#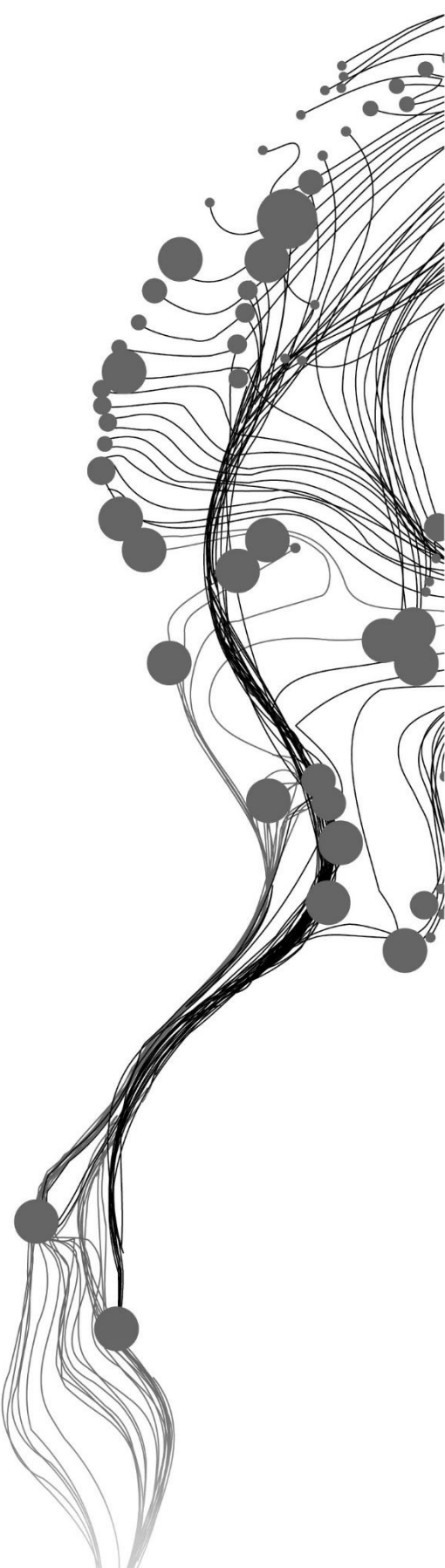 EARLY DETECTION OF *HERACLEUM MANTEGAZZIANUM* (GIANT HOGWEED) BASED ON LEAF SPECTRAL CHARACTERISTICS FROM UAV IMAGES USING SVM AND OBIA TECHNIQUES

SHEEBA LAWRENCE

SUPERVISORS:

dr. ir. W. Bijker

dr. V.A. Tolpekin

EARLY DETECTION OF *HERACLEUM MANTEGAZZIANUM* (GIANT HOGWEED) BASED ON LEAF SPECTRAL CHARACTERISTICS FROM UAV IMAGES USING SVM AND OBIA TECHNIQUES

SHEEBA LAWRENCE

Enschede, The Netherlands, February, 2019

ABSTRACT

Plant invasions pose a great threat to biodiversity, human health and economies. There have been many efforts taken worldwide to eradicate invasive species. Remote Sensing has been successfully used in the past two decades as a tool for monitoring invasive plants. *Heracleum mantegazzianum* also known as Giant Hogweed, is an invasive species that was introduced to Europe in the 19th century as an ornamental plant. Previous studies detect *Heracleum mantegazzianum* during the flowering and early fruiting season using its distinct giant umbels, which is from June to August. This limits the period of data acquisition. *Heracleum mantegazzianum* depends entirely on seed dispersal for its spread, which takes place after the fruiting season from August to October. Detecting the plant in flowering phase also limits the time window for eradication measures. Hence, in this research, both pixel-based approach (Support Vector Machine or SVM) and object-based approach (Object-Based image Analysis or OBIA) on high-resolution data acquired form Unmanned Aerial Vehicles (UAV) were carried out to detect the plant by using its leaves. SVM uses only spectral information while OBIA uses spectral and spatial information. Both the methods performed well, but SVM was time effective even though it produced salt and pepper effect in the resulting map. The accuracy of OBIA was lesser in low flying heights due to the effect of shadow, and in that case, a targeted analysis focusing on the shadow region was carried out to increase the accuracy. The high overall accuracy, between 91% and 97%, which was achieved for both the methods, indicates that early detection is possible by using the spectral information of the leaves of the plant. As a possible third approach, this study also revealed that UAV images could be used as an input for deep learning based mobile application (ObsIdentify for plants). ObsIdentify for plants mobile application was developed to identify plants in the field based on their photo and uses a database with field photos as training data.

Keywords: *Heracleum mantegazzianum,* Giant Hogweed, Unmanned Aerial Vehicle, Support Vector Machine, Object-Based Image Analysis, ObsIdentify

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Motivation and Problem Statement

Heracleum mantegazzianum also known as Giant Hogweed is an invasive plant species, which was introduced as an ornamental plant in the 19th century to Europe from the western Caucasus Mountains (Otte & Franke, 1998). According to the Government of Canada (2004), "Invasive species are plants, animals, and micro-organisms introduced by human action outside their natural past or present distribution whose introduction or spread threatens the environment, the economy, or society, including human health". The European Commission has listed a number of invasive alien species of concern called the Union list, which are a threat to the natural ecosystem. Alien species that turn into invasive are one of the main reasons for the loss of biodiversity across the planet (Early et al., 2016). The Union list consists of 23 invasive plants, and Giant Hogweed is one among them (European Commission, 2015). Giant Hogweed is phototoxic, which means when the contacted skin is exposed to sunlight it can cause severe skin inflammations called phytophotodermatitis. The symptoms range from painful blisters to full chemical burn (Moenickes & Thiele, 2013). The blisters leave a black or purplish scar, which can last for several years. Eye exposure to the plant sap can also lead to temporary blindness (ISAP, 2018). Giant Hogweed reproduces only by seeds (Perglová et al., 2007) and produces tens of thousands of seeds per year which are dispersed to new places through water and end up spreading miles away from the mother plant (Jakuboski, 2011). The Giant Hogweed colonies are known to extend up to 2 km along the river corridor and sometimes even 1 ha of adjacent habitat (Caffrey, 2001). The plants flower mostly in their third or fourth year (Caffrey, 1999). The flowering period of Giant Hogweed is from late June to July (Moravcová et al., 2005) and the seed dispersal starts from August and lasts until October (Nielsen et al., 2005). Giant Hogweed solely depends on seed dispersal for its spread and hence, it is logical to remove the plants before the seed dispersal stage to stop the spreading.

The first step of invasive plant control is to identify the species. To manage the invasions, it is necessary to monitor the species, and its spread regularly and effectively. In the Netherlands, the Dutch Botanical Research Foundation called Floron uses volunteers to identify and map the distribution of the invasive plants. The volunteers cover 800-1000 km² every year. Each species has its own set of characteristics which the volunteers use as a guide to identify them in the field. Remote Sensing has been successfully used in large-scale invasive plant monitoring in the last two decades (Huang & Asner, 2009). Images from airborne sensors were used in previous studies in the Czech Republic (Müllerová et al., 2013; Müllerová et al., 2005) for the detection of *Heracleum mantegazzianum*. The Giant Hogweed often grows under trees, along with the roads, waterways or at the edge of the forest. Earlier studies using images from airborne sensors had the limitation of not able to map the plant growing under tree canopies. Unmanned Aerial Vehicles (UAVs) can be an answer to this problem since the cameras of UAVs can be set in a specific angle to capture images of Giant Hogweed which grows under the trees. UAVs are state of the art technology derived from their military ancestors. However, booming civilian applications have prompted wider applications (Casagrande et al., 2018). Also, the advancement of imaging sensors in recent years has paved the way for ultra-high resolution monitoring with less than 10 cm of ground sampling distance (GSD) (Gao et al., 2018). Moreover, UAVs also allow for flexibility in flight schedules which makes them an easy option. Identifying the plant through volunteers is a labour intensive and tedious process. With successful detection of the species, the spread can be mapped, and the infected sites can be easily monitored using the UAV images. If the location of the alien plant can be detected accurately, suitable controlling measures which are sustainable to the environment as well as regulated by the European Union (European Union, 2012) can be undertaken.

Giant Hogweed produces big white umbrella-like composite flowers called umbels, which are up to 80 cm wide at their mature stage (Moravcová et al., 2005). The huge size of the umbels, which are distinct from the nearby vegetation, was used to detect Giant Hogweed in various studies. Huang and Asner (2009) have reviewed the application of Remote Sensing to various invasive plants from spectral, spatial and temporal perspectives. As a part of their research, the authors reviewed a study by Müllerová et al., (2005) on using aerial photos as a tool for assessing the regional dynamics of *Heracleum mantegazzianum*. According to the review, in panchromatic and colour-infrared (CIR) aerial photographs the brightness of *Heracleum mantegazzianum* was much higher than the nearby vegetation not only during the flowering but also in early fruiting season, which is from June to August. This brightness was attributed to the distinct structure of fruiting umbels. Hence, this distinct spectral characteristic of umbels along with the size of the umbels was the focus of Giant Hogweed detection. This limited the data acquisition to coincide with flowering and early fruiting period (Müllerová et al., 2013) in almost all studies related to Giant Hogweed. The flowering period is relatively concentrated on a short period for Giant Hogweed while other species can have flowering period spread over longer periods (Perglová et al., 2006). However, using the leaves instead of umbels can provide the user with the flexibility of the timing of data acquisition. In addition, it can also provide a buffer window for eradication measures. However, this may be possible only if the Giant Hogweed leaves are spectrally distinct from the nearby vegetation and background (Jones et al., 2011; Maheu-Giroux & De Blois, 2005; Müllerová et al., 2005). In addition, the data should also provide enough spectral and spatial information for accurate detection using only Giant Hogweed's leaves. The existing studies fail to explain whether Giant Hogweed leaves have distinct spectral characteristics for a successful detection from its neighbouring vegetation. This may be one of the reasons for not using the leaves for detection in earlier studies. However, investigating the plant's presence using a number of combination of spectral bands can help to identify a solution for the issue. This can be possible with UAVs as modern UAVs provide the flexibility to interchange sensors according to the user requirements (Mufalli et al., 2012). Perhaps the most serious issue would be whether the spectral information of the leaves would be enough for early detection. In the event of failure using only the spectral information, other methods should be investigated, which adds spatial context along with spectral context for the detection process.

Conventional or pixel-based classification approaches (Bayesian or artificial neural networks) uses spectral information of the remotely sensed data. However, using high-resolution data like UAV data causes an increase in between-class spectral confusion and within-class spectral variation (Mathieu et al., 2007). Support Vector Machines (SVM) are machine learning based nonparametric classifiers which have been used in various Remote Sensing classification studies, and they are reported to produce more accurate results than traditional methods (Huang et al., 2002; Kavzoglu & Colkesen, 2009). However, if the spectral information is not enough for successful detection, an object-based approach (Object-Based Image Analysis or OBIA) can be implemented which uses spectral, spatial and temporal information (Blaschke, 2010). Also, OBIA is suitable for high-resolution imagery and provides a solution for salt and pepper effect from traditional pixel-based approaches (Whiteside et al., 2011).

In addition, Waarneming.nl, a website which shows the different sightings of plant and animal species of the Netherlands, has launched a mobile application called 'ObsIdentify or plants' along with the Observation International organisation. The app is an initiative by the organisation which helps to recognise various plants and insects through pictures taken by mobile phones. The app uses deep learning to identify the species. It will be an interesting study to see whether a deep learning based app which was trained using mobile pictures taken in the field, can recognise Giant Hogweed using UAV images. The results can be used for further studies.

1.2. Research Identification

1.2.1. Research Objectives

The primary objective of this research is to see whether early detection of Giant Hogweed is possible by using the spectral information of the leaves from UAV data, acquired using a multispectral sensor that provides a choice of spectral bands. This can be further divided into the following sub-objectives

1. To see whether early detection of Giant Hogweed is possible by using its spectral information.
2. To assess whether adding spatial information to spectral information can increase the accuracy of detection.
3. To check whether UAV images can be used as an input for a deep learning based mobile application.
4. To examine whether different altitudes of the UAV flight or shadow can influence the accuracy of the detection for all the methods.

1.2.2. Research Questions

The following are the research questions that need to be answered for the sub-objectives above.

Sub Objective 1:

- Is early detection of Giant Hogweed possible by using only the spectral information of leaves?

- What is the accuracy of the SVM method?

Sub Objective 2:

- Which extra parameters (shape, size and or texture) are used in addition to spectral information? Did it increase the accuracy of the detection in the OBIA method?

- What is the accuracy of OBIA method?

Sub Objective 3:

- Can UAV images be used as an input for a deep learning based species recognition mobile application (ObsIdentify for plants)?

Sub Objective 4:

- How does the altitude of the flight influence the accuracy of detection of Giant Hogweed in SVM?

- How does the altitude of the flight influence the accuracy of detection of Giant Hogweed in OBIA?

- How does the altitude of the flight influence the accuracy of detection of Giant Hogweed in the mobile application?

- Does the shadow affect the accuracy of detection in OBIA?

- Does the shadow affect the accuracy of detection in SVM?

- Between the two methods, which is more efficient in terms of accuracy and computation time?

1.2.3. Innovation aimed at

Giant Hogweed is usually detected using its umbels by Remote Sensing methods. However, detecting the plant at its flowering stage limits the data acquisition period. In addition, the invasive control cannot be fully effective because the buffer period between flowering and seed dispersal stage is very less. According to the Manual guideline for the management and control of Giant Hogweed (Nielsen et al., 2005) the plant must be eradicated in its early stage to prevent the spread of the species efficiently. Hence, this research aims at

detecting the plants at an early stage by implementing algorithms specifically designed to use the spectral information of the leaves of the plant.

1.3. Thesis structure

This thesis consists of seven chapters. Chapter 1 briefly discusses the motivation and problem that will be addressed in this research. Chapter 2 reviews the previous studies related to this research. Chapter 3 describes the specification of the data used in this research. Chapter 4 explains the methodology of performing SVM and OBIA while chapter 5 explains the results obtained from SVM and OBIA along with the accuracy assessment of both the methods. Chapter 6 discusses the accuracy assessment, strength and weakness of both methods. Chapter 7 gives a short conclusion on this research along with the future recommendations.

2. LITERATURE REVIEW

This chapter begins with a brief review of related work in line with this research and is followed by the explanation of using UAV over other airborne sensors continued by reviews of the two methods, which was used in this research namely: Object-Based Image Analysis (OBIA) and Support Vector Machines (SVM).

2.1. Related work

There have been only a handful of studies conducted in the past for the detection of *Heracleum mantegazzianum* using Remote Sensing. In one of the studies conducted by Müllerová et al. (2005), aerial photos have been used as a source of data for assessing the regional dynamics of *Heracleum mantegazzianum*. The study area contains ten heavily invaded sites in the Czech Republic. The aerial photographs dated between 1947 (before the invasion started) and 2000 were analysed. The images were obtained during the flowering, and early fruiting stage (June – August). With the help of huge umbels, the plants were spotted in the imagery and digitised manually. The rate of spread of the invasion over 50 years was calculated by performing ANCOVA (Analysis of Covariance) over the digitised data. The study used aerial images only from June – August to detect the plant with the help of the giant umbels.

In another study, the same author tries to review the various effects of data resolution and classification methods in the detection of *Heracleum mantegazzianum* (Müllerová et al., 2013) over a 50 year invasion period (1962 – 2010). This study uses various types of aerial images (panchromatic, multispectral, colour orthophoto) and satellite images with different spectral and spatial resolution from the period of July to August. This study also focuses on the detection of the plant during its flowering stage. Both, the pixel-based approach (supervised and unsupervised) and the object-based approach were incorporated in the study. In OBIA, multi-resolution segmentation was used to segment the image, and the segmented image objects were classified using rule-based classification with conditions related to spectrum, shape, texture and context. The success of detection was attributed to data acquired at correct phenological stages (i.e. flowering stage), and the detection was mainly based on umbels. The umbels are composed of buds, flowers and ripening seeds (Perglová et al., 2006), which leads to spectral heterogeneity. This was attributed as a reason for the low accuracy of pixel-based classification. The object-based approach reached higher accuracy compared with the pixel-based approach because it employed additional parameters like shape, texture and context. The study also had a limitation of not able to detect Hogweed under the tree canopy. The study emphasises on the period of acquisition of data to coincide with flowering or early fruiting.

Another study by Michez et al., (2016) uses UAV imagery to map riparian zone invasive species, and *Heracleum mantegazzianum* is one among them. The study used only visible and near-infrared orthophotos. The data were acquired during the flowering phase of the plant with a GSD of 0.05 m. OBIA was performed for the segmentation process using eCognition software. The classification was performed using supervised Random Forest algorithms. The study achieved better accuracy than the previous study by Müllerová et al. (2013). The research used manually delineated umbels as reference data for the OBIA analysis.

Another study by Müllerová et al., (2017) explains the importance of timing of monitoring using two invasive plants *Heracleum mantegazzianum* and *Fallopia japonica*. This study also was set up on the fact that Giant Hogweed has a distinct flowering phase and uses both UAV and satellite data. The UAV was used to capture the data in RGB and near-infrared bands. The study reviews both the pixel based, object-based (OBIA) approach and machine learning methods and used data from July and September (2013 – 2016), i.e. the

flowering and post-flowering period. The accuracy of detection was very high during the flowering period (up to 100%), and it dropped to 60% in the post-flowering period. In addition, the high resolution of the UAV images reduced the accuracy of object-based classification results in the study; this was because the classification was based on giant umbels. As the spatial resolution increased, instead of consistent white dots, the umbels were visible as separate flowers, which decreased the classification accuracy.

