Extraction of morphological and spectral features of potato plants from high resolution multispectral images

Jorge Luis Rodríguez Galvis

Universidad Nacional de Colombia
Facultad de Ciencias Agrarias
Bogotá D.C., Colombia
2020

Extraction of morphological and spectral features of potato plants from high resolution multispectral images

Jorge Luis Rodríguez Galvis

A Thesis submitted as partial fulfilment of the degree of Master in Geomatics
at Universidad Nacional de Colombia

Director:
Ph.D. Ivan Alberto Lizarazo Salcedo
Codirector:
Ph.D. Flavio Augusto Prieto Ortíz

Research line:
Geospatial technologies

Research groups:
Análisis Espacial del Territorio y del cambio global (AET-CG)
Grupo de Automática de la Universidad Nacional (GAUNAL)

Universidad Nacional de Colombia
Facultad de Ciencias Agrarias
Bogotá D.C., Colombia
2020

Abstract

This work studies the use of spectral and morphological features in the evaluation and detection of potato late blight using very high resolution multispectral images captured by Unmanned Aerial Vehicles (UAV). Traditional late blight detection and mapping methods are time-consuming, require great human effort and, in many cases, are subjective. The study of the geometric and spectral characteristics of potato plants by means of UAV can contribute to improving the efficiency of the field detection systems that are currently used. This research seeks to contribute to the determination of the capture, processing and analysis methods of the data acquired through UAV in a way that provides producers with reliable tools for the improvement and management of their crops in an agile and efficient way. The approach of this study integrates morphological operations and evaluates the performance of five machine learning algorithms: Random Forest (RF), Gradient Boosting classifier (GBC), Support Vector Classifier (SVC), Linear Support Vector Classifier (LSVC) and K- Nearest Neighbours (KNN) to detect late blight areas. The main components of the proposed approach are: (i) radiometric and geometric correction of raw images; (ii) elimination of bare soil by applying a thresholding technique; (iii) the generation of spectral indices; (iv) the construction of morphological features of the plants; (v) a supervised classification procedure using ML algorithms; and (vi) use of pre-trained models to classify a new data set. The performance of the method is evaluated on two dates in an experimental potato field. The results showed that the LSVC and RF classifiers performed the best in terms of accuracy and execution time metrics. The study showed that the proposed method allows the detection of late blight with little human intervention.

Acknowledgements

First, I would like to thank my supervisors, professor Iván Lizarazo and Flavio Prieto, who provided me with detailed feedback and guided me through this project. Secondly, I want I thank Prof. Luis Ernesto Rodriguez for inviting me to participate in the potato crop nutrient program. I gratefully acknowledge the strong sense of commitment and hard work from the SIGA research team at Universidad Nacional, in particular the performance of the UAV's pilot, flight engineer, and crew. Also, I thank the reviewers' comments which have helped me to enhance the quality of the document.

I want to especially thank Víctor Angulo for his support in planning the field work, as well as for his suggestions and recommendations throughout the development of this project. Also, to Omar Cristancho for his valuable help regarding agronomy issues and his support in the field missions.

Contents

List of Figures

List of Tables

Glossary

AI Artificial Intelligence. 13

ANN Artificial Neural Network. 14

AUC Area Under the Curve. 58, 62

CIRE Red-edge Chlorophyll index. 29

CNN Convolutional Neural Networks. 2

CV Computer vision. 13, 14

DIP Digital Image Processing. 13

DN Digital Number. 8

DNA Deoxyribonucleic acid. 3, 7

DSMs Digital Surface Models. 17

EVI Enhanced vegetation index. 12, 29

EVI2 Two-band enhanced vegetation index. 29

G-loss Generalisation Loss. 43, 70, 72

GBC Gradient Boost Classifier. VII, 42, 55, 58, 62, 65, 68

GNDVI Green normalized difference vegetation index. 12, 29

GSD Ground Sample Distance. 34, 35

IoU Intersection-over-Union. 44

KNN K-Nearest Neighbours. 43, 55, 58, 59, 62, 67

LA Leaf Area. 16

LAI Leaf area index. 12, 29

LB Late Blight. 1, 4, 5, 46, 54, 72

LBB Leaf Bounding Box. 16

LL Leaf Length. 16

LN Leaf Number. 16

LSVC Linear Suppor Vector Classification. V, 40, 43, 46, 48, 49, 50, 51, 53, 55, 56, 58, 59, 62, 67, 74

LW Leaf Width. 16

LWLR Leaf Width-Lenght Ratio. 16

MAE Mean Absolute Error. 2, 62

ML Machine Learning. VIII, 2, 70, 71, 74

MST Minimum Spanning Trees. 15

MV Machine vision. 13, 14

NDVI Normalized Difference Vegetation Index. 12, 29

NDVIRE Red-edge normalized difference vegetation index. 29

PA Plant Area. 16

PAR Photosynthetically Active Radiation. 12

PBB Plant Bounding Box. 16

PDE Partial Differential Equation. 14

PH Plant Height. 16

PR Pattern Recognition. 13

RF Random Forest. V, 2, 17, 29, 37, 42, 46, 48, 50, 52, 55, 59, 62, 65

RMSE Root Mean Squared Error. 62

RV Radiometric Values. 8

SAVI Soil adjusted vegetation index. 12, 29

SCP Semi-Automatic Classification Plugin. 44

SfM Structure from Motion. 1

SVC Support Vector Classification. 42, 55, 56, 62

TLF Total Leaf Expansion. 16

UAV Unmanned Aerial Vehicle. 1, 2, 4, 5, 12, 16, 17, 27, 70, 72

VIs Vegetation Indices. 12, 29, 31, 37

1 Introduction

The study of the characteristic features of plants based on images has great potential in tasks of detailed characterization, leaf segmentation or in the identification of mutations. However, the development of capture methods and analysis techniques poses new challenges, such as data acquisition in the field, where conditions are not controlled and the handling of large volumes of information derived from capture processes is required (Granier and Vile, 2014). This project evaluates the use of multispectral images of very high spatial resolution in determining geometric and spectral features in a variety of potato, susceptible to Late Blight (LB) (*Phytophthora infestans*), with different agronomic management. Different means of image capture will be used, including an Unmanned Aerial Vehicle (UAV) to capture very high resolution images that were subjected to a photogrammetric adjustment based on Structure from Motion (SfM) algorithms to obtain orthomosaics, to which machine vision methods were applied to extract differential characteristics between potato plants, in terms of spectral responses and morphological patterns. In this document, the problem under study is described, a literature review is made, the objectives, the work method, the conducted results, the work schedule and the project resources are described.

Precision agriculture and plant phenotyping as information and technology based fields require incorporating new approaches into traditional monitoring and rating systems to accomplish different needs and challenges (Mahlein, 2016a). Assessment of diseases in crops based on remote sensing techniques allows consistent monitoring of field crop health (Shakoor et al., 2017). Conventional methods are based on intensive field sampling, where samples are collected to analyse the attributes of the crop. Traditional inspection is usually performed by trained experts according to characteristic plant disease symptoms or evident signs of a pathogen (Mahlein, 2016a). However, a significant human effort is needed to successfully cover a large area (EPPO, 2007). Current technological advancements, including wireless sensor networks, proximal sensor and remote sensors, are potentially important tools for farmers switching to environmentally friendly farming methods and cost-effective farming practices (Srbinovska et al., 2015; Tokekar et al., 2016).

Late blight (LB), is a mayor disease in potato crops, caused by *Phytopththora infestans* (Mont.) de Bary, which damages the plant foliage, leading to low productivity and economic losses (Jiang et al., 2012; Arora et al., 2014; Kuhl et al., 2007). Significant efforts have been made to improve potato crop resistance Śliwka et al. (2012), and implement socio-ecological-system approaches for disease management (Pacilly et al., 2016). However, severity of the LB seems to be increasing in different potato growing areas due to *Phythoptora infestans* ability to adapt and increase tolerance to the most potent fungicides (Arora et al., 2014; Wiik, 2014).

Real time Polymerase Chain Reaction (PCR) technique for detection of *Phytopththora infestans* oospores in soil have been developed (Hussain et al., 2014). This technique gives

farmers a reliable tool for identifying the location of the pathogen and the corresponding management strategies. While such approach is accurate, it relies on sophisticated equipment and expertise, which are very expensive for smallholder farmers in developing countries.

Assessment of LB disease impact is usually performed by periodic surveys in the field through visual inspection on the growing crops (EPPO, 2007). Such assessments are often inaccurate and subjective (Sugiura et al., 2016). Traversing the entire crop field for visual inspection of LB presence involve significant time and human effort. It requires walking in certain patterns, searching for specific symptoms of disease in stems, leaves or roots (Forbes et al., 2014).

The use of unmanned aerial vehicles (UAV) in agriculture poses important advantages as they are often low cost platforms that can be equipped with RGB, multispectral, or hyperspectral cameras for the rapid acquisition of high resolution images for crop loss assessment (Zhao and Jiang, 2010). UAVs have been used in agriculture for water stress assessment (Gago et al., 2015), monitoring of NDVI index during the wheat growth (Hassan et al., 2018), assessment of crop hail damage in potato crops (Zhou et al., 2016), diagnosis of nitrogen status (Liu et al., 2018a) and rice lodging assessment ((Liu et al., 2018b)).

In recent years, several image-based techniques for monitoring diseases severity using imagery from UAVs have been proposed. Sugiura et al. (2016) proposed a phenotyping system for mapping potato LB through analysis of pixel change between consecutive periods of image capture. Although this approach is accurate, it is necessary to have images from at least two periods to detect affected zones. Duarte-Carvajalino et al. (2018) performed the assessment of LB severity in potato from UAV-based high resolution multispectral images acquired with a low-cost camera and machine learning algorithms. More recently Franceschini et al. (2019) evaluated the feasibility of using UAV multispectral imagery for early detection and severity assessment of late blight in potato crops, their experimental results showed that optical data acquired at canopy level with centimetre resolution has the potential to provide useful information for detecting late blight incidence and assessing its severity in early stages of disease development.

Vegetation indices as well as supervised and unsupervised classification methods and clustering such as artificial neural networks, decision trees, support vector machines or k-means, have been used in precision agriculture and can be effective for detection, identification, and quantification of plant diseases from sensor data (Franke and Menz, 2007; Zhang et al., 2019; Bagheri, 2020). It is usual to separate plants from other objects when assessing plants diseases. Threshold-based approaches have been used to separate plants from bare soil, many colour-index based approaches for plant segmentation make use of either zero threshold or a threshold based on Otsu's method (Hamuda et al., 2016a).

ML methods have recently been applied in the study of LB in potato crops. Duarte-Carvajalino et al. (2018) carried out an evaluation of the LB in 14 different potato genotypes using multispectral images captured with UAV and ML methods, such as multi layer perceptron, deep learning Convolutional Neural Networks (CNN), support vector regression, and RF. Their results suggested that it is possible to replace the visual estimation of LB disease severity with ML algorithms, as the Mean Absolute Error (MAE) found was 11.72%, which was found acceptable.

1.1 Problem statement

Kumar et al. (2015) and Singh et al. (2016), estimate that agricultural production must double by 2025 if the demand of the world's growing population is to be met. Genetic studies have great potential in understanding crop yield and resistance to different types of stress through the use of technological resources such as genotyping and sequencing, however, the success of the application of genomics depends on the ability to accurately assess complex plant traits (Kumar et al., 2015). Plant phenotyping is the quantitative evaluation of the traits of a genotype given an environment and experiment, in terms of its scalar (plant height), multivalued (chemical and Deoxyribonucleic acid (DNA) transcription) characteristics, even those based on in images, which includes measurements made directly, as well as indirect measurements (Bolger et al., 2017). Studying how phenotypes change in plants of the same species in different environments is essential to understand their operation, this study is based on the detailed classification of plant properties, such as the number of leaves, their arrangement, the state of maturation or the similarity with another variety (Minervini et al., 2016).

Difficulties in the study of phenotypes in the field limit the analysis of quantitative genetic characteristics, in particular those related to the resistance of plants to stress and their yield (Araus and Cairns, 2014). Plant phenotyping has emerged in recent years due to the advancement of high-performance sensors for image capture (Coppens et al., 2017), since its use avoids the application of invasive or destructive study methods. Image-based plant phenotyping techniques have gained ground in tasks where identification is needed for detailed categorization procedures: such as segmentation of leaves in different crops or in identification of mutations (Minervini et al., 2016). However, the development of digital image analysis methods and techniques in phenotyping studies poses new challenges, one of them is the handling of large volumes of information derived from data capture processes (Coppens et al., 2017).

1.2 Conceptual Approach

The study of diseases in potato crops has been of great importance in the world, in fact, the Royal Horticultural Society, founded in 1805, was created after considerable losses in crops between the years 1770 and 1800. Its main The purpose was to acquire wild potato varieties from which they have been derived from the affected varieties and begin to reproduce them in order to study the problems and suffered during that time (John Reader, 2009). The inspection of potato crops is a highly studied subject, there are currently standards such as that of the European and Mediterranean Plant Protection Organization (EPPO), which establishes the inspection procedures that must be carried out in the different stages of potato cultivation: before planting the seed, during the growth of the plants to maturity, during harvest and post-harvest. During the growth and maturation stage of the plants in the crop, the inspection is done visually, the phytopathologists must check the leaves, stems and tuberous roots. To do this, they must walk the crop and follow different inspection schemes established for each plantation (EPPO, 2007).

Optical imaging sensors capable to measure reflectance, temperature or fluorescence have been used for assessment of plant diseases in different crops. Those optical sensors for plant disease assessment comprise RGB sensors, which can be found in almost every mobile phone, computer or tablet, multispectral sensors, hyperspectral sensors, thermal sensors and fluorescence imaging devices. Although there are a variety of sensors available that can be used in agriculture and plant disease detection, by now, there are no sensors capable of detecting plant diseases by itself and most of the sensors can not measure plant physiological parameters directly (Mahlein, 2016b), so, interpretation and development of methods for data analysis is crucial.

Sugiura et al. (2016) developed a field phenotyping system to estimate the resistance to LB infestations in potato crops. The developed system uses the processing of aerial images obtained through the use of unmanned aerial vehicles to estimate the severity of infections in plant leaves. The developed system proved to be an alternative to traditional manual measurement methods, which can offer rapid assessments of the condition of crops. The study showed that aerial images obtained with UAV are an effective way to capture the spatial and temporal variability of the state of crops.

1.3 Research question

Taking into account the above, this research project aims to answer the following question: *What are the spectral and morphological features that allow estimating the severity level of late blight (Phytophthora infestans) in potato crops from multispectral images acquired using an UAV platform?*

1.4 Objectives

1.4.1 General objective

To develop a workflow for extraction of spectral and morphological traits of potato crops from multispectral images of high resolution acquired using unmanned aerial vehicles to assess the severity of late blight disease.

1.4.2 Specific objectives

1. To identify and extract the spectral traits of potato plants that allow estimating the damage caused by late blight in the potato crop.
2. To identify and extract the morphological traits of potato plants that allow estimating the damage caused by late blight, combining techniques of computer vision and pattern recognition.
3. To combine spectral and morphological traits of potato plants that allow estimating the damage caused by late blight disease.

1.5 Research importance

Improving potato crops may be the only way to combat late blight (Struik, 2010). The study of the geometric and spectral features of potato plants by means of UAV can contribute to improve the efficiency of field phenotyping systems currently used in potato improvement. This research can contribute to the determination of the methods of capture, processing and analysis of the data acquired through UAV in a way that provides producers with reliable tools for the improvement and management of their crops in an agile and efficient way.

1.6 Scope of the study

The aim of this study is to determine which spectral and morphological traits allow the evaluation of LB in an experimental potato crop from high resolution multispectral images acquired by using an UAV.

1.7 Structure of the thesis

This document is organised as follows. chapter 2 presents the state of the art for this work, chapter 3 describes the data and method used to classify LB, as well as healthy potato (HP) plants in experimental plots. section 3.4 presents the radiometric correction, photogrammetry processing and subset generation of the proposed method. section 3.5 presents the bare soil removal, the training of the ML algorithms and the classification of the datasets. chapter 4 contains the results of this study. chapter 5 deploys the discussion of the proposed method. Finally, chapter 6 presents the conclusions.

1.8 Publications

1. **Jorge Rodríguez**, Iván Lizarazo, Flavio Prieto, Victor Angulo-Morales, Assessment of potato late blight from UAV-based multispectral imagery, Computers and Electronics in Agriculture, Volume 184, 2021, 106061, ISSN 0168-1699, doi: 10.1016/j.compag.2021.106061.
2. V. Angulo, **J. Rodriguez**, E. Gaona, F. Prieto and I. Lizarazo, 2020. "A Supervoxel-Based Approach for Leaves Segmentation of Potato Plants from Point Clouds," IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 2020, pp. 4902-4905, doi: 10.1109/IGARSS39084.2020.9324365.
3. Angulo-Morales, V., **Rodríguez-Galvis, J.**, Lizarazo-Salcedo, I., and Gaona-García, E.: A NOVEL METHOD FOR ESTIMATION OF STRUCTURAL CHANGES IN POTATO CROPS FROM UAV-BASED DIGITAL SURFACE MODELS, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-1-2020, 269–276, DOI: 10.5194/isprs-annals-V-1-2020-269-2020, 2020.
4. V. A. Morales, **J. Rodriguez Galvis**, E. G. Garcia and I. Lizarazo Salcedo, "Estimation of Individual Potato Plants Area and Volume From UAV-based Multispectral Images,"

IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 6259-6262. DOI: 10.1109/IGARSS.2019.8898199

2 State of the Art

2.1 Introduction

Potato (*Solanum tuberosum subsp. tuberosum*), is a herbaceous crop that belongs to the Solanaceae family, which grows to a height of approximately 1 meter, with elongated pinnate leaves and flowers that can be: white, purple, pink or bluish with fruits that can be slightly yellow or green of approximately one inch (OECD, 1997). It is estimated that the first potato plants were selected and cultivated between 6000 and 10,000 years. Worldwide, the potato crop is the fourth most important after corn, rice and wheat (Namugga et al., 2017). In Colombia, it is sown mainly in the department of Cundinamarca (45.54 %), followed by Boyacá (25.08 %), and Nariño (19.28 %).[1]

Field phenotyping is one of the main components in crop breeding as the phenotype of a plant is the expression of the interaction between genetic and environmental factors and is reated to important yield traits and resistance to different sources of stress. (Sankaran et al., 2015). Plant phenotyping is the quantitative evaluation of the traits of a given genotype in a given environment or experiment, ranging from scalar (plant height), multi value (chemical or DNA transcription), to images and include directly measured attributes and those derived from analysis (Bolger et al., 2017).

Field phenotyping is gaining ground as the only approach capable of providing the required performance in terms of number of plants or populations, as well as providing an accurate description of trait expression in real world crops. However, to date, most fieldwork-based phenotyping systems have focused on rapid assessment of individual sets of traits, such as vegetation indices. (White et al., 2012).

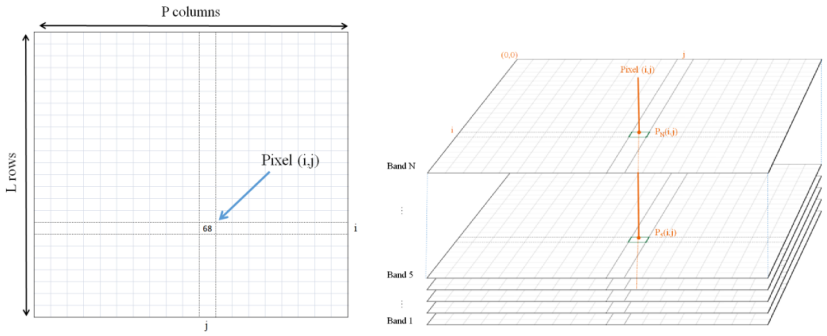
Unlike animals, which generally maintain a very similar structure regardless of the environment in which they develop, plants can show different architectures depending on the environmental conditions in which they develop (Tardieu et al., 2017). Thus, the same genotype can give rise to different phenotypes. The phenotype is defined by some ecologists as a particular trait that a plant can present, such as the Specific Leaf Area (SLA), or as relationships between several characteristics, such as the Leaf Area Index (LAI). However, these definitions may be too simple since they pose a one-to-one relationship between genes and phenotypes, when in fact the same gene can generate different phenotypic traits or a phenotypic trait can be generated by different groups of genes. (Kumar et al., 2015).

[1] Source: Boletin Mensual INSUMOS Y FACTORES ASOCIADOS A LA PRODUCCIÓN AGROPECUARIA. Enero 2017 – Num 15. DANE

2.2 Crop spectral traits

2.2.1 Multispectral image

An image is a table or matrix of pixels organised in rows and columns, where each pixel is a numerical value known as Digital Number (DN) or radiometric value, representing the average level of energy coming from the surface of the area covered by the pixel Figure 2.1a, (Ose et al., 2016).



(a) (b)

Figure 2.1: Multispectral image. (a) pixel inside an image, (b) multispectral image. From Ose et al. (2016)

A multispectral image is made up of several channels or bands, in which each pixel of the image contains information captured at different wavelengths (Figure 2.1b).

2.2.2 Radiometric correction

In an image, errors can occur in the Radiometric Values (RV) of the pixels, these errors can come from several sources: instrumental failures, illumination or reflectance of the captured objects that affect the accuracy of the quantitative measurements obtained. An example of the errors found in the images is random noise, which can be produced by electronic interference, intermittency in the transmission of the signal, data storage, among others. This type of noise can be reduced by applying smoothing filters to the image (Jones and Vaughan, 2010; Kamberova and Bajcsy, 2006).

Similarly, the angle of incidence of light on an object, and its reflectance, affect the amount of energy that the sensor receives. This makes, for example, two images of the same area captured at a different time or date look very different even though the spectral conditions in the area have not changed. The effects produced by the angle of incidence of the energy can be attenuated by dividing the value of the pixel by the cosine of the angle of

incidence with respect to the normal. Reflectance is generally converted to *reflectance at the surface* or *reflectance at the satellite* by implementing correction algorithms, although this reflectance does not necessarily accurately represent the true reflectance of the object at the time of his catch due to additional interactions (Figure 2.2) of the electromagnetic radiation with the atmosphere (Jones and Vaughan, 2010).

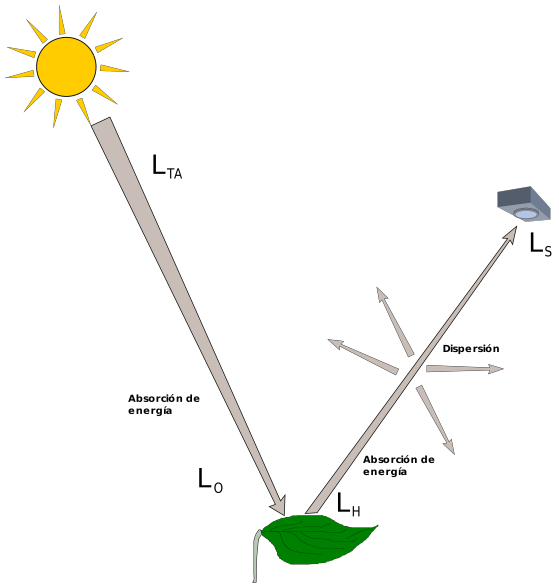


Figure 2.2: Relationship between the radiance received at the sensor L_S, the radiance reflected at the surface L_H and the radiance at the top of the atmosphere L_{TA} and that received at the surface of the object L_O. Adapted from Jones and Vaughan (2010)

2.2.3 Geometric correction

In an image there are different types of distortions, caused by factors such as the orientation of the camera at the time of capture, the variation in flight height, the relief of the capture area, among others. These distortions can be greater as the greater the distance from the sensor to the surface to be captured. There are two classes of geometric errors: systematic and unsystematic, the systematic ones generally come from the capture equipment, they are predictable and are found in all the images captured by the same sensor. Unsystematic errors almost always occur in a single image and must be corrected individually by the user (Jones and Vaughan, 2010).

2.2.4 Optical spectrum

According to directive 2006/25/EC of the European Parliament and Council, the optical spectrum is defined as a portion of radiation belonging to the electromagnetic spectrum, comprised

between 100 nm and 1 mm of wavelength. The optical spectrum is divided into: ultraviolet radiation, visible radiation and infrared radiation[2]. Jones and Vaughan (2010) define the optical region as all wavelengths of the electromagnetic spectrum, from ultraviolet to near and mid-infrared, between 250 nm and 3000 nm. On the other hand, the National Aeronautics and Space Administration (NASA), places the optical spectrum in the region between 10^{21} Hz (0.0001 nm approximately) and the 100 GHz (3 mm approximately)[3].

For the purposes of this project, when talking about the optical spectrum, reference will be made to the range between 250 nm and 3000 nm, following what is proposed by Jones and Vaughan (2010) (Figure 2.3).

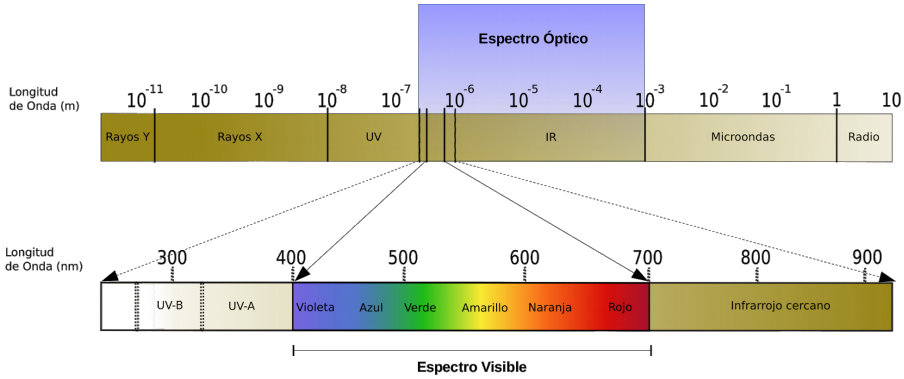


Figure 2.3: Optical spectrum and visible spectrum within the electromagnetic spectrum. Adapted from Jones and Vaughan (2010)

2.2.5 Spectral traits

The traits of a plant can be related to the wavelength of the reflected or absorbed energy. How each plant reflects the energy in each of the wavelengths allows us to differentiate them. However, in the visible and near infrared wavelengths, one of the major limitations is the change in lighting conditions at the time of image capture, for which radiometric correction must be performed on the acquired images. (Sankaran et al., 2015).

The interaction of electromagnetic energy with the leaves of plants depends not only on the wavelength but on the diversity of structures of the leaves, their chemical properties, age, thickness, water content, among others, these factors affect the spectral reflectance of the leaf

[2]DIRECTIVE 2006/25/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2006. On the mínimum health and safety requirements regarding the eposure of workers to risks arising from physical agents (artificial optical radiation) (19th individual Directive within the meaning of Article 16(1) of Directive 89/391/EEC)

[3]Optical Spectrum, https://www.nasa.gov/directorates/heo/scan/engineering/technology/txt_opticalcomm.html, Last Updated: June 26, 2018, Editor: Ashley Campbell.

