Fortin, J.A. Bayesian updating of land-cover estimates in a data-rich environment. Remote Sensing of Environment 2016, 186, 234-249.

Coltin, B.; McMichael, S.; Smith, T.; Fong, T. Automatic boosted flood mapping from satellite data. International Journal of Remote Sensing 2016, 37, 993-1015.

Conners, R.W.; Trivedi, M.M.; Harlow, C.A. Segmentation of a high-resolution urban scene using texture operators. Computer vision, graphics, and image processing 1984, 25, 273-310.

Daubenmire, R. Mountain topography and vegetation patterns. Northwest Science 1980, 54, 146-152.

Davis, J.; Goadrich, M. The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd international conference on Machine learning; pp. 233-240.

Desclée, B.; Bogaert, P.; Defourny, P. Forest change detection by statistical object-based method. Remote Sensing of Environment 2006, 102, 1-11.

Dong, J.; Xiao, X.; Menarguez, M.A.; Zhang, G.; Qin, Y.; Thau, D.; Biradar, C.; Moore III, B. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. Remote sensing of environment 2016, 185, 142-154.

E Nyland, K.; E Gunn, G.; I Shiklomanov, N.; N Engstrom, R.; A Streletskiy, D. Land Cover Change in the Lower Yenisei River Using Dense Stacking of Landsat Imagery in Google Earth Engine. Remote Sensing 2018, 10, 1226.

European-space-agency. Copernicus Data Access Policy. Availabe online: https://www.copernicus.eu/en/about-copernicus/international-cooperation (accessed on 2019-05-23).

Fischer, G.; Nachtergaele, F.O.; Prieler, S.; Teixeira, E.; TÃ³th, G.; van Velthuizen, H.; Verelst, L.; Wiberg, D. Global Agro-ecological Zones (GAEZ v3. 0)-Model Documentation. 2012.

GÃ³mez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. ISPRS Journal of Photogrammetry and Remote Sensing 2016, 116, 55-72.

Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment 2017, 202, 18-27.

Hachigonta, S.; Nelson, G.C.; Thomas, T.S.; Sibanda, L.M. Southern African agriculture and climate change: a comprehensive analysis; Intl Food Policy Res Inst: 2013; Vol. 3.

Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.; Tyukavina, A.; Thau, D.; Stehman, S.; Goetz, S.; Loveland, T.R. High-resolution global maps of 21st-century forest cover change. science 2013, 342, 850-853.

Hao, P.; Zhan, Y.; Wang, L.; Niu, Z.; Shakir, M. Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA. Remote Sensing 2015, 7, 5347-5369.

Haralick, R.M.; Shanmugam, K. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics 1973, 610-621.

Huang, H.; Chen, Y.; Clinton, N.; Wang, J.; Wang, X.; Liu, C.; Gong, P.; Yang, J.; Bai, Y.; Zheng, Y. Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine. Remote Sensing of Environment 2017, 202, 166-176.

Inglada, J.; Vincent, A.; Arias, M.; Tardy, B.; Morin, D.; Rodes, I. Operational high-resolution land cover map production at the country scale using satellite image time series. Remote Sensing 2017, 9, 95.

Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. Remote sensing of Environment 2008, 112, 3833-3845.

Kamwi, J.M.; Chirwa, P.W.; Manda, S.O.; Graz, P.F.; Kätsch, C. Livelihoods, land use and land cover change in the Zambezi Region, Namibia. Population and Environment 2015, 37, 207-230.

Latham, J.; Cumani, R.; Rosati, I.; Bloise, M. Global land cover share (GLC-SHARE) database beta-release version 1.0-2014. FAO: Rome, Italy 2014.

LeCun, Yann, and Yoshua Bengio. "Convolutional networks for images, speech, and time series." The handbook of brain theory and neural networks 3361.10 (1995): 1995.

Lobell, D.B.; Thau, D.; Seifert, C.; Engle, E.; Little, B. A scalable satellite-based crop yield mapper. Remote Sensing of Environment 2015, 164, 324-333.