It can be inferred that all the previous studies used the huge distinct white umbels as a source of detection of the Giant Hogweed. The studies that used OBIA as an approach for the detection also used umbels. The studies gave good accuracy for the detection of the images only from peak flowering periods. The previous studies, which used airborne sensors, had the limitation of not mapping Giant Hogweed growing under trees. It can also be seen that the studies that used UAV as a source of data (Michez et al., 2016; Müllerová et al., 2017) used only the visible and near-infrared bands for the analysis. In addition, using high-resolution data decreased the accuracy of detection when using both pixel-based and object-based approach because of the high spectral variability of umbels.

2.2. Technical Background

2.2.1. Importance of UAV in this study

The identification, distribution and the estimation of the cover of invasive species are traditionally surveyed manually with the help of surveyors (Lillesand et al., 2014). The advancement in satellite remote sensing techniques led to various opportunities to map and monitor invasive species (Royimani et al., 2018). There are many studies in the past, which used Remote Sensing as a tool for the management of invasive plant control. These studies have used various sources of Remote Sensing data typically from satellite or aircraft. Recent studies have been using data from UAV because they are known to provide a high spectral and spatial resolution (Dvořák et al., 2015; Müllerová et al., 2017; Manfreda et al., 2018) and can also be adjusted according to the requirements of the user. The manned airborne platforms, in principle, can give high-resolution imagery. However, they have many disadvantages like operational complexity, safety, logistics and cost. On the other hand, recent advancements in UAVs have created a perfect alternative monitoring platform that provides a range of applications with a relatively small investment. They offer not only high-resolution imagery but also high versatility, adaptability, and flexibility (Manfreda et al., 2018). Also, UAVs are one of the options that can map understory beneath canopies, which is important since *Heracleum mantegazzianum* is known to grow also under tree canopies.

Multispectral sensors such as Parrot Sequoia coupled with UAVs have been used in various vegetation mapping applications such as classification (Ahmed et al., 2017), monitoring (Assmann et al., 2019). Parrot Sequoia provides four separate bands Green, Red, Red Edge, near infrared along with an RGB band. Parrot Sequoia analyses the vitality of plants by capturing the amount of light they absorb and reflect. Hence, this research uses UAV coupled with Parrot Sequoia multispectral sensor for acquiring the data.

2.2.2. SVM and OBIA

This study aims to use the spectral characteristics of Giant Hogweed leaves to separate it from its neighbourhood vegetation for early detection through image classification. Image classification algorithms can be performed using either supervised or unsupervised classification approaches. SVM is a supervised machine learning algorithm which is based on statistical learning theory, used for classification problems (Vapnik, 2006). Supervised classification of remote sensing data can be challenging because most of the methods require a large number of training samples (Mountrakis et al., 2011). SVM is more appealing in the remote sensing field because of its ability to often produce higher classification accuracy by using only small training data sets (Mantero et al., 2004). SVM has been successfully used in many image classification problems for invasive plant studies but mostly using satellite imageries (Calleja et al., 2019; Gavier-Pizarro

et al., 2012). SVM is a pixel-based approach, and the classifier uses the spectral characteristics of individual pixels for classification. Hence, it is used as an approach in this research to detect Giant Hogweed using its spectral characteristics.

Object-Based Image Analysis (OBIA) is selected as the other method to compare with SVM results. It is selected over traditional pixel-based approaches because the pixel-based classification does not yield appropriate results when classifying high-resolution imagery. This is because the internal variability within homogeneous classes increases as the resolution increases which in turn decreases the spatial separability in the spectral data space of the classes. This internal variability leads to reduced accuracies in pixel-based classification such as K-means, Fuzzy C means (Chakraborty et al., 2017). Object-Based Image Analysis groups the pixels together based on their spectral, spatial and temporal similarity (Blaschke, 2010). OBIA segments the image into groups of contiguous pixels called 'image objects' that are further classified according to the spectral variable, size, shape, texture, spatial relationship (Blaschke, 2010; Müllerová et al., 2013). Some studies in the past have used OBIA as an approach to detect Giant Hogweed (Michez et al., 2016; Müllerová et al., 2013) but the studies were based on segmenting the giant umbels.

This work explores the ways for early detection of Giant Hogweed before the flowering stage by using the spectral characteristics of the leaves. This study will also be using UAV as a source of data so that plants growing under the tree canopies can also be detected. The study also aims to use different spectral bands (Green, Red, Red Edge and near-infrared) to see whether the leaves of Giant Hogweed are spectrally distinct from neighbourhood vegetation, rather than only using near infrared and visible bands like the previous studies.

3. DATASET

3.1. Data Acquisition

The study area is a small area close to the village of Lonneker in the Netherlands. The data used in this study were captured using UAV and was acquired on June 28, 2018, with varying altitudes 11 m, 12 m, and 14 m. The sensor used was Parrot Sequoia, a multispectral sensor, which recorded images in RGB, Green, Red, Red Edge and Near-infrared (NIR) bands. The Giant Hogweed was growing along an unpaved road under the trees. There was a maize field on the other side of the road. Figure 3.1 shows all the available bands with the presence of Hogweed being highlighted by the red ellipses.
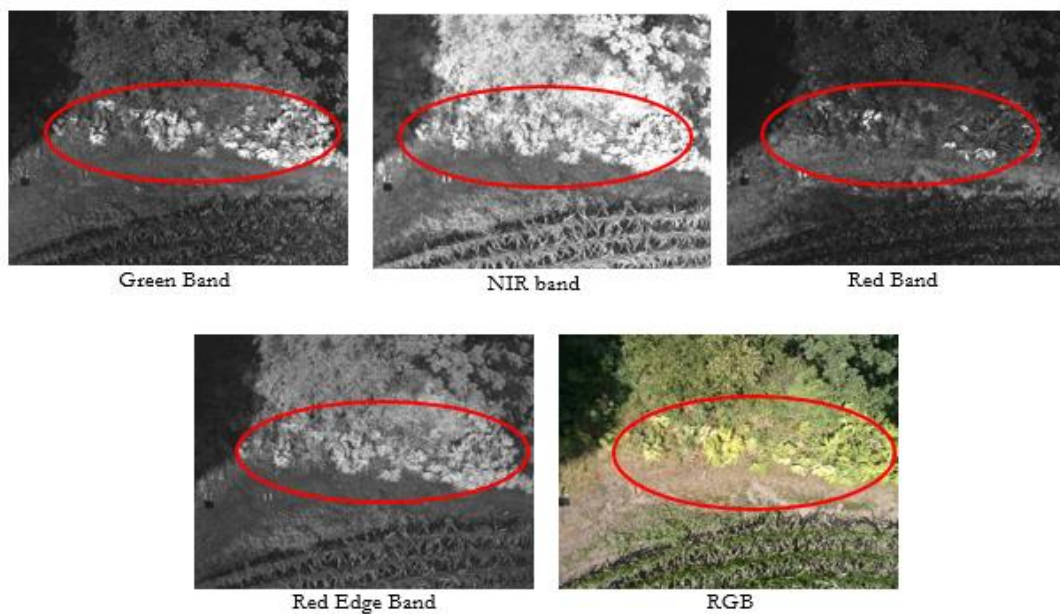
Figure 3.1 Raw UAV image tiles of different bands. The red ellipses show the position of the Giant Hogweed, which is more distinct in Green, Red edge & NIR bands.

It can be seen from Figure 3.1 that Giant Hogweed is spectrally distinct from the nearby vegetation in the Green band. The whole vegetation in the scene is prominent in both NIR and Red Edge bands. In order to get clearer pictures of Giant Hogweed under the trees, the UAV camera was tilted to an angle of 11 degrees. The angle tilt was measured using a protractor as shown in Figure 3.2. The black part was the camera mount, which was connected to the UAV through the white metal piece.

Figure 3.2 Measuring the tilt of the camera mounted in UAV using a protractor. The black part is where the camera was located.

The total number of images acquired in each flying height can be seen in Table 3.1.

Table 3.1 Total number of images for each flying height including all the bands namely; RGB, Red, Red Edge, Green and NIR.

Flying height (in m)	No. of images (including all bands)
11	65
12	255
14	105

The images were taken continuously every few seconds. The Parrot Sequoia sensor comes with a sunshine sensor along with a multispectral sensor, which is used to calibrate the images depending on the different variations of sunlight (MicaSense, 2017). The wavelength and bandwidth of all the bands of the images are tabulated in Table 3.2.

Table 3.2 Wavelength and bandwidth of all the image bands from the Parrot Sequoia multispectral sensor.

Band	Wavelength (in nm)	Bandwidth (in nm)
Green	550	40
Red	660	40
Red Edge	735	10
NIR	790	40

3.2.   Software Used

The pre-processing of the UAV images was done using pix4D mapper, a photogrammetry software. It was also used to create orthomosaics of UAV images. In addition, ERDAS IMAGINE 2018 was used to create layer stacks of different spectral bands of orthomosaics. eCognition Developer 9.3 was used for the object-based approach. eCognition allows the user to export the classified features as raster or vector. SVM was carried out in the RStudio (R Core Team, 2018) using the package e1071 (Meyer et al., 2018). The classification code for SVM was obtained from the EOS department at ITC. The accuracy assessment for OBIA was carried out in ArcMap 10.6. The final output maps were also created using ArcMap 10.6.

4. METHODS

4.1. Overview

The overview of the methodology adopted for this research is shown in the flowchart as seen in Figure 4.1.
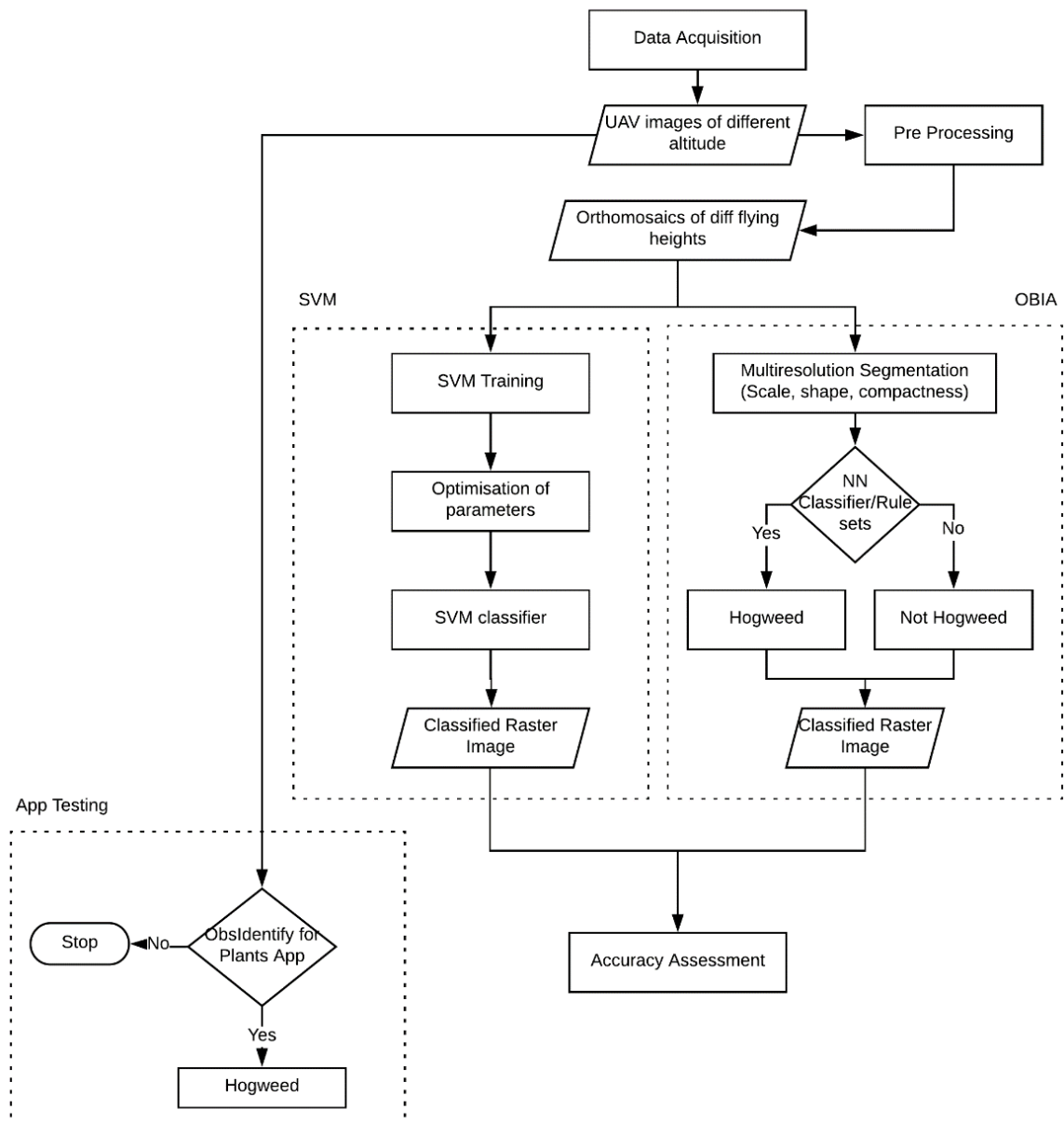
Figure 4.1 Workflow of the methodology adopted for this research.

4.2. Data Pre-Processing

The images acquired from the UAVs are pre-processed first before any further analysis. Pix4D has a dedicated template for Parrot Sequoia multispectral sensor and provides automatic radiometric calibration. The orthomosaics are created from the Digital Surface Model (DSM), which is, in turn, is created from the

point cloud. The pre-processing phase consists of orthorectification of the imagery and mosaicking the individual images of a particular height. The process was repeated for all the flying heights and each band. The final ortho mosaics were used for further analysis.

4.3.   Object-Based Image Analysis

The image analysis workflow can be divided into two steps: segmentation and classification. The software offers a collection of algorithms for both processes, which are tailored for different aspects of image analysis. The main concept is that the important semantic information, which is necessary to interpret an image, is not represented as individual pixels but rather as image objects. Unlike a pixel-based approach which classifies the individual pixel, eCognition classifies the image object as a whole which is output from the segmentation step (Trimble, 2012). The rule sets and applications developed for a specific task can be reused for larger areas which can be used for automating the image analysis process (Trimble, 2015). eCognition allows the user to export the classified features as raster or vector.

4.3.1.   Segmentation

Image segmentation is the process of dividing an image into homogeneous regions of a set of pixels which have similar features like intensity, colour, variance, texture (Tamaki et al., 1999). First, the images are segmented into image objects, which is a group of homogeneous pixels. These image objects will be later classified. Multi-resolution segmentation was selected for the segmentation process, which is a Bottom-up segmentation. Bottom-up segmentation segments the image into homogeneous regions in terms of colour, texture and other image-based criteria. A recognition process is then used to group these regions into image objects (Borenstein & Ullman, 2008). The segmentation process in eCognition is defined by three parameters namely scale, shape and compactness. The scale value defines how large the image objects should be. The higher the value for scale parameter, the larger the image objects and vice versa. The values of shape and compactness parameters can be anywhere from 0 to 1. The shape parameter determines the influence of colour in the segmentation process, the higher the value lower the influence. The compactness parameter determines how compact the image objects should be. The higher the value, the more compact the image objects. The image objects are displayed on top of the imagery in the eCognition software at the end of the segmentation process.

The optimum values of the parameters for each flying height were determined by the trial and error method. The segmentation results were validated visually for each set of values of the parameters. The validation was performed by seeing whether the image objects were perfectly aligned with the real-world features. The process was repeated until the image objects resemble with the real-world features.

4.3.2.   Classification

Classification is the process of labelling the image objects into user-defined classes. The segmented image objects were classified to one of the two classes (Hogweed, Not Hogweed) for further analysis. The classes were defined as Hogweed and Not Hogweed. The background classes excluding Hogweed were merged to form one class and was named Not Hogweed. This was done to exclude misclassification error within background classes, since detecting Hogweed is the main objective of this study. OBIA requires training samples for successful classification. Sufficient amount of samples were selected for each of the two classes from the successfully segmented image objects. The samples were selected from user interpretation of high-resolution imagery on which the segmentation results were overlaid. The image objects created from the segmentation process can be classified using two different classification methods: (1) Nearest Neighbour classification or classification based on samples and (2) Rule-based classification or classification based on prior knowledge stored as rule sets (Schöpfer et al., 2010).

Nearest Neighbour classification is a supervised classification method which uses fuzzy rules (Wei et al., 2005). The classifier classifies image objects based on the selected samples within a defined feature space. With the help of selected samples and defined statistics, the segmented image objects were classified to either Hogweed or Not Hogweed class. Next, the statistics are defined. The nearest neighbour classifies the image objects based on the resemblance to the samples and the defined statistics. The classifier assigns the image objects to the classes, which are closer to the classes of the selected samples. Rule-based classification provides a way to overcome the spectral similarity of two different classes (Schöpfer et al., 2010) by integrating prior knowledge of the target classes. The classification results are then exported as shapefile and raster file for accuracy assessment analysis.