(ρ_λ), its absorbance (α_λ) and its spectral transmissivity (τ_λ) (Jones and Vaughan, 2010).

The reflection of energy in the leaf is a complex phenomenon because only a part of the energy is reflected from the surface of the leaf, the greater portion of the incident energy penetrates the structures of the leaf where an important portion of the visible energy is absorbed by chloroplasts but a greater amount of energy in the infrared range is reflected or scattered towards the leaf surface or in a wide variety of directions between cells and intercellular spaces, as shown in Figure 2.4 (Jones and Vaughan, 2010).

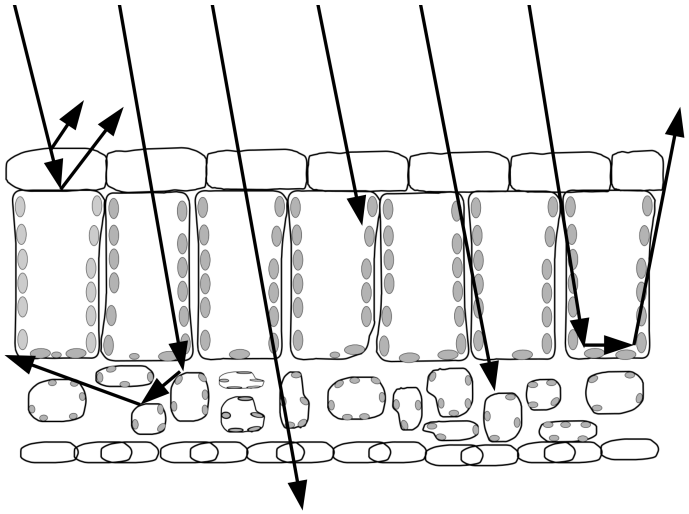
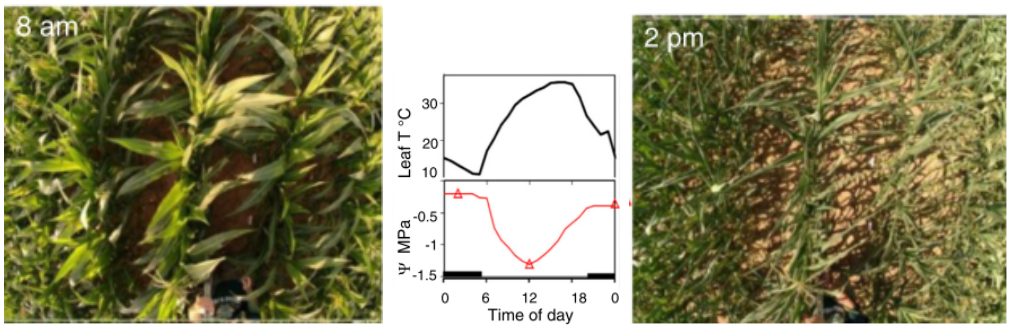


Figure 2.4: Diagram of the typical cross-section of a dicotyledonous leaf, illustrating the multiple paths that electromagnetic radiation can follow in its interaction with the leaf. Adapted from Jones and Vaughan (2010)

Crop improvement requires a new emphasis on phenotyping of specific and well-defined physiological characteristics, which is why it requires the union of classic techniques of crop improvement, as well as several levels of phenotyping that allow identifying which genotypes express the desired traits for a particular crop. This is why high-yield phenotyping systems can be very useful in this field (Ghanem et al., 2015).

The possibility of making measurements that can be reproduced over time and with different platforms requires standardised processes, which include: calibration of the sensors, selection of the angles of capture, and selection of the hours of the day in which the observations should be made. For example, the Figure 2.5 shows the change in the appearance of corn plants in the same day, when the conditions of temperature and humidity change. In the images you can see how the change in these variables has a great impact on the appearance of plants, from having an appearance of vigour in the morning to have an appearance of wilt in the afternoon (Tardieu et al., 2017).



(a) (b) (c)

Figure 2.5: Corn plant caught at two different times on the same day. (a) Image captured at 8 am, with a temperature of 11^o C and a water potential in the leaf of 0 MPa, (b) evolution of the temperature and the potential of water in the leaf during the day, (c) Image of the same area captured at 2 pm with a temperature of 36^o C and a water potential in the leaf of -1.5 MPa, with -1.5 being a value close to the lethal values for most of the plants. Taken from (Tardieu et al., 2017).

Sugiura et al. (2016) developed a technique to estimate the severity of disease incidence in potato crops, specifically late blight. This technique was based on the use of RGB aerial images acquired in the field by means of a UAV, the experimental cultivation studied in this work consisted of 262 blocks of 0.75 m $\times$ 3 m with only one variety and the same type of treatment. For the study, images were captured during 11 periods: July 17, July 18, July 23, July 25, July 26, July 27, July 30, August 2, August 3, August 7 and August 15 (Figure 2.6).

The results showed that the estimate of the area affected by the disease from the RGB aerial images in each period, with respect to the area estimated with field observations, had a correlation of 0.77, which was acceptable for growers. The estimation method from images acquired with a UAV turned out to be more efficient than with the conventional visual estimation method.

Vegetation Indices (VIs) are a well-established method for estimating biomass (Bendig et al., 2015), and can be good indicators of crop disease severity. Normalized Difference Vegetation Index (NDVI), Green normalized difference vegetation index (GNDVI), Soil adjusted vegetation index (SAVI), Enhanced vegetation index (EVI) are considered typical VIs related to biomass, Photosynthetically Active Radiation (PAR), and Leaf area index (LAI). Those VIs have been used in precision agriculture tasks, proving to be very useful to detect the severity level of crop diseases (Zhao et al., 2020).

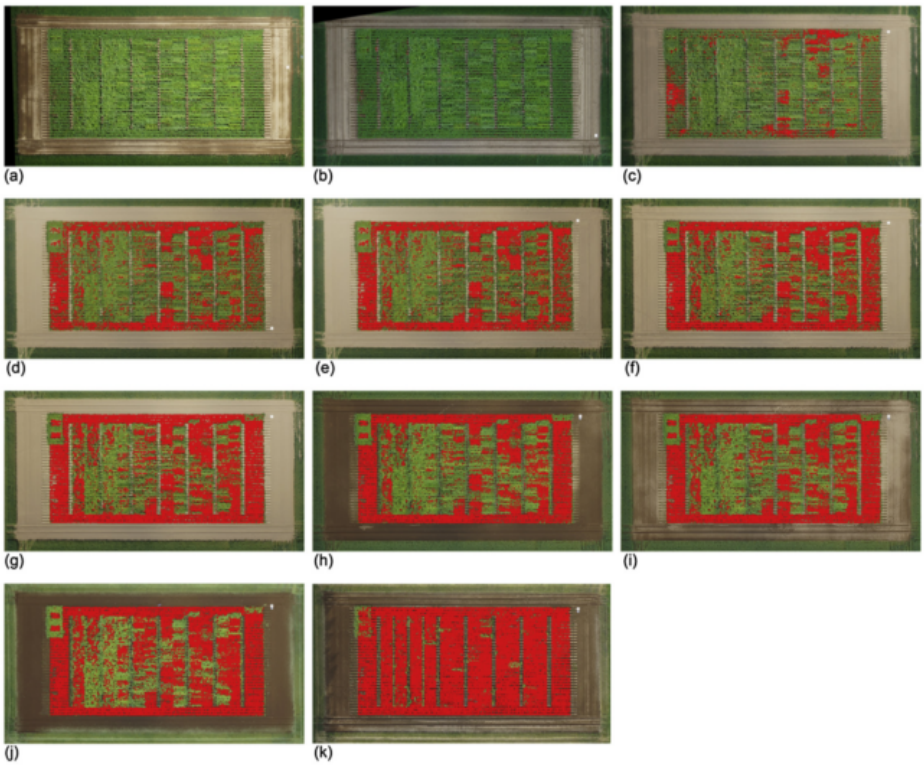


Figure 2.6: Progress of the damage produced by the development of late blight, the red zones show the affected areas detected by digital processing of the aerial images acquired for the periods: July 17, July 18, July 23, July 25, July 26, July 27, July 30, August 2, August 3, August 7 and August 15. From Sugiura et al. (2016)

2.3 Crop morphological traits

2.3.1 Computer vision

Vision is a task of processing information from images, but it is also a process of representing that information. In the case of human vision, it is usually unlikely that we will stop to try to understand these two aspects in depth. However, since the emergence of machines capable of acquiring and processing images, understanding in depth not only these two aspects, but many more of the world around us, became necessary (Marr, 1982).

It is common to find in the literature the terms: Computer vision (CV) and Machine vision (MV) as synonyms. However, Batchelor (2012) states that *Machine Vision* should be recognised as a field substantially different from CV. The author suggests that CV, Artificial Intelligence (AI), Pattern Recognition (PR) and Digital Image Processing (DIP) contribute together to the field of MV.

Batchelor (2012) define MV as an area that " ... has to do with the engineering of

electronic, mechanical, optical and integrated software systems that allow the examination of objects and materials, human artefacts and manufacturing processes so that defects are detected and the quality, operational efficiency and safety of processes and products are improved. It is also used in the control of machines used in the manufacturing process ... ". The definition specifies that MV requires various areas, such as mechanics, optics, lighting, sensors, electronics, signal processing, digital image processing, industrial engineering, among others.

Thus, although MV and CV are in charge of processing images within electronic devices, in the case of CV it is explicit that the processing of images in this field takes place on a computer while the MV allows such processing to be carried out on specialised electronic devices or electro-optical equipment. Similarly, CV has much more general purposes, in contrast to MV, whose objectives are focused on the processes carried out within the industrial sector. The objective of the CV is the identification and interpretation of the objects present in the images, while that of the MV is not so much the identification as such of an object but, more specifically, the determination of possible defects in the manufacture of that object (Batchelor, 2012)

Image segmentation

It is the division of the pixels in an image into groups so that the grouped pixels share certain characteristics (Batchelor, 2012). There are different methods for image segmentation: thresholding, Edge based, Region based, clustering based, watershed based, Partial Differential Equation (PDE), Artificial Neural Network (ANN) based (Kaur and Kaur, 2014) and graph segmentation methods (Camilus and Govindan, 2012). In this study we used two segmentation methods, thresholding and graph based segmentation.

- **Thresholding methods**

  Thresholding based methods are widely used in particular in tasks of image binarisation, where the objective is to separate objects from the background Hamuda et al. (2016b). These methods divide the image pixels with respect to their intensity level and are used over images having lighter objects than background. The values of thresholds can be computed with the help of the peaks of the image histograms, or can be selected in a manual way. These methods can also be classified into three types: global thresholding, variable thresholding and multiple thresholding (Kaur and Kaur, 2014).

  1. **Global thresholding.** In this method, a value L is selected and is constant for the whole image:

  $$s(x,y) = \begin{cases} 1, \text{if } i(x,y) > L \\ 0, \text{if } i(x,y) \leq L \end{cases} \tag{2.1}$$

  where, $i(x,y)$ is the original image and $s(x,y)$ is the segmented image.

  2. **Variable thresholding.** In this method, the value L can vary through the image. There are also two types of variable thresholding:

- *Local threshold.* The value of L depends on the neighbourhood of x and y.
- *Adaptive threshold.* The value L is a function of x and y.

3. **Multiple Thresholding.** In this method, there are multiple threshold values like L0 and L1. By using these output image can be computed as:

$$s(x,y) = \begin{cases} m, \text{if } i(x,y) > L1 \\ n, \text{if } i(x,y) \leq L1 \\ o, \text{if } i(x,y) \leq L0 \end{cases} \tag{2.2}$$

- **Graph based segmentation**

  Graph-based segmentation methods treats an image as a graph in which vertices are composed of pixels. Each edge has a weight determined, in most cases, based on the vertices it relates.

  If we consider an image as a graph G, the idea is to find a set of sub-graphs $\{SG_1, SG_2, ..., SG_n\}$ from G such that for all $k \in \{1, 2, ..., n\}, \forall i, j$ and $i \neq j$, $w_i, w_j \in SG_k$ with walks between w_i and w_j (Camilus and Govindan, 2012).

  Felzenszwalb and Huttenlocher (2004) method is an efficient graph-based segmentation method that uses Minimum Spanning Trees (MST), which is a subgraph that links all the vertices of the graph and contains only a single path between two vertices. The method assumes that edges between vertices in the same segment have low weights than edges between vertices in different segments.

Mathematical morphology

Mathematical morphology comprises the study of geometric structures present in images. The analysis of the geometric structures present in the images is carried out through the use of set theory, in which the main operations are erosion and dilation (Conci et al., 2018).

A dilation of a binary image $A = \{a_0, a_1, a_2, ..., a_n\}$, containing n foreground pixels, by a structuring element B is given by Equation 2.3. Figure 2.7d shows an example of the dilation of an image A by a structuring element B (Figure 2.7b).

$$D^B(A) = \bigcup_{b \in B} A_b \tag{2.3}$$

Erosion of a binary image is the opposite operation to dilation, defined by Equation 2.4. Figure 2.7c shows an example of the dilation of an image A by the same structuring element B (Figure 2.7b)

$$E^B(A) = \bigcap_{b \in B} A_{-b} \tag{2.4}$$

From a combination of erosion and dilation it is possible to generate operations such as opening (Figure 2.7e), and closing (Figure 2.7f).

2.3.2 Morphological traits

The study of the geometric features of plants has been useful to find differences between plants grown in different environments. Mishra et al. (2012), for example, carried out the systematic comparison of the leaves of *Arabidopsis thaliana* plants, which have been cultivated both in the laboratory and under field conditions, finding that the leaves of the plants that grew in the laboratory, under different conditions, they shared much longer leaves than those grown in the field.

Different authors have studied phenotypic traits based on the morphology of plants in the laboratory, such as Leaf Length (LL), Leaf Width (LW)), Leaf Width-Lenght Ratio (LWLR), Total Leaf Expansion (TLF), NLeaf Number (LN), Plant Area (PA), Leaf Area (LA), Plant Height (PH), Leaf Bounding Box (LBB) and Plant Bounding Box (PBB). Mishra et al. (2012) studied the phenotypic plasticity of *Arabidopsis thaliana* under different growth regimes in climate chambers and in field conditions. Their results showed that plants grown indoors have enlarged leaves, different leaf shapes and longer petioles, also, they found that the photoperiod is the main determinant of leaf size and shape.

White et al. (2012) focused on the simultaneous use of proximal sensors to record spectral reflectance, canopy temperature, and plant architecture to assess traits, such as adaptations to water deficits or heat stress during a single diurnal cycle and quantifying stress recovery. Their work allowed defining key criteria, experimental approaches, equipment and data analysis tools required for robust high-throughput field-based phenotyping.

Granier and Vile (2014) review different studies and highlight the latest advances in plant multi-trait phenotyping and discuss future needs to ensure the best use of quantitative data in plant phenotyping. They highlight that it is necessary to take into account environmental and temporal variations of the phenotype and its integration at different levels, such as subcellular, cellular, tissue, organ or whole-plant level.

An et al. (2016) introduce an automated high-throughput phenotyping workflow using low-cost imaging systems which they implemented in laboratory and in field conditions to study the same phenotypes. They found that young leaves mainly grow by elongation, although, in later developmental stages, total leaf expansion slows relative to leaf-width growth increases.

Minervini et al. (2016) presents a collection of datasets of raw and labelled images of rosette plants and define a set of computer vision and classification tasks for analysing their data. Also, they show exemplary use cases and results on some tasks carried on with parts of these data. They conclude that computer vision and machine learning are ideally suited to help in plant phenotyping tasks.

Pradal et al. (2017) presents an infrastructure for plant phenotyping capable of measure traits associated to the plants' adaptation to climate change. Those traits include the leaf appearance, duration of phenological phases, plant grow rate in terms of area and volume, plant organ expansion and plant morphology.

Hu et al. (2018) studied a new method for plant height estimation in a sorghum breeding trial. Their work integrated an RGB camera and an UAV to capture images that were

processed to generate Digital Surface Models (DSMs). Three methods were used to process those DSMs to estimate the height of the plants. Their experimental results showed that the proposed method performed similar to the manual measurements, allowing to estimate plant height accurately.

Recently, Li et al. (2019) carried out a study in which they estimated the emergence of plants in a potato crop from RGB aerial images captured with a UAV. In this work, the plants in the study area were separated from the soil by applying a combined segmentation method using Excess Green Index (Woebbecke et al., 1995) and the Otsu's threshold method (Otsu, 1979). In this study were also calculated 6 morphological traits from each plant and then classified using RF to estimate the number of plants in emergence. To do this, they analysed 3 field experimental crops, and they found that UAV-based RGB images allowed to estimate the number of plants in emergence stage with an r^2 of 0.96.

Table Table 2.1 shows a summary of the morphological traits worked in these studies. Except for Mishra et al. (2012), the other studies have been based on the use of high resolution images.

Table 2.1: Morphological traits reported in literature.

Rasgos Morfológicos																**Planta**	**Autor**
LL	**LW**	**LWLR**	**LA**	**LN**	**TLF**	**PA**	**PH**	**PP**	**LBB**	**PBB**	**LPBB**	**LWRB**	**CA**	**SO**			
•	•	•	•			•										*Arabidopsis thaliana*	Mishra et al. (2012)
				•			•									*Solanum tuberosum*	White et al. (2012)
			•	•		•										*Arabidopsis thaliana*	Granier and Vile (2014)
•	•				•											*Arabidopsis thaliana*	An et al. (2016)
									•	•						*Arabidopsis thaliana*	Minervini et al. (2016)
						•	•									No menciona	Pradal et al. (2017)
							•									*Sorghum*	Hu et al. (2018)
		•				•		•		•	•	•	•	•	•	*Solanum tuberosum*	Li et al. (2019)

LL: Leaf length
LW: Leaf width
LWLR: Leaf width-length ratio
LA: Leaf area
LN: Leaf number
TLF: Total leaf area
PP: Plant area
PA: Plant height
PH: Plant perimeter
LBB: Leaf bonding box
PBB: Plant bounding box
LPBB: Plant bounding box long side
LWRB: Length-width bounding box ratio
CA: Area of the convex hull
SO: Convex hull area - canopy area ratio

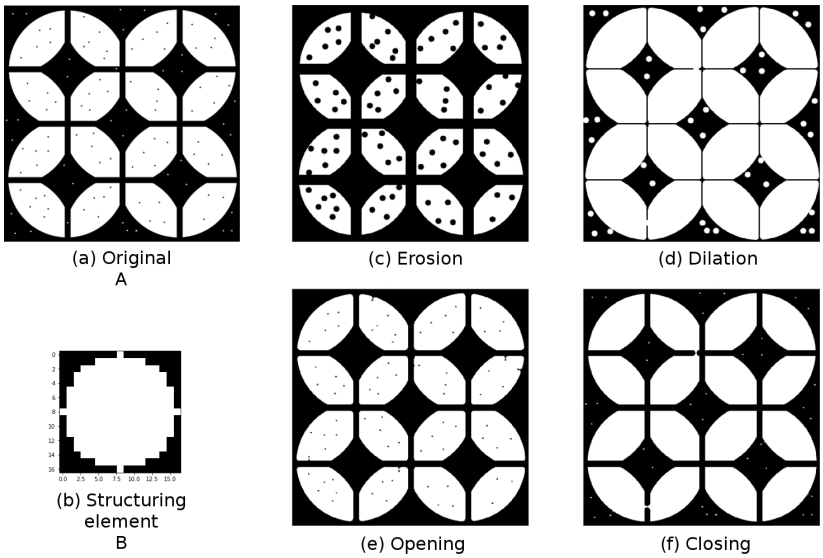


Figure 2.7: Some basic morphological operations.

3 Data and methods

3.1 Study area

The study area is a 1920 sq.m. potato field located in Subachoque, Colombia (Figure 3.1). This field is part of an experimental plot designed to evaluate the potato *Diacol capiro* variety response to different nutrient treatments (Table 3.1). As the weather conditions favoured the appearance of LB disease in several plots, the project had the opportunity to monitor the crop disease development. The experimental plot was inspected every 15 days during the entire crop life span (i.e. 120 days from planting to maturity).

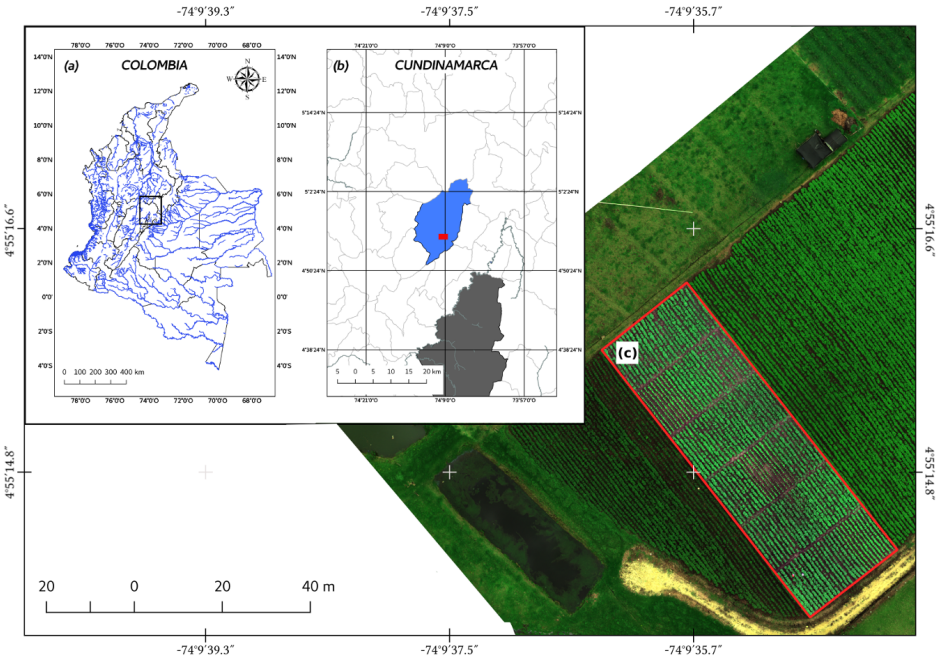


Figure 3.1: Location of the study area, a potato experimental plot in Subachoque, Cundinamarca (Colombia): (a) Municipalities of Colombia (black colour) and main rivers (blue colour); (b) study area (red colour), municipalities (light grey colour), Bogota D.C. capital city of Colombia (dark grey colour) and Subachoque municipality (blue colour area). (c) The experimental potato plot area is enclosed by the polygon highlighted in red colour.

Table 3.1: Treatments used in the experimental potato crop.

Treatment	N	P	K	Total Kg/Ha
T1	150	200	150	840.7
T2	150	200	300	1090.7
T3	250	200	150	1058.1
T4	250	200	300	1308.1
T5*	0	0	0	0
T6	200-309-267-86(CaO)-43(MgO)			1800

3.2 Data acquisition

Eight-minute UAV flights were performed over the potato field 72 days after planting and 86 days after planting, which corresponds to the time when blackish/brown lesions with white sporulation where clearly visible on leaves and stems of the experimental crop (Schumann et al., 2000). Flight altitude over terrain was 40 m above ground level. A MicaSense RedEdge camera was setup on a Tarot 680PRO hexacopter to obtain aerial images with approximately 4 cm pixel size.

Figure 3.2 shows the distribution of the experimental blocks and treatments of the experimental potato crop. T5 corresponds to the control treatment blocks and T6 corresponds to the farmer's usual treatment. Overall the control treatment blocks were in a growth stage below the average growth stage of the rest of the experimental blocks.

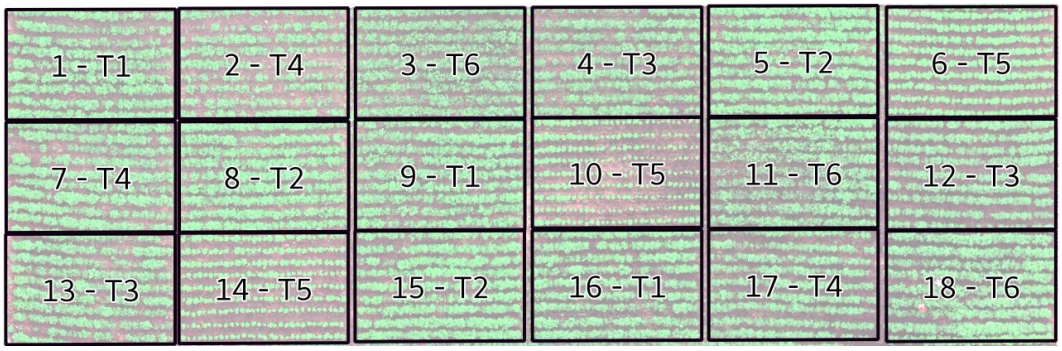


Figure 3.2: Distribution of experimental plots and treatments. The first number indicates the block number and "TN" indicates the treatment, where N = 1,2,... 6.

High resolution multispectral images were acquired at 40 m altitude above the ground surface at 11:00 am local time (GMT-5). Photogrammetric flight mission, shown in Figure 3.3a, was carried on according to techniques described by Pepe et al. (2018). Photographs overlap followed UAV photogrammetric standards suggested by Eisenbeiß et al. (2009), i.e. mean forward overlap was 80%, and mean sidelap was 60%. MicaSense RedEdge is a non-metric digital sensor with 4.8 mm × 3.6 mm size and 5.4 mm focal length Micasense (2018). Each multispectral image acquired five bands as described in Table 3.2. In this work, camera bands 4 and 5 have been reset according to the following order: Blue (B); Green (G); Red

(R); Red edge (RE) and Near infrared (NIR).

Table 3.2: Micasense RedEdge camera band specifications.

Band	Name	Centre Wavelength (nm)	Bandwidth (nm)
1	Blue	475	20
2	Green	560	20
3	Red	668	10
4	Near infrared	840	40
5	Rededge	717	10

Georeferencing was achieved by positioning eight ground control points (GCPs) throughout the potato experimental field. The points were signalised with circle-shaped targets. Raw GNSS measurements of GCPs were collected and stored using a Leica GPS System 500 GNSS receiver for post-mission processing. This receiver allows to calculate baselines with a precision of up to 5-10 mm + 1ppm.