Loveland, T.R.; Dwyer, J.L. Landsat: Building a strong future. Remote Sensing of Environment 2012, 122, 22-29.

Mäkelä, H.; Pekkarinen, A. Estimation of timber volume at the sample plot level by means of image segmentation and Landsat TM imagery. Remote Sensing of Environment 2001, 77, 66-75.

Mardani, M.; Fujii, Y.; Saito, T. Detection and Mapping of Hairy Vetch in Images Obtained by UAVs, In Proceedings of the International Workshop on Image Electronics and Visual Computing, Da Nang, Vietnam, 2017.

Matton, N.; Canto, G.; Waldner, F.; Valero, S.; Morin, D.; Inglada, J.; Arias, M.; Bontemps, S.; Koetz, B.; Defourny, P. An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series. Remote Sensing 2015, 7, 13208-13232.

Mokarram, M.; Sathyamoorthy, D. Modeling the relationship between elevation, aspect and spatial distribution of vegetation in the Darab Mountain, Iran using remote sensing data. Modeling Earth Systems and Environment 2015, 1, 30.

Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 2011, 66, 247-259.

Mumby, P.; Green, E.; Edwards, A.; Clark, C. The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. Journal of Environmental Management 1999, 55, 157-166.

Myroslava Lesiv, S.F., Ian McCallum, Ian McCallum, Martin Herold, Jean-Francois Pekel, Marcel Buchhorn, Bruno Smets,. Evaluation of ESA CCI prototype land cover map at 20m; International Institute for Applied Systems Analysis: 2017.

Nations, F.a.A.O.o.t.U. ISO19144-2: Geographic information - Classification systems â€" Part 2: Land Cover Meta Language (LCML). International Organization for Standardization (ISO): 2012.

Nations, U. Resolution adopted by the General Assembly on 25 September 2015. In Transforming Our World: The 2030 Agenda for Sustainable Development, United Nations New York, NY: 2015.

Nations, U. World population prospects: The 2015 revision. United Nations Econ Soc Aff 2015, 33, 1-66.

Patel, N.N.; Angiuli, E.; Gamba, P.; Gaughan, A.; Lisini, G.; Stevens, F.R.; Tatem, A.J.; Trianni, G. Multitemporal settlement and population mapping from Landsat using Google Earth Engine. International Journal of Applied Earth Observation and Geoinformation 2015, 35, 199-208.

Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. Nature 2016, 540, 418.

Ridder, R.M. Global forest resources assessment 2010: options and recommendations for a global remote sensing survey of forests. FAO For. Resour. Assess. Programme Work. Pap 2007, 141.

Rouse Jr, J.W.; Haas, R.; Schell, J.; Deering, D. Monitoring vegetation systems in the Great Plains with ERTS. 1974.

Shah, V.P.; Younan, N.H.; King, R.L. An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets. IEEE transactions on geoscience and remote sensing 2008, 46, 1323-1335.

Sidhu, N.; Pebesma, E.; Câmara, G. Using Google Earth Engine to detect land cover change: Singapore as a use case. European Journal of Remote Sensing 2018, 51, 486-500.

Stibig, H.J.; Belward, A.; Roy, P.; Rosalinaâ€¢ Wasrin, U.; Agrawal, S.; Joshi, P.; Hildanus; Beuchle, R.; Fritz, S.; Mubareka, S. A landâ€¢ cover map for South and Southeast Asia derived from SPOTâ€¢ VEGETATION data. Journal of biogeography 2007, 34, 625-637.

The United Nations, FAO. Land Cover Atlas of Lesotho. FAO Headquarters, Rome, 2017.

UN FAO. GeoNetwork opensource portal to spatial data and information. Available online: http://www.fao.org/geonetwork/srv/en/main.home (accessed on 2019-07-30).

United-Nations. Sustainable Development Goals indicators. Availabe online: https://unstats.un.org/sdgs/metadata/files/Metadata-02-04-01.pdf (accessed on 2019-05-23).

Vancutsem, C.; Marinho, E.; Kayitakire, F.; See, L.; Fritz, S. Harmonizing and combining existing land cover/land use datasets for cropland area monitoring at the African continental scale. Remote Sensing 2013, 5, 19-41.

Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. Chemometrics and intelligent laboratory systems 1987, 2, 37-52.

Woodcock, C.E.; Allen, R.; Anderson, M.; Belward, A.; Bindschadler, R.; Cohen, W.; Gao, F.; Goward, S.N.; Helder, D.; Helmer, E. Free access to Landsat imagery. Science 2008, 320, 1011-1011.

Xiong, J.; Thenkabail, P.S.; Gumma, M.K.; Teluguntla, P.; Poehnelt, J.; Congalton, R.G.; Yadav, K.; Thau, D. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS Journal of Photogrammetry and Remote Sensing 2017, 126, 225-244.

Xue, J.; Su, B. Significant remote sensing vegetation indices: a review of developments and applications. Journal of Sensors 2017, 2017.

Yan, L.; Roy, D. Automated crop field extraction from multi-temporal Web Enabled Landsat Data. Remote Sensing of Environment 2014, 144, 42-64.

Zhang, Q.; Li, B.; Thau, D.; Moore, R. Building a better urban picture: Combining day and night remote sensing imagery. Remote Sensing 2015, 7, 11887-11913.

Chapter 6: CONCLUSION AND FUTURE OF RESEARCH

6.1 Conclusion

We conclude the study in this section. In this thesis, we have discussed unexplored problems and challenges on aerial image analysis for agro-environmental studies.

In chapter 2, we studied the feasibility of role base image analysis method to detect the hairy vetch in UAV aerial imageries. Also, we compared the texture properties of different plants in study area. By comparing the results of color filtering and texture analysis we can realize the potential of proposed method to implement an automatic algorithm to detect and map the hairy vetch (or other plants) in aerial imagery.

In chapter 3, we take look at challenges and aspects of UAV aerial image mosaicking. We try to understand the important points and main problems in stitching aerial images form our study area. The results showed that the low altitude of flight can significantly affect the quality of stitching algorithms and therefore applying more sophisticated methodology that take advantage of camera position system is required. Therefore, we moved forward with applying available software.

In chapter 4, we proposed a machine learning based methodology for detecting 6 different land surface types including the two target plants. The methodology resulted in 94% accuracy in detecting the land cover types and was tested on the study areas for 2 sequenced years. The results reveal that the color features can detect the hair vetch with higher quality however it is not successful in detect and discriminate plants in both years and the texture features have higher ability in detecting the plant types in both years.

Chapter 5 is dedicated to applying the idea of the propose methodology in previous chapter on satellite driven aerial images. We implemented the methodology and modified it in order to provide the landcover products of the Lesotho in southern Africa. The method was taking advantage of open access satellite data, open access geospatial analysis platform, Google Earth Engine, and the machine learning technology to provide accurate landcover products in reasonable time in national level of Lesotho.

Finally, in chapter 6 we move forward to more advance AI techniques for landcover mapping. In this study we proposed a neural network in order to increase the accuracy of landcover products in previous chapter. A fully connected dense network was trained and tested on the same set of features. The result opens the door for future studies in this area.

6.2 Future of research

The results of chapter 2,3, and 4 showed that while the methodology is successful in discriminate the type of the plants in aerial imagery the problem or the dilemma of ground truth data accuracy occurs in this study too. In machine learning the accuracy and the quantity of labeled data or ground truth data has direct effect on which path should be selected to solve the problem, weather it should be a feature learning methodology or a feature engineering one. Therefore, this thesis suggests a deep and comprehensive research on a methodology that can automatically populates the ground truth data.

The chapter 5 and 6 clear the fact that the age of role-based methodologies in the field of remote sensing is covered and the future researches must focus on AI methodologies to obtain more reliable results. The ability of our methodology in successfully performing on both UAV and the satellite imagery indicates that although the targets are slightly different in those data the effect of fine grain texture analysis is the key in classifying the type of land surface and spectral analysis alone is not adequate for the state of art methodologies. Therefore we suggest that the future studies develop methodologies that are capable of learning or extracting highly complex textures as well as spectral characteristics of data.