4.4.   Support Vector Machines

The main goal of the SVM classifier is to determine the optimal hyperplane that maximises the margin between two classes (Kavzoglu & Colkesen, 2009) therefore minimising the confusion between classes. The training data sets are used to determine optimum hyperplane. Support vectors are the critical elements of the training data set which define the hyperplane and are the extreme data points in the data set which are used to define the hyperplane. SVM uses a set of mathematical functions called kernels to transfer the input data into the desired form. The performance of SVM is related to the choice of the kernel function. Radial basis function (RBF) kernels and polynomial kernels are the two commonly used kernels for the classification of remotely sensed images (Huang et al., 2002; Pal & Mather, 2005). The main limitation of using SVM is the need to select suitable parameter values for the kernels which in turn improves the capability of well-trained SVMs to accurately predict the unknown data (Tso & Mather, 2009). The polynomial kernel has three parameters excluding the cost factor (offset, scaling and degree) while RBF has only one parameter excluding the cost factor (gamma). The number of parameters influences the complexity of the model, and hence, RBF is selected for this research since the model can be calibrated relatively quicker as opposed to the other kernel.

Training datasets for all the flying heights were created as a shapefile based on visual interpretation of high-resolution imagery. In order to train the SVM effectively, the training samples were created with all the land cover classes in the study area namely: Hogweed, Sand, Vegetation and Maize. Gamma (γ) and cost (C) are the two parameters which need to be defined for an RBF kernel. Intuitively, the γ parameter defines how far the influence of tthe raining sample reaches. For a small value the decision boundary is linear, and as the value increases, it leads to overfitting of the decision boundary (Ben-Hur & Weston, 2010). The C parameter is an error penalty to ensure satisfactory classification results (Tso & Mather, 2009). It is a tradeoff between the correct classification of training samples against the maximisation of margin between two classes. The lower the value, the higher the margin, the larger the value, the smaller the margin.

The goal is to find the optimum value for both the parameters, which in turn gives the best classification results. The training samples for the study area had four classes (Hogweed, Vegetation, Sand, and Maize). Since this research aims only at the accurate detection of Hogweed, precision, recall and F_1 score for the class Hogweed was calculated, and F_1 score was used as a measure to find the optimum parameters and the misclassification (error) of Vegetation, Sand and Maize classes were excluded. Precision and recall are based on the measure of how relevant the classification is. Precision (p) is a measure, which is computed by the ratio of the number of correctly predicted positive observations, i.e. true positives (T_p) to the total predicted positive observations, i.e. the sum of true positives (T_p) and false positives (F_p). It is calculated as follows:

$$p = \frac{T_p}{T_p + F_p}$$

Recall (*r*) is a measure which is computed by the ratio of correctly predicted positive observations to all the observations in the reference class, i.e. the sum of true positives (T_p) and false negatives (F_n) (Ghorbani, Ebadi, & Sedaghat, 2019). It is calculated as follows:

$$r = \frac{T_p}{T_p + F_n}$$

High precision relates to a low false positive rate while high recall relates to the low false negative rate. Furthermore, a high score for both precision and recall shows that the classifier is returning accurate results (high precision) with a majority of all positive results (high recall). The F_1 score, which is the weighted harmonic average of both precision and recall, was used as a measure for the optimum parameter selection of the RBF kernel. The value takes both false positives and false negatives into account. The F_1 score is a measure between 0 to 1, with 0 being worst and 1 being the best. It is calculated as follows:

$$F_1 = 2 \frac{p \cdot r}{p + r}$$

According to Hsu et al., (2003), the practical method for finding the optimum C and γ parameter is to select values with an exponentially growing sequence. They also suggested a range of values for both parameters, which was used in this study as well. The value of C parameter ranges from 2^{-5}, 2^{-3} ...2^{15} while the value of γ parameter ranges from 2^{-15}, 2^{-13}…2^3. Various pairs of values needs to be tested by running the process in a loop to select the optimum value. The set of values, which gives the maximum F_1 score for individual flying heights, were selected as the optimum values for the C and γ parameters.

4.5.   Accuracy Assessment

To know the success of the classification carried out and to find if the objectives of the analysis are met, it is essential to perform an accuracy assessment. It can be done visually by comparing the classification results, but for reliable assessment, we need a method, which can quantify the errors. Accuracy assessment is a process to compare the classified image to another dataset which is considered as accurate or ground truth data (Esri, 2018).

The confusion or error matrix is one of the widely promoted and often used accuracy assessment methods, which was used in this research as an accuracy measure for both the methods. It is a cross-tabulation of the classified class label against that observed in the ground or reference data (Foody, 2002). The reference data can be derived from interpreting high-resolution imagery, existing classified imagery, or GIS data layers. In this research, point shapefiles were created in Arc GIS using the high-resolution imagery as reference data. The reference points were created randomly using the 'Create Random points' tool in Arc GIS to avoid bias. Stratified sampling was used for selecting validation points from the reference data. The Giant Hogweed has a distinguishable fern-like leaf structure compared to the nearby vegetation, and it is very prominent in the green band. With the help of these criteria, the reference points were classified to one of the four classes (Hogweed, Maize, Sand, Vegetation) using the high-resolution UAV imagery by visual interpretation. The reference point in the classes Maize, Sand, and Vegetation, were merged together and labelled as Not Hogweed for OBIA accuracy assessment.

The confusion matrix was created from the reference data and classified results for further assessment. The measures of error in the confusion matrix were also calculated for each class. Bakx et al., (2013) explain the different components of the confusion matrix as follows. The column in the confusion matrix sums up to the total number of labelled reference pixels per class, and the row sums up to the total number of pixels labelled by the classifier. With the help of these, the errors of omission and errors of the commission can be defined. Errors of omission represent the values that belong to a class but were predicted by the classifier

in a different class. The error of omission relates to the column in the confusion matrix and is a corollary to the Producer's Accuracy. The producer's accuracy is the map accuracy, which shows how often the real features on the ground are correctly shown on the classified map as well. Errors of commission represent the values that are predicted to be in a class by the classifier, but they belong to another class. The error of commission relates to the row in the confusion matrix and is a corollary to the user's accuracy. User's accuracy tells us how reliable is the classification map is to the real world.

4.6.   Testing of ObsIdentify for Plants Application

ObsIdentify for plants is a mobile application for Android users, which was launched in 2017. The app uses deep learning to identify various wild plants from the Netherlands and Belgium using mobile pictures. Deep learning is an application of neural networks, and it consists of several layers of nodes, usually more than 4 between the input and output (Arel et al., 2010). The main disadvantage of deep learning is that it requires a large amount of data for training, but that does not pose as a problem for the app since it uses photos from websites Waarneming.nl and Waarnemingen.be to train the software. Users can upload pictures of various flora and fauna along with the location where it was sighted in the websites. The sites are maintained by volunteers. Around 6,000,000 – 7,000,000 photos are uploaded each year in the waarneming.nl website (Waarneming.nl, 2019). The app is trained not only using the perfect pictures from the field but also using the bad pictures (Strien, 2018). As the volume of the dataset increases, the performance of deep learning increases (Talha et al., 2019). When the app was released, it could identify about 1,300 wild plants from both countries. A beta version was introduced in 2018 with an expanded library of about 2200+ species including mosses (Hogeweg, 2018).

The app is very straightforward to use. Users can either take photos using the mobile camera or upload it from their album directly. The app has a species list that lists all the species, which the app can identify along with their basic information. The main interface (Figure 4.2) of the app is simple and efficient. The app also allows the user to crop the image to focus more on the target species that needs to be identified. The UAV images of different flying heights were uploaded, and the area with Giant Hogweed was cropped using the crop feature from the app. After uploading the picture, the 'ID' option was clicked, to identify the species. The app displays the results as a list of species names, which it thinks that closely resembles the input image. Moreover, there is also a small pie chart next to the species name, which indicates the measure of the likelihood of that species. The recognition process completely works offline. Users can also upload more than one image; for example, different images showing different angles of the same plant. This method was often found to improve recognition (Herremans, 2017).
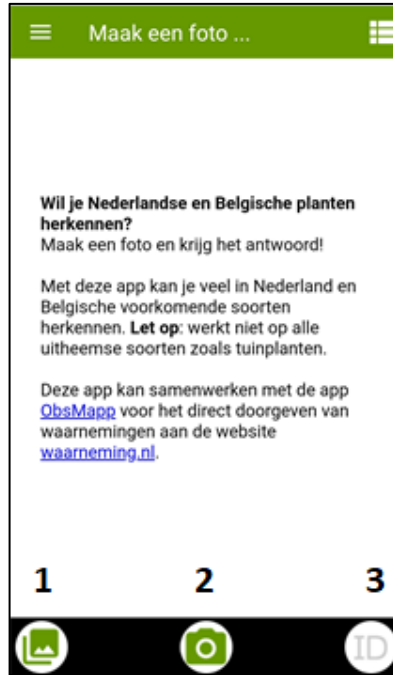
Figure 4.2 The main interface of the ObsIdentify for Plants mobile application. The options in the bottom from left are 1. To upload pictures from the mobile album, 2. To directly click a picture with the mobile camera 3. The ID button, which is used as a start button for the process after uploading the picture.

5. RESULTS

5.1.   Data Pre Processing

The layer stack of flying height 11 m, 12 m and 14 m created using the orthophoto mosaics of Green, NIR, and Red Edge bands can be seen in Figure 5.1.
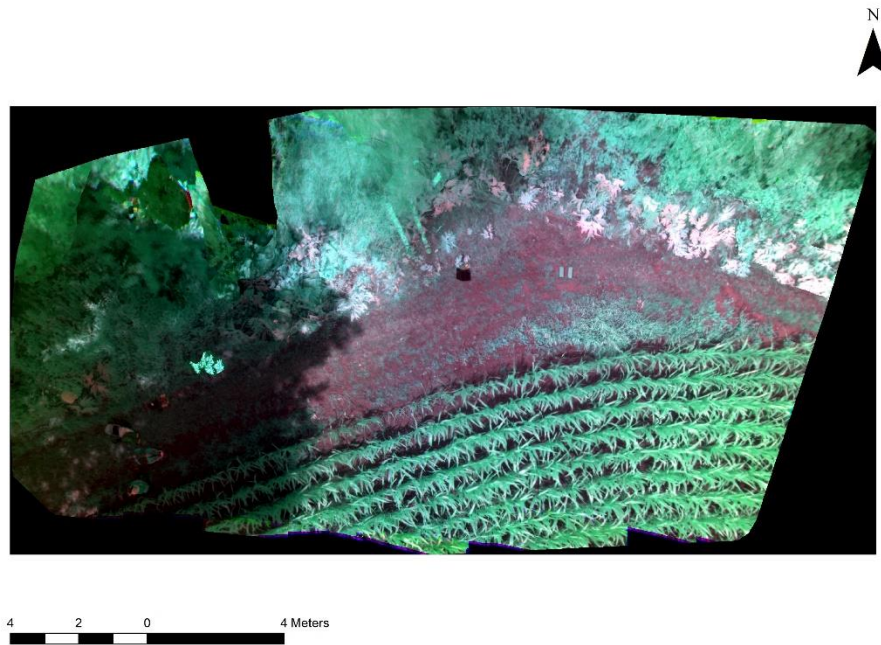
(a)
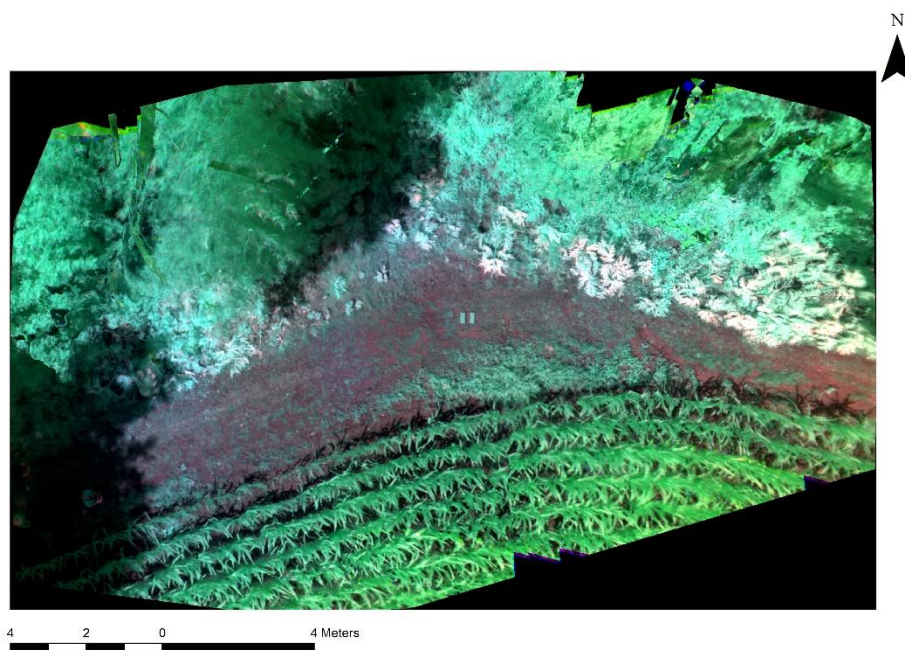
(b)
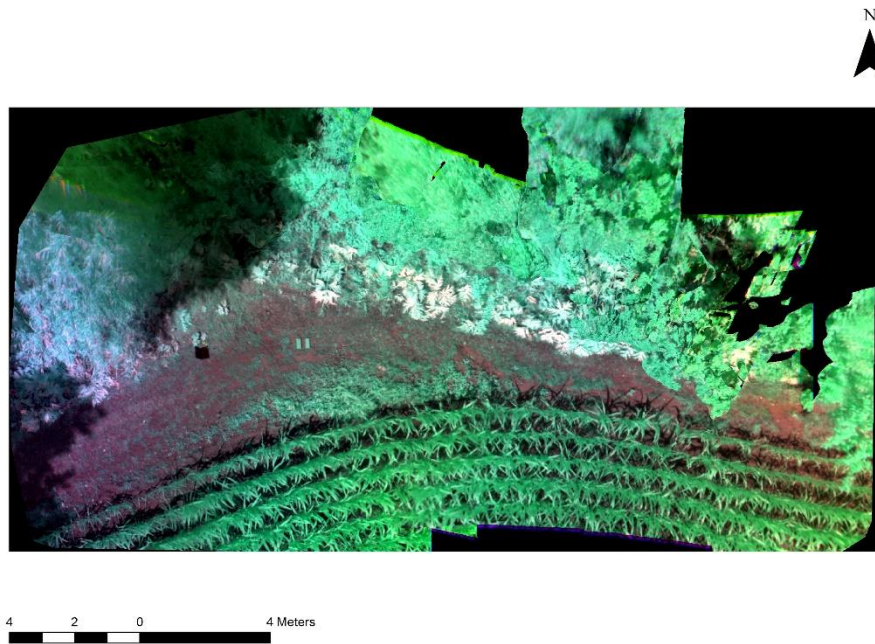
(c)

Figure 5.1 Layer stack of spectral bands Red Edge, NIR and Green bands of 11 m (a) 12 m (b) and 14 m (c) flying heights. The bands Red Edge, NIR and Green are displayed in Red, Green and Blue respectively.

The GSD for orthomosaics were calculated by default by the software, and they were around 0.01 m for all three flying heights with only a small variation in decimal. It can be seen in the layer stack of flying height 11 m (Figure 5.1 a) that almost half of the study area was under the effect of shadow. As the flying height increases the effect of shadow decreases and the area coverage increases. The area of the three flying height scenes can be seen in Table 5.1.

Table 5.1 The area of the study area for all three flying heights shown in square metres.

Flying height (in m)	Area (in sq. m)
11	275.4
12	279.5
14	308.5

5.2. Object-Based Image Analysis

5.2.1. Segmentation

eCognition allows using multiple bands to perform segmentation. Different combinations of image bands were used, and it was found that the segmentation carried out using a combination of Green, NIR and Red Edge bands gave image objects, which closely resembled the real-world features.

Since in the Green band, Giant Hogweed is more prominent than the nearby vegetation (Figure 3.1), a weight of 2 was given for Green band during the segmentation process while the other two bands, NIR and Red Edge were given a weight of 1. The segmentation analysis showed promising results at each stage when

the values of the parameters were changed. The process was carried out separately for different heights instead of using the same parameters for all the heights. At first, a value of 100 was given for the scale parameter, which resulted in small image objects. The image objects were small compared to the real-world features; hence, the value of the scale parameter was increased to accommodate the shape of the hogweed. Values from 600 to 700 accommodated the size of the hogweed image objects for all flying heights.