Markers comprised a yellow-coloured circle (0.10 m diameter) inside a black-coloured circle (0.30 m diameter). Figure 3.3(a) shows the flight lines for each mission performed over the study area. Flight path is defined by a series of way-points depicted as white markers. After finishing the photogrammetric mission, the UAV heads to its home point (depicted as the eleven white marker close to the lower left corner). Figure 3.3(c) illustrates markers design, which makes them easy for GCP identification from aerial images. Interested readers can observe photographs from fieldwork on the study area on this link: https://github.com/jorlrodriguezg/jorlrodriguezg.github.io/tree/master/Images/field

3.3 Methods

In this work, we used a five-stage method for assessment of LB:

(i) data acquisition;

(ii) data processing for converting raw images to an orthorectified mosaic representing surface reflectance;

(iii) data analysis for background removal and crop classification into healthy and diseased categories;

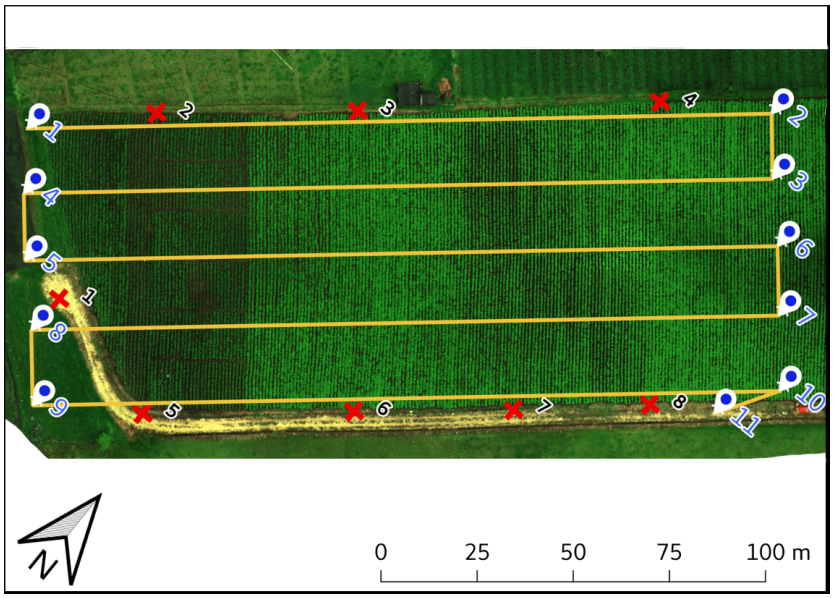
(iv) model training and supervised classification stage; and

(v) use of trained models to classify a new dataset.

Figure 3.4 illustrates the workflow conducted at each stage. The following sections provide a detailed description of every step.

Field assessment of late blight

Experimental crop area consisted of an array of 18 blocks of 12 m × 8 m inside a field of 77 m × 24 m (Figure 3.6). Each block had 9 rows arranged along the long side of the field with



(a)

(b)

(c)

Figure 3.3: Data capture: (a) Flight path (yellow lines), GCPs (red markers) and way points (white markers). The home point is identified by way point eleven. (b) UAV used to acquire the multispectral images. (c) Marker used for signaling GCPs.

an area of 216 sq.m.. The space between rows was 1m and each row contained approximately 30 seed tubers of the same variety.

Disease severity was evaluated as the percentage of foliage area that was infected, on a scale from 0% to 100%, following the standard procedure described by Forbes et al. (2014), where researchers visually estimate the percentage of total leaf area that is affected by the disease by comparing the green and non-green portions, assuming late blight is the only or dominant foliage disease. Although 44 zones were evaluated at plant level to be used as training zones for the ML algorithms, severity of LB was evaluated at block level as suggested by Forbes et al. (2014) due to the high number of plants in each block.

Different LB severity levels were observed in the blocks of the experimental plot. Four infection levels were assessed following the severity characteristics used by Bock et al. (2010). Figure 3.5 shows example plants from blocks at different levels of infection. Level 1 (LB1),

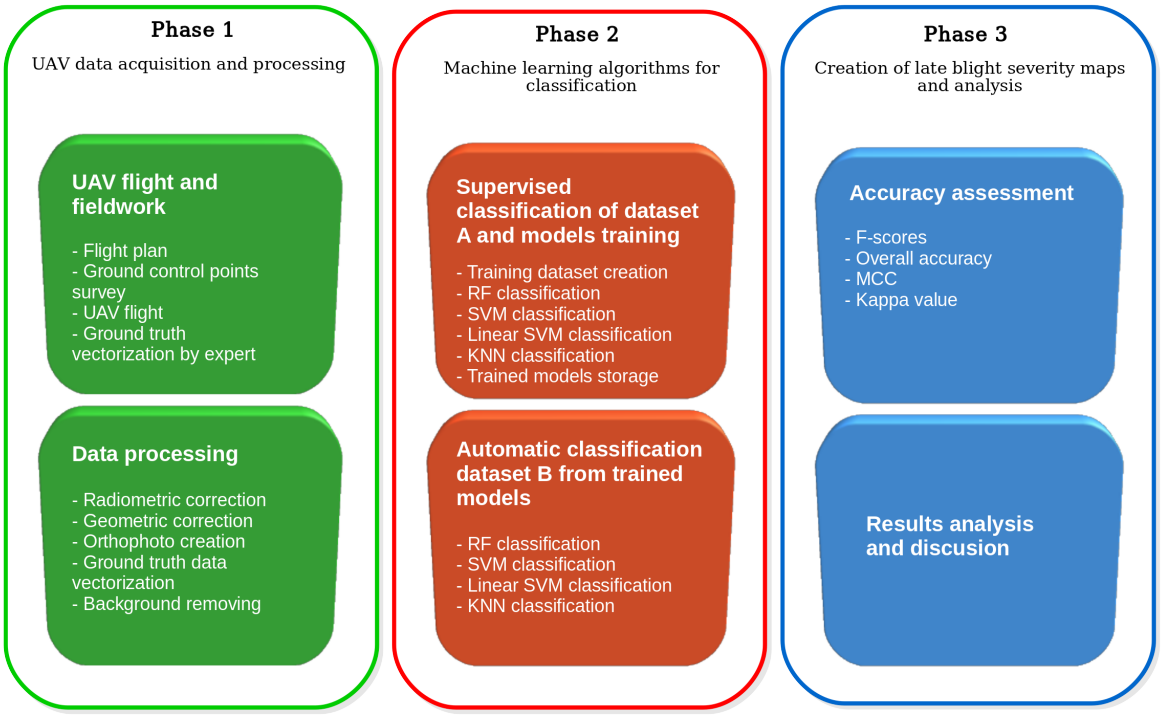


Figure 3.4: Workflow of the research conducted for assessing LB from UAV-based imagery.

trace - 25% leaf area infected (Figure 3.5a); Level 2 (LB2), 26 - 50% leaf area infected (Figure 3.5b); Level 3 (LB3), 51 - 75% leaf area infected (Figure 3.5c); Level 4 (LB4), > 75% leaf area infected (Figure 3.5d). A Level 0 (LB0), indicates an apparently infection-free block. Blocks in the potato experimental crop ranged between LB1 and LB4 for the two datasets.

3.4 Data pre-processing

The data pre-processing of the proposed method consist of two phases, (i) radiometric correction; and (ii) photogrammetric processing.

3.4.1 Radiometric correction

Radiometric variations of the same object can occur in an image. These variations can come from various sources: instrumental faults, lighting changes caused by weather conditions or reflectance of captured objects that affects the accuracy of the obtained quantitative measurements (Jones and Vaughan, 2010; Kamberova and Bajcsy, 2006).

Weather conditions during the photogrammetric missions were highly unstable and small cumulus were predominant in the sky. Moreover, winds of approximately 6 m/s created conditions for rolling cloud cover, making illuminance suddenly change during the flights. As a result, the study area illumination was heterogeneous and radiation received at the sensor

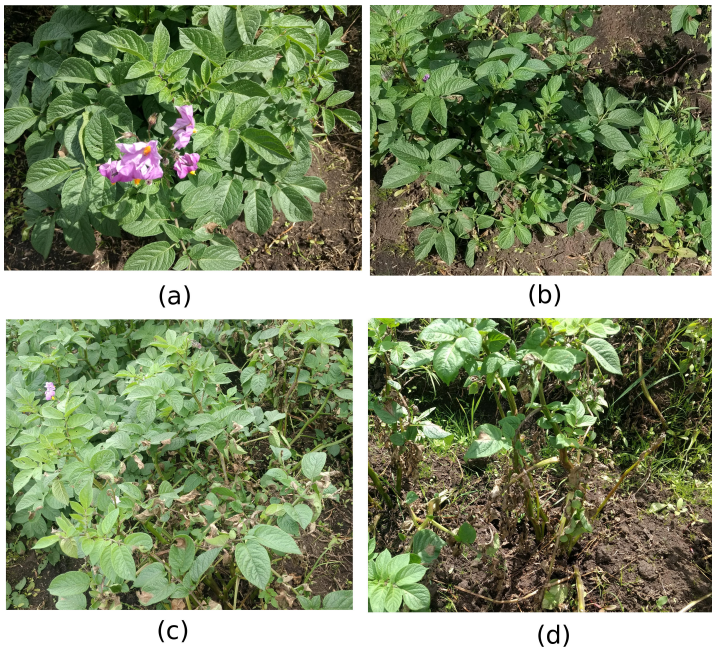


Figure 3.5: Plants belonging to blocks evaluated as LB1, LB2, LB3 and LB4. (a) Potato leaves with LB at infection level 1; (b) Potato leaves with LB at infection level 2; (c) Potato leaves with LB at infection level 3; and (d) Potato leaves with LB at infection level 4.

changed almost at each point of image acquisition.

According to Mamaghani and Salvaggio (2019), radiometric correction was achieved by using parameters from the Micasense RedEdge multispectral camera and its Digital Light Sensor (DLS). The camera sensor measures spectral radiance coming from the field surface, while DLS measures down-welling spectral irradiance.

Spectral radiance was obtained by using Equation 3.1:

$$L_i = V_i(x,y)\frac{a_{1,i}}{g_i}\frac{p_i(x,y) - p_{BL_i}}{t_{e,i} + a_{2,i}y - a_{3,i}t_{e,i}y}, \tag{3.1}$$

where p is the normalised raw pixel value, obtained by division of the raw pixel by 2^N, where N is the number of bits in the image; p_{BL} is the normalised black level value; $a_{1,i}$, $a_{2,i}$, $a_{3,i}$ are the radiometric calibration coefficients; $V_i(x,y)$ is the vignette polynomial function for pixel location (x,y); g_i is the sensor gain setting; $t_{e,i}$ is the image exposure time; x, y are the pixel column and row number, respectively; and L_i is the spectral radiance in W/m2/sr/nm.

The vignette map, is represented as

$$V(x,y) = \frac{I(x,y)}{k_i} \tag{3.2}$$

where $I(x,y)$ is the intensity pixel value at (x,y) and k_i is a correction factor

$$k_i = 1 + k_{0,i}r_i + k_{1,i}r_i^2 + k_{2,i}r_i^3 + k_{3,i}r_i^4 + k_{4,i}r_i^5 + k_{5,i}r_i^6 \tag{3.3}$$

where $k_{0,i}$ to $k_{5,i}$ are polynomial correction coefficients and r_i is the distance of the pixel to the vignette centres.

$$r_i = \sqrt{(x - c_{x,i})^2 + (y - c_{y,i})^2} \tag{3.4}$$

where $c_{x,i}$ and $c_{y,i}$ represent the vignette centre and i denotes the spectral band number.

Surface reflectance was obtained by using the At-Altitude Radiance Ratio (AARR) technique, which uses information stored by the DLS

$$\rho_i = \frac{L_{s,i}}{DLS_i} \tag{3.5}$$

$$DLS_i = \frac{E'_{solar,i}}{\pi} cos(\sigma')\tau_i + L_{\downarrow solar,i} \tag{3.6}$$

where DLS_i is the downwelling light sensor radiance recorded by the MicaSense RedEdge; ρ_i is the reflectance factor; $L_{s,i}$ is the band effective spectral irradiance; $E'_{solar,i}$ is the spectral exoatmospheric solar irradiance; σ is the solar zenith angle; τ_i is the spectral transmission from space to the UAV; $L_{\downarrow solar,i}$ is the solar scattered down-welling sky radiance propagating towards the UAV and i denotes the spectral band number.

3.4.2 Photogrammetric processing

Orthomosaics for every date were obtained performing an automated procedure described by Küng et al. (2011) which implements the following steps:

- Determination of characteristic points in each image. Then, these points go through a matching process to find tie points between images and calculate the orientation parameters of the camera at the time of capture.
- Point cloud densification based on estimated image orientation parameters.
- Generation of a mesh representing the surface of the surveyed area.
- Generation of the Digital Surface Model (DSM) of the surveyed area.
- Orthorectification and mosaicking of the multispectral images to generate an orthophoto of the surveyed area.

3.5 Dataset preparation

The orthophoto in the study area contains reflectance values coming from potato plants, bare soil and weeds present in the study area, as well as small elements such as pebbles and stones.

At this stage we analyse the influence of these elements on images and histograms, in order to be able to remove them from the image since, for the purpose of the study, elements different than potato plants represent noise that affects the results.

3.5.1 Subset Generation

The generated orthophoto covered an area of 3.2 Ha, which includes the experimental plot and a bigger area where a different variety of potato was located. This variety is a yellow potato, which means the tuber root colour is yellow. In contrast, the potato from the experimental plot was from one variety of white potato. The interest area was clipped using a reference polygon (Figure 3.1).

Experimental crop area consisted of an array of 18 blocks of 12 m × 8 m inside a field of 77 m × 24 m (Figure 3.6). Each block had 9 rows arranged along the long side of the field with an area of 216 sq.m.. The space between rows was 1m and each row contained approximately 30 seed tubers of the same variety.

3.6 Data analysis

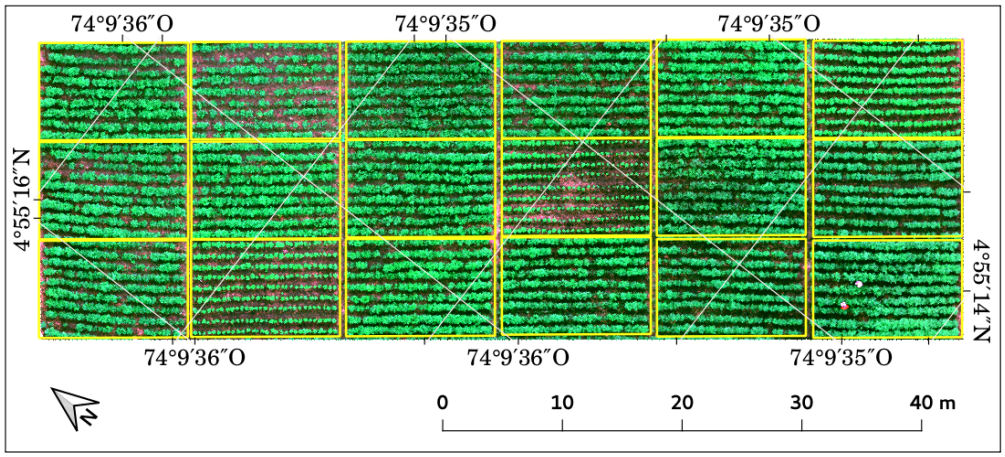
3.6.1 Background Removal

Background removal is intended for deleting information corresponding to soil surface, which is considered irrelevant for the assessment and detection of LB in the experimental plot. This step aims to create a multispectral image with vegetative structures only. Those vegetative structures correspond to potato leaves, stems and weed. Spectral responses of the multispectral images showed clear separation between bare soil and other objects (Healthy potato plants, LB potato plants and weeds, Figure 3.7g) in the NIR band. However, for the dataset B separation of weeds from other objects (Healthy potato plants and LB potato plants) in the same band was not clear, therefore, manual adjustment of the threshold was necessary to improve the separation of potato plants from weeds.

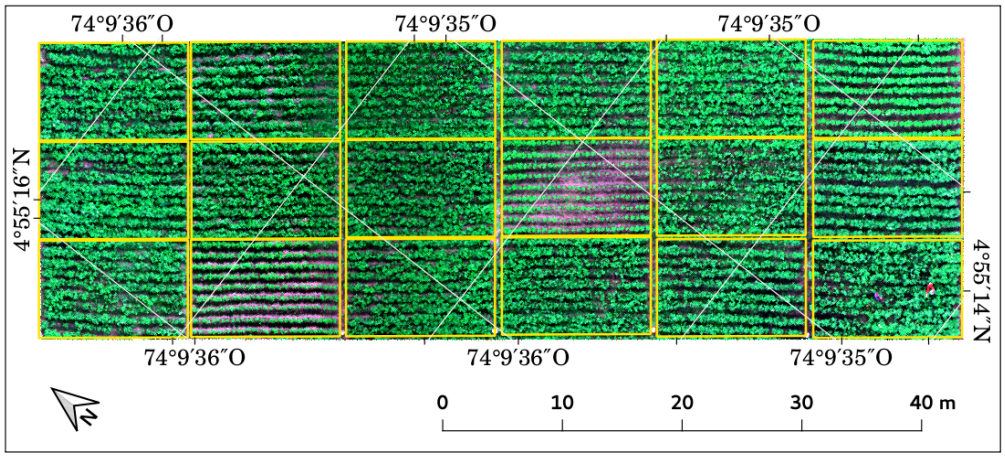
Thresholding

A thresholding technique reduces a grey-level image into an image where objects and background are represented by two levels: a binary image (Glasbey, 1993). Due to the difference in reflectance between the soil and the potato plants in the NIR band (Figure 3.7g), this band was used to separate the potato plants from the bare soil.

The first step was to analyse multispectral bands of the orthomosaics by plotting their histograms. The histograms analysed were classified into two groups: those belonging to images without presence of weeds in the crop and those belonging to images in which the presence of weeds was intermediate to high. Figure 3.7 shows the histogram for the NIR band in the multispectral orthomosaic for the two categories under study, as well as ground reference images of the two categories. G, RE and NIR bands have right-skewed bimodal histograms in the first category. The category with weeds presence had right-skewed histogram without



(a)



(b)

Figure 3.6: Layout of the experimental crop; (a) dataset A; (b) dataset B. The yellow lines indicate the division of the field into 18 experimental blocks.

a clear valley (Figure 3.7f). The method used to separate vegetation from bare soil was to apply Otsu's thresholding algorithm (Otsu, 1979) to find an optimal value to be used for segmentation and, then adjust the threshold value, if necessary, to improve separation of the potato plants from bare soil as well as separation from weeds. This step required threshold adjustment to update segmentation limit for dataset B where weeds had higher density. A similar method was used by Li et al. (2019) to extract potato plants from images captured with an UAV.

Image masking

Figure 3.8a shows the near infrared image before the background removing, at this stage there is presence of soil and weeds. To remove background from original multispectral image we multiply each band for the binary image (Figure 3.8b) obtained in the thresholding step using

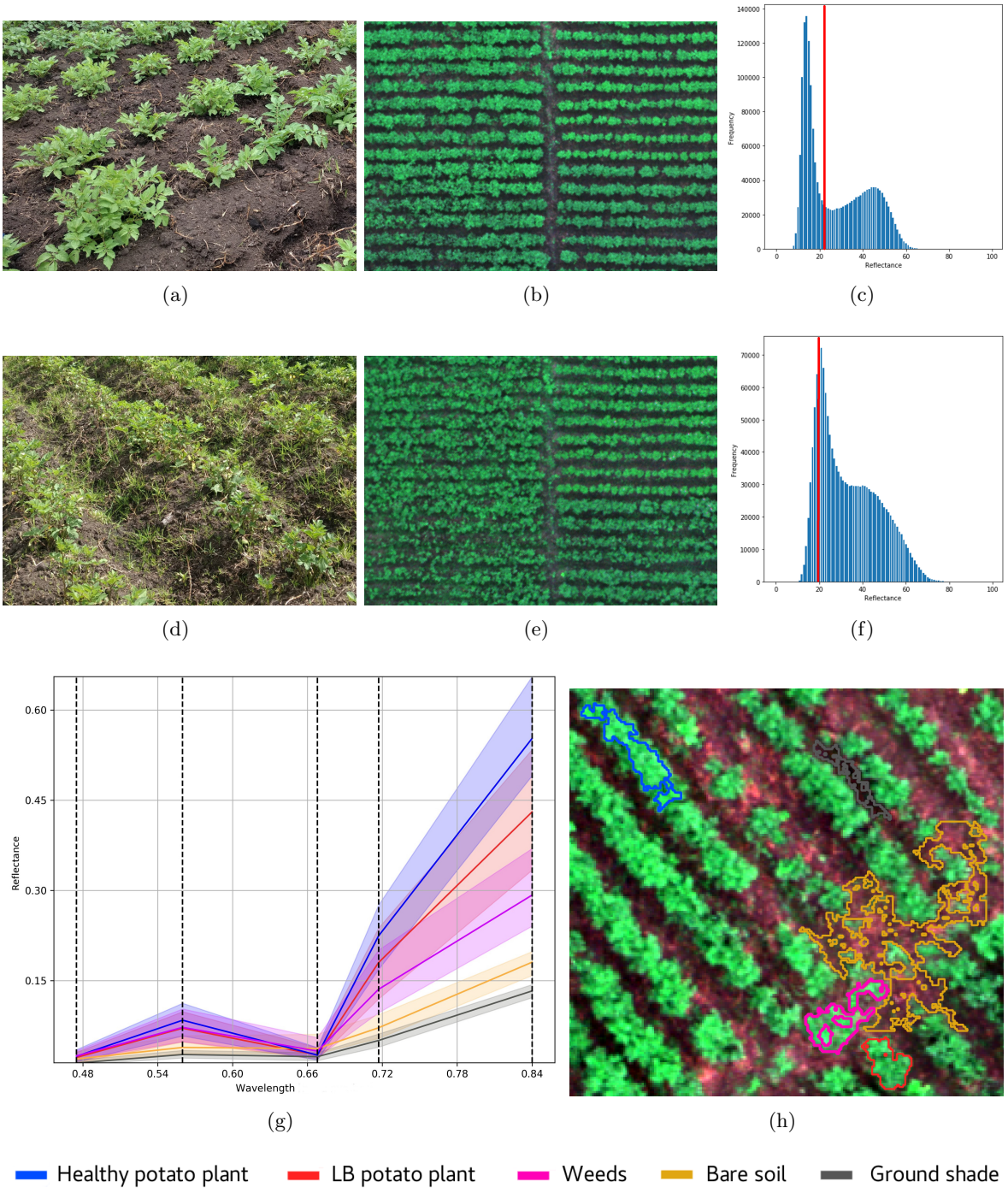


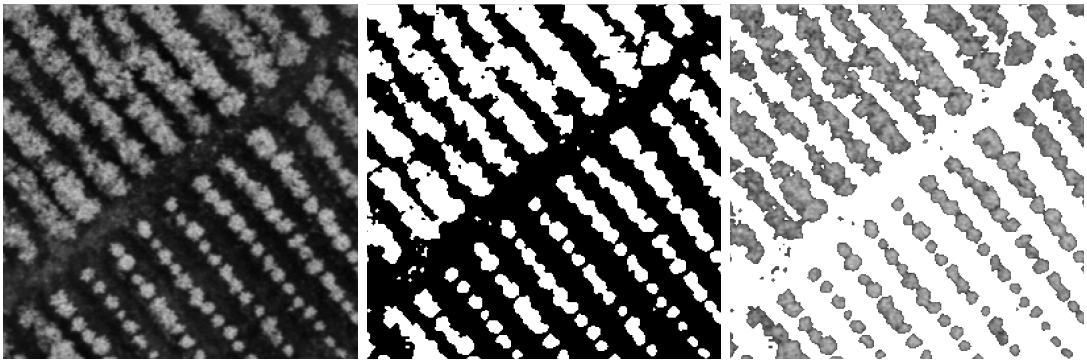
Figure 3.7: Ground and aerial view of the potato experimental crop, histograms for the NIR band and spectral responses of five objects identified in the orthophoto. (a) ground reference for images without weeds; (b) aerial image reference for images without weeds; (c) Near infrared histogram for images without weeds; (h) ground reference for weeds presence; (i) aerial image reference for weeds presence; (f) Near infrared histogram for images with weeds presence; (g) reflectance values of five objects in the orthophoto (Healthy potato plants, LB potato plants, weeds, bare soil and ground shade); and (h) the five objects of interest in the orthophoto.

Equation 3.7.

$$C_i(x, y) = A_i(x, y)B(x, y), \tag{3.7}$$

where $C_i(x, y)$ is the resulting pixel value at position (x, y) for i band without background, $A_i(x, y)$ is the original pixel value at position (x, y) for i band of the multispectral image and $B(x, y)$ is the binary image pixel value at position (x, y) created in the thresholding process.

Figure 3.8c illustrates how looks the infrared band from the multispectral image after the masking step. It can be seen how the bare soil surface does not appear any more in this image.



(a) (b) (c)

Figure 3.8: Images from background removal step: (a) Original band; (b) Binary image used as mask for ground removal; (c) band without bare soil and weeds.

3.6.2 Spectral traits

We used a three-stage method for assessment of LB using spectral traits:

(i) data processing for VIs calculation;

(ii) data analysis for model training and supervised classification stage;

(iii) use of trained models to classify a new dataset.

Figure 3.9 illustrates the workflow conducted at each stage. Eight broadband VIs where derived from the multispectral images Table 3.3. SIX VIs, including SAVI, NDVI, GNDVI, EVI, Two-band enhanced vegetation index (EVI2) and LAI were based on visible and NIR reflectance, and the other two, VIs Red-edge normalized difference vegetation index (NDVIRE) and Red-edge Chlorophyll index (CIRE), were based on the red edge and NIR reflectance.

Traits importance for variable selection was calculated using RF. Variable selection using RF has been used before in general classification tasks Speiser et al. (2019), to improve the land cover classification in urbanized coastal areas, and is widely used in pattern recognition Zhang and Yang (2020) . There are two main variable selection objectives: (i) to find important

variables highly related to the response variable for interpretation purposes; (ii) to find a few variables sufficient to a very good explanatory prediction of the response variable Genuer et al. (2010). In this study we carried on variable selection for the second reason.