The values for shape and compactness were started from a minimum value of 0.2, and it was observed that when the values were low, it yielded many small meaningless image objects (Figure 5.2) because the influence of colour is high with low values. This is known as over-segmentation of the image. Over-segmentation is when the parameters for segmentation produce smaller image objects than the actual features (Haque et al., 2016). To avoid over-segmentation, the values were increased, and the segmentation outputs were observed. As the values increased, it resulted in less meaningless image objects. Values from 0.6 to 0.8 were optimum for less over-segmentation of the image for all the flying heights.
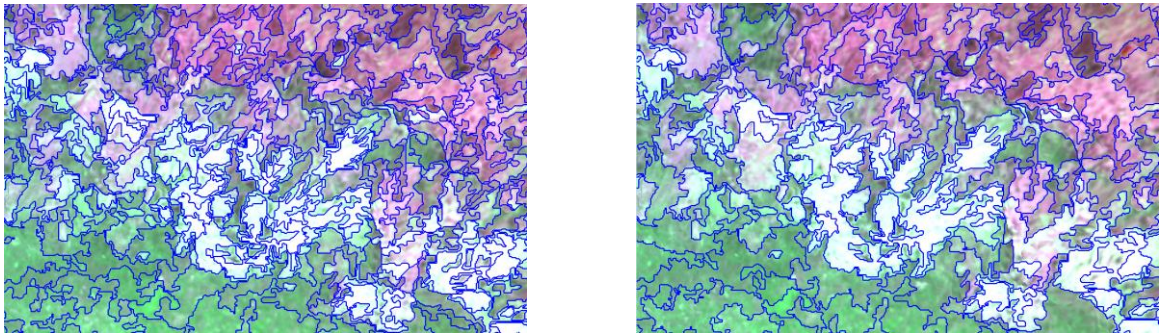
Figure 5.2 Left, over-segmented image objects. Right, segmentation with meaningful image objects. The segmentation results are overlaid on a layer stack image whose combination is Green, NIR and Red Edge bands in eCognition software.

Another main reason for controlling over-segmentation was because it also resulted in many small image objects for the flying height of 11 m, in the maize region which was spectrally similar to the Hogweed (Figure 5.3). As the value for shape and compactness were increased, it resulted in less spectrally similar image objects in the maize region (Figure 5.4). This inconsistency, which was found only in 11 m flying height may be due to the high spectral variability of high-resolution data in low flying heights.
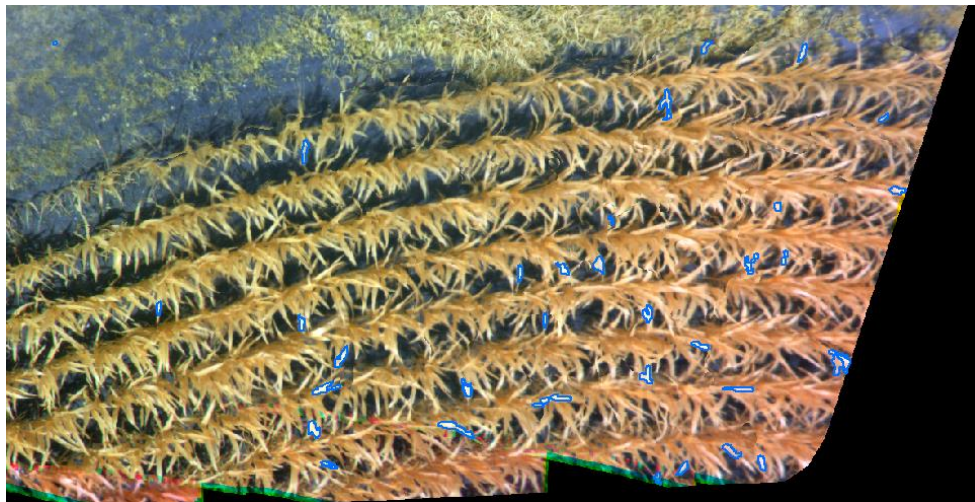
Figure 5.3 Several small Not Hogweed image objects, but spectrally similar to Hogweed in the maize region. These image objects are shown as blue polygons overlaid on flying height 11 m. These small image objects were due to over-segmentation.
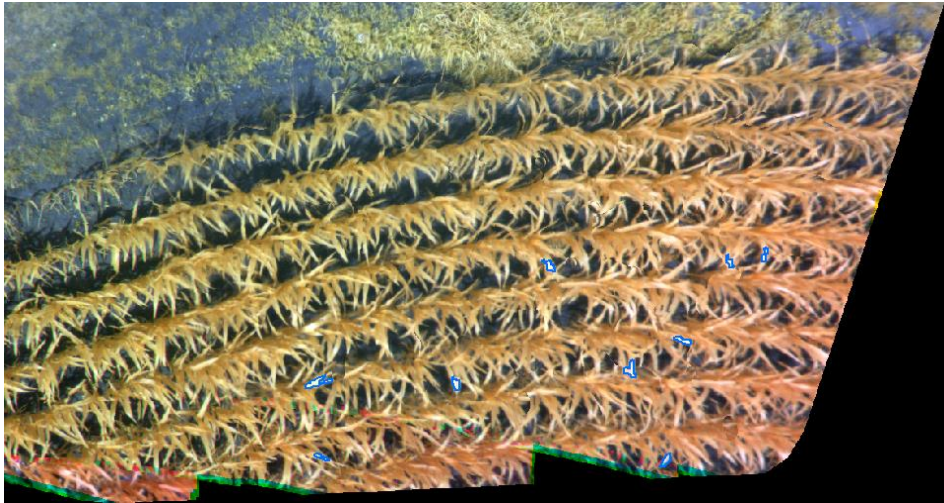
Figure 5.4 Not Hogweed image objects in the maize region overlaid as blue polygons on flying height 11 m. These image objects are spectrally similar to Hogweed. These are fewer than in Figure 5.3, due to the increase of shape and compactness values.

For each combination of values for the parameters, the changes in the segmentation results were observed. The combination of parameters, which gave image objects that were similar to the shape of the Hogweed with less over-segmentation, were tabulated as optimum values. The optimum values for the parameters for all the three flying heights are as tabulated in Table 5.2.

Table 5.2 Optimum values for the parameters scale, shape and compactness of the Segmentation process of Object-Based Image Analysis performed in eCognition software for all flying heights.

Height (in m)	Scale	Shape	Compactness
11	700	0.6	0.6
12	700	0.8	0.8
14	600	0.8	0.8

As can be seen from Table 5.2, the scale value decreases with an increase in flying height. This is because as the flying height increases, the spatial resolution decreases which means the size of the real-world feature, measured in pixels, decreases with an increase in flying height. It can also be seen that the value of shape and compactness were less for flying height 11 m when compared with 12 m and 14 m. As the value of shape decreases the influence of colour decreases in the segmentation process. The value was reduced to successfully segment image objects in the shadow region since half of 11 m flying height was covered with shadow. However, a trade-off was made for the value that gave few misclassified image objects, which were spectrally similar to Hogweed in Maize region (see Figure 5.3 and Figure 5.4) and also segmented Hogweed image objects in shadow region. However, this value was not efficient for the entire shadow region. It did not successfully segment all the image objects in the shadow region. This is because of the effect of shadow in the flying height 11 m, which led to under-segmentation in the shadow region. This under-segmentation could be attributed to the spectral similarity within the image objects in the shadow region. Under-segmentation is a segmentation inaccuracy which refers to creating too fewer segments (Möller et al., 2007). Hence, the shadow area was segmented separately from the sunlit region, and the sunlit region was masked (Figure 5.5), and the segmentation was performed separately on the shadow region. The optimum values for scale, shape, and compactness were 500, 0.3 and 0.9. The values for parameters scale, shape, and

compactness had to be changed because segmentation in the shadow region is different from the sunlit region. The decrease in shape and scale values were to overcome the under-segmentation.
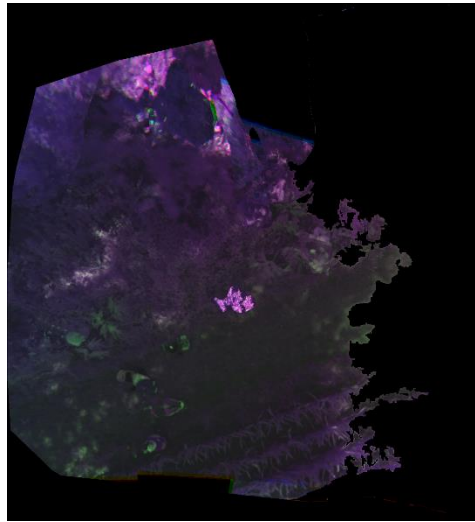
Figure 5.5 Segmented shadow region of 11 m flying height seperated from the sunlit region for further targeted analysis

5.2.2. Classification

The classification step was carried out by selecting samples from the successfully segmented image objects. The idea is that classifier will use the selected samples to classify the entire image. The parameters selected to classify the image objects using the Nearest Neighbour classification method were layer mean, brightness, standard deviation, hue, saturation and intensity. The rule-based classification was used to reclassify the image objects from other classes, which were misclassified as Hogweed. Although after increasing the values for shape and compactness parameters to control over-segmentation in flying height 11 m, there were still some small meaningless image objects in the Not Hogweed class, which were classified as Hogweed due to the spectral similarity. Image objects having an area of lesser than 100-pixel were reclassified as Not Hogweed from Hogweed class. The following steps were used to define the value of the area-based rule. The area of the image objects in the maize field that were spectrally similar to Hogweed was identified using the image object information feature in eCognition. The threshold value was defined in such a way that the value does not affect the correctly classified Hogweed image objects. This rule-based classification was done only for 11 m flying height as the other flying heights 12 m, and 14 m did not have small image objects as explained in the section 5.2.1. The classified features were evaluated visually before doing further accuracy assessment. This visual assessment was done by exporting the classified features as shapefile and overlaying them on top of the high-resolution UAV imagery (Figure 5.6)
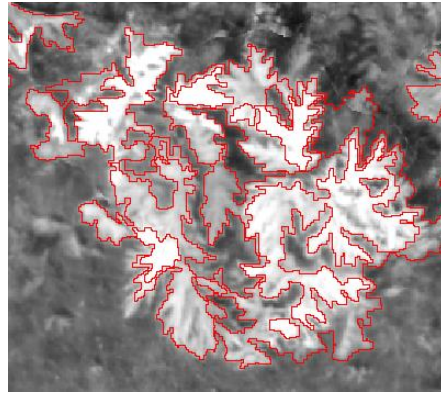
Figure 5.6 Classified shapefile output from eCognition, which was overlaid on the Green Band of the UAV image for evaluation of the classification method in ArcGIS software. The image shows the Hogweed class.

The final classified results were used to make maps showing Hogweed in black colour and Not Hogweed in white colour for all flying heights as shown in figures 5.7 – 5.9. The Hogweed was located in the north side of the study area in a long stretch as seen in the maps. Also, as the flying height increases, the size of the Hogweed leaves decreases, which can be seen when comparing the map of 11 m flying height with a map of 14 m flying height. Giant Hogweed's leaves have a fern-like structure that was distinct from the nearby vegetation and was preserved in the segmentation (Figure 5.6) and hence can be seen in all the final classified maps too.

Figure 5.7 Classified raster image for 11 m flying height using Object-Based Image Analysis where Hogweed class is represented as black and Not Hogweed class is represented as white.
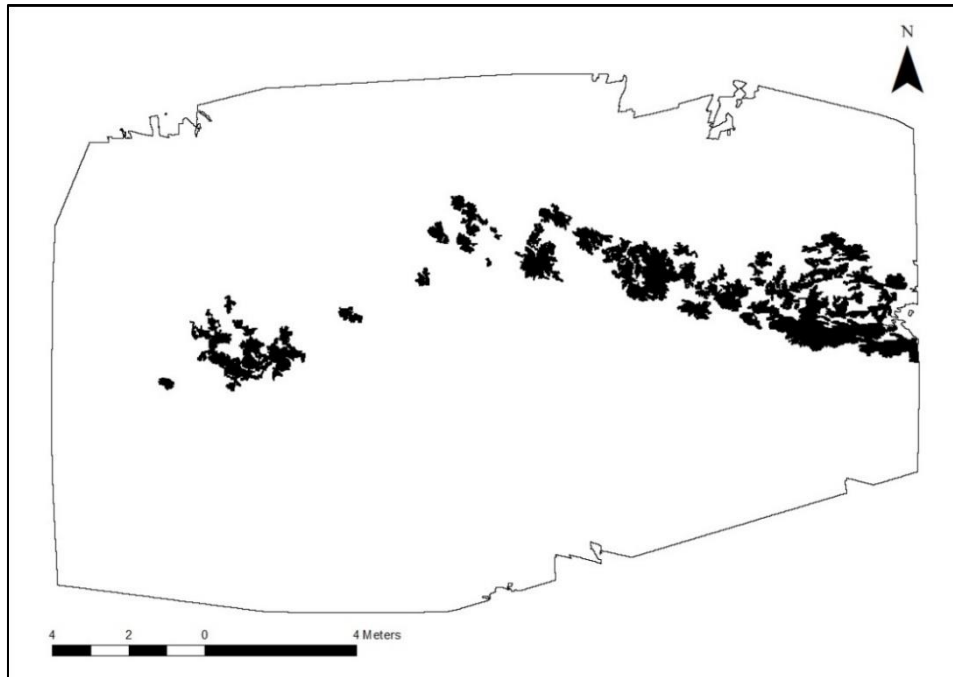
Figure 5.8 Classified raster image for 12 m flying height using Object-Based Image Analysis where Hogweed class is represented as black and Not Hogweed class is represented as white.
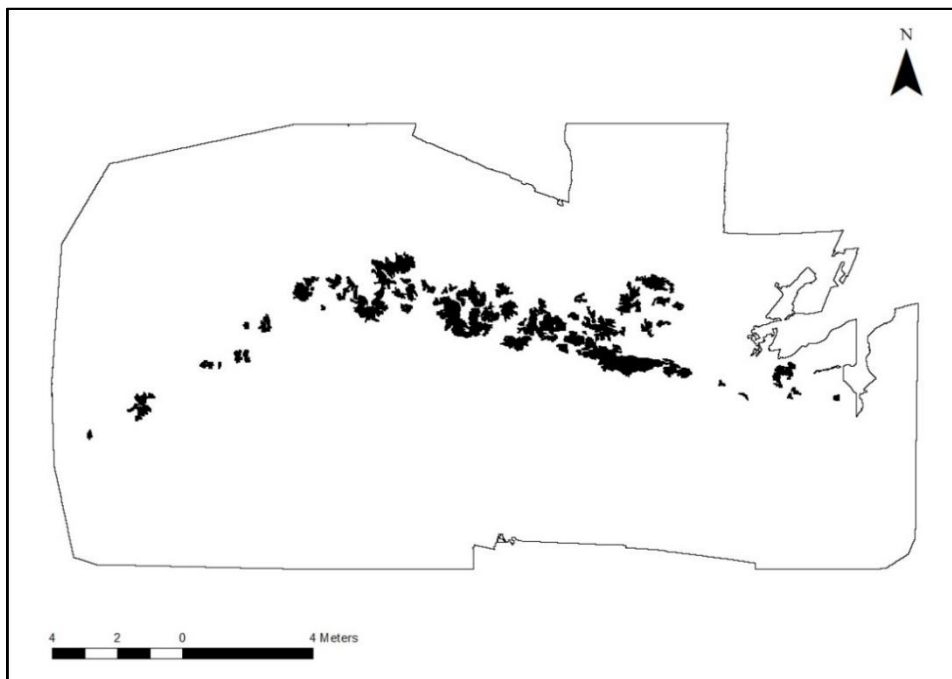
Figure 5.9 Classified raster image for 14 m flying height using Object-Based Image Analysis where Hogweed class is represented as black and Not Hogweed class is represented as white.

It can be seen from Figure 5.7 that OBIA failed to classify the Hogweed image objects in the shadow region. The spectral variability between the shaded region and unshaded region for the same real-world features are high (Lichtblau & Oswald, 2019) which resulted in affecting the classification. Hence, a targeted segmentation was done only for the shadow region and samples from the shadow region for both the classes

were selected for classification, and the classified raster, where Hogweed is in black and Not Hogweed is in white for only the shadow region can be seen in Figure 5.10.
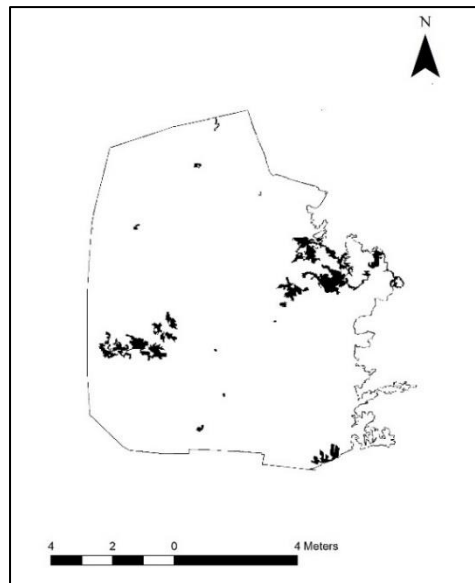
Figure 5.10 Classified raster image for 11 m flying height for the shadow region using Object-Based Image Analysis where Hogweed class is represented as black and Not Hogweed class is represented as white.
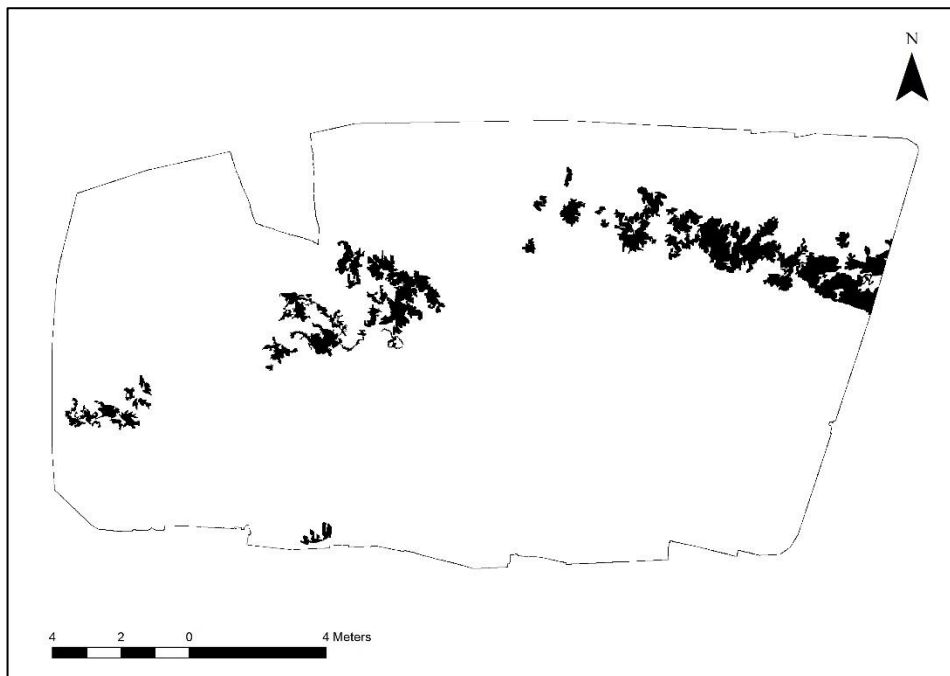
Figure 5.11 Classified raster image for 11 m flying height with Hogweed image objects from shadow and sunlit region, using Object-Based Image Analysis where Hogweed class is represented as black and Not Hogweed class is represented as white.

The Hogweed objects classified from the shadow region and sunlit region were merged to create a final classified map for 11 m flying height as seen in Figure 5.11. There were some misclassified image objects, but they were fewer in this particular case. A rule classification based on the area was not used to reclassify misclassified image objects, because they were almost similar in size to the Hogweed image objects.

5.3. Support Vector Machines

The main goal is to identify the best values for C and γ that will help the classifier to classify the data accurately. The F_1 score for each combination of parameters was tabulated. The pair of values, which gave the best F_1 score, was finalised as the optimum parameter set for a specific flying height. The results are tabulated in Table 5.3.

Table 5.3 Optimum values of parameters cost and gamma for Radial Basis Kernel of Support Vector Machine method for all the flying heights with their respective precision, recall and F_1 score.

Height (in m)	Cost	Gamma	Precision	Recall	F_1 score
11	128	1	0.84	0.8936	0.8659
12	0.5	2	1	0.9	0.9473
14	4	0.5	0.9	0.9	0.9000

The result was a classified binary raster image, which was used to make maps showing Hogweed class in black colour and Not Hogweed class in white colour. The classified raster images for all the flying heights are shown in the figures 5.12 – 5.14. Similar to OBIA the size of the Hogweed visibly decreased as the flying height increased. In addition, the distinct fern-like structure of the Giant Hogweed's leaves was preserved in SVM as well.
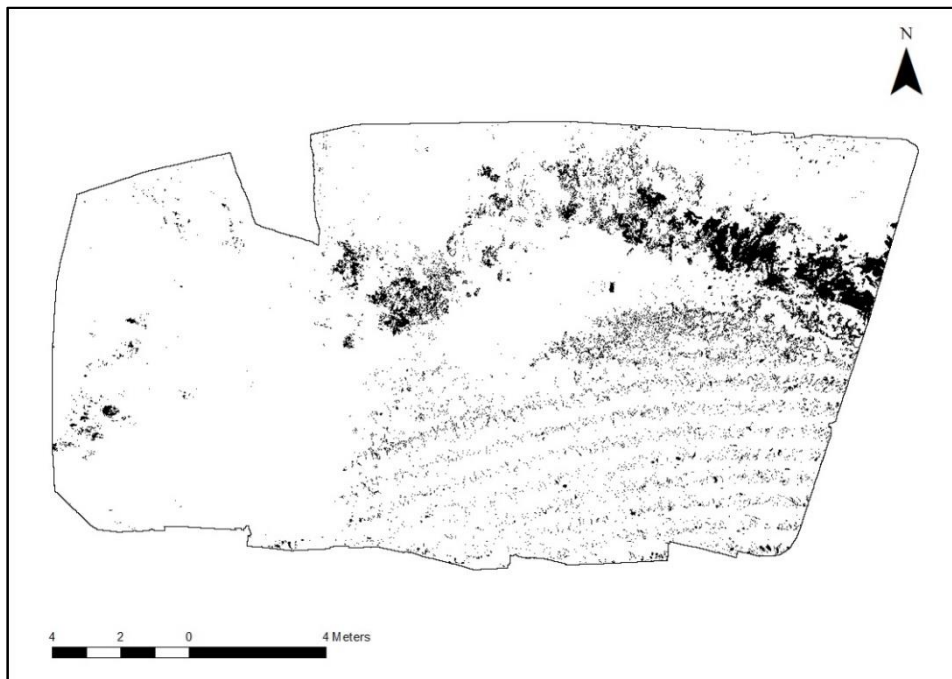
Figure 5.12 Classified raster image for 11 m flying height using SVM, where Hogweed class is represented as black and Not Hogweed class is represented as white.

It is evident from the classified image for flying height 11 m (Figure 5.12) that Not Hogweed classes were also classified as Hogweed. SVM classifies the data according to the spectral information and hence spectrally similar pixels from other classes such as Maize, Sand and Vegetation were also classified as

Hogweed. There were some misclassified pixels in the sand class because of the presence of grass in the unpaved road.
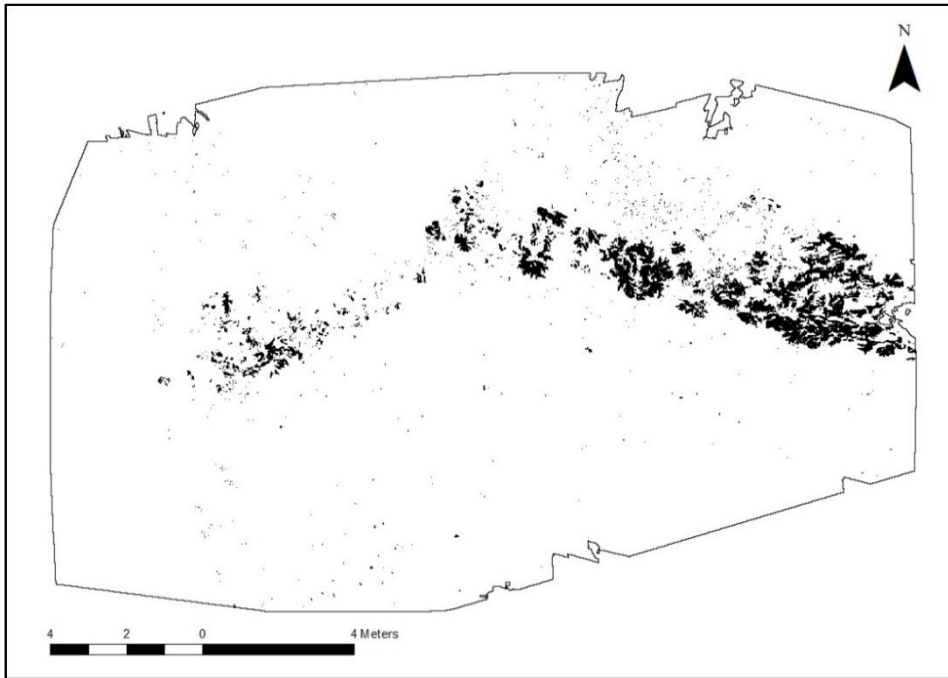
Figure 5.13 Classified raster image for 12 m flying height using SVM, where Hogweed class is represented as black and Not Hogweed class is represented as white.
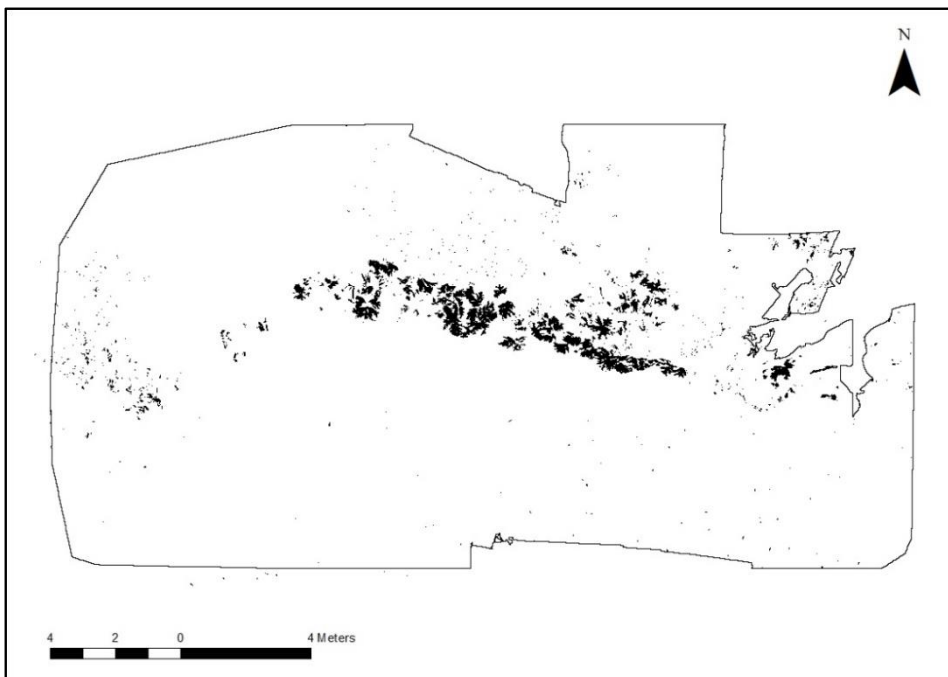
Figure 5.14 Classified raster image for 11 m flying height using SVM, where Hogweed class is represented as black and Not Hogweed class is represented as white.

The salt and pepper effect, which is common to a pixel-based approach when applied to high-resolution images, can be seen in all three flying heights. Among all the three flying heights, the 11 m has the most misclassified pixels; this was mainly due to the effect of shadow. Compared to the other flying heights, 11 m has most of the study area covered with shadow, which resulted in training samples being selected from the shadow region as well. Hence, the process was repeated for 11 m flying height without training samples for Hogweed from the shadow region. The process was again run in a loop; the set of C and γ values with the highest F_1 score was selected. The resulting output raster image had visibly less misclassified pixels (Figure 5.15). The optimum C and γ were 4096 and 8 respectively, and the F_1 score for the combination was 0.7804. The F_1 score for 11 m flying height without training samples from the shadow region is much lower than the one, which had samples from the shadow region. This is because the classifier did not classify the Hogweed in the shadow region due to the absence of training samples in the shadow region.
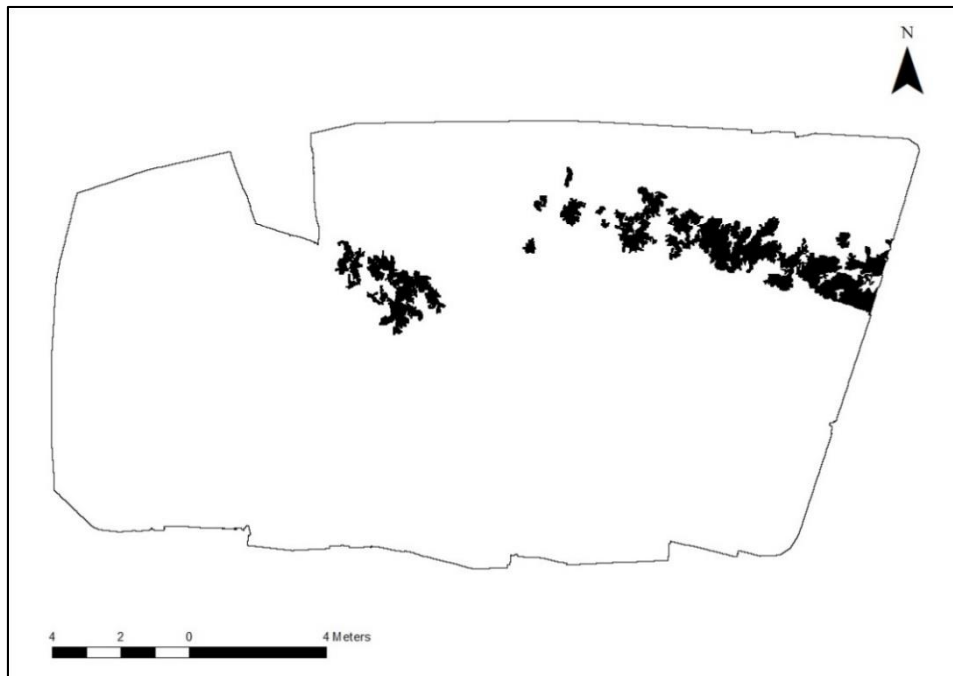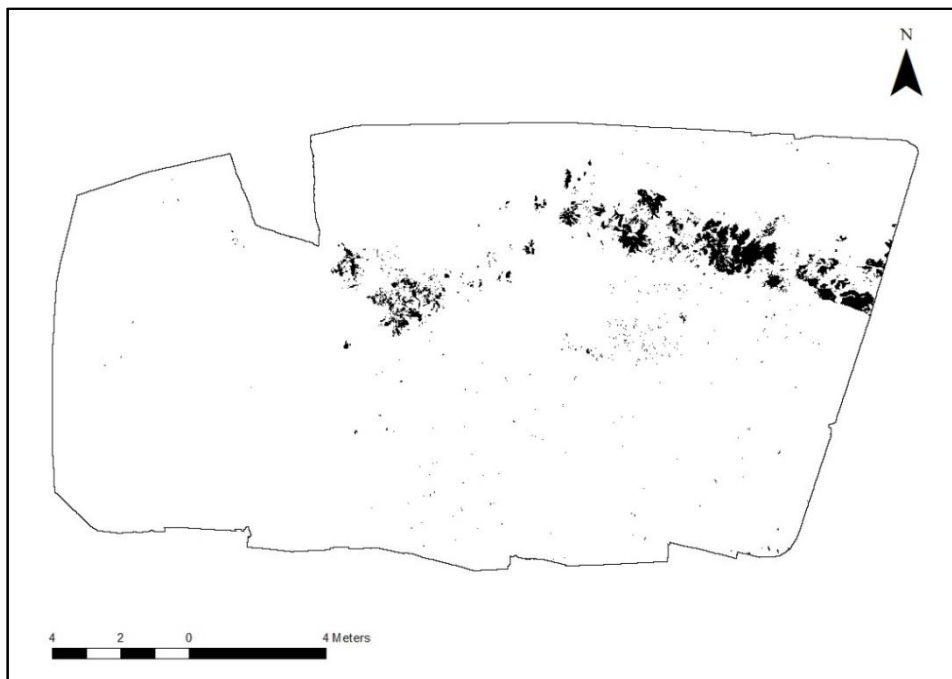

Figure 5.15 Classified a raster image for 11 m flying height using SVM without training samples of Hogweed from the shadow region. Hogweed class is represented as black, and Not Hogweed class is represented as white.

When comparing the classified image raster for flying height 11 m with samples from shadow region (Figure 5.12) and without samples from shadow region (Figure 5.15), the number of misclassified pixels in other classes has reduced visibly. A possible explanation for this might be that the misclassified pixels were spectrally similar to Hogweed in the shadow region.

5.4. Accuracy Assessment

5.4.1. Object-Based Image Analysis

The confusion matrix for the predicted class and the actual class was created for all the flying heights to quantify the errors, and they are shown in tables 5.4 – 5.6. The table shows the Producer's Accuracy, User's Accuracy and the Overall Accuracy, which are used as measures for the assessment of the methods.

Table 5.4 Confusion Matrix for flying height 11 m, which is for the accuracy assessment for the Object-Based Image Analysis method.

Confusion Matrix		Actual class			Error of Commission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**31**	0	31	0%	100%
	Not Hogweed	19	**150**	169	11.24%	88.76%
	Total	50	150	**200**		
Error of Omission		38%	0%			
Producer's Accuracy		62%	100%			
Overall Accuracy		91%				

Table 5.5 Confusion Matrix for flying height 12 m, which is used for the accuracy assessment for the Object-Based Image Analysis method.

Confusion Matrix		Actual class			Error of Omission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**42**	0	42	0%	100%
	Not Hogweed	8	**150**	158	5.06%	94.94%
	Total	50	150	**200**		
Error of Commission		16%	0%			
Producer's Accuracy		84%	100%			
Overall Accuracy		96%				

Table 5.6 Confusion Matrix for flying height 14 m, which is used for the accuracy assessment for Object-Based Image Analysis method.

Confusion Matrix		Actual class			Error of Commission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**43**	0	43	0%	100%
	Not Hogweed	7	**150**	157	4.06%	95.94%
	Total	50	150	**200**		
Error of Omission		14%	0%			
Producer's Accuracy		86%	100%			
Overall Accuracy		97%				

A total of 200 reference points were selected for the accuracy assessment of Object-Based Image Analysis. In the above table, the diagonal elements show the number of points that were correctly classified. The overall accuracy using OBIA was 91%, 96% and 97% for 11 m, 12 m, and 14 m respectively. Overall

accuracy was calculated by dividing the number of correctly classified points by the total number of points. From the confusion matrix, it can be inferred that the overall error was skewed in such a way that all the flying heights had omission error, but no commission error, i.e. Hogweed class was classified as Not Hogweed class, but there was no Not Hogweed class that was classified as Hogweed class. To understand more about the cause of this skewness the errors were checked by overlaying them on high-resolution imagery. While plotting the errors, it was found out that the effect of shadow was the main reason for most of the misclassified points. This was due to the spectral similarity between the Not Hogweed and Hogweed class in the shadow region, which resulted in the inefficiency of OBIA classifier to classify Hogweed image objects correctly. In general, the 12 m and 14 m study area did not have any commission error. However, it should be noted that the rule-based classification helped to eradicate the commission error in flying height 11 m. It can also be seen that the accuracy increases as the flying height increases. The increase in accuracy was due to the decrease in error of omission with increasing flying heights that was mainly due to the decrease in the effect of shadow in 12 m and 14 m flying heights when compared with 11 m flying height.