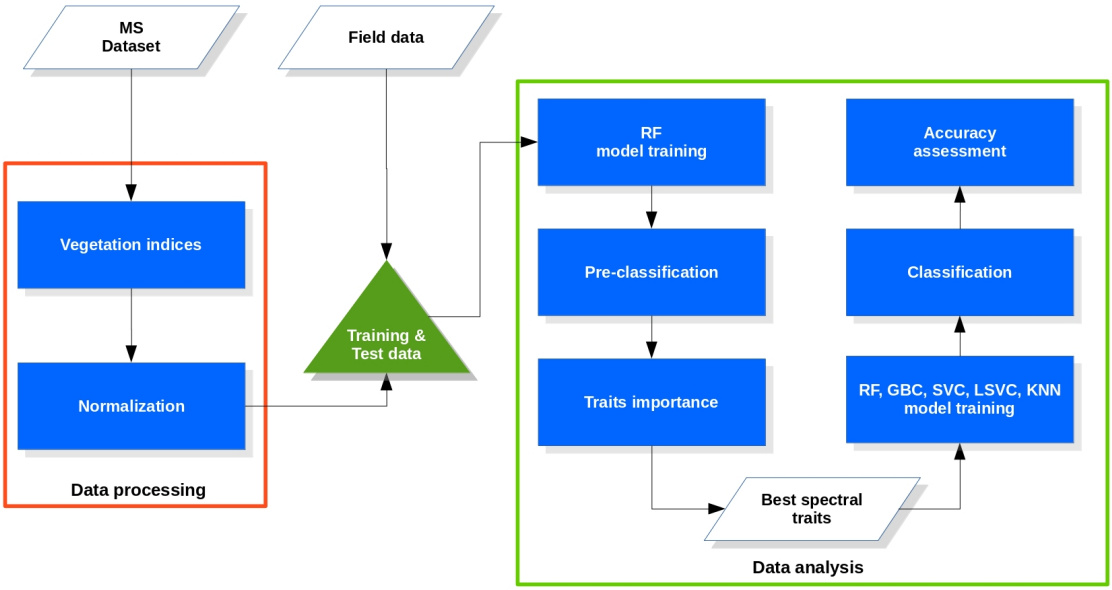
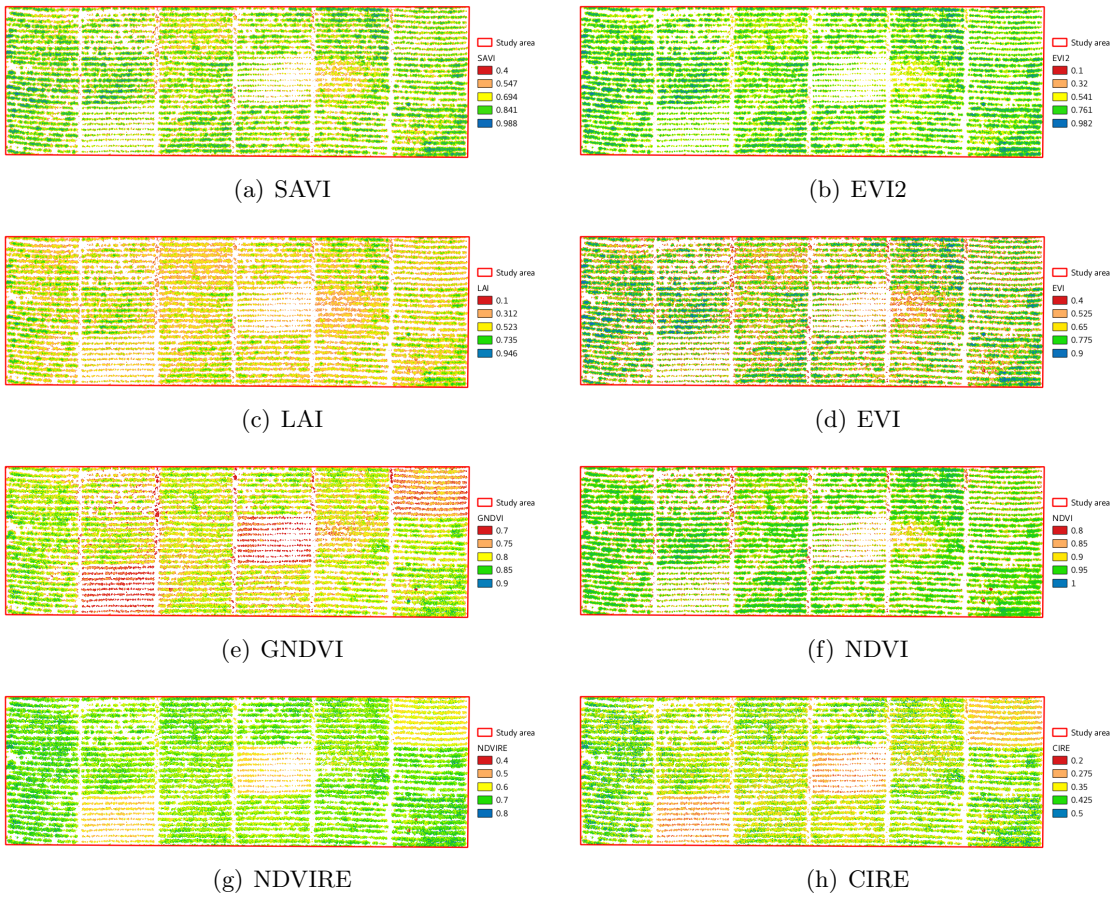


Figure 3.9: Workflow of the method conducted for assessing LB from UAV-based spectral traits.

Table 3.3: Vegetation indices selected in this study; NIR, RE, R and G represent the surface reflectance of near infrared, red-edge, red and green bands of MicaSense multispectral image, respectively.

Index	Name	Formula	References
SAVI	Soil adjusted vegetation index	$\frac{NIR-R}{(NIR+R+L)}(1+L)$	Huete (1988)
EVI2	Two-band enhanced vegetation index	$2.5 * \frac{NIR-R}{(NIR+2.4R+1)}$	Jiang et al. (2008)
LAI	Leaf area index	$-(\frac{1}{k})ln(a(1-bEVI2))$	Liu et al. (2012)
EVI	Enhanced vegetation index	$2.5 * \frac{NIR-R}{(NIR+6R-7.5B+1)}$	Huete et al. (2002)
GNDVI	Green normalized difference vegetation index	$\frac{NIR-G}{NIR+G}$	Sankaran et al. (2018)
NDVI	Normalized difference vegetation index	$\frac{NIR-R}{NIR+R}$	Rouse J.˜W. et al. (1974)
$NDVI_{RE}$	Red-edge normalized difference vegetation index	$\frac{NIR-RE}{NIR+RE}$	Gitelson and Merzlyak (1994)
CI_{RE}	Chlorophyll index Red-edge	$\frac{NIR}{RE}-1$	Gitelson et al. (2003)



(a) SAVI (b) EVI2

(c) LAI (d) EVI

(e) GNDVI (f) NDVI

(g) NDVIRE (h) CIRE

Figure 3.10: Vegetation indices for Dataset A.

Figure 3.10 and Figure 3.11 show the VIs derived for each dataset. It is possible to see huge differences in the red edge based VIs for the two periods under study.

3.6.3 Morphological traits

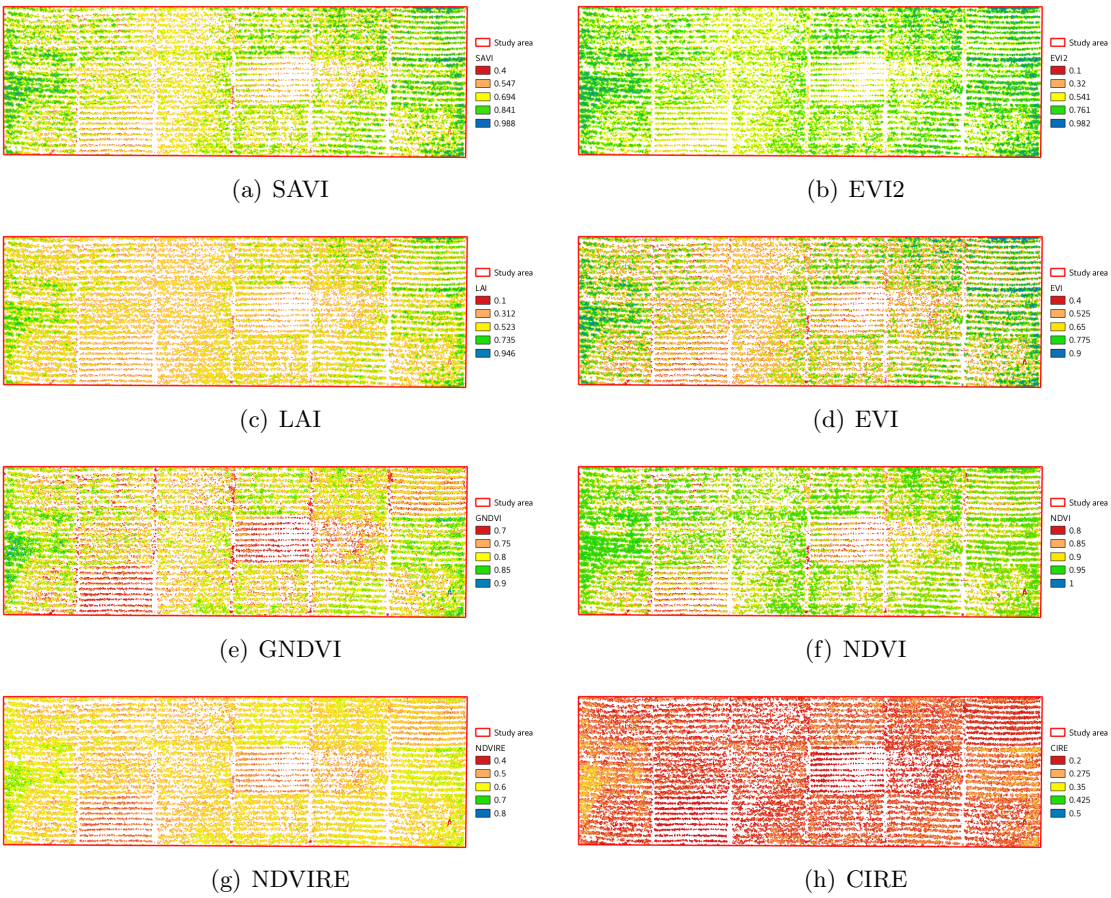
We used a three-stage method for assessing of LB using morphological traits:

(i) data processing for morphological traits extraction;

(ii) data analysis for model training and supervised classification stage;

(iii) use of trained models to classify a new dataset.

Figure 3.12 illustrates the workflow conducted at each stage.



(a) SAVI (b) EVI2

(c) LAI (d) EVI

(e) GNDVI (f) NDVI

(g) NDVIRE (h) CIRE

Figure 3.11: Vegetation indices for Dataset B.

Morphological traits extraction

In this work, two different segmentation methods were used, the first method was the Otsu method, used with the objective of removing the soil (bottom) in the orthomosaics. The second method was Felzenszwalb and Huttenlocher (2004) method, used with the aim of performing an over-segmentation (generating superpixels from the orthomosaic from which the background was previously removed), so that the regions found by this last method could be used to extract morphological the features.

To be able to derive morphological traits from the multispectral images, first it was necessary to create an RGB composited image, this image composition was created by enhancing the segmented bands NIR, Red and Green of the mutispectral image. The image enhancing consisted of an stretch to minimum and maximun reflectance values of each band. The RGB composite was then oversegmented using Felzenswalb method. This method measures the evidence for a boundary between two regions by comparing two quantities: one based on intensity differences across the boundary, and the other based on intensity differences between neighbouring pixels within each region.

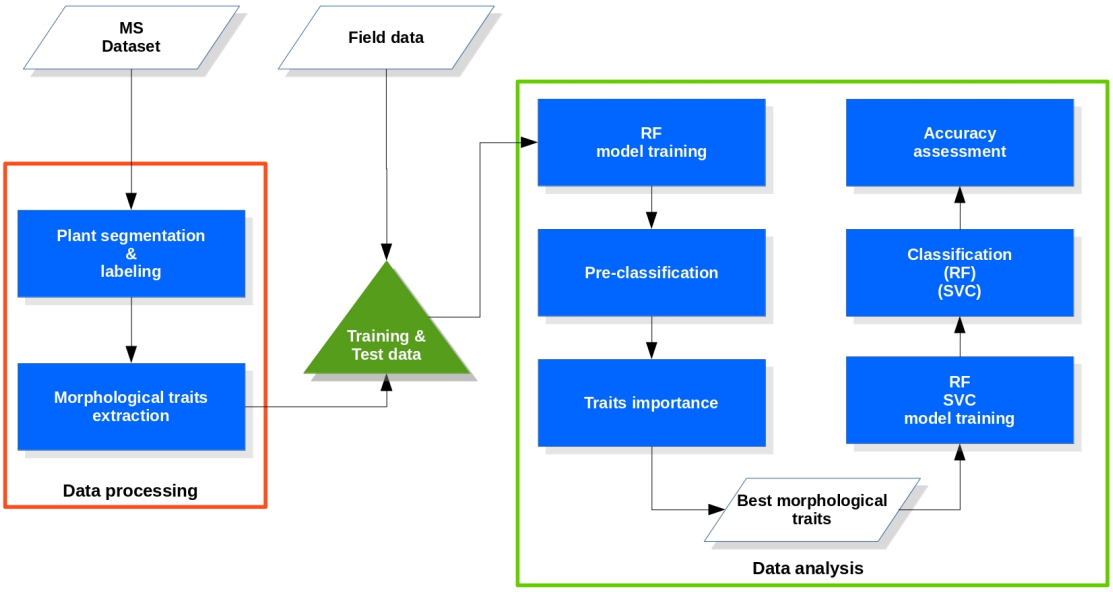
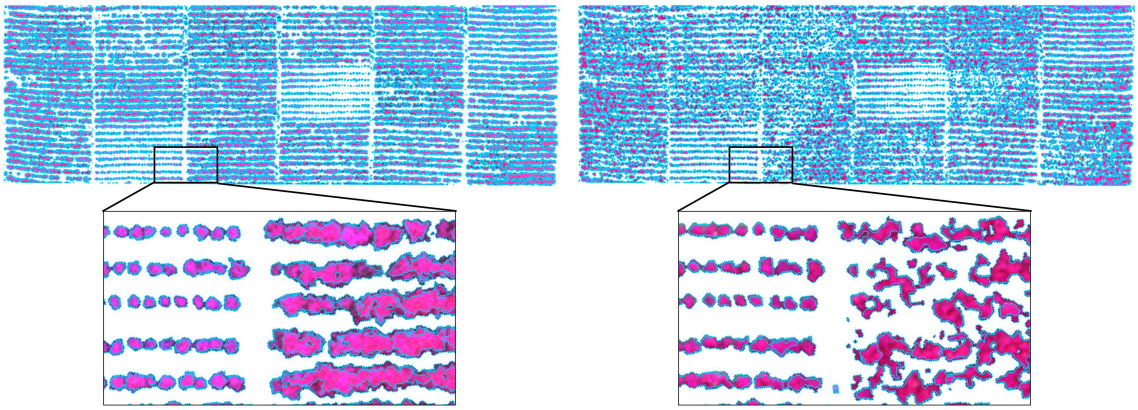


Figure 3.12: Workflow conducted for assessing LB from UAV-based morphological traits.

Although most of the studies presented in the chapter 2 were based on the use of images to derive morphological traits, it should be noted that, with the exception to the study carried on by Li et al. (2019), those studies were based on very high-resolution images captured with proximal sensors at very short distances that allowed for the extraction of features at the leaf level. In this study, the spatial resolution only allowed to work at the plant level, when it was possible to separate individual plants and at the row-level where the complex overlapping of the plants made it difficult to separate them.

We used the Felzenszwalb and Huttenlocher (2004) segmentation method available in scikit-image libraries Van Der Walt et al. (2014) to process the image obtained after the background removing step. This method produces an oversegmentation, also called super-pixels, of a multi-channel image using a fast, minimum spanning tree based clustering on the image grid. The function has a parameter called scale that sets an observation level. Higher scale means less and larger segments. There is also a parameter called sigma that is the diameter of a Gaussian kernel, used for smoothing the image prior to segmentation. Segment size within an image can vary greatly depending on local contrast. For RGB images, the algorithm computes a separate segmentation for each channel and then combines these. The combined segmentation is the intersection of the separate segmentations on the colour channels (Figure 3.13).



(a) (b)

Figure 3.13: Oversegmentation of datasets A and B using Felzenszwalb and Huttenlocher (2004) method. (a) dataset A; and (b) dataset B. Blue polygons represent the boundaries of the segments (superpixels) found for each dataset. These superpixels then serve as a basis for the extraction of morphological traits.

Plant area

The amount of space enclosed by the plant boundary (Figure 3.14b) calculated using the Equation 3.8.

$$PA = \sum_{i}^{N} P_i g^2 \tag{3.8}$$

where P_i is a pixel belonging to the plant region, N is the number of pixels in the plant region and g is the image Ground Sample Distance (GSD).

Plant perimeter

Length of the plant boundary (red polygon in Figure 3.14a) which approximates the contour as a line through the centres of border pixels using a 4-connectivity (orange polygon in Figure 3.14c).

Plant area-perimeter ratio

The relation between the plant area and the plant perimeter in pixels.

$$APR = \frac{PA}{PP} \tag{3.9}$$

Convex hull area

Number of pixels of convex hull image, which is the smallest convex polygon that encloses the region (Figure 3.14d).

Plant Solidity

Plant area over area of the convex hull area.

$$PS = \frac{PA}{CHA} \tag{3.10}$$

Plant Orientation

Angle between the 0th axis (rows) and the major axis of the ellipse that has the same second moments as the region, ranging from -pi/2 to pi/2 counter-clockwise.

Plant Major axis length

The length of the major axis of the ellipse that has the same normalised second central moments as the region (orange line in Figure 3.14f).

Plant Minor axis length

The length of the minor axis of the ellipse that has the same normalised second central moments as the region (orange line in Figure 3.14g).

Plant Minor axis length / Major axis length ratio

The relation between the minor axis length and the major axis length of the ellipse that has the same normalised second central moments as the region.

Plant Bounding box area

Area of the bounding box of the plant region calculated using Equation 3.11.

$$PBA = \sum_{i}^{N} \sum_{j}^{M} P_{i,j} g^2 \tag{3.11}$$

where $P_{i,j}$ is a pixel belonging to the plant bounding box region, N is the number of rows in the plant bounding box, M is the number of columns in the plant bounding box, and g is the image GSD.

Plant Eccentricity

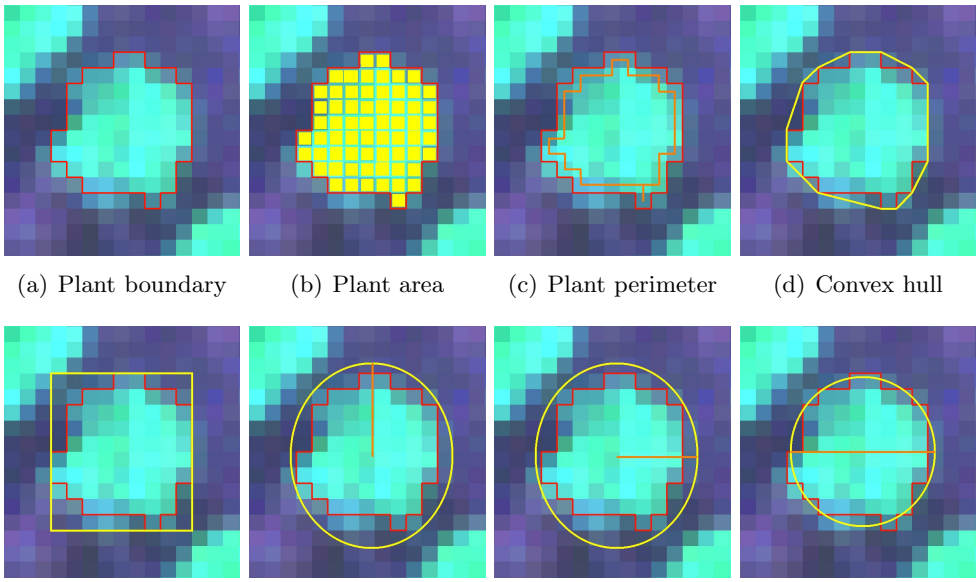
Eccentricity of the ellipse that has the same second-moments as the region. The eccentricity is the ratio of the focal distance (distance between focal points) over the major axis length. The value is in the interval $[0, 1)$. When it is 0, the ellipse becomes a circle.

Plant Equivalent diameter

It is the diameter of the circle which have the same area of the plant region (orange line in Figure 3.14h).

Plant extent

The plant area over bounding box area.



(a) Plant boundary (b) Plant area (c) Plant perimeter (d) Convex hull

(e) Bounding box (f) Major axis length (g) Minor axis length (h) Equivalent diameter

Figure 3.14: Main plant morphological traits. (a) plant perimeter (red polygon); (b) convex hull (yellow polygon); (c) Bounding box (yellow polygon); (d) Major axis length (orange line); (e) minor axis length (orange line); and (f) equivalent diameter (orange line).

Figure 3.15 shows the different stages necessary to derive what we called synthetic bands of morphological traits. Once the oversegmentation of the RGB image was carried out, the resulting image contains superpixels (Figure 3.15a), for each one of those superpixels we calculated the traits presented before (i.e. area, perimeter), the result is a dictionary containing the traits for each object (Figure 3.15c). The next step was to create new images and to assign for each pixel of an object the value of its morphological trait (Figure 3.15d), then, each new

image became a morphological band (Figure 3.15e) and was stacked in a n-dimensional array. Then, traits importance for variable selection was calculated using RF similar to the method used for spectral traits.

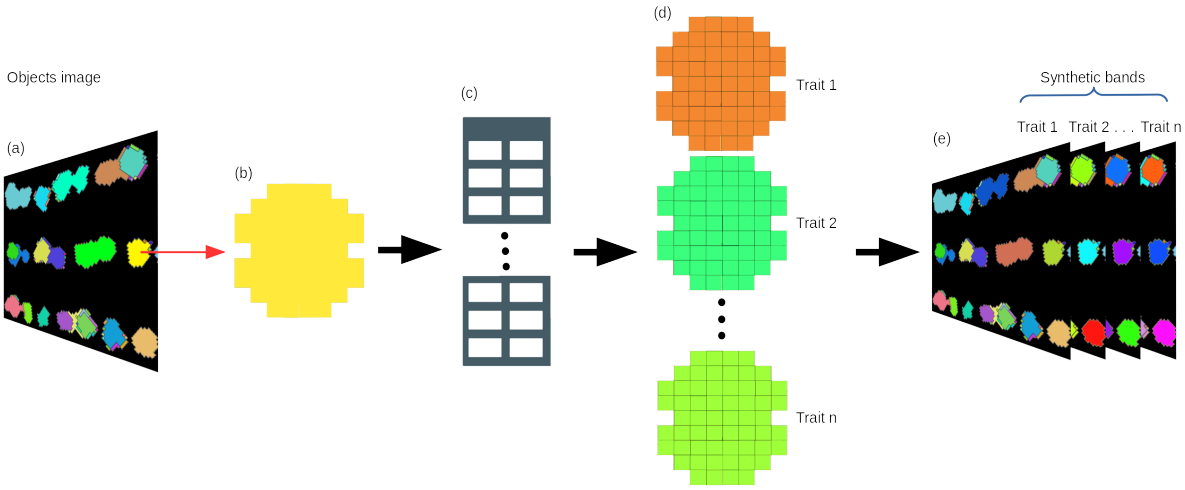


Figure 3.15: Construction of morphological traits bands. (a) objects image; (b) object representing a plant; (c) morphological plant traits table; (d) assignment of table features to each pixel; and (e) stacking of morphological bands.

Figure 3.16 and Figure 3.17 show the morphological bands derived from the multispectral bands by using Felzenszwalb and Huttenlocher (2004) method.

3.6.4 Combined spectral and morphological traits

The major challenge when combining spectral and morphological traits is the differences between the nature of spectral and morphological data. Spectral bands and vegetation indices are data based on pixels while morphological bands are region-based data. However, we tackled down this issue by creating synthetic morphological bands. This bands had the same spatial resolution as the spectral bands, so, stacking into the same n-dimensional array was possible. The resulting array had 20 dimensions, 8 belonging to the VIs and 12 belonging to the morphological traits. Figure 3.18 shows the conceptual approach at the object level and band level of the adopted strategy for combining spectral and morphological traits.

Spatial and spectral information have been combined before for dimension reduction Xu et al. (2019). This combination is based on the idea that spatial and spectral information can be taken into account through the spectral constraint of the superpixel regions. Superpixel segmentation is used here to profit from the spatial neighbourhood structure as it can segment images adaptively, matching the spatial characteristics.

3.6.5 Training data

The training zones, shown in Figure 3.19e, were vectorized from the information collected in the field by visual inspection of the experimental crop. Zones corresponding to diseased a and

(a) Area

(b) Perimeter

(c) Area/perimeter ratio

(d) Solidity

(e) Orientation

(f) Major axis length

(g) Minor axis length

(h) Minor/Major axes ratio

(i) Bounding box area

(j) Eccentricity

(k) Equivalent diameter

(l) Extent

Figure 3.16: Morphological traits dataset A.

(a) Area

(b) Perimeter

(c) Area/perimeter ratio

(d) Solidity

(e) Orientation

(f) Major axis length

(g) Minor axis length

(h) Minor/Major axes ratio

(i) Bounding box area

(j) Eccentricity

(k) Equivalent diameter

(l) Extent

Figure 3.17: Morphological traits dataset B.

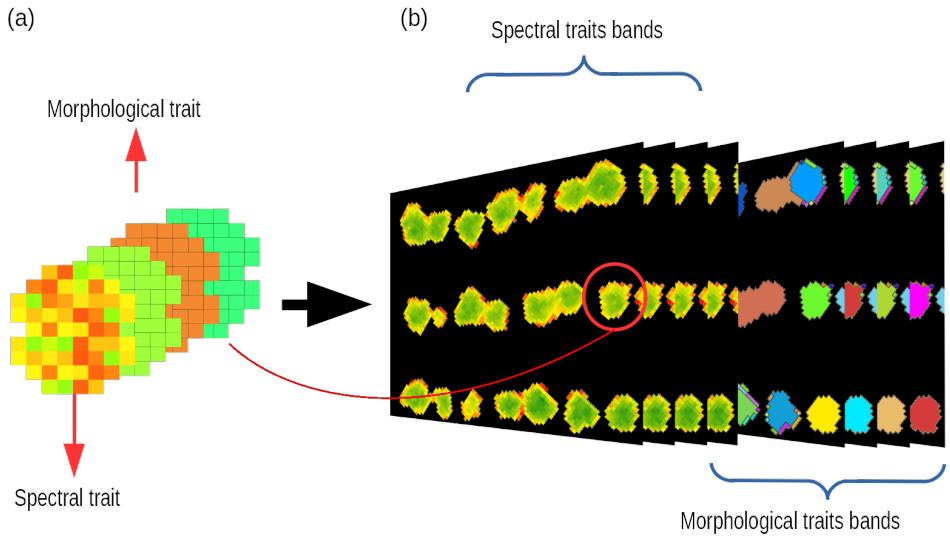


Figure 3.18: Representation of combined spectral and morphological traits and bands. (a) object level spectral and morphological traits; (b) n-dimensional array stacking spectral and morphological traits bands.

healthy plants were identified in the field, and then the pixels belonging to those regions were identified in the multispectral images. This process required image-enhancing techniques to improve visualisation of the regions of interest and further confirmation based on the mean spectral responses of those zones in the images. Figure 3.19a and Figure 3.19b show the box-plot for each multispectral band and the mean spectral responses for the training zones. It can be seen that regions belonging to diseased plants share similar spectral responses in the orthophoto with a mean reflectance in the red edge and the near infrared lower than those regions were healthy potato plants were identified. Although individual spectral responses of healthy potato plants and LB potato plants overlapped at certain point, averaged spectral responses of diseased and healthy plants responses showed a clear separation.

Figure 3.19c and Figure 3.19d show ground-based images of two plants, identified in the field as a diseased and a healthy plant, respectively. In total 44 polygons were used to build the models: 23 for plants with evidence of LB and 21 for healthy plants. The reference polygons were then rasterized and randomly divided into two sets, a training set made up of 75% of the data and a test set made up of the remaining 25% of the data. Performance of the classifiers during the training was evaluated through the Receiver Operating Characteristic (ROC) curves (Omar and Ivrissimtzis, 2019). The area under the ROC curve (AUC), is a common metric that represents the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance (Fawcett, 2006).

3.6.6 Classification methods

Four classification methods where tested: Random Forest (RF), Gradient Boosting Classifier (GBC), Support Vector Classifier (SVC), LSVC and K-Nearest Neighbors Classifier (KNN).

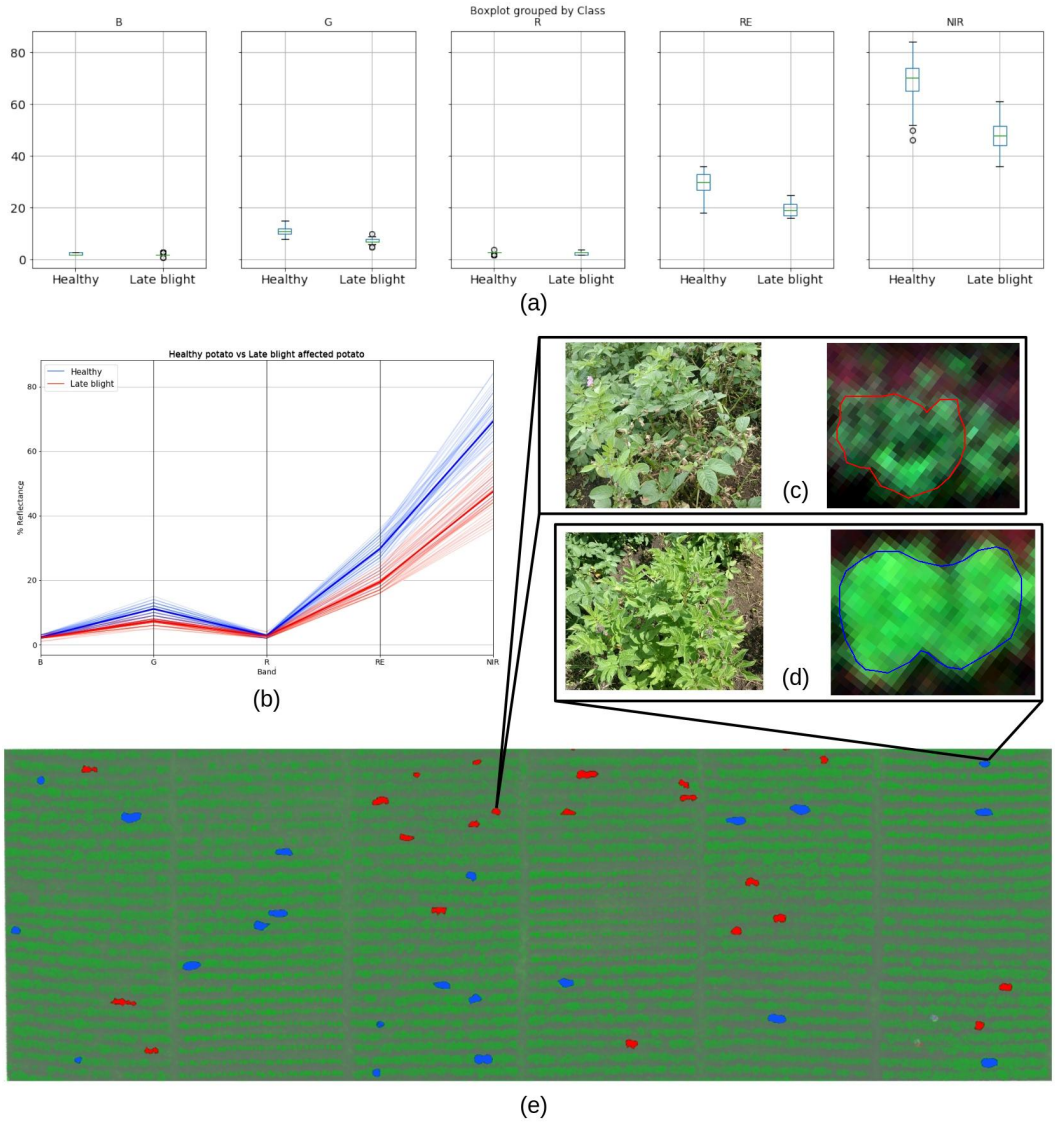


Figure 3.19: Spectral responses of healthy and diseased potato plants belonging to the training zones based on field visual inspection of the experimental crop. (a) Box-plot displaying the distribution of healthy and LB plants spectral responses; (b) spectral responses of the healthy and LB plants for the five bands of the Micasense camera; (c) potato plant with LB and its orthophoto view; (d) healthy potato plant and its orthophoto view; and (e) training zones based on field visual inspection of the experimental crop. LB zones are shown in red colour and healthy plants in blue colour.

Random forest

RF is a machine learning algorithm which combines hundreds of decision trees where each tree rely upon the values of a random vector sampled independently (Breiman, 2001). Prediction is made by aggregating the predictions of the ensemble by majority vote for classification or averaging for regression. RF does not require values to follow a particular statistical distribution as it is a non-parametrical method (Puissant et al., 2014; Rodriguez-Galiano et al., 2012).

The RF trained model was based on 500 trees, the criterion for measure the quality of a split was Gini impurity, the max depth of the trees was 4, the minimum number of samples required to split an internal node was 2, the minimum number of samples required to be at a leaf node was 2, the maximum number of features to consider when looking for the best split was set as the square root of the number of features. Bootstrap samples were used when building trees instead of the whole dataset, out-of-bag samples were used to estimate the generalisation accuracy. Weights for each class were not used to train the model. AUC of the RF model was 0.92.

Gradient Boosting Classifier

GBC is a decision trees-based classification method that has been used in image classification for which it is not required to assume any distribution of the data to be classified, although it is a highly sensitive method to training data that is incorrectly labelled or that presents ambiguity between classes and to unbalanced datasets which can affect the classification results (Lawrence et al., 2004).

The trained model obtained from GBC was based on 75% random samples from the entire training dataset, the learning rate was 0.8, 500 boosting stages were performed, sub-sample was set to 1, the minimum number of samples required to split an internal node was 4, the minimum number of samples required to be at a leaf node was 1, maximum depth of the individual regression estimators was 4. The AUC of the GBC model was 0.91.

C-Support Vector Classification

Support Vector Classification (SVC) is a classification method created initially for binary classification but has been extended to multi-class problems (Hsu and Lin, 2002). For multi-class classifications, several binary classifiers must be built, which makes its execution demanding in terms of processing capacity, making it more suitable for small datasets (Hsu and Lin, 2002; Chang and Lin, 2011).

The main parameters of the trained model obtained from SVC were set up as follows: regularisation parameter of 1, the Radial Basis Function (RBF) kernel was used, the gamma value was set to $\frac{1}{n\sigma_n^2}$, where σ_n^2 is the variance of the training input samples and n is the number of features. The AUC of the SVC model was 0.91.

Linear SVC

In this method, a random number generator is used to select characteristics when fitting the model. As a result, it is possible to get slightly different results for the same dataset (Varoquaux et al., 2015). Unlike SVC, LSVC does not use a kernel for model training, a feature that allows training with much larger data sets using a linear classifier (Fan et al., 2008).

The main parameters of the trained model obtained from LSVC were set up as follows: regularisation parameter of 1, tolerance for stopping criteria of 1e-4, multi-class strategy was one-vs-rest (ovr), so if the training data contain more than two classes the training method is one class versus the rest of classes. The maximum number of iterations was 1000. The AUC of the LSVC model was 0.92.

K-Nearest Neighbours

K-Nearest Neighbours (KNN) is a method of data classification that can be supervised or unsupervised, unlike RF, GBC, SVC and LSVC, it belongs to the group of instance-based classifiers, that is, KNN does not create a model from training data but stores instances of those data and classifies the whole dataset by majority vote of the nearest neighbours of each point (Lee et al., 2012; Varoquaux et al., 2015). The KNN algorithm used implements two data structures to handle training data and compute neighbours, BallTree (Omohundro, 1989) and KDTree (Moore, 1991). It is also possible to compute neighbours by brute-force (Varoquaux et al., 2015).

The main parameters of the trained model obtained from KNN were set up as follows: number of neighbours of 5, all points in each neighbourhood were weighted equally, the most appropriate KNN algorithm (BallTree, KDTree or brute-force) for the training was selected based on the values passed to fit method, the leaf size passed to either BallTree or KDTree algorithms in case of being selected was 30, the lower parameter for the Minkowski metric was 2 and the distance metric to use for the tree was minkowski. The AUC of the KNN model was 0.91.

3.6.7 Classification of dataset B using pre-trained models

The RF, GBC, SVC, LSVC and KNN models trained with dataset A were stored for use in dataset B. Dataset B was processed following the steps of the data analysis stage. After that, the pre-trained models were loaded and image classification was performed directly without going through the training and fitting stages of the models. This allows to establish the level of transferability of the ML models.

Transferability of the ML models was measured by using the metric Generalisation Loss (G-loss) introduced by Klemenjak et al. (2019). This metric links the accuracy of the classification results obtained for the ML methods applied to the datasets A and B according to the Equation 3.12.

$$\text{G-loss} = \left(1 - \frac{OA_b}{OA_a}\right) 100\% \tag{3.12}$$

where, OA_a is the accuracy of the classification results of the dataset A and OA_b is the accuracy of the classification results of the dataset B.

3.6.8 Accuracy Assessment

Ground truth

The reference classifications (Figure 3.20) of the study area were provided by experts from the research group, who prepared them based on field observations, spectral responses of healthy and diseased plants and visual image interpretation. It should be noted that the assessment of diseased areas based on image interpretation is a very difficult task. This is due to different factors such as the mixture of healthy and diseased areas in the same pixel, the occlusion of diseased areas in the lower parts of the canopy by healthy vegetation in the upper layers of the canopy, the occurrence of subtle symptoms in plants, especially when the severity levels are low, among others. Evaluation of the classification models in this scenario can lead to unreliable results. To be able to compare the performance of the ML algorithms, the predicted percentage of late blight severity on the experimental plot was obtained for each experimental block, excluding the areas between blocks, which contained weeds that could affect the results.

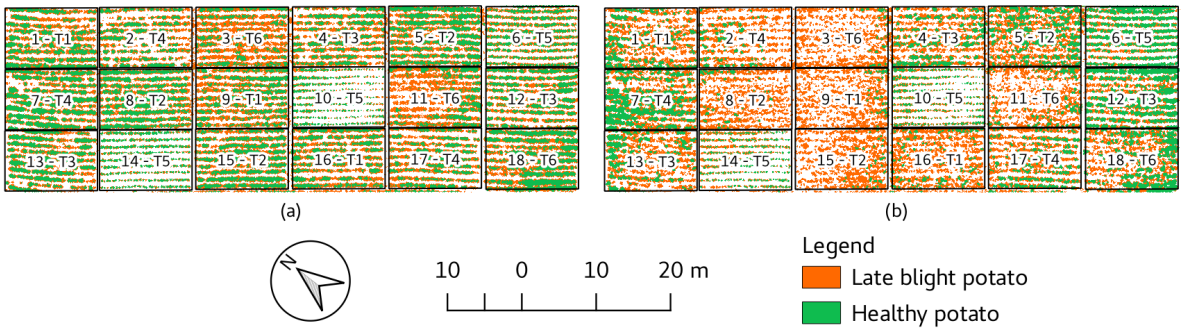


Figure 3.20: Reference classifications. (a) dataset A; (b) dataset B.

Background removing quality

The quality of the segmentation results was evaluated using reference segmentation of the experimental plot. The elaboration of the reference segmentation was carried out by manual editing of the two data sets. In this work we use the QGIS Semi-Automatic Classification Plugin (SCP) (Congedo, 2021) to generate polygons of interest that correspond to the areas of bare soil and weeds, and then, with the same tool, proceed to eliminate said areas of one of the mosaic bands until only the areas belonging to potato plants were present. Finally that band was used to generate a binary image that was used as ground truth. The Intersection-over-Union (IoU), also known as the Jaccard index was used (Equation 3.13). IoU is the area

of intersection between the predicted segmentation and the reference segmentation, divided by the joint area between the predicted segmentation and the reference segmentation (Drass et al., 2020).

$$Q = \frac{A \cap B}{A \cup B} \tag{3.13}$$

Thematic accuracy

Thematic quality of classification was evaluated using reference classifications of the experimental plot. The confusion matrix between reference classification and each classification result from ML methods was calculated. Based on the confusion matrix the following metrics were derived: overall accuracy (OA) (Equation 3.14), precision (Equation 3.15), recall (Equation 3.16) and F-score (Equation 3.17), (Tharwat, 2018). Matthews correlation coefficient (MCC) (Equation 3.18) was also calculated as it offers higher confidence whit unbalanced data sets (Tharwat, 2018).

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.14}$$

$$Precision = \frac{TP}{TP + FP} \tag{3.15}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.16}$$

$$F - score = \frac{2TP}{FP + 2TP + FN} \tag{3.17}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{3.18}$$

where TP, FP, TN and FN are the true positive, false positive, true negative and false negative samples, respectively.

The *scikit-learn* library (Varoquaux et al., 2015), was used to train the models, perform the classification and assess accuracy in this work. A prototype of the method was implemented in python and is available at the following link: `https://jorlrodriguezg.github.io/`

4 Results

4.1 Background removing

Figure 4.1a and Figure 4.1b show the orthophotos obtained for the two datasets: (a) dataset A; (b) dataset B. Visual assessment of potato LB from orthophotos is possible, although, it requires significant effort. It can be seen that there are important changes between the two dates. On dataset A the crop rows are clearly defined for the entire crop even in those areas where LB was evident. On the contrary, on dataset B only in the control block crop rows were certainly defined while for most of the crop area rows did not have a consistent pattern.

Figure 4.1c and Figure 4.1d show the orthophotos before the background removal step. After background removal (Figure 4.1i and Figure 4.1j), those areas affected by LB became evident just by the identification of the erratic pattern found in the crop rows. This pattern is produced by the loss of leaves in the potato plants caused by the disease.

4.2 Spectral traits

Figure 4.2 shows the 2D projection of the decision boundaries for the case of the red and infrared bands. Although it should be highlighted that these projections may not appropriately represent the final decision boundaries and surely vary when taking into account the 5 dimensions of the multispectral image, they do provide an approximation to the way in which the data could be grouped into LB plants and Healthy plants.

Figure 4.3 corresponds to the classification results when using spectral traits for each dataset under study: dataset A and dataset B. (a) dataset A, RF classification result; (b) dataset A, GBC classification result; (c) dataset A, SVC classification result; (d) dataset A, LSVC classification result; (e) dataset A, KNN classification result; (f) dataset B, RF classification result; (g) dataset B, GBC classification result; (h) dataset B, SVC classification result; (i) dataset B, LSVC classification result; (j) dataset B, KNN classification result. In both cases it is possible to see that it appears to be an overestimation of the LB presence within the experimental crop.

Figure 4.4 shows the variable importance measures from RF when using only spectral traits, 5, 6, 7 and 8 corresponds to SAVI, EVI2, LAI and NIR respectively which had the highest measures. However, the accuracy results were better when using only SAVI and EVI2 bands as input in the classification stage.

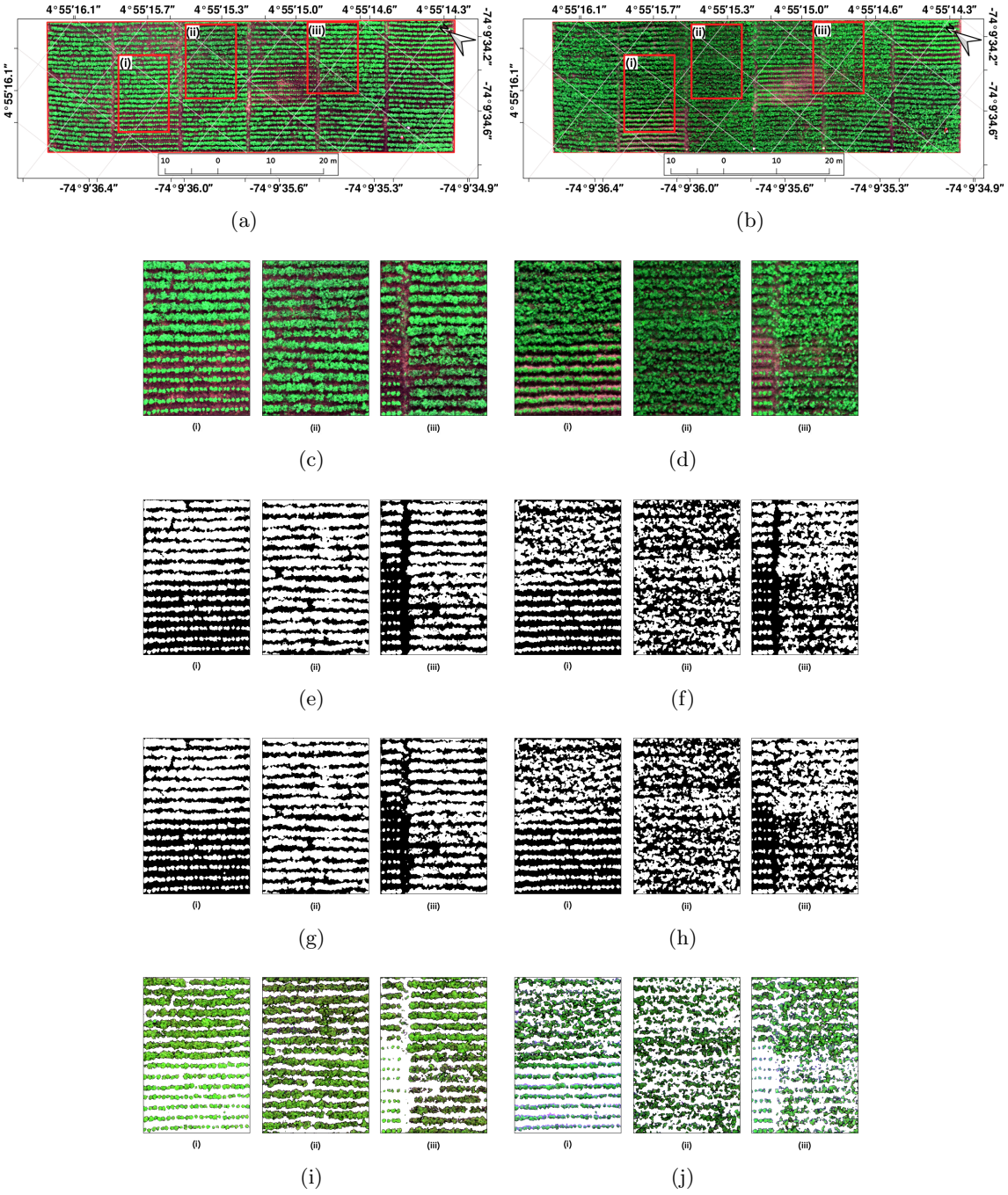
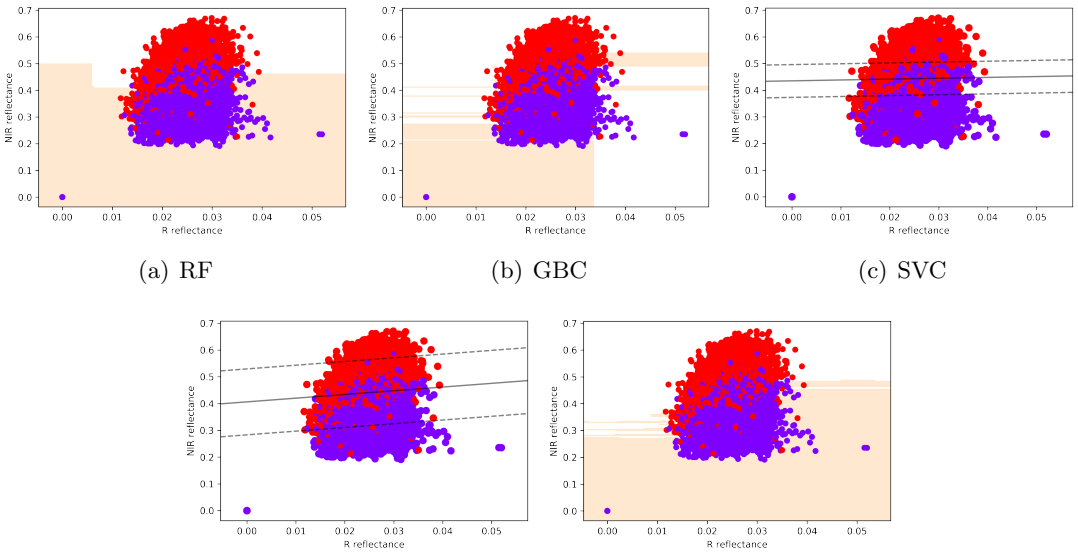


Figure 4.1: RGB orthophotos obtained for each period time under study before and after background removal. The red polygons in (a) and (b) indicate the areas shown in higher resolution below each orthophoto. (a) dataset A, before background removal; (b) dataset B, before background removal; (c) close-up of dataset A, before background removal; (d) close-up of dataset B, before background removal; (e) close-up of dataset A manual segmentation reference; (f) close-up of dataset B manual segmentation reference; (g) close-up of dataset A segmentation result; (h) close-up of dataset B segmentation result; (i) close-up of dataset A final segmentation result yield by (g); and (j) close-up of dataset B final segmentation result yield by (h).



(a) RF (b) GBC (c) SVC

(d) LSVC (e) KNN

Figure 4.2: 2D projection of the decision boundary plots for each classification method, here we only consider bands red and near infrared. (a) RF; (b) GBC; (c) SVC; (d) LSVC; and (e) KNN. The purple dots correspond to LB and the red dots to healthy plants.

4.3 Morphological traits

Figure 4.5 corresponds to the classification results when using morphological traits for each dataset under study: dataset A and dataset B. (a) dataset A, RF classification result; (b) dataset A, GBC classification result; (c) dataset A, SVC classification result; (d) dataset A, LSVC classification result; (e) dataset A, KNN classification result; (f) dataset B, RF classification result; (g) dataset B, GBC classification result; (h) dataset B, SVC classification result; (i) dataset B, LSVC classification result; (j) dataset B, KNN classification result. It is difficult to the naked eye to see a coincidence with the ground truth.

Figure 4.6 shows the variable importance measures from RF when using only morphological traits. The three most important variables from this measures were area, equivalent diameter and minor axis length. However, after repeated experiments, varying the combination of the morphological traits, we found that the morphological traits that gave the higher accuracy results were: area, eccentricity and extent.

4.4 Combined spectral and morphological traits

Figure 4.7 corresponds to the classification results when combining spectral and morphological traits for each dataset under study: dataset A and dataset B. (a) dataset A, RF classification result; (b) dataset A, GBC classification result; (c) dataset A, SVC classification result; (d)

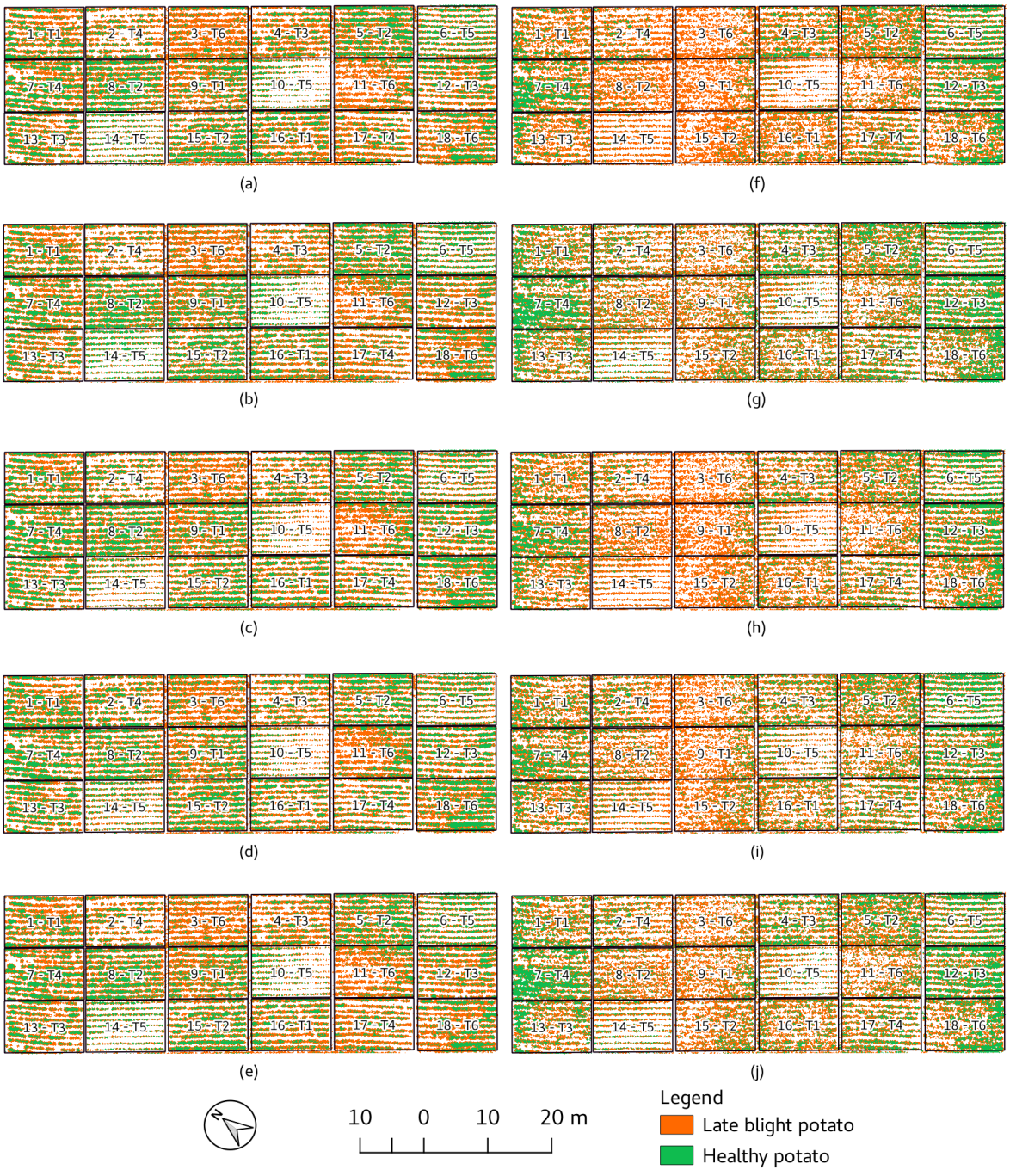
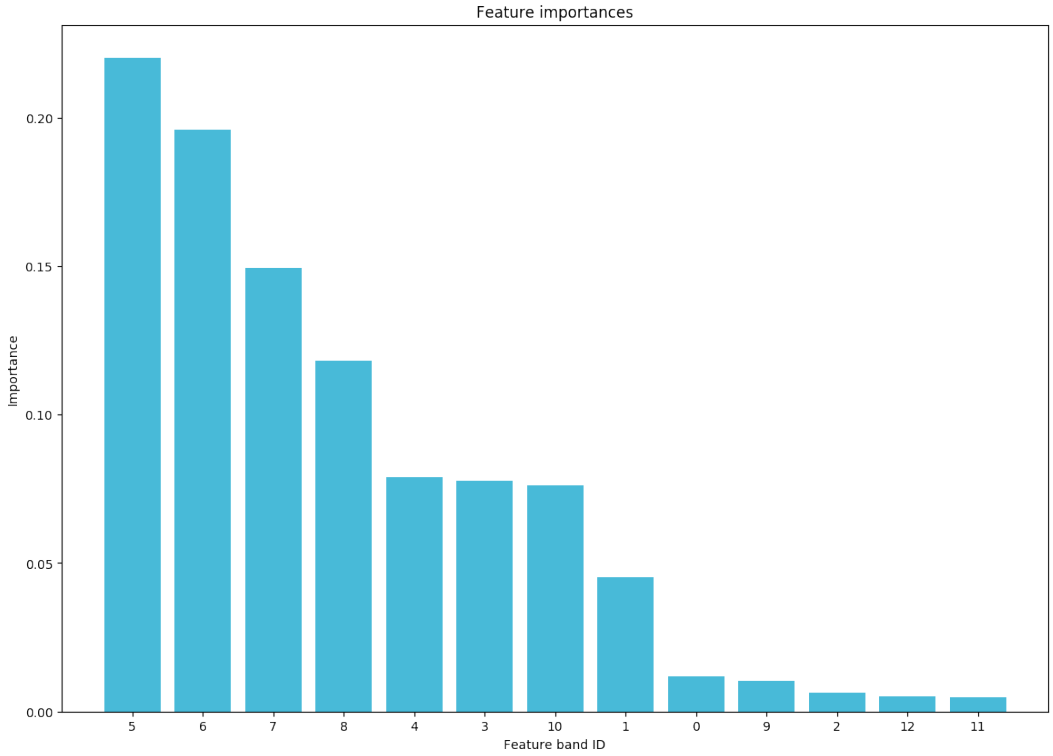


Figure 4.3: Classification results when using spectral traits for each dataset under study: dataset A and dataset B. (a) dataset A, RF classification result; (b) dataset A, GBC classification result; (c) dataset A, SVC classification result; (d) dataset A, LSVC classification result; (e) dataset A, KNN classification result; (f) dataset B, RF classification result; (g) dataset B, GBC classification result; (h) dataset B, SVC classification result; (i) dataset B, LSVC classification result; (j) dataset B, KNN classification result.