The accuracy assessment for the classified image with Hogweed image objects merged from both the shadow and sunlit region Figure 5.11, can be seen in Table 5.7.

Table 5.7 Confusion Matrix for flying height 11 m, which consists of Hogweed image objects from both shadow region and sunlit region.

Confusion Matrix		Actual class			Error of Commission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**40**	0	40	0%	100%
	Not Hogweed	10	**150**	160	6.25%	93.75%
	Total	50	150	**200**		
Error of Omission		20%	0%			
Producer's Accuracy		80%	100%			
Overall Accuracy		**95%**				

The confusion matrix of flying height 11 m, which did not have image objects from shadow region (Table 5.4) was compared with the confusion matrix of flying height 11 m which had image objects from both sunlit and shaded region (Table 5.7). It can be seen that the Producer's Accuracy for the Hogweed class increased from 62% to 80% and the overall accuracy increased from 81% to 95%. This was because of the addition of Hogweed image objects from the shadow region decreased the error of omission from 38% to 20% in Table 5.7, which in turn increased the Overall Accuracy.

5.4.2. Support Vector Machine

The confusion matrix was used as a measure for accuracy assessment of the SVM algorithm for consistency. The confusion matrix with all the four classes namely: Hogweed, Not Hogweed, Sand, and Maize along with simplified confusion matrix with only two classes Hogweed and Not Hogweed for all the flying heights are shown in the tables 5.8 – 5.13. The simplified confusion matrix was added in addition to show the Overall Accuracy excluding the misclassification error from background classes (Sand, Vegetation, and Maize), which will be used as a measure for assessment of the method. The table also shows the Producer's Accuracy, User's Accuracy, which will also be used as measures for the assessment of the methods.

Table 5.8 Confusion Matrix for flying height 11 m, which was used for the accuracy assessment of the SVM method that shows all four training classes.

Confusion Matrix		Actual class					User's Accuracy
		Hogweed	Not Hogweed	Sand	Maize	Total	
Predicted class	Hogweed	**42**	0	2	3	47	89%
	Not Hogweed	0	**41**	1	27	69	59.42%
	Sand	4	1	**37**	6	48	77.08%
	Maize	4	8	10	**14**	36	38.89%
	Total	50	50	50	50	**200**	
Producer's Accuracy		84%	82%	74%	28%		
Overall Accuracy		67%					

Table 5.9 Simplified Confusion Matrix for flying height 11 m, which was used for the accuracy assessment of the SVM method that shows only Hogweed and Not Hogweed class.

Confusion Matrix		Actual class			Error of Commission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**42**	5	47	11%	89%
	Not Hogweed	8	**145**	153	5.23%	94.77%
	Total	50	150	**200**		
Error of Omission		16%	3%			
Producer's Accuracy		84%	97%			
Overall Accuracy		94%				

Table 5.10 Confusion Matrix for flying height 12 m, which was used for the accuracy assessment of the SVM method that shows all four training classes.

Confusion Matrix		Actual class					User's Accuracy
		Hogweed	Not Hogweed	Sand	Maize	Total	
Predicted class	Hogweed	**47**	1	1	1	50	94%
	Not Hogweed	2	**32**	1	6	41	78.05%
	Sand	0	0	**43**	0	43	100%
	Maize	1	17	5	**43**	66	65.15%
	Total	50	50	50	50	**200**	
Producer's Accuracy		94%	64%	56%	86%		
Overall Accuracy		83%					

Table 5.11 Simplified Confusion Matrix for flying height 12 m, which was used for the accuracy assessment of the SVM method that shows only Hogweed and Not Hogweed class.

Confusion Matrix		Actual class			Error of Commission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**47**	3	50	6%	94%
	Not Hogweed	3	**147**	150	2%	98%
	Total	50	150	**200**		
Error of Omission		6%	2%			
Producer's Accuracy		94%	98%			
Overall Accuracy		97%				

Table 5.12 Confusion Matrix for flying height 14 m, which was used for the accuracy assessment of the SVM method that shows all four training classes.

Confusion Matrix		Actual class					User's Accuracy
		Hogweed	Not Hogweed	Sand	Maize	Total	
Predicted class	Hogweed	**45**	3	0	2	50	90%
	Not Hogweed	1	**30**	6	11	48	62.50%
	Sand	3	6	**43**	4	56	76.79%
	Maize	1	11	1	**33**	46	71.74%
	Total	50	50	50	50	**200**	
Producer's Accuracy		90%	60%	86%	66%		
Overall Accuracy		76%					

Table 5.13 Simplified Confusion Matrix for flying height 14 m, which was used for the accuracy assessment of the SVM method that shows only Hogweed and Not Hogweed class.

Confusion Matrix		Actual class			Error of Commission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**45**	5	50	10%	90%
	Not Hogweed	5	**145**	150	3.33%	96.67%
	Total	50	150	**200**		
Error of Omission		10%	3%			
Producer's Accuracy		90%	97%			
Overall Accuracy		95%				

The same 200 reference points, which were selected for OBIA accuracy assessment, were also used for accuracy assessment of the SVM method to avoid bias. The reference points were spread over all the four classes. It can be seen from the simplified confusion matrix tables (Table 5.9, Table 5.11, and Table 5.13) that unlike OBIA, the errors were skewed in both directions. Even though the misclassification was evident in the Maize region for the 11 m flying height in Figure 5.12, the number of false positives were only 5 (Table 5.9). The most relevant classification was chosen according to the F_1 score, which takes both False Positives and False Negatives into account. Hence, the final classification has the least False Positives and False Negatives even though there was a visible misclassification spread over other classes. This inconsistency might be due to overfitting of the SVM model. Generally, a higher value of C results in overfitting of the SVM model. The C parameter controls the trade-off between classifying training point correctly and having a smooth decision boundary. From Table 5.3, it can be inferred that the value of C is higher for flying height 11 m when compared with 12 m and 14 m. According to Liu and Gillies (2016) in machine learning, overfitting happens when the algorithm tries to maximise the inter-class discrimination with less training samples. The overfitting of the SVM model was seen only in 11 m. This inconsistency can be attributed to the presence of training samples in the shadow region, which were spectrally similar to other classes in the study area, which resulted in the classifier maximising the inter-class discrimination for better accuracy. This led to overfitting of the model, which led to fewer errors even though the misclassification was evident. The overall accuracy using the SVM method is 94%, 97% and 95% for 11 m, 12 m and 14 m respectively.

The confusion matrix and the simplified confusion matrix for flying height 11 m without training samples from the shadow region are shown in the tables 5.14 and 5.15.

Table 5.14 Confusion Matrix for flying height 11 m, which was used for the accuracy assessment of SVM method that shows all four training classes and has no training samples for the Hogweed class from the shadow region.

Confusion Matrix		Actual class					User's Accuracy
		Hogweed	Not Hogweed	Sand	Maize	Total	
Predicted class	Hogweed	**32**	0	0	0	32	100%
	Not Hogweed	5	**35**	10	24	74	47.3%
	Sand	7	1	**32**	3	43	74.42%
	Maize	6	14	8	**23**	51	45.1%
	Total	50	50	50	50	**200**	
Producer's Accuracy		64%	70%	64%	46%		
Overall Accuracy		61%					

Table 5.15 Simplified Confusion Matrix for flying height 11 m, which was used for the accuracy assessment of SVM method that shows only Hogweed and Not Hogweed class and has no training samples for the Hogweed class from the shadow region.

Confusion Matrix		Actual class			Error of Commission	User's Accuracy
		Hogweed	Not Hogweed	Total		
Predicted class	Hogweed	**32**	0	32	0%	100%
	Not Hogweed	18	**150**	168	17.17%	82.29%
	Total	50	150	**200**		
Error of Omission		36%	0%			
Producer's Accuracy		64%	100%			
Overall Accuracy		**91%**				

The removal of training samples of Hogweed class from the shadow region in flying height 11 m decreased the Producer's accuracy from 84% to 64% and the overall accuracy from 94% to 91%. This was because the classifier was not able to classify the Hogweed in the shadow region due to the absence of training samples. However, Table 5.9 and Table 5.15 shows that the error of commission has decreased from 11% to 0%. A possible explanation is that the misclassified pixels were spectrally similar Hogweed in shadow region. Hence, the removal of Hogweed training classes from the shadow region resulted in 0% commission error.

5.5. Testing of ObsIdentify for Plants Application

RGB bands of UAV images with different flying heights were given as an input for the ObsIdentify for plants mobile application. The images were cropped using the crop feature from the app to focus on leaves and leaves with the umbels region. Images with good exposure of sunlight as well as in shadow region were selected as input images to understand whether the effect of shadow would be a problem in deep learning based mobile application as well.

UAV images with low flying heights like 5m, 8m and 11 m, which had both shadow region (Figure 5.16) and good exposure of sunlight (Figure 5.17), were recognised as *Heracleum mantegazzianum* by the application with one image only. Moreover, the application had no difficulty in identifying images with only leaves and images with leaves and umbels.

As the flying height increased (12 m and 14 m), the app had some difficulty in recognising the images and the results were always two to three species with *Heracleum mantegazzianum* being the first choice, second choice or third choice. However, when one or more images of the same scene were added the app recognised images as *Heracleum mantegazzianum* without any difficulty. In summary, these results conclude that only flying height influences the accuracy of recognition of the species whereas neither the effect of shadow nor the parts of plant influence the accuracy of recognition. The selection of a good subset of the image that has only Giant Hogweed is also equally important, i.e. the image for the detection should be cropped in such a way that Giant Hogweed covers most of the region than the nearby vegetation.
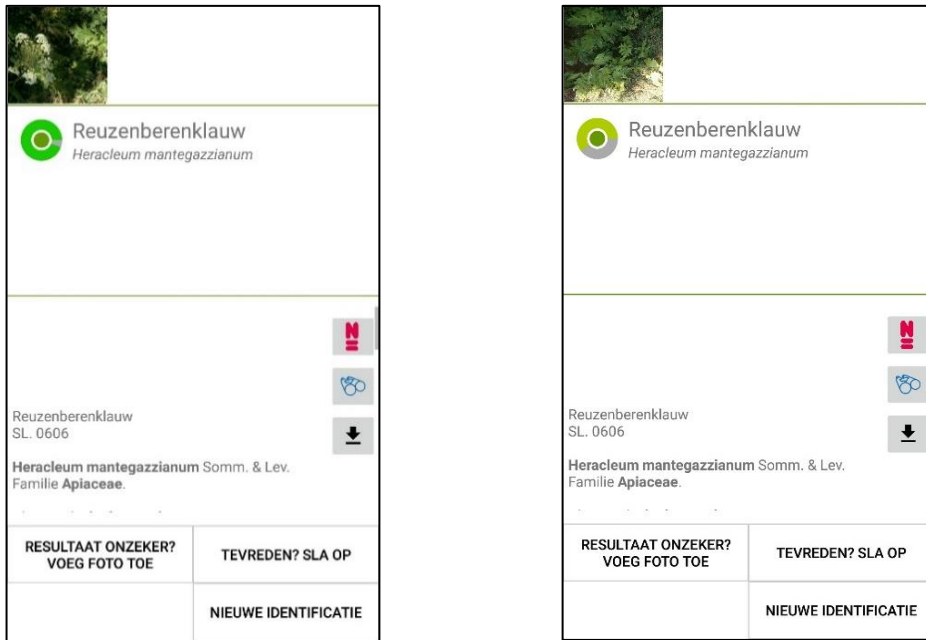
Figure 5.16 Screenshots of app results of flying height 8m in the shadow region that shows the maximum likelihood of *Heracleum mantegazzianum*. Left: image with umbels and leaves. Right: image with only leaves.
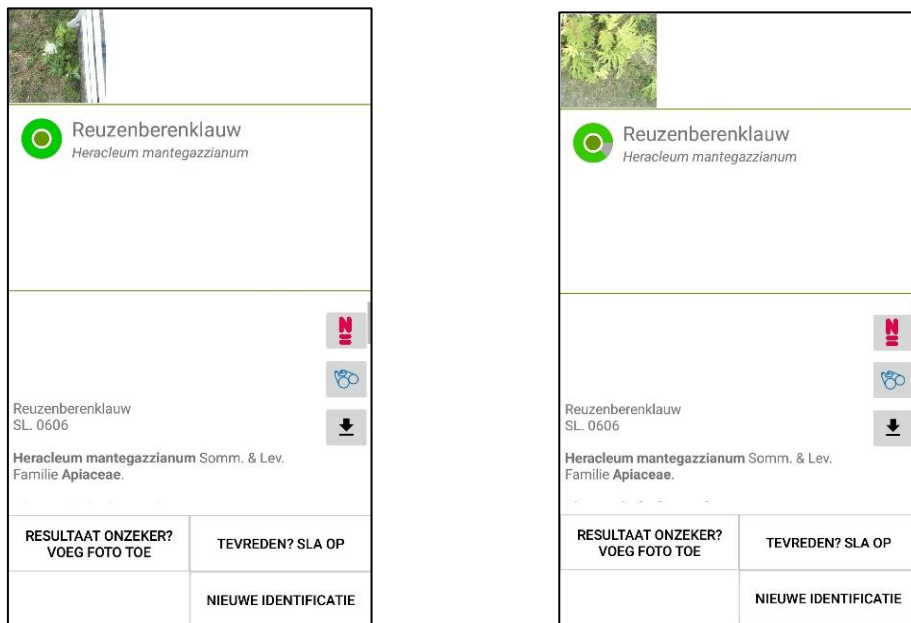
Figure 5.17 Screenshots of app results of flying height 5m with bright sun exposure that shows the maximum likelihood of *Heracleum mantegazzianum*. Left: image with umbels and leaves. Right: image with only leaves.
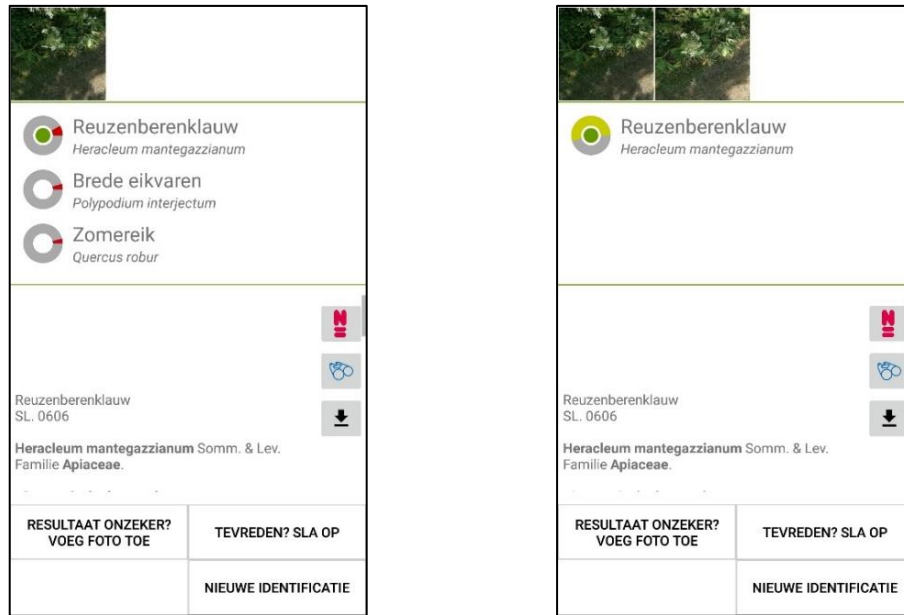
Figure 5.18 Left: Screenshot of app results of flying height 14 m with umbels and leaves in the shadow region that shows three possibilities of species with *Heracleum mantegazzianum* being the first choice. Right: Screenshot of app results of flying height 14 m with umbels and leaves in the shadow region with two different images of the same scene that shows a maximum likelihood of *Heracleum mantegazzianum*.
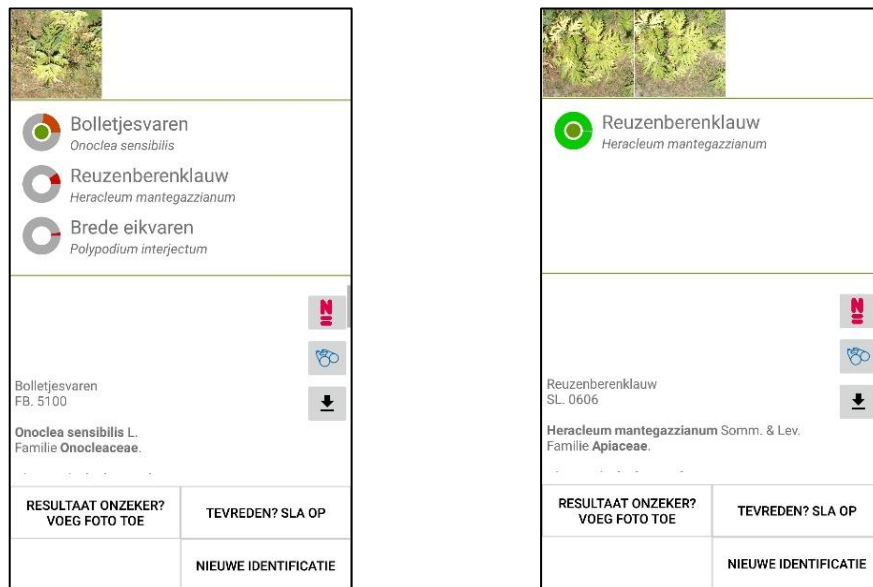
Figure 5.19 Left: Screenshots of app results of flying height 14 m with leaves in bright sun exposed region that shows three possibilities of species with *Heracleum mantegazzianum* being the second choice. Right: Screenshots of app results of flying height 14 m with leaves in the shadow region with two different images of the same scene that shows a maximum likelihood of *Heracleum mantegazzianum*.