ID	**0**	**1**	**2**	**3**	**4**	**5**	**6**
Trait	Blue	Green	Red	Red edge	NIR	SAVI	EVI2

ID	**7**	**8**	**9**	**10**	**11**	**12**
Trait	LAI	EVI	GNDVI	NDVI	$NDVI_{RE}$	CI_{RE}

Figure 4.4: Spectral traits - Variable important measures from RF for spectral traits.

dataset A, LSVC classification result; (e) dataset A, KNN classification result; (f) dataset B, RF classification result; (g) dataset B, GBC classification result; (h) dataset B, SVC classification result; (i) dataset B, LSVC classification result; (j) dataset B, KNN classification result.

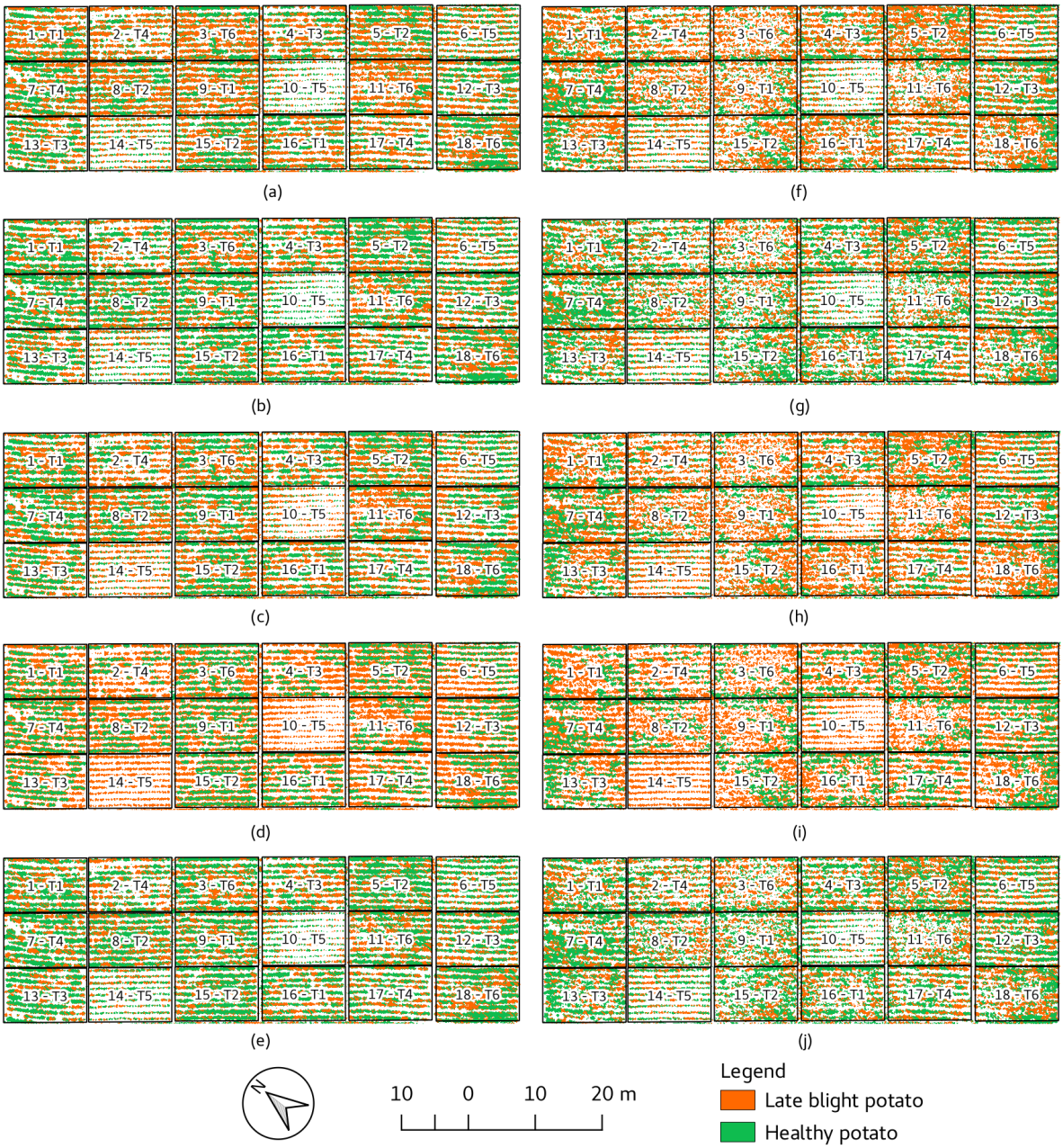
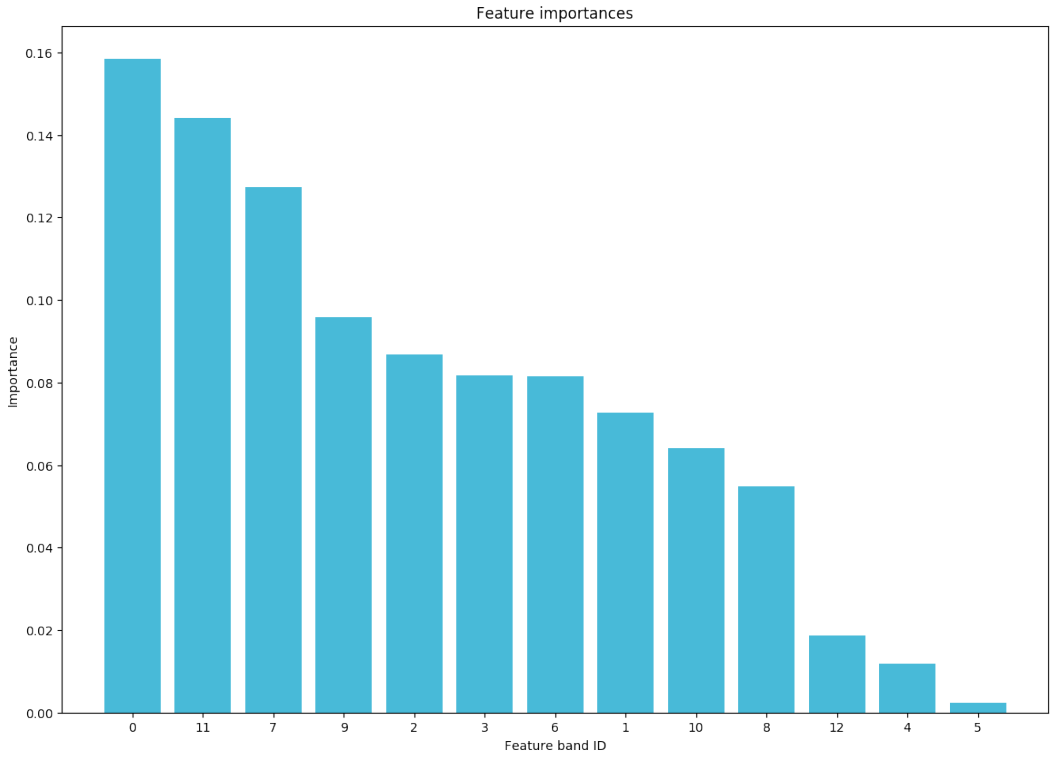


Figure 4.5: Classification results when using morphological traits for each dataset under study: dataset A and dataset B. (a) dataset A, RF classification result; (b) dataset A, GBC classification result; (c) dataset A, SVC classification result; (d) dataset A, LSVC classification result; (e) dataset A, KNN classification result; (f) dataset B, RF classification result; (g) dataset B, GBC classification result; (h) dataset B, SVC classification result; (i) dataset B, LSVC classification result; (j) dataset B, KNN classification result.



ID	0	**1**	**2**	**3**	**4**	**5**	**6**
Trait	Area	Perimeter	Area/Perimeter	Convex hull area	Solidity	Orientation	Major axis length
ID	**7**	**8**	**9**	**10**	**11**	**12**	
Trait	Minor axis length	Minor/Mayor axis ratio	Bounding box area	Eccentricity	Equivalent diameter	Extent	

Figure 4.6: Morphological traits - Feature importance measures from RF.

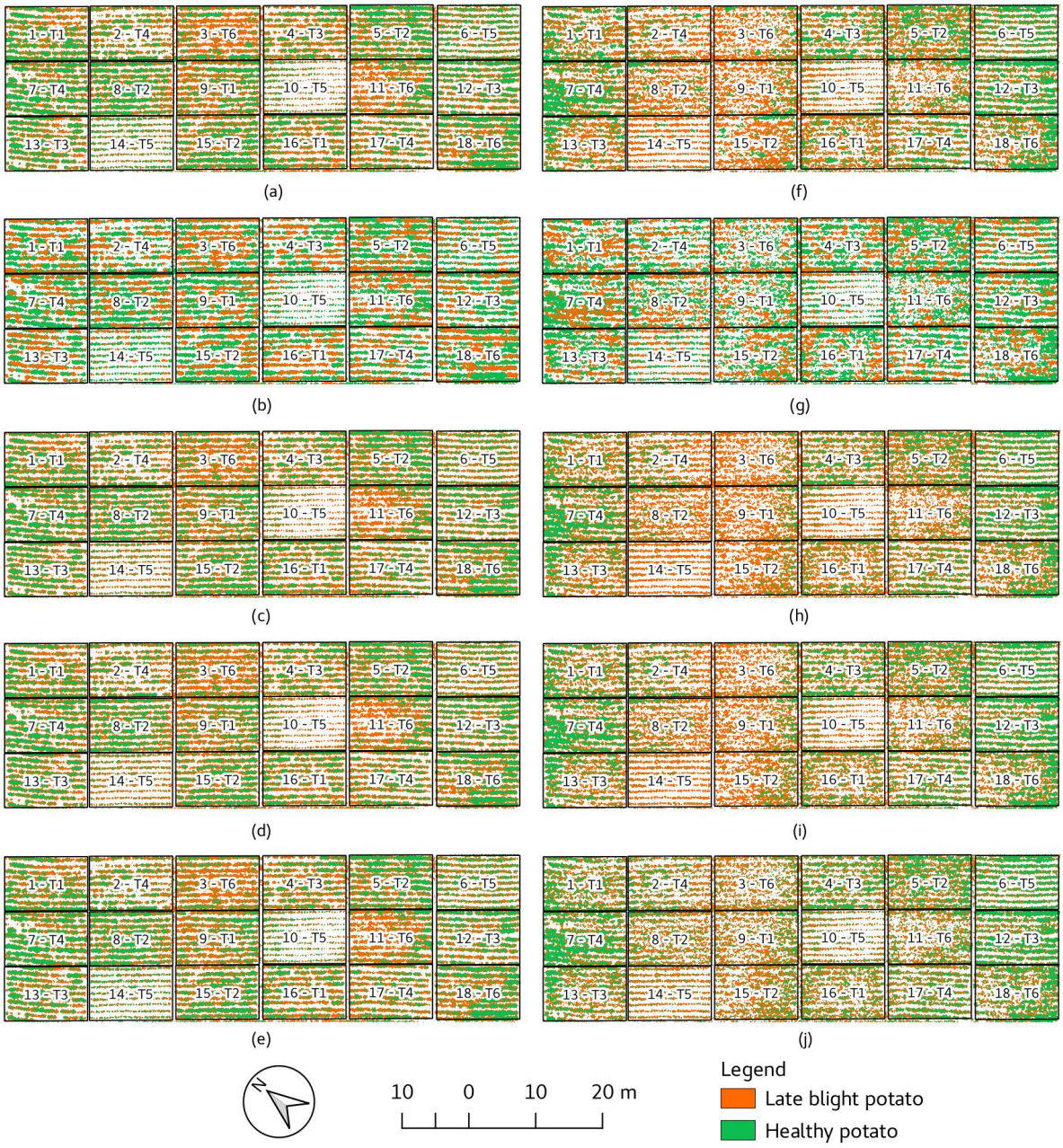


Figure 4.7: Classification results when combining spectral and morphological traits for each dataset under study: dataset A and dataset B. (a) dataset A, RF classification result; (b) dataset A, GBC classification result; (c) dataset A, SVC classification result; (d) dataset A, LSVC classification result; (e) dataset A, KNN classification result; (f) dataset B, RF classification result; (g) dataset B, GBC classification result; (h) dataset B, SVC classification result; (i) dataset B, LSVC classification result; (j) dataset B, KNN classification result.

5 Discussion

The aim of this thesis was to assess potato LB from UAV-based multispectral imagery through the extraction of spectral and morphological traits and the use of machine learning algorithms. Although the use of spectral traits and the combination of spectral and morphological traits showed very promising results, there is still work to do to improve the results when only morphological traits are used.

This chapter is organised as follows: section 5.1 contains the analysis of background removing, section 5.2 presents the analysis of the use of spectral traits for assessing LB, section 5.3 presents the analysis of the use of morphological traits for assessing LB, section 5.4 presents the analysis of the use of combined spectral and morphological traits for assessing LB, section 5.5 presents the general discussion, section 5.6 summarise the accomplished objectives, section 5.7 deploys the advantages of the methodology, and section 5.8 deploys the contributions to knowledge in Geomatics.

5.1 Background removing

Table 5.1 shows the accuracy results for the background removing step. The process of background removing was aimed at suppressing bare soil from the image and reducing "noise" before classifying the data. However, since the background itself was not our object of interest it was removed based on the Otsu's threshold algorithm which was manually adjusted to improve the plant separation. For dataset A, where weeds were marginal inside the experimental blocks, adjusting the threshold value manually led to oversegmentation and deterioration of the segmentation results. However, for dataset B emerging weeds required to adjust the threshold obtained by Otsu's algorithm. Although adjusting the threshold value helped to remove most of the weeds present in the orthophoto for the two datasets, there was still presence of weeds in the segmented result. Overall weeds remnants presence was predominant only along a narrow zone dividing the experimental blocks where cultural management such as hill-up was not carried on. Thus, for this study those remnants were considered negligible. In addition, to be able to compare the performance of the ML algorithms, the predicted percentage of late blight severity on the experimental plot was obtained for each experimental block, excluding the areas between blocks, which contained remnants of weeds from the background removing step, that could affect the severity estimation results.

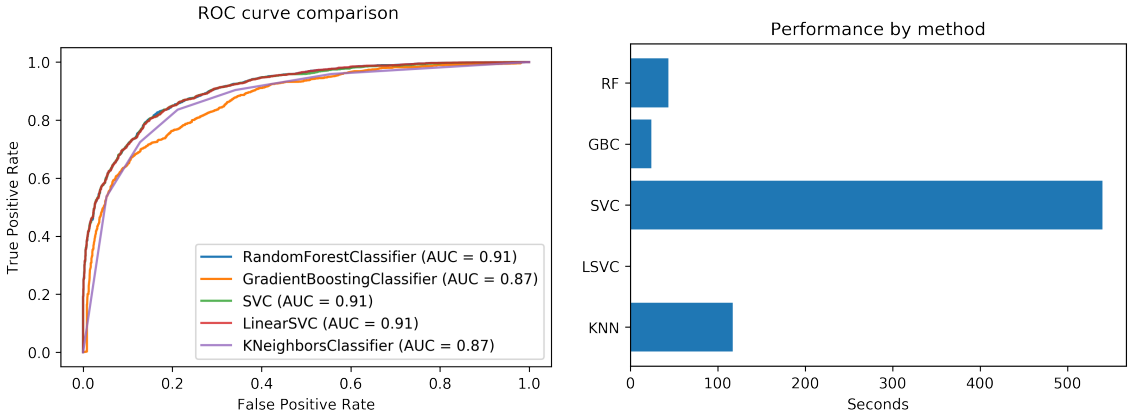
Applying a segmentation technique may significantly reduce the amount of data to be processed. Furthermore, most of the soil surface may be removed by this technique leading to an improvement in the processing time because of the filtering of less relevant data. However, as weed presence increments inside the crop, a thresholding method becomes less efficient and manual input is necessary to achieve the separation task between potato plants and the

Table 5.1: Segmentation accuracy for datasets A and B.

Dataset	IoU
Dataset A	97.7%
Dataset B	97.6%

background.

5.2 Spectral traits



(a) (b)

Figure 5.1: Spectral traits - Performance by classification method on the training dataset. (a) ROC curves for RF, GBC, SVC, LSVC and KNN; and (b) processing time for the ML algorithms.

Figure 5.1a shows the ROC curves used to measure how well the trained models fitted the training data. It can be seen that the models had a similar behaviour, in particular, the AUC is equal to 0.91 for RF, SVC and LSVC methods. By far the method that required the longest processing time was SVC, the run time was approximately 539.8 seconds, by comparison, KNN which was the second method at run time took approximately 117 seconds. RF and GBC had shorter execution times, with 43 seconds and 24 seconds respectively. The most efficient method in terms of execution time was LSVC with a time of 0.07 seconds (Figure 5.1b). It can be seen in Figure 4.2, that even when data are grouped in a very similar way, that is, red dots and purple dots are almost identical in all the plots, when it comes to the decision boundaries for each method, there are clear differences.

Table 5.2 shows thematic accuracy metrics: the classification method that achieves the best results, compared to the reference classifications, for the dataset A was SVC with OA = 0.897 and MCC=0.795 and for the dataset B was RF with OA = 0.822 and MCC=0.604. RF and SVC metrics were close enough that two decimal places did not allow to differentiate them, so three decimal places were needed to find the highest metrics. Overall, the main

Table 5.2: Spectral traits - F-Scores per class, overall accuracy and MCC obtained from assessment of image classifications versus reference classification. Bold values correspond to the highest OA, F-Scores and MCC.

Dataset	Class	RF			GBC			SVC			LSVC			KNN		
	category	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
Dataset A	LB	0.849	0.941	0.893	0.756	0.755	0.755	0.876	0.920	**0.897**	0.917	0.815	0.863	0.786	0.806	0.796
	Healthy	0.937	0.840	0.886	0.765	0.767	0.766	0.919	0.875	**0.897**	0.840	0.930	0.882	0.809	0.790	0.799
	OA		0.890			0.761			**0.897**			0.873			0.798	
	MCC		0.784			0.522			**0.795**			0.751			0.595	
Dataset B	LB	0.822	0.923	**0.870**	0.795	0.708	0.749	0.824	0.895	0.858	0.833	0.829	0.831	0.810	0.759	0.784
	Healthy	0.823	0.642	**0.721**	0.562	0.673	0.613	0.777	0.658	0.712	0.697	0.703	0.700	0.612	0.680	0.644
	OA		**0.822**			0.695			0.810			0.784			0.731	
	MCC		**0.604**			0.369			0.576			0.531			0.43	

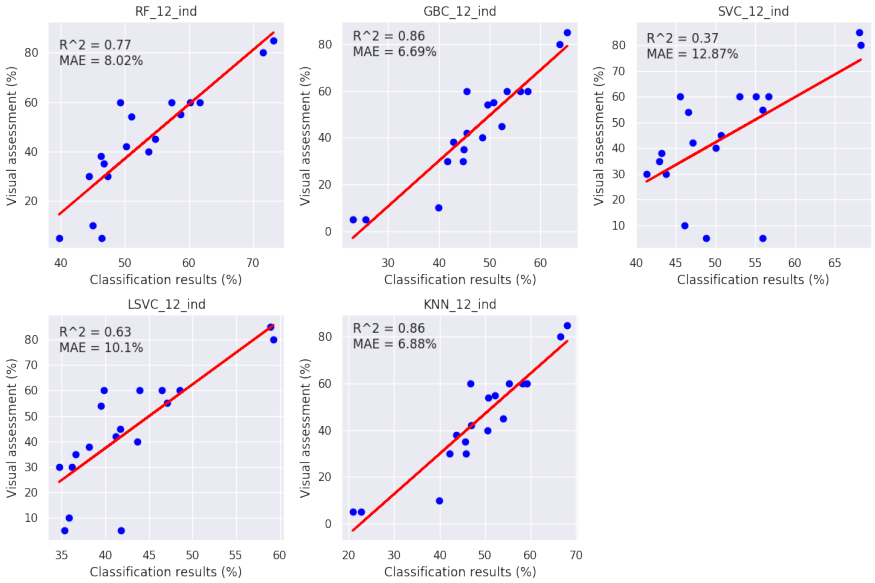
*Here we used three decimal places because the metrics were close for various classifiers, so two decimal places did not allow to differentiate them.

difference between RF and SVC in this study was the processing time. Precision and recall in Table 5.2 did not show significant evidence of difference in the capability of the ML algorithms to classify healthy plants than diseased plants for the dataset A. However, for the dataset B those metrics suggest a better capability of the ML algorithms to classify diseased plants than healthy plants.

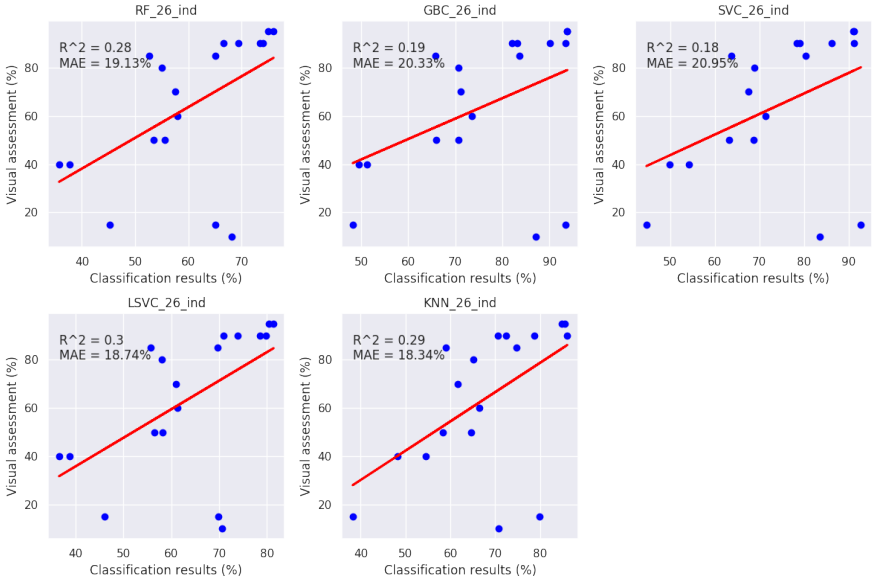
Figure 5.2 shows the correlation between infection level from classification results and visual assessment methods for each ML algorithm under study when using spectral traits. Figure 5.2a shows the correlation between infection level from classification results and visual assessment for the Dataset A, it is possible to see that GBC and KNN showed the best performance for the Dataset A with $R^2 = 0.86$ and a $MAE = 6.69\%$ and $R^2 = 0.86$ and a $MAE = 6.88\%$ respectively. Figure 5.2b shows the correlation between infection level from classification results and visual assessment for the Dataset B, it is possible to see that LSVC and KNN showed the best performance for the Dataset B with $R^2 = 0.30$ and a $MAE = 18.74\%$ and $R^2 = 0.29$ and a $MAE = 18.34\%$. Even though correlation graphics had a good MAE and R^2 for RF, GBC and KNN with MAE = 8.02% and $R^2 = 0.77$, MAE = 6.69% and $R^2 = 0.86$ and, MAE = 6.88% and $R^2 = 0.86$ respectively, it should be noted that, for our data, in all cases ML algorithms tend to overestimate the severity of late blight disease at the low level of infection (LB1). In general, results showed a better performance of the ML algorithms for advanced levels of infection (LB3 and LB4). Blocks with higher visual score had similar predicted values of infection.

Table 5.3 contains the RMSE for each ML and dataset studied. It can be seen that GBC and KNN showed the lower RMSE for the Dataset A with $RMSE = 8.48$ and $RMSE = 8.52$ respectively. Also, LSVC and KNN had the lower RMSE for the Dataset B with $RMSE = 24.94$ and $RMSE = 25.12$ respectively.

For the dataset A, SVC and LSVC overestimated the level of infection of blocks 6, 10 and 14 which corresponds to control blocks (T5) without any application of fertilisers. Those blocks were in a stage of development earlier than all the remaining experimental blocks and its spectral response was in average lower, in particular for dataset B, in which all the classifiers overestimated the level of affectation of those control blocks. Visual scoring and classification results for blocks with T6 showed the higher level of infection followed by T1, T4, T3 and T2, this suggest a relationship between the treatment and the response of potato plants to LB that should be further explored in future works.



(a) Dataset A - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.



(b) Dataset B - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.

Figure 5.2: Spectral traits - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study. (a) Correlation between infection level from image processing and visual assessment for the Dataset A; and (b) Correlation between infection level from classification results and visual assessment for the Dataset B. The solid red line shows the regression line. GBC and KNN showed the best performance for the Dataset A with $R^2 = 0.86$ and a $MAE = 6.69\%$ and $R^2 = 0.86$ and a $MAE = 6.88\%$ respectively. LSVC and KNN showed the best performance for the Dataset B with $R^2 = 0.30$ and a $MAE = 18.74\%$ and $R^2 = 0.29$ and a $MAE = 18.34\%$.

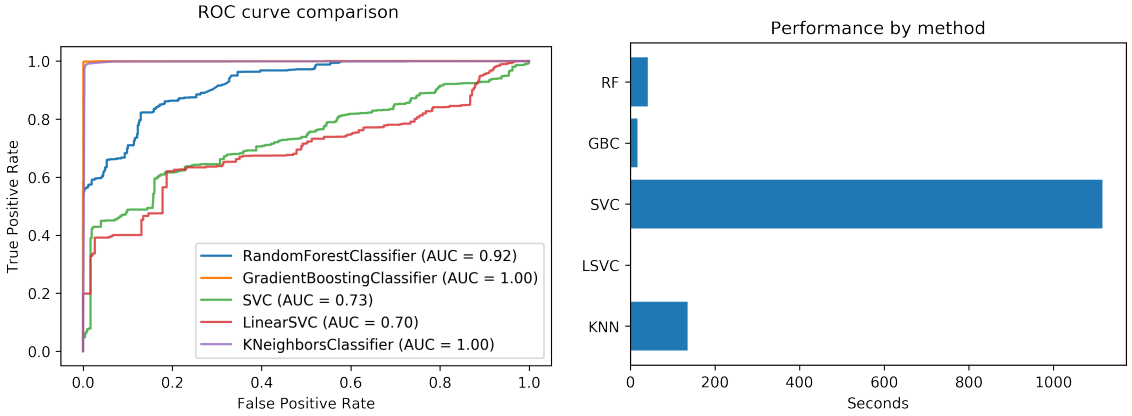
Table 5.3: Spectral traits - Mean absolute error, root mean squared error and coefficient of determination obtained from assessment of image classifications versus visual assessment.

ML algorithm		Dataset	
		A	**B**
RF	MAE	8.02	19.13
	RMSE	10.84	25.33
	R^2	0.77	0.28
GBC	MAE	6.69	20.33
	RMSE	8.48	26.73
	R^2	0.86	0.19
SVC	MAE	12.87	20.95
	RMSE	18.23	26.98
	R^2	0.37	0.18
LSVC	MAE	10.10	18.74
	RMSE	13.99	24.94
	R^2	0.63	0.30
KNN	MAE	6.88	18.34
	RMSE	8.52	25.12
	R^2	0.86	0.29

5.3 Morphological traits

Figure 5.3a shows the ROC curves used to measure how well the trained models fitted the training data. It can be seen that the models had a different behaviour, it can be seen that GBC and KNN had an Area Under the Curve (AUC) equal to one, which suggest over-fitting issues. By far the method that required the longest processing time was SVC, the run time was approximately 1100 seconds, by comparison, KNN which was the second method at run time took approximately 150 seconds. RF and GBC had shorter execution times, with 43 seconds and 24 seconds respectively. The most efficient method in terms of execution time was LSVC with a time of 0.1 seconds.

In this study, when using only morphological traits, we experimented varying the number of selected variables to be used to train the model. When using the first 6 bands found with RF, the AUC for random forest rose to 0.97 and for LSVC to 0.96, the validation against the reference classification also increased, although to a greater extent for LSVC than for random forest. When using the first 5 bands found with RF, the AUC for random forest fell to 0.96 and for LSVC to 0.92. In contrast, the validation against the reference classification improved, finding: Accuracy: 0.745 and MCC: 0.506 for RF and Accuracy: 0.816 and MCC: 0.653 for LSVC. When using the first 4 bands found with RF, the AUC for random forest dropped to 0.95 and for LSVC to 0.91. In contrast, the validation against the reference classification improved, finding: Accuracy: 0.749 and MCC: 0.517 for RF and Accuracy: 0.839 and MCC: 0.692 for LSVC.



(a) (b)

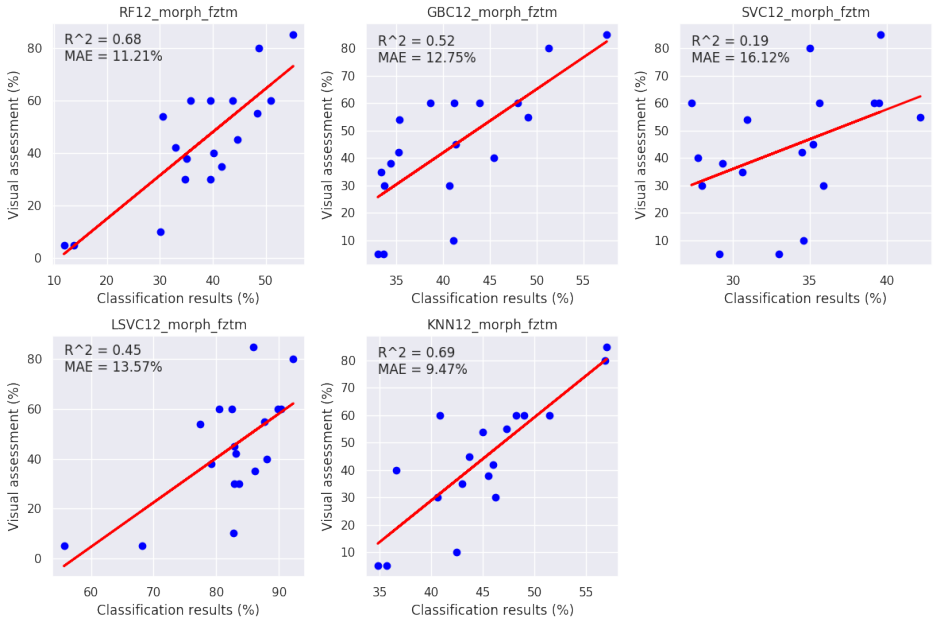
Figure 5.3: Morphological traits - Performance by classification method on the training dataset. (a) ROC curves for RF, GBC, SVC, LSVC and KNN; (b) Execution time for each classifier.

Table 5.4 contains the RMSE for each ML and dataset studied. It can be seen that KNN showed the lower RMSE for the Dataset A with $RMSE = 12.98$. Also, RF had the lower RMSE for the Dataset B with $RMSE = 19.78$.

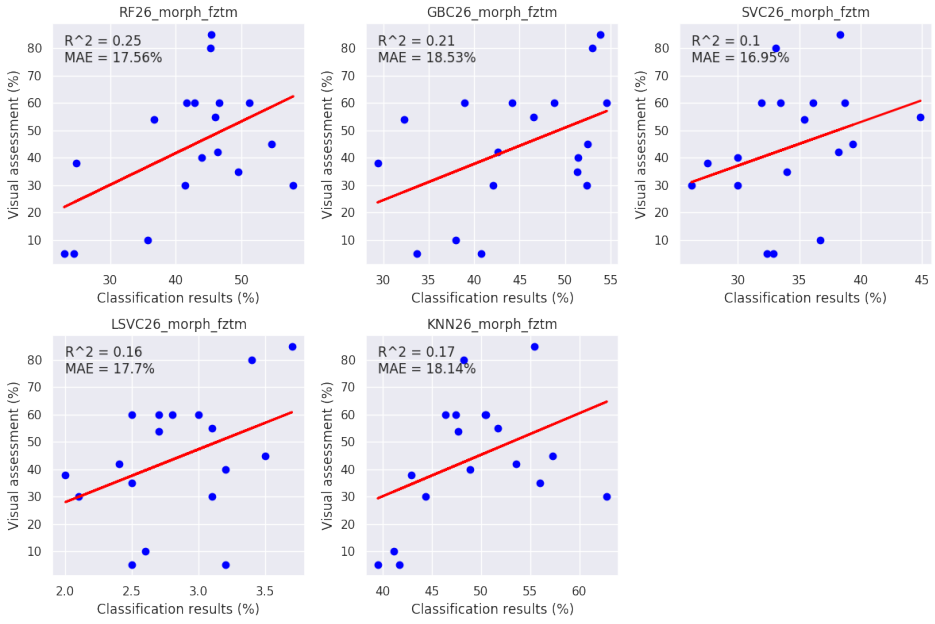
Table 5.5 shows thematic accuracy metrics when using morphological traits: the classification method that achieves the best results, compared to the reference classifications, for the two datasets was LSVC with OA = 0.681 and MCC=0.375 for the dataset A and OA = 0.612 and MCC = 0.165. Although precision and recall did not show significant evidence of difference in the capability of the ML algorithms to classify healthy plants than diseased plants for the dataset A for the dataset B those metrics suggest a better capability of the ML algorithms to classify diseased plants than healthy plants, similar to the findings when using only spectral traits.

Figure 5.4 shows the correlation between infection level from classification results and visual assessment methods for each ML algorithm under study when using morphological traits. Figure 5.4a shows the correlation between infection level from classification results and visual assessment for the Dataset A, it is possible to see that KNN showed the best performance for the Dataset A with $R^2 = 0.69$ and a $MAE = 9.47\%$. Figure 5.4b shows the correlation between infection level from classification results and visual assessment for the Dataset B, it is possible to see that RF showed the best performance for the Dataset B with $R^2 = 0.25$ and a $MAE = 17.56\%$.

Experiments showed that by working with the most important variables calculated from RF the performance of the classification results versus the ground truth and the visual scoring were lower than the performance achieved using other morphological traits. Thus, variable importance measures from RF can be unreliable when working with morphological traits only. This results seems to agree with the reported bias in importance measures from RF by Strobl et al. (2007).



(a) Dataset A - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.



(b) Dataset B - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.

Figure 5.4: Morphological traits - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study. (a) Correlation between infection level from image processing and visual assessment for the Dataset A; and (b) Correlation between infection level from classification results and visual assessment for the Dataset B. The solid red line shows the regression line. KNN showed the best performance for the Dataset A with $R^2 = 0.69$ and a $MAE = 9.47\%$. RF showed the best performance for the Dataset B with $R^2 = 0.25$ and a $MAE = 17.56\%$.

Table 5.4: Morphological traits - Mean absolute error, root mean squared error and coefficient of determination obtained from assessment of image classifications versus visual assessment when using morphological traits.

ML algorithm		Dataset	
		A	**B**
RF	MAE	11.21	17.56
	RMSE	12.98	19.78
	R^2	0.68	0.25
GBC	MAE	12.75	18.53
	RMSE	15.65	20.33
	R^2	0.52	0.01
SVC	MAE	16.12	16.95
	RMSE	20.68	21.65
	R^2	0.19	0.10
LSVC	MAE	13.57	17.70
	RMSE	17.01	20.99
	R^2	0.45	0.16
KNN	MAE	9.47	18.14
	RMSE	12.73	20.81
	R^2	0.69	0.17

Table 5.5: Morphological traits - F-Scores per class, overall accuracy and MCC obtained from assessment of image classifications versus reference classification. Bold values correspond to the highest OA, F-Scores and MCC.

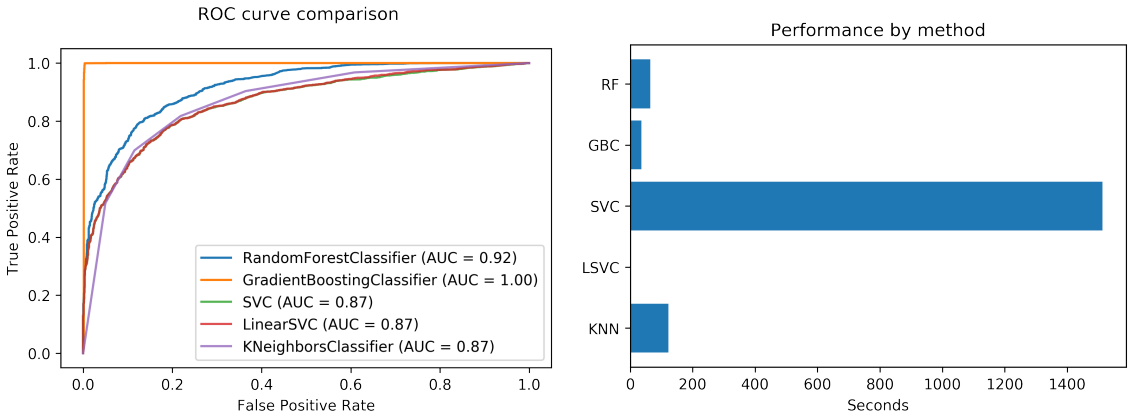
Dataset	**Class**	**RF**			**GBC**			**SVC**			**LSVC**			**KNN**		
	category	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
Dataset A	LB	0.499	0.585	0.539	0.534	0.482	0.506	0.542	0.590	0.565	0.642	0.788	**0.707**	0.545	0.470	0.505
	Healthy	0.525	0.439	0.478	0.546	0.597	0.570	0.571	0.521	0.545	0.740	0.579	**0.650**	0.552	0.625	0.586
	OA		0.510			0.541			0.555			**0.681**			0.549	
	MCC		0.024			0.079			0.112			**0.375**			0.096	
Dataset B	LB	0.666	0.655	0.661	0.659	0.478	0.554	0.683	0.690	0.686	0.703	0.683	**0.693**	0.668	0.480	0.559
	Healthy	0.400	0.411	0.405	0.373	0.557	0.447	0.433	0.425	0.429	0.460	0.484	**0.472**	0.381	0.573	0.458
	OA		0.568			0.506			0.595			**0.612**			0.513	
	MCC		0.066			0.033			0.115			**0.165**			0.051	

*Here we used three decimal places because the metrics were close for various classifiers, so two decimal places did not allow to differentiate them.

5.4 Combined spectral and morphological traits

Figure 5.5a shows the ROC curves used to measure how well the trained models fitted the training data. It can be seen that the models had a similar behaviour, in particular, the AUC is greater than 0.8 for RF, SVC and LSVC methods. The highest AUC was 0.83 for RF, followed by SVC and LSVC with an AUC of 0.82. By far the method that required the longest processing time was SVC, the run time was approximately 1414.14 seconds, by comparison, KNN which was the second method at run time took approximately 127.10 seconds. RF and GBC had shorter execution times, with 46.77 seconds and 27.69 seconds respectively. The

most efficient method in terms of execution time was LSVC with a time of 0.09 seconds. The curve shape and the AUC also suggest that there were over-fitting issues for GBC.



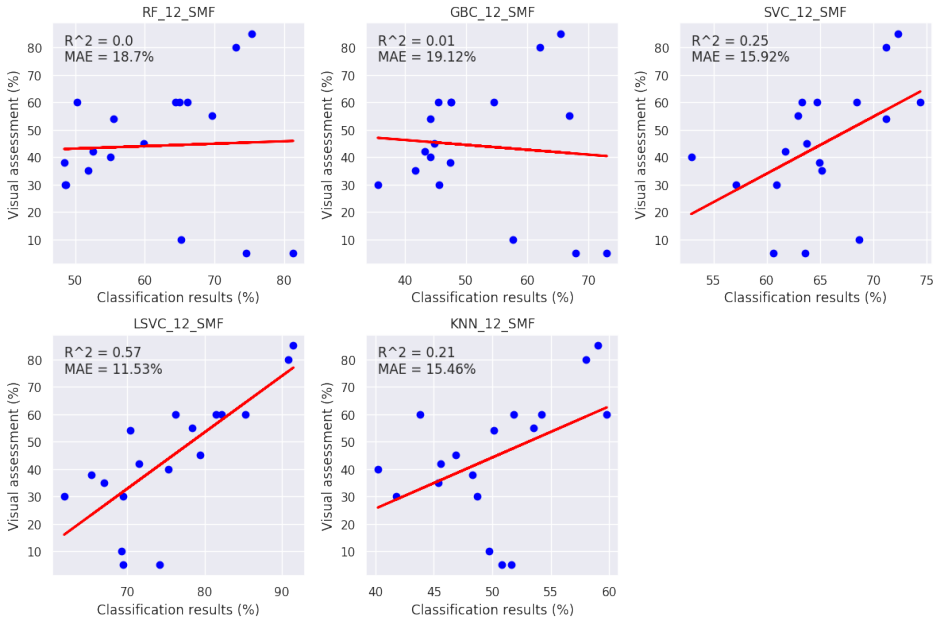
(a) (b)

Figure 5.5: Combined spectral and morphological traits - Performance by classification method on the training dataset. (a) ROC curves for RF, GBC, SVC, LSVC and KNN; (b) Execution time for each classifier.

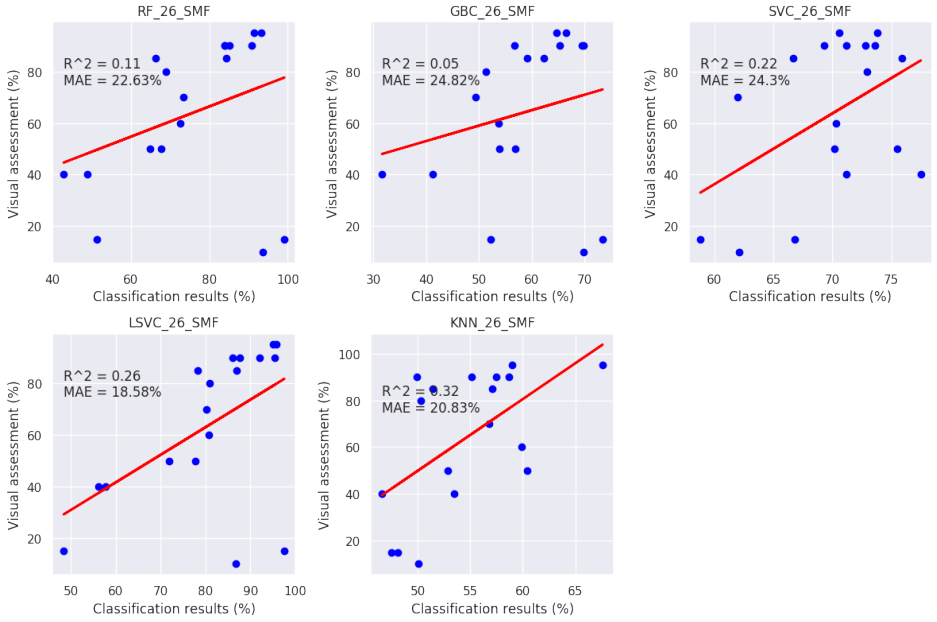
Table 5.6 shows mean absolute error, root mean squared error and coefficient of determination obtained from assessment of image classifications versus visual assessment when using combined spectral and morphological traits. It should be noted that even though the coefficients R^2 are lower than those obtained when using only spectral traits, the MAE and Root Mean Squared Error (RMSE) are significantly better in this approach. Furthermore, The difference between the dataset A and the dataset B decrease to close values for RF, GBC and LSVC.

Table 5.7 shows thematic accuracy metrics when using combined spectral and morphological traits: the classification method that achieves the best results, compared to the reference classifications, for the two datasets was SVC with OA = 0.885 and MCC=0.785 for the dataset A and OA = 0.817 and MCC = 0.598. Precision and recall did not show significant evidence of difference in the capability of the ML algorithms to classify healthy plants than diseased plants for the two datasets.

Figure 5.6 shows the correlation between infection level from classification results and visual assessment methods for each ML algorithm under study when using combined spectral and morphological traits. Figure 5.6a shows the correlation between infection level from classification results and visual assessment for the Dataset A, it is possible to see that LSVC showed the best performance for the Dataset A with an $R^2 = 0.57$ and a $MAE = 11.53\%$. Figure 5.6b shows the correlation between infection level from classification results and visual assessment for the Dataset B, it is possible to see that KNN showed the best performance for the Dataset B with an $R^2 = 0.32$ and a $MAE = 20.83\%$.



(a) Dataset A - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.



(b) Dataset B - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.

Figure 5.6: Combined spectral and morphological traits - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study. (a) Correlation between infection level from image processing and visual assessment for the Dataset A; and (b) Correlation between infection level from classification results and visual assessment for the Dataset B. The solid red line shows the regression line. LSVC showed the best performance for the Dataset A with a $R^2 = 0.57$ and a $MAE = 11.53\%$. KNN showed the best performance for the Dataset B with a $R^2 = 0.32$ and a $MAE = 20.83\%$.

Table 5.6: Combined spectral and morphological traits - Mean absolute error, root mean squared error and coefficient of determination obtained from assessment of image classifications versus visual assessment.

ML algorithm		Dataset	
		A	**B**
RF	MAE	5.89	7.08
	RMSE	7.89	9.39
	R^2	0.74	0.79
GBC	MAE	9.16	9.95
	RMSE	10.52	11.66
	R^2	0.50	0.67
SVC	MAE	5.54	18.18
	RMSE	7.79	19.70
	R^2	0.75	0.02
LSVC	MAE	5.09	6.39
	RMSE	6.48	8.58
	R^2	0.81	0.82
KNN	MAE	5.62	17.30
	RMSE	7.57	18.92
	R^2	0.75	0.11

Table 5.7: Combined spectral and morphological traits - F-Scores per class, overall accuracy and MCC obtained from assessment of image classifications versus reference classification. Bold values correspond to the highest OA, F-Scores and MCC.

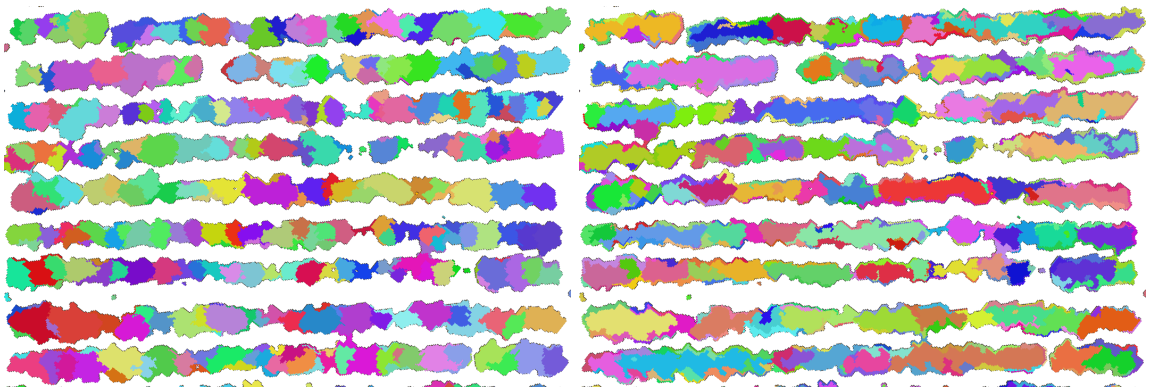
Dataset	**Class**	**RF**			**GBC**			**SVC**			**LSVC**			**KNN**		
	category	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
Dataset A	LB	0.805	0.907	0.853	0.496	0.475	0.485	0.819	0.981	**0.893**	0.813	0.983	0.890	0.733	0.826	0.777
	Healthy	0.899	0.790	0.841	0.516	0.536	0.526	0.978	0.792	**0.875**	0.980	0.784	0.871	0.811	0.712	0.758
	OA		0.847			0.507			**0.885**			0.881			0.768	
	MCC		0.701			0.012			**0.785**			0.780			0.541	
Dataset B	LB	0.836	0.779	0.807	0.663	0.458	0.542	0.849	0.870	**0.859**	0.871	0.799	0.833	0.835	0.670	0.773
	Healthy	0.647	0.727	0.685	0.375	0.583	0.456	0.756	0.722	**0.738**	0.686	0.788	0.734	0.563	0.763	0.648
	OA		0.760			0.503			**0.817**			0.795			0.703	
	MCC		0.495			0.040			**0.598**			0.572			0.415	

*Here we used three decimal places because the metrics were close for various classifiers, so two decimal places did not allow to differentiate them.

5.5 General discussion

One of the most difficult tasks was to separate each plant within the rows of the crop, the initial objective was to find a method that would allow, in the most approximate way possible, the identification and separation of each plant. Initially, the Quickshift segmentation (Figure 5.7a) method was used, whose segmentation results seem to fit well to the contour of the plants. However, the results in terms of accuracy of the classification using this type of segmentation were not satisfactory. Contrary to expectations, the segmentation that allowed the best results was Felzenswalb (Figure 5.7b). Although this method produces regions that do not necessarily

correspond to individual plants, these regions seem to better represent the changes between healthy vegetation and diseased vegetation.



(a) (b)

Figure 5.7: Plant segmentation method comparison. (a) Quickshift; and (b) Felzenswalb.

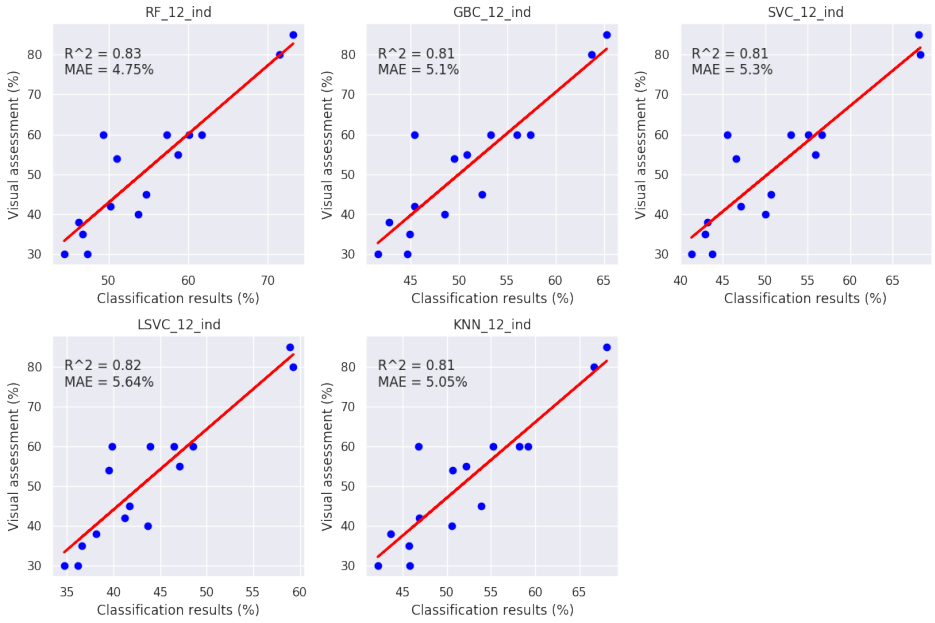
5.5.1 Influence of the control blocks in the results

The potato experimental crop consisted of an array of blocks with different nutrient treatments. Three out of eighteen blocks were established as control blocks. Those blocks of the experimental plot were in a state of development behind the other blocks, their spectral response in general was also below the other blocks even though the severity of the disease in those blocks was low. In general, in the 3 approaches of this study, these blocks strongly influenced the accuracy results, particularly in the comparison of classification results versus visual assessment.