6. DISCUSSION

In this chapter, a comparison of the accuracy assessment of the two methods adopted along with a detailed discussion on the results of both the methods, explaining the strength and weakness is presented.

6.1.   Comparison of OBIA and SVM accuracy assessment

This section discusses the comparison of the accuracy assessment of both methods adopted for this research namely: Object-Based Image Analysis and Support Vector Machine. The Producer's Accuracy (PA) for the Hogweed class and the OA (Overall Accuracy) from the confusion matrix are used as a measure for the assessment of both methods. The PA for Hogweed class is only discussed because of its relevance to this research and hence the PA for Not Hogweed class is not discussed. The OA and PA for both methods can be seen in Table 6.1.

Table 6.1 Producer's Accuracy of Hogweed class and Overall Accuracy for all the flying heights of OBIA and SVM including the results of the targeted analysis of shadow and sunlit areas for 11 m flying height as well.

Flying height (in m)	OBIA				SVM			
	Shadow + Sunlit		Only sunlit		Shadow + Sunlit		Only sunlit	
	PA	OA	PA	OA	PA	OA	PA	OA
11	80%	95%	62%	91%	84%	94%	64%	91%
12	84%	96%	-	-	94%	97%	-	-
14	86%	97%	-	-	90%	95%	-	-

It can be inferred from Table 6.1 that the OA of all the flying heights for both the methods is above 90%. SVM returned an error of commission less than 11% for all flying heights whereas, OBIA returned 0% error of commission. Introducing an area rule effectively reduced the error of commission to 0% in 11 m flying height of OBIA method. The low error of commission for both methods further supports that the spectral information of Giant Hogweed is distinct from its neighbourhood vegetation for successful detection. This distinct spectral information helped both the classifiers to accurately separate Hogweed from Not Hogweed. However, it should also be noted that the misclassification error within Not Hogweed class (Maize, Vegetation and Sand) was not included since the main objective of this research was to accurately detect Giant Hogweed from its neighbouring vegetation.

In terms of PA for Hogweed class, surprisingly SVM shows better results than OBIA. It was also found out that the effect of shadow plays an important role when classifying the image using spectral information. The inefficiency of OBIA classifier in the shadow region led to a PA of only 62%. Hence, a targeted OBIA analysis for the shadow region was done only for 11 m flying height. The performed targeted analysis using OBIA for 11 m flying height was successful because it increased the PA for the Hogweed class substantially from 62% to 80%. Also, the effect of shadow resulted in misclassification in the SVM method too since the spectral information of Hogweed in the shadow region is not distinct from its nearby vegetation. However, the SVM classifier classified Hogweed image objects in shadow region because of overfitting. Hence, a targeted analysis was carried out by removing the samples from the shadow region, which resulted in a PA of 64%. This results further support that the spectral variability between Hogweed image objects in the

shadow region and the sunlit region is very high. It can also be seen from the above table that the performance of both classifiers was almost similar when samples from the shadow region were not included for the training, which shows that for higher PA, the region affected by shadow should be treated separately from the sunlit region .

The effect of shadow is a universal problem, and it was encountered in one of the previous studies too. The change in luminosity and the presence of shadow due to the position of the sun caused high spectral variability within the classes which resulted in the decrease of classification accuracy (Michez et al., 2016).

6.2.   Strength and weakness of OBIA

OBIA was the first method adopted in the methodology because an object-based approach is more suitable for high-resolution data than the pixel-based approach. OBIA reduces within-class spectral variation that removes salt and pepper effects in images, which is a typical phenomenon in pixel-based approach (D. Liu & Xia, 2010), also seen in the SVM results of this research. Defining the rule sets for segmentation and classification is easy and straightforward. Even though defining the rule sets is easy, the optimum parameter is defined by a repeated trial and error method, which is very time-consuming. Moreover, the results of segmentation need to be checked by the user for every set of parameter values, which can lead to ambiguity, as there is no set of formal rules for a good segmentation (Hay & Castilla, 2006). Segmentation errors are one of the main disadvantages of OBIA segmentation. Two types of errors that often exist in a segmentation process are over-segmentation and under-segmentation, which was seen in this research as well. A study by Liu and Xia (2010), which assesses the advantages and limitations of OBIA for high-resolution data, points out that under-segmentation introduces error to the subsequent classification process no matter which classifier is used. In addition, features extracted from over-segmented and under-segmented segments do not represent the properties of real-world objects (e.g. shape and area). The effect of shadow in the flying height 11 m resulted in under-segmentation, which resulted in the inefficiency of OBIA to classify Hogweed image objects from the shadow region. This resulted in a very low PA of only 54% for flying height 11 m. By performing targeted analysis in the shadow region and sunlit region separately, the PA was increased from 54% to 72% for flying height 11 m.

One main advantage of using OBIA would be that it allows the user to add spatial information along with the spectral information for a more refined classification. In this research, an area rule was introduced to reclassify misclassified image objects, which were spectrally similar to Hogweed image objects. The misclassified image objects were from the maize region of the study area and were very small, and hence the rule was applicable. However, the area rule cannot be used if the misclassified image objects are similar in size to Hogweed image objects.

Moreover, OBIA uses segmented image objects as samples for the classification process. This can lead to a bias in the final classification results depending on the area of the study and the total number of image objects segmented. For example, selecting 50 samples from a total of 100 segmented image objects is not an efficient classification. For this reason, in this research, a total of 10 – 20 image objects were selected as samples for the OBIA classification.

The classification results from OBIA gave an OA of 97% for the 14 m flying height with a PA (Hogweed class) of 86%. It can be seen from Table 6.1 the classifier gives better accuracy as the flying height increases. The variability of high-resolution data decreases with an increase in height, which also is a factor that helped the software efficiently classify Giant Hogweed in the higher-flying heights. The main issue identified from using OBIA classifier was that the error of omission was high (38%) in the flying height 11 m, which was due to the effect of shadow and high spectral variability. The targeted analysis shows that separate analysis on the shadow region could be an answer for the inefficiency of OBIA classifier in the shadow region. Hence, the effect of shadow was not a limitation in this research. The spectral variability also resulted in

over-segmentation, which was successfully eradicated using an area rule. The spectral variability of 11 m is high compared to other flying heights even though the GSD of all three flying heights are almost similar. This is because the orthophotos were resampled from UAV images and the UAV images from low flying heights have more spectral variability when compared to the data from higher-flying heights. The resampling method resulted in almost similar GSD but retained high spectral variability for the lower flying height (11 m).

6.3. Strength and weakness of SVM

Several studies have shown that an object-based approach is desirable for high-resolution data, but interestingly SVM outperformed OBIA in this research in many aspects. The values from Table 6.1 shows the overview of the comparison between both the methods. It can be seen that the PA for the Hogweed class is higher for all flying heights in the SVM method. OBIA classifier did not return any false positives while SVM returned. This was due to the salt and pepper effect of SVM. However, the false positives did not affect the OA drastically since they were only less than 5 misclassified points. This shows that the SVM classifier was efficient in separating Hogweed pixels from Not Hogweed pixels.

Optimisation of parameters for the SVM model is one of the main limitations for using SVM. A kernel with fewer parameters was chosen in this study due to time constraint. Furthermore, processing the entire image in a loop for the parameter optimisation is time-consuming. A small portion of the image, which contains all the classes, can be chosen for parameter optimisation to overcome the situation. This, in turn, will speed up the process significantly.

One of the main advantages of using SVM is that it requires no human interference for the classification process. The process is automated, unlike OBIA which needs a user to validate both segmentation and classification results. This makes SVM less biased than OBIA. However, it should also be noted the overfitting of the SVM model is a disadvantage. Overfitting of the model was seen in flying height 11 m. However, increasing the size of the training samples could help in overcoming the overfitting of the model. In addition, like OBIA the sunlit region and shadow region can be analysed separately and hence, this is not a major weakness.

SVM gives salt and pepper effect as any pixel-based approach would. SVM like any other pixel-based approach considers only spectral properties and does not consider the relationship between spatial and spectral properties, which results in salt and pepper effect. However, this cannot be deemed as a limitation since there are several additional ways to eradicate the salt and pepper effect. Post classification smoothing algorithms such as majority filter (Huang et al., 2004) or likelihood class filter (Su, 2016) can be used to overcome the effect. This might increase the steps in the overall SVM classification process but also the OA would increase for all heights since the process would effectively decrease the commission error. Interestingly, the effect of shadow did not affect the accuracy for SVM as OBIA did, for flying height 11 m because the model tried to maximise the inter-class separability. Although it resulted in overfitting of the model, the PA and OA were as high as 84% and 94% respectively. The comparison of results suggests that SVM can be preferred for analysing shadow prone area with additional measures taken for overfitting as suggested.

6.4. Strength and Weakness of using UAV

The methodology as explained in chapter 3 was carried out using the spectral characteristics of the leaves of Giant Hogweed while the umbels were not used for detection at any point in this study. The success of the detection could be attributed to the use of different spectral bands like Green, NIR, and Red Edge instead of using only visible and NIR bands like in the previous studies (Michez et al., 2016; Müllerová et

al., 2017). As discussed in the introduction section UAVs are a great option since Giant Hogweed usually grows under trees and UAVs can be successfully used for understory mapping.

Although UAVs are the best option for the source of data, they also have some limitations. The area that can be covered by UAVs is significantly smaller than airborne sensors. Hence, UAVs are much suitable for monitoring the efficiency of eradication efforts or mapping a small infected area (Dvořák et al., 2015). In the Netherlands, UAVs are prohibited to be flown over a group of people or connected buildings. Giant Hogweed was introduced as an ornamental plant in Europe, and hence they can be seen in urban areas too. However, due to the law UAVs cannot be flown in an urban area, which is a major limitation for data acquisition in urban areas infected by Giant Hogweed.

6.5.   ObsIdentify for plants

ObsIdentify for plants, the deep learning based mobile application showed promising results in the detection of Giant Hogweed using UAV images. The application, which was trained using mobile pictures from the waarneming.nl website, had no difficulty in identifying UAV images. Furthermore, the app produced accurate results using images with only leaves and images with shadow. A possible explanation for the accurate detection of images with shadow can be attributed to the training of the application using bad pictures along with good pictures as explained in the methodology chapter. The algorithm has been trained to detect images with shadow as well. This opens up many opportunities for early detection of Giant Hogweed since the shadow is not a problem and only leaves are sufficient for accurate detection. Furthermore, deep learning could be a solution for the detection of images with shadow, and the data can be easily acquired from the waarneming.nl website for training the algorithm. The detection of Giant Hogweed using the ObsIdentify algorithm could be automated for further detection studies, but the main issue is feeding a tile of the image with very little to no neighbouring vegetation. As discussed in the results chapter, careful subset selection plays an important role in the accuracy of detection. Giant Hogweed had huge leaves and was growing away from the dense understory in this research study area, which made the subset selection easier. Moreover, it was done manually. If the Giant Hogweed were growing in a dense understory, the subset selection would have been difficult. It can be seen from the results of SVM and OBIA that as the flying height increases the overall accuracy decreases due to the decrease in spectral variability and effect of shadow. Whereas in the mobile application the accuracy of detection increases as the flying height decreases. One possible implication could be that the algorithm was trained using mobile pictures, which are normally taken at a very close range.

Future studies using the ObsIdentify algorithm is therefore suggested because of the promising results and the wide opportunities. The algorithm is already trained for all the wild plants in the Netherlands and Belgium and the results from this research show that UAV images can be used as an input. The algorithm can be used not only for Giant Hogweed detection but for other invasive plants detection as well. However, further studies involving the automation of subset selection for the detection should be undertaken.

7. CONCLUSION AND RECOMMENDATIONS

7.1. Conclusion

The main objective of this study was to find whether an early detection is possible by using the spectral characteristics of Giant Hogweed leaves, from its nearby vegetation. The results of this study show that early detection is possible with the help of spectral characteristics of leaves. This allows the detection of Giant Hogweed before the flowering period and does not limit the data acquisition to only flowering period. The second major finding was that the spatial context (OBIA) is not necessary for early detection; spectral information of Giant Hogweed's leaves are more than enough for successful detection using SVM. Despite the small study area, the study certainly adds to our understanding that spectral information of Giant Hogweed's leaves can be used as important information for early detection. The study was successful since Giant Hogweed was not spectrally similar to nearby vegetation. This may not be true for all cases. Hence, this is a limitation of this study. Further studies, which takes this variable into account, will need to be undertaken. As explained before the mapping of invasive species in the Netherlands, is done by volunteers. With the help of UAV images, the detection can be done effectively with less labour work. One of the more significant findings to emerge from this study is that UAV images can be used as an input for the deep learning based mobile application, ObsIdentify for plants and shadow does not pose as a problem. This would be a fruitful area for further work.

7.2. Research Questions: Answered

This section discusses the answers to the research questions from section 1.2.

1. Is early detection of Giant Hogweed possible by using the spectral information of leaves?
    The high OA of above 90% (see section 5.4.2) of SVM method, which uses only spectral information, shows that the early detection is possible by using only spectral information of Giant Hogweed's leaves.

2. What is the accuracy of the SVM method?
    The OA of SVM is 94%, 97% and 95% for 11 m, 12 m and 14 m flying heights respectively. In addition, the OA for the SVM method without training samples from the shadow region is 91%

3. Which extra parameters (shape, size and or texture) is used in addition to spectral information? Did it increase the accuracy of the detection in the OBIA method?
    OBIA adds spatial information in addition to spectral information for the classification; hence there was no salt and pepper effect like SVM. An area rule was also introduced, which resulted in zero commission error; however, there was no significant increase in OA when compared with SVM. The results of this research also proved that spectral information is enough for successful detection.

4. What is the accuracy of OBIA method?
    The OA of OBIA is 91%, 96% and 97% for 11 m, 12 m and 14 m respectively. In addition, the OA for 11 m with Hogweed image objects from shadow and sunlit region is 95%.

5. Can UAV images be used as an input for a deep learning based species recognition mobile application (ObsIdentify for plants)?
    Yes, UAV images can be used as an input for deep learning based mobile app, which was trained using pictures from mobile.

6. How does the altitude of the flight influence the accuracy of detection of the plant in SVM?
   From the results, it can be seen that the altitude does not influence the accuracy of detection. All three flying heights have similar OA.

7. How does the altitude of the flight influence the accuracy of detection of the plant in OBIA?
   The over-segmentation of OBIA suggest that the flying height influences the accuracy of detection. This influence can be attributed to the inefficiency of OBIA handling spectral variability within classes as the flying height decreases. However, additional measures like rule-based classification can be used to overcome the effect of flying height.

8. How does the altitude of the flight influence the accuracy of detection of the plant in the mobile application?
   The altitude of the flight plays an important role in the detection of the plant using ObsIdentify. From the results, it can be seen that the app successfully detected Giant Hogweed in low flying heights and as the flying height increased the app had difficulty. This is due to the training of the algorithm using mobile pictures from the website waarneming.nl, which are always taken at a close range.

9. Does the shadow affect the accuracy of detection in OBIA?
   The effect of shadow is a weakness for OBIA classifier since it leads to under-segmentation, which in turn leads to the inefficiency of classifier classifying Hogweed image objects in the shadow region. However, separate analysis on shadow and the sunlit region was found to increase the accuracy.

10. Does the shadow affect the accuracy of detection in SVM?
    Overall, the results indicate that shadow does affect the accuracy even though the degree of influence is not as significant as OBIA.

11. Does the shadow affect the accuracy of detection in the mobile application?
    In low flying heights like less than 11 m, the shadow does not affect the accuracy of detection. However, as the flying height increases the shadow also influences the accuracy detection along with flying height. This is because the images used for training the algorithm includes bad pictures and pictures taken in the shaded region.

12. Between the two methods, which is more efficient in terms of accuracy and computation time?
    Both methods give almost similar accuracy bur from the results and discussion chapter; it can be seen that SVM is the more efficient than OBIA in this research because,
    - The PA for Hogweed class is high in SVM for all flying heights.
    - Fully automated classification with little human interference.
    - Less overall computation time and less bias in final classification when compared with OBIA.
    - SVM classifies Hogweed in shadow region unlike OBIA which needs targeted analysis

7.3. Recommendation

The following suggestions are recommended for future work.

- The study should be repeated using different other plants to see whether there are any plants, which is spectrally similar to Giant Hogweed's leaves. In that case, a new parameter must be defined which can help the user to differentiate Giant Hogweed from other plants.
- In addition, a natural progression of this work would be to find a solution as a part of the analysis for the effect of shadow rather than targeted analysis. For example, a spatial context can be defined

which can help Hogweed image objects from both shadow and sunlit region to be classified without segmenting the study area into two separate regions.

- A further study on using the algorithm from deep learning based app ObsIdentify could be very fruitful research for invasive plant detection as discussed in the discussion chapter.

LIST OF REFERENCES

Ahmed, O. S., Shemrock, A., Chabot, D., Dillon, C., Williams, G., Wasson, R., & Franklin, S. E. (2017). Hierarchical land cover and vegetation classification using multispectral data acquired from an unmanned aerial vehicle. *International Journal of Remote Sensing*, *38*(8–10), 2037–2052. https://doi.org/10.1080/01431161.2017.1294781

Arel, I., Rose, D., & Karnowski, T. (2010). Deep machine learning-A new frontier in artificial intelligence research. *IEEE Computational Intelligence Magazine*, *5*(4), 13–18. https://doi.org/10.1109/MCI.2010.938364

Assmann, J. J., Kerby, J. T., Cunliffe, A. M., & Myers-Smith, I. H. (2019). Vegetation monitoring using multispectral sensors — best practices and lessons learned from high latitudes. *Journal of Unmanned Vehicle Systems*, *7*(1), 54–75. https://doi.org/10.1139/juvs-2018-0018

Bakx, W., Janssen, L., Schetselaar, E., Tempfli, K., Tolpekin, V., & Westinga, E. (2013). The core of GIScience. In V. A. Tolpekin & A. Stein (Eds.), *The core of GIScience: a systems based approach* (pp. 205–225). University of Twente Faculty of Geo-Information and Earth Observation (ITC). Retrieved from https://www.itc.nl/library/education/core-of-giscience/

Ben-Hur, A., & Weston, J. (2010). A User's Guide to Support Vector Machines. In O. Carugo & F. Eisenhaber (Eds.), *Data Mining Techniques for the Life Sciences. Methods in Molecular Biology (Methods and Protocols)* (pp. 223–239). Humana Press. https://doi.org/10.1007/978-1-60327-241-4_13

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004

Borenstein, E., & Ullman, S. (2008). Combined top-down/bottom-up segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(12), 2109–2125. https://doi.org/10.1109/TPAMI.2007.70840

Caffrey, J. M. (1999). Phenology and long-term control of Heracleum mantegazzianum. *The International Journal of Aquatic Sciences*, *415*, 223–228. https://doi.org/10.1023/A:1003854221931

Caffrey, J. M. (2001). Management of giant hogweed (Heracleum mantegazzianum) in an Irish River catchment. *Journal of Aquatic Plant Management*, *39*, 28–33. Retrieved from http://www.apms.org/japm/vol39/v39p28.pdf

Calleja, F., Ondiviela, B., Galván, C., Recio, M., & Juanes, J. A. (2019). Mapping Estuarine Vegetation Using Satellite Imagery: The case of the invasive species Baccharis halimifolia at a Natura 2000 site. *Continental Shelf Research*, *174*, 35–47. https://doi.org/10.1016/J.CSR.2019.01.002

Casagrande, G., Sik, A., & Szabó, G. (2018). *Small Flying Drones Applications for Geographic Observation*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-66577-1_1

Chakraborty, D., Singh, S., & Dutta, D. (2017). Segmentation and classification of high spatial resolution images based on Hölder exponents and variance. *Geo-Spatial Information Science*, *20*(1), 39–45. https://doi.org/10.1080/10095020.2017.1307660

David, M., Evgenia, D., Kurt, H., Andreas, W., & Leisch, F. (2018). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-0. Retrieved from https://cran.r-project.org/package=e1071

Dvořák, P., Müllerová, J., Bartaloš, T., & Brůna, J. (2015). Unmanned aerial vehicles for alien plant species detection and monitoring. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *XL-1/W4*, 83–90. https://doi.org/10.5194/isprsarchives-XL-1-W4-83-2015

Early, R., Bradley, B. A., Dukes, J. S., Lawler, J. J., Olden, J. D., Blumenthal, D. M., … Tatem, A. J. (2016). Global threats from invasive alien species in the twenty-first century and national response capacities. *Nature Communications*, *7*(1), 12485. https://doi.org/10.1038/ncomms12485

Esri. (2018). Accuracy Assessment for Image Classification. Retrieved November 12, 2018, from http://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/accuracy-assessment-for-image-classification.htm

European Commission. (2015). List of invasive alien species of Union concern. Retrieved January 24, 2019, from http://ec.europa.eu/environment/nature/invasivealien/list/index_en.htm

European Union. (2012). EUR-Lex Access to European Union law. Retrieved March 13, 2019, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32014R1143

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*.

Elsevier. https://doi.org/10.1016/S0034-4257(01)00295-4

Gao, J., Liao, W., Nuyttens, D., Lootens, P., Vangeyte, J., Pižurica, A., … Pieters, J. G. (2018). Fusion of pixel and object-based features for weed mapping using unmanned aerial vehicle imagery. *International Journal of Applied Earth Observation and Geoinformation*, *67*, 43–53. https://doi.org/10.1016/j.jag.2017.12.012

Gavier-Pizarro, G. I., Kuemmerle, T., Hoyos, L. E., Stewart, S. I., Huebner, C. D., Keuler, N. S., & Radeloff, V. C. (2012). Monitoring the invasion of an exotic tree (Ligustrum lucidum) from 1983 to 2006 with Landsat TM/ETM+ satellite data and Support Vector Machines in Córdoba, Argentina. *Remote Sensing of Environment*, *122*, 134–145. https://doi.org/10.1016/j.rse.2011.09.023

Ghorbani, F., Ebadi, H., & Sedaghat, A. (2019). Geospatial Target Detection from High-Resolution Remote-Sensing Images Based on PIIFD Descriptor and Salient Regions. *Journal of the Indian Society of Remote Sensing*, 1–13. https://doi.org/10.1007/s12524-019-00944-4

Government of Canada. (2004). Ontario's Invading Species Awareness Program. Retrieved July 26, 2018, from http://www.invadingspecies.com/

Haque, M. E., Al-Ramadan, B., & Johnson, B. A. (2016). Rule-based land cover classification from very high-resolution satellite image with multiresolution segmentation. *Journal of Applied Remote Sensing*, *10*(3), 036004. https://doi.org/10.1117/1.JRS.10.036004

Hay, G., & Castilla, G. (2006). Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT). In *Proceedings of 1st International Conference on Object-based Image Analysis (OBIA 2006)* (Vol. XXXVI-4/C4, p. 3). Salzburg University, Austria. Retrieved from https://www.isprs.org/proceedings/XXXVI/4-C42/

Herremans, M. (2017). Nieuwe app is insectenkenner in je broekzak. Retrieved January 23, 2019, from https://www.natuurpunt.be/nieuws/nieuwe-app-insectenkenner-je-broekzak-20170915

Hogeweg, L. (2018). Ervaringen met planten. Retrieved January 23, 2019, from https://forum.waarneming.nl/smf/index.php?topic=412192.0

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2016). A Practical Guide to Support Vector Classification. *BJU International*, *101*(1), 1396–1400. https://doi.org/10.1177/02632760022050997

Huang, C., & Asner, G. P. (2009). Applications of remote sensing to alien invasive plant studies. *Sensors (Switzerland)*. Molecular Diversity Preservation International. https://doi.org/10.3390/s90604869

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*(4), 725–749. https://doi.org/10.1080/01431160110040323

Huang, H., Legarsky, J. J., Gudimetla, S., & Davis, C. H. (2004). Post-classification smoothing of digital classification map of St. Louis, Missouri. In *IEEE International IEEE International IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004* (Vol. 5, pp. 3039–3041). IEEE. https://doi.org/10.1109/igarss.2004.1370338

ISAP. (2018). Giant Hogweed – Ontario's Invading Species Awareness Program. Retrieved July 26, 2018, from http://www.invadingspecies.com/giant-hogweed/

Jakuboski, S. (2011). CAUTION: Giant Hogweed might be growing in your own backyard. Retrieved August 26, 2018, from https://www.nature.com/scitable/blog/green-science/caution_giant_hogweed_might_be

Jones, D., Pike, S., Thomas, M., & Murphy, D. (2011). Object-based image analysis for detection of Japanese Knotweed s.l. taxa (polygonaceae) in Wales (UK). *Remote Sensing*, *3*(2), 319–342. https://doi.org/10.3390/rs3020319

Kavzoglu, T., & Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, *11*(5), 352–359. https://doi.org/10.1016/J.JAG.2009.06.002

Lichtblau, E., & Oswald, C. J. (2019). Classification of impervious land-use features using object-based image analysis and data fusion. *Computers, Environment and Urban Systems*, *75*, 103–116. https://doi.org/10.1016/J.COMPENVURBSYS.2019.01.007

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2014). *Remote sensing and image interpretation*. Wiley. Retrieved from https://books.google.nl/books?hl=en&lr=&id=AFHDCAAAQBAJ&oi=fnd&pg=PA1&ots=0Crtf ZfCcu&sig=9Xg1WPT8-z9D2ws3lEDZ69IntN0&redir_esc=y#v=onepage&q&f=false

Liu, D., & Xia, F. (2010). Assessing object-based classification: Advantages and limitations. *Remote Sensing*

Letters, *1*(4), 187–194. https://doi.org/10.1080/01431161003743173

Liu, R., & Gillies, D. F. (2016). Overfitting in linear feature extraction for classification of high-dimensional image data. *Pattern Recognition*, *53*, 73–86. https://doi.org/10.1016/j.patcog.2015.11.015

Maheu-Giroux, M., & De Blois, S. (2005). Mapping the invasive species Phragmites australis in linear wetland corridors. *Aquatic Botany*, *83*(4), 310–320. https://doi.org/10.1016/j.aquabot.2005.07.002

Manfreda, S., McCabe, M., Miller, P., Lucas, R., Pajuelo Madrigal, V., Mallinis, G., … Toth, B. (2018). On the Use of Unmanned Aerial Systems for Environmental Monitoring. *Remote Sensing*, *10*(4), 641. https://doi.org/10.3390/rs10040641

Mantero, P., Moser, G., & Serpico, S. B. (2004). Partially supervised classification of remote sensing images using SVM-based probability density estimation. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(3), 327–336. https://doi.org/10.1109/WARSD.2003.1295212

Mathieu, R., Aryal, J., & Chong, A. K. (2007). Object-based classification of ikonos imagery for mapping large-scale vegetation communities in urban areas. *Sensors*, *7*(11), 2860–2880. https://doi.org/10.3390/s7112860

MicaSense. (2017). What does the Sunshine (Irradiance) Sensor on Sequoia do? – MicaSense Knowledge Base. Retrieved February 4, 2019, from https://support.micasense.com/hc/en-us/articles/216533688-What-does-the-Sunshine-Irradiance-Sensor-on-Sequoia-do-

Michez, A., Piégay, H., Jonathan, L., Claessens, H., & Lejeune, P. (2016). Mapping of riparian invasive species with supervised classification of Unmanned Aerial System (UAS) imagery. *International Journal of Applied Earth Observation and Geoinformation*, *44*, 88–94. https://doi.org/10.1016/j.jag.2015.06.014

Moenickes, S., & Thiele, J. (2013). What shapes giant hogweed invasion? Answers from a spatio-temporal model integrating multiscale monitoring data. *Biological Invasions*, *15*(1), 61–73. https://doi.org/10.1007/s10530-012-0268-z

Möller, M., Lymburner, L., & Volk, M. (2007). The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation*, *9*(3), 311–321. https://doi.org/10.1016/j.jag.2006.10.002

Moravcová, L., Perglová, I., Pyšek, P., Jarošík, V., & Pergl, J. (2005). Effects of fruit position on fruit mass and seed germination in the alien species Heracleum mantegazzianum (Apiaceae) and the implications for its invasion. *Acta Oecologica*, *28*(1), 1–10. https://doi.org/10.1016/j.actao.2005.01.004

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*. Elsevier. https://doi.org/10.1016/j.isprsjprs.2010.11.001

Mufalli, F., Batta, R., & Nagi, R. (2012). Simultaneous sensor selection and routing of unmanned aerial vehicles for complex mission plans. *Computers and Operations Research*, *39*(11), 2787–2799. https://doi.org/10.1016/j.cor.2012.02.010

Müllerová, J., Brůna, J., Bartaloš, T., Dvořák, P., Vítková, M., & Pyšek, P. (2017). Timing Is Important: Unmanned Aircraft vs. Satellite Imagery in Plant Invasion Monitoring. *Frontiers in Plant Science*, *8*, 887. https://doi.org/10.3389/fpls.2017.00887

Müllerová, J., Pergl, J., & Pyšek, P. (2013). Remote sensing as a tool for monitoring plant invasions: Testing the effects of data resolution and image classification approach on the detection of a model plant species Heracleum mantegazzianum (giant hogweed). *International Journal of Applied Earth Observation and Geoinformation*, *25*(1), 55–65. https://doi.org/10.1016/j.jag.2013.03.004

Müllerová, J., Pyšek, P., Jarošík, V., & Pergl, J. (2005). Aerial photographs as a tool for assessing the regional dynamics of the invasive plant species Heracleum mantegazzianum. *Journal of Applied Ecology*, *42*(6), 1042–1053. https://doi.org/10.1111/j.1365-2664.2005.01092.x

Nielsen, C., Ravn, H. P., Nentwig, W., Wade, M., & Nentwig, W. (2005). *The Giant Hogweed Best Practice Manual Guidelines for the management and control of an invasive weed in Europe About the publication Editors: Citation*. Retrieved from www.giant-alien.dk

Otte, A., & Franke, R. (1998). The ecology of the Caucasian herbaceous perennial Heracleum mantegazzianum Somm. et Lev. (Giant Hogweed) in cultural ecosystems of Central Europe. *Phytocoenologia*, *28*(2), 205–232. https://doi.org/10.1127/phyto/28/1998/205

Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, *26*(5), 1007–1011. https://doi.org/10.1080/01431160512331314083

Perglová, I., Pergl, J., & Pysek, P. (2007). Reproductive ecology of Heracleum mantegazzianum. In P. Pyšek, M. J. W. Cock, W. Nentwig, & H. P. Ravn (Eds.), *Ecology and management of giant hogweed (Heracleum mantegazzianum)* (pp. 55–73). https://doi.org/10.1079/9781845932060.0055

Perglová, I., Pergl, J., & Pyšek, P. (2006). Flowering phenology and reproductive effort of the invasive alien plant Heracleum mantegazzianum. *Preslia*, *78*(3), 265–285. Retrieved from http://www.ibot.cas.cz/preslia

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.r-project.org/

Royimani, L., Mutanga, O., Odindi, J., Dube, T., & Matongera, T. N. (2018). Advancements in satellite remote sensing for mapping and monitoring of Alien Invasive Plant species (AIPs). *Physics and Chemistry of the Earth, Parts A/B/C*. https://doi.org/10.1016/j.pce.2018.12.004

Schöpfer, E., Lang, S., & Strobl, J. (2010). Segmentation and object-based image analysis. In T. Rashed & C. Jürgens (Eds.), *Remote Sensing of Urban and Suburban Areas* (pp. 181–192). Springer Netherlands. https://doi.org/10.1007/978-1-4020-4385-7_10

Strien, W. van. (2018). The flora and fauna in the pocket - Bionieuws. Retrieved February 11, 2019, from https://bionieuws.nl/artikelen/de-flora-en-fauna-in-de-broekzak/

Su, T. C. (2016). A filter-based post-processing technique for improving homogeneity of pixel-wise classification data. *European Journal of Remote Sensing*, *49*, 531–552. https://doi.org/10.5721/EuJRS20164928

Talha, M., Ali, S., Shah, S., Khan, F. G., & Iqbal, J. (2019). Integration of Big Data and Deep Learning. In *Deep Learning: Convergence to Big Data Analytics* (pp. 43–52). Springer, Singapore. https://doi.org/10.1007/978-981-13-3459-7_4

Tamaki, T., Yamamura, T., & Ohnishi, N. (1999). *Image segmentation and object extraction based on geometric features of regions* (Vol. 3653). Retrieved from http://www.ohnishi.nuie.nagoya-u.ac.jp/

Trimble. (2012). Introduction — eCognition Community. Retrieved January 27, 2019, from http://community.ecognition.com/home/introduction

Trimble. (2015). eCognition Developer. Retrieved January 27, 2019, from http://www.ecognition.com/suite/ecognition-developer

Tso, B., & Mather, P. M. (2009). *Classification methods for remotely sensed data*. CRC Press. Retrieved from https://epdf.tips/classification-methods-for-remotely-sensed-data62966.html

Vapnik, V. N. (2006). The Nature of Statistical Learning Theory. *Technometrics*, *38*(4), 409. https://doi.org/10.2307/1271324

Waarneming.nl. (2019). Statistieken. Retrieved January 24, 2019, from https://waarneming.nl/stats/

Wei, W., Chen, X., & Ma, A. (2005). Object-oriented information extraction and application in high-resolution remote sensing image. In *2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.* (Vol. 6, pp. 3803–3806). https://doi.org/10.1109/IGARSS.2005.1525737

Whiteside, T. G., Boggs, G. S., & Maier, S. W. (2011). Comparing object-based and pixel-based classifications for mapping savannas. *International Journal of Applied Earth Observation and Geoinformation*, *13*(6), 884–893. https://doi.org/10.1016/j.jag.2011.06.008