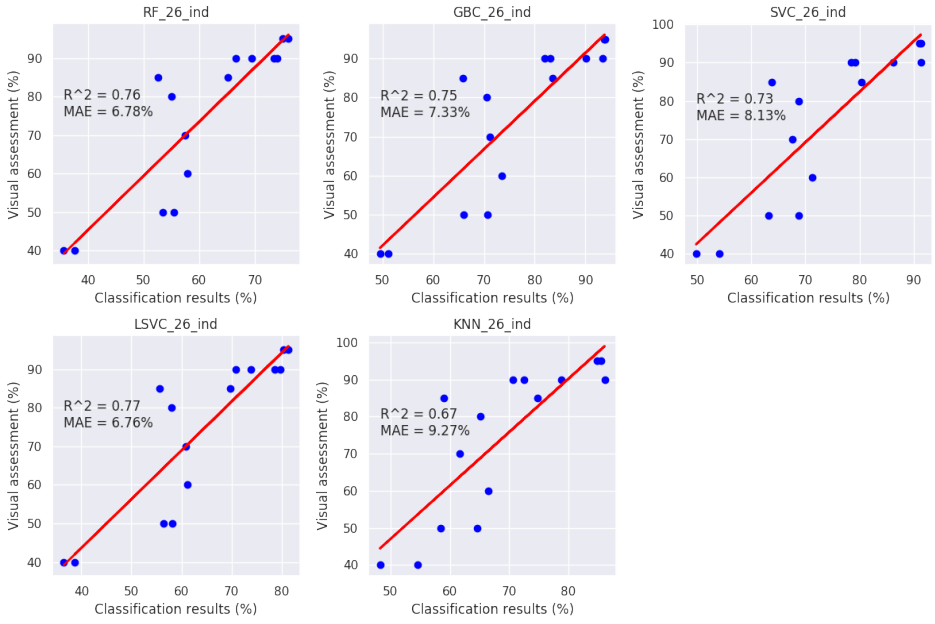
When comparing the results of classification versus visual evaluation without taking into account the control blocks, it is possible to see that in all 3 cases the relationship improves. Even though the relationships improve, there is still low performance when using only morphological traits.

Figure 5.8 shows the correlation between infection level from classification results and visual assessment methods for each ML algorithm under study when using spectral traits without taking into account the control blocks. Figure 5.8a shows the correlation between infection level from classification results and visual assessment for the Dataset A, it is possible to see that RF showed the best performance for the Dataset A with an $R^2 = 0.83$ and a $MAE = 4.75\%$. Figure 5.8b shows the correlation between infection level from classification results and visual assessment for the Dataset B, it is possible to see that LSVC showed the best performance for the Dataset B with an $R^2 = 0.77$ and a $MAE = 6.76\%$.

Figure 5.9 shows the correlation between infection level from classification results and visual assessment methods for each ML algorithm under study when using morphological traits without taking into account the control blocks. Figure 5.9a shows the correlation between infection level from classification results and visual assessment for the Dataset A, it is possible to see that GBC showed the best performance for the Dataset A with an $R^2 = 0.7$ and a



(a) Dataset A - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.



(b) Dataset B - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.

Figure 5.8: Spectral traits, no control blocks - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study excluding control blocks. (a) Correlation between infection level from image processing and visual assessment for the Dataset A; and (b) Correlation between infection level from classification results and visual assessment for the Dataset B. The solid red line shows the regression line. RF showed the best performance for the Dataset A with a $R^2 = 0.83$ and a $MAE = 4.75\%$. LSVC showed the best performance for the Dataset B with a $R^2 = 0.77$ and a $MAE = 6.76\%$.

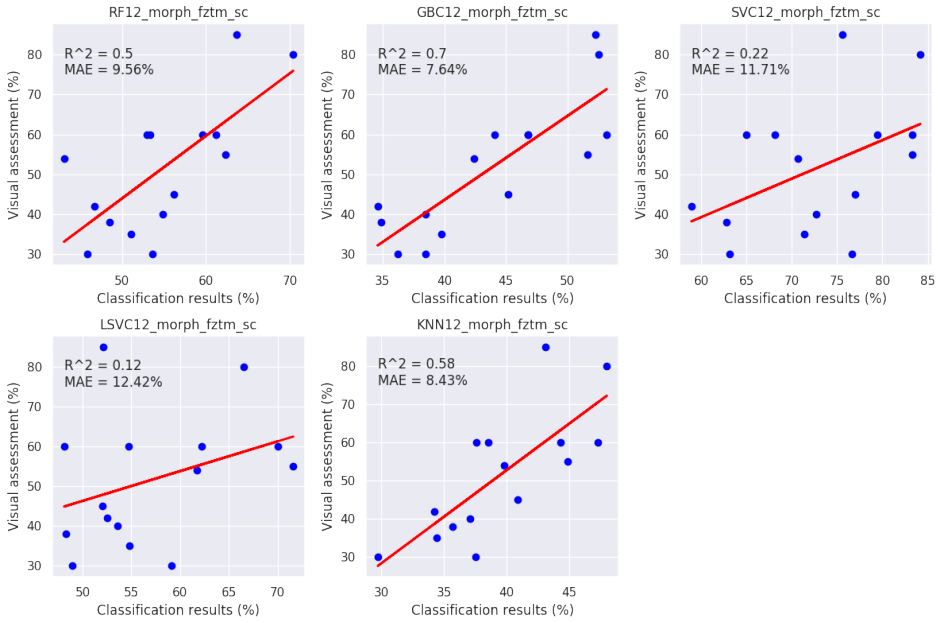
$MAE = 7.64\%$. Figure 5.9b shows the correlation between infection level from classification results and visual assessment for the Dataset B, it is possible to see that KNN showed the best performance for the Dataset B with an $R^2 = 0.48$ and a $MAE = 10.31\%$.

Figure 5.10 shows the correlation between infection level from classification results and visual assessment methods for each ML algorithm under study when using combined spectral and morphological traits without taking into account the control blocks. Figure 5.10a shows the correlation between infection level from classification results and visual assessment for the Dataset A, it is possible to see that LSVC showed the best performance for the Dataset A with an $R^2 = 0.81$ and a $MAE = 5.09\%$. Figure 5.10b shows the correlation between infection level from classification results and visual assessment for the Dataset B, it is possible to see that LSVC showed the best performance for the Dataset B with an $R^2 = 0.82$ and a $MAE = 6.39\%$.

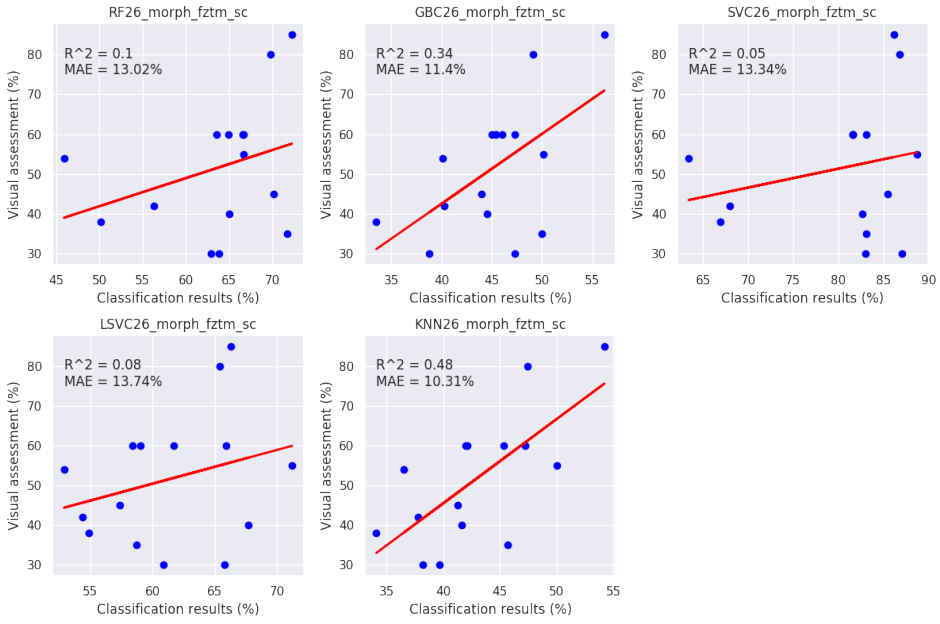
Two out of three of similar studies summarised in Table 5.8 did not implemented ML algorithms for assessing of potato diseases (Sugiura et al., 2016; Gibson-Poole et al., 2017). In the remaining study Duarte-Carvajalino et al. (2018) carried out an evaluation of the LB in 14 different potato genotypes using multispectral images captured with an UAV and ML methods such as multi-layer perceptron, deep learning CNN, support vector regression, and RF. The method proposed by Duarte-Carvajalino et al. (2018) requires that the user performs the manual extraction of each subset used for model training, which is time consuming and may introduce bias in the process. The method proposed here requires less time since the longest task is the creation of training polygons and once the training models are created, it is possible to reuse them to perform the classification of new datasets. With an OA of 87.8%, MCC equal to 0.768 and F-scores for LB and HP of 88.4% and 86.3% respectively, the proposed method can be considered satisfactory for the evaluation of LB in potato crops when the trained models are fitted using training data from the same date. Although more studies are required to ensure that models are transferable to a later date, the results obtained, OA of 79.9%, MCC equal to 0.556 and F-scores for LB and HP of 86.0% and 64.0% respectively, seems to be promising.

Table 5.8: Comparative table of similar studies on potato with respect to the proposed method.

Reference	**Disease**	**Machine learning/ Deep Learning algorithm used**	**Issues**	**Performace** (Best algorithm)
Sugiura et al. (2016)	LB	Machine learning algorithms were not used.	Sparse acquisition of images caused a slight decline in correlation between the AUDPCs from the image processing and visual assessment	RMSE =14.7 % $R^2 = 0.77$
Gibson-Poole et al. (2017)	Blackleg	Machine learning algorithms were not used.	High covariance when processing the digital surface model due to lack of GCP.	User accuracy = 65% Producer accuracy = 73% OA =87% Kappa = 0.61
Duarte-Carvajalino et al. (2018)	LB	Multilayer perceptron, Support vector regression, CNN and RF.	It is required that the user manually cut images from each plot to train the algorithms	MAE = 11.72%
Proposed method	LB	RF, GBC, SVC, LSVC and KNN	Find an appropriate adjustment value for threshold to allow accurate separation between the background and potato plants when weed presence increase inside the crop.	OA = 89.7 % MCC = 79.5 % F-score LB = 0.897 F-score HP = 0.897 $R^2 = 0.83$ $MAE = 4.75\%$

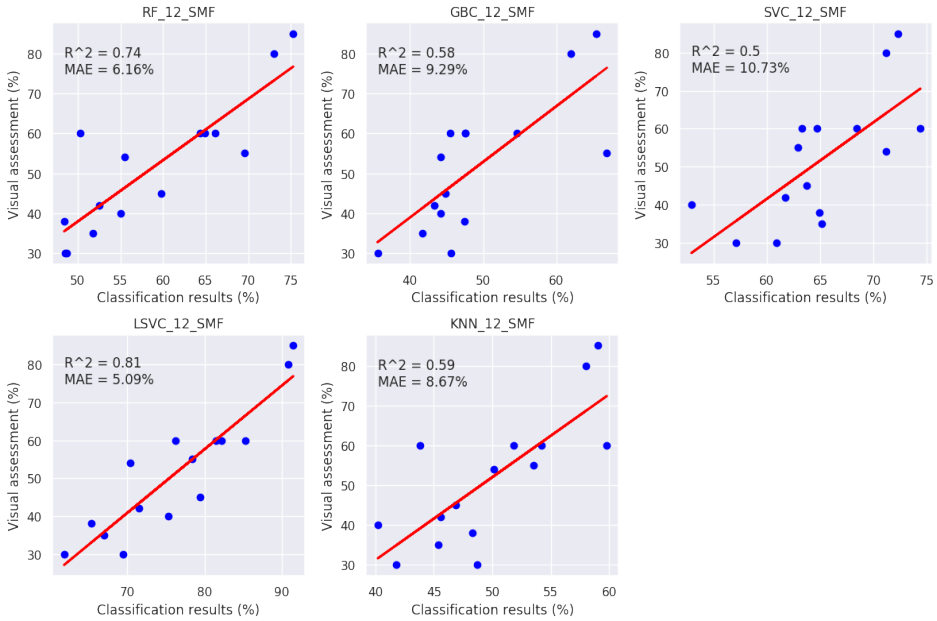


(a) Dataset A - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.

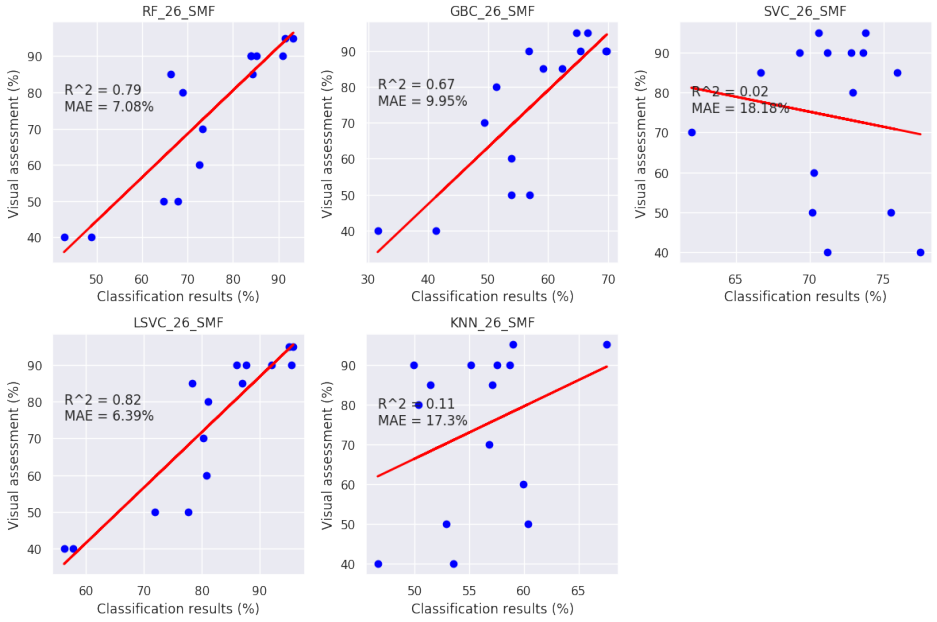


(b) Dataset B - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.

Figure 5.9: Morphological traits, no control blocks - Correlation between infection level from classification results using only morphological traits and visual assessment methods for each ML algorithm under study excluding control blocks. (a) Correlation between infection level from image processing and visual assessment for the Dataset A; and (b) Correlation between infection level from classification results and visual assessment for the Dataset B. The solid red line shows the regression line. GBC showed the best performance for the dataset A with a $R^2 = 0.7$ and MAE = 7.64%.



(a) Dataset A - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.



(b) Dataset B - Correlation between infection level from classification results and visual assessment methods for each ML algorithm under study.

Figure 5.10: Combined spectral and morphological traits, no control blocks - Correlation between infection level from classification results using only morphological traits and visual assessment methods for each ML algorithm under study excluding control blocks. (a) Correlation between infection level from image processing and visual assessment for the Dataset A; and (b) Correlation between infection level from classification results and visual assessment for the Dataset B. The solid red line shows the regression line. LSVC showed the best performance for datasets A and B with a $R^2 = 0.81$ and MAE = 5.09% and $R^2 = 0.82$ and MAE = 6.39% respectively.

An important limitation of this study was the lack of data for low levels of infection beyond the control blocks. As mentioned above the first acquisition of multispectral images was carried on several days after first symptoms were identified in the experimental crop and the disease rapidly spread over the field, so, at the time of the first field data capture was done, most of the experimental blocks had reached infection level LB2 and above and LB1 was present only in the control blocks. Even though there was a lack of data for LB1, blocks 5, 8, 7, and 15 in dataset A were in an LB2 stage and predicted values by ML algorithms RF, GBC, and KNN were similar to the visual scoring, although slightly higher. This suggests a good capability to differentiate LB potato plants from healthy potato plants when the infection level reached LB2 or above.

5.5.2 Transferability of the ML models

Table 5.9 shows the generalisation loss calculated for each approach and ML method evaluated in this study. When using only morphological traits the performance of the ML methods is low, so it is possible to see negative G-loss percentages caused by higher accuracy values in the performance of the dataset B, which can be misleading. However, it can be seen that, when using only spectral traits, G-loss is higher (8.9%) compared to the G-loss when using combined spectral and morphological traits (7.4%). This suggests that there is a slight improvement in the performance of the classifiers when including combined traits.

Table 5.9: G-loss (%) calculated for each approach and ML method studied.

Approach	**ML method**					**Mean**
	RF	GBC	SVC	LSVC	KNN	
Spectral traits	7.6	8.7	9.7	10.2	8.4	8.9
Morphological traits	-11.4	6.5	-7.2	10.1	6.6	0.9
Combined traits	10.3	0.8	7.7	9.8	8.5	7.4

5.6 Accomplishment of the objectives

The aim of this project was to respond to the research question: *What are the spectral and morphological features that allow estimating the severity level of late blight (Phytophthora infestans) in potato crops from multispectral images acquired using an UAV platform?*

In chapter 4, we presented the variable importance measures for spectral traits and morphological traits. According to those results:

5.6.1 Spectral traits that allow estimating the severity level of LB

The experimental results and the selection of variables carried out with the data obtained in this work allowed us to find the spectral features that are most important in the evaluation of LB in potato crops. Finding the most important features can make information processing more efficient since the number of variables to be included in demanding processes such as

those in which machine learning methods are applied is reduced. In addition, the vegetation indices are easy and quick to calculate. In this work, the spectral features found to be most important were:

1. SAVI.
2. EVI2.
3. LAI.

Note: responds to specific objective 1.

5.6.2 Morphological traits that allow estimating the severity level of LB

The most important morphological features found in this work, according to the experimental results were:

1. Area.
2. Equivalent diameter.
3. Minor axis length.

We found that the area is an important factor in our data since those affected areas present defoliation, this makes that when removing the soil in the images, these areas look fragmented. That is, the superpixels calculated in the zones with disease tend to be smaller than in the zones of healthy plants as it can be seen in Figure 3.13 In the case of the equivalent diameter, this can be related to the area, areas that have a smaller equivalent diameter could be related to the presence of the disease. In the case of the length of the minor axis, we found that those over-segmented areas corresponding to areas with the presence of the disease were characterized by a very irregular geometry (very elongated areas), which means that there is a significant difference in the length of the axes of the approximate ellipse to these regions.

Note: responds to specific objective 2.

5.6.3 Combined spectral and morphological traits

- Slight improvement in the accuracy of the classification results vs ground truth and visual scoring. The experimental results showed that the evaluation of the accuracy of the classification results, when combined features were used, improved in particular in dataset B.
- Significant improvement in the transferability of the ML models. The experimental results showed that the G-loss metric improved from 8.9% when only spectral features were used to 7.4% when spectral and morphological features were used combined.

Note: responds to specific objective 3.

5.7 Advantages and limitations of the methodology

5.7.1 Advantages

1. Allows repetitive coverage of an agricultural field. Besides the field inspection of the potato crop, in this work, labours related to the acquisition of the multispectral images showed to be efficient as the time necessary to set up the UAV for flight was as short as 15 minutes and the use of rechargeable batteries made easy to repeat the missions in the event of an unforeseen event.
2. Allows for easy collection of data. In this work, using a UAV to acquire multispectral images showed to be an efficient method of data acquisition as each flight mission covered 3.2 Ha in approximately eight minutes, and with a spatial resolution of approximately 3.2 cm pixel size proving to have huge potential in crop scouting tasks.
3. Improve transferability of the ML models. G-loss metric showed that there was an improvement in the classification results when we used combined spectral and morphological traits.

5.7.2 Disadvantages

- Large amount of data to be processed. The multispectral images acquired with UAV for each flight of 1 hectare to 40 meters above the ground can have an approximate weight of 5 Gb with the sensor used in this study and the processed files (i.e. DSM, orthomosaic) can have an approximate weight of 3Gb.
- Requires high-performance computing. The raw images acquired in the field must go through a photogrammetric adjustment process that allows the derivation of products such as digital surface models or orthomosaics. Furthermore, they must be post-processed using machine learning methods. These processes require high computing capacity and the use of graphic accelerators that not all computers have.
- Need to finding an appropriate adjustment value for the threshold to allow accurate separation between the background and the potato plants when weed presence increase inside the crop.

5.8 Contributions to knowledge in Geomatics

This thesis contributes to the knowledge in Geomatics in the following aspects:

- It integrates concepts and techniques used in computer vision in the processing of multispectral images acquired through remote sensors that can serve as a basement for other investigations.
- It implements and evaluates a computational workflow for the processing and analysis of optical data acquired from UAV platforms to assess the severity of LB in potato crops, published in a public repository and available to be used in other projects.

- It evaluates the level of complementary between multispectral traits, widely used in remote sensing studies, and morphological traits, widely used in computer vision studies, to determine the presence of a disease affecting an important crop in Colombia.

6 Conclusions

In this study, experimental results showed that the ML algorithms which gave the best results, according two the accuracy metrics were SVC and RF. For LSVC algorithm, overall accuracy and Matthews correlation coefficient for the dataset A were 0.897 and 0.795 respectively and for the dataset B Overall accuracy and Matthews correlation coefficient were 0.799 and 0.556 respectively and for RF algorithm, overall accuracy and Matthews correlation coefficient for the dataset A were 0.890 and 0.784 respectively and for the dataset B Overall accuracy and Matthews correlation coefficient were 0.786 and 0.520 respectively. Although the results in terms of accuracy metrics for dataset B are not as good as those for dataset A, the results are promising when considering that the models used for the classification of this dataset were pre-trained with data acquired in a different date under dissimilar lighting and weather conditions.

When using spectral traits only the bands which allow better accuracy results were SAVI, EVI2 and LAI. When using morphological traits only the bands which allow better accuracy results were area, equivalent diameter and minor axis lenght. However, when using combined spectral and morphological traits, accuracy results improved when all the bands were used in he classification stage. This suggest that there is still work to do to understand the contribution of the traits to the correct identification of diseased plants when working with combined morphological and spectral traits.

Combination of spectral and morphological traits demonstrate to achieve the better the lower MAE for both dataset A and dataset B, in contrast, when using only spectral traits the results are good only for the dataset A. This suggest that morphological traits have the potential to improve transferability of the ML models trained, however, further experiments should be carried on.

Main limitation of the proposed approach relates to the need of finding an appropriate adjustment value for the threshold which allows accurate separation between the background and potato plants when weed presence increase inside the crop. In this study, there was no need to adjust the threshold in the dataset A, where the presence of weeds was marginal as the Otsu's threshold algorithm was able to find the appropriate value for separation of plants from bare soil. However, as weed presence increased in the dataset B, it was necessary to use Otsu's algorithm with and adjusted threshold to allow accurate separation. Although adjusting the threshold value helped to remove most of the weeds present in the orthophoto for the two datasets, there was still the presence of weeds in the segmented result. However, weeds remnants presence were predominant between the experimental blocks where cultural management such as hill-up was not carried on, so, for this study those remnants were considered negligible, even though, additional experiments need to be conducted to better understand the influence of weed presence into the quality of background removal from images and accuracy of the infection level.

Bibliography

An, N., Palmer, C. M., Baker, R. L., Markelz, R. J., Ta, J., Covington, M. F., Maloof, J. N., Welch, S. M., and Weinig, C. (2016). Plant high-throughput phenotyping using photogrammetry and imaging techniques to measure leaf length and rosette area. *Computers and Electronics in Agriculture*, 127:376–394.

Araus, J. L. and Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, 19(1):52–61.

Arora, R. K., Sharma, S., and Singh, B. P. (2014). Late blight disease of potato and its management. *Potato Journal*, 41(1):16–40.

Bagheri, N. (2020). Application of aerial remote sensing technology for detection of fire blight infected pear trees. *Computers and Electronics in Agriculture*, 168(September 2019):105147.

Batchelor, B. G., editor (2012). *Machine Vision Handbook*. Springer London, London.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M. L., and Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, 39:79–87.

Bock, C. H., Poole, G. H., Parker, P. E., and Gottwald, T. R. (2010). Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Critical Reviews in Plant Sciences*, 29(2):59–107.

Bolger, M., Schwacke, R., Gundlach, H., Schmutzer, T., Chen, J., Arend, D., Oppermann, M., Weise, S., Lange, M., Fiorani, F., Spannagl, M., Scholz, U., Mayer, K., and Usadel, B. (2017). From plant genomes to phenotypes. *Journal of Biotechnology*.

Breiman, L. (2001). Random Forest. *Machine Learning*, 45:5–32.

Camilus, K. S. and Govindan, V. K. (2012). A Review on Graph Based Segmentation. *I.J. Image, Graphics and Signal Processing*, (June):1–13.

Chang, C. C. and Lin, C. J. (2011). LIBSVM: A Library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–39.

Conci, A., Rodrigues, O., and Liatsis, P. (2018). Morphological classifiers. *Pattern Recognition*, 84:82–96.

Congedo, L. (2021). Semi-Automatic Classification Plugin.

Coppens, F., Wuyts, N., Inzé, D., and Dhondt, S. (2017). Unlocking the potential of plant phenotyping data through integration and data-driven approaches. *Current Opinion in Systems Biology*, 4:58–63.

Drass, M., Berthold, H., Müller-Braun, S., Seel, M., König, M., Hof, P., Schneider, J., and Oechsner, M. (2020). Semantic Segmentation with Deep Learning: Detection of Cracks at the Cut Edge of Glass (under review). *Glass Structures & Engineering*.

Duarte-Carvajalino, J. M., Alzate, D. F., Ramirez, A. A., Santa-Sepulveda, J. D., Fajardo-Rojas, A. E., and Soto-Suárez, M. (2018). Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. *Remote Sensing*, 10(10).

Eisenbeiß, H., Zurich, E. T. H., Eisenbeiß, H., and Zürich, E. T. H. (2009). *UAV photogrammetry*. Number 18515.

EPPO (2007). General crop inspection procedure for potatoes. Technical report, European and Mediterranean Plant Protection Organization.

Fan, R. E., Chang, K. W., Hsieh, C. J., Wang, X. R., and Lin, C. J. (2008). LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9(2008):1871–1874.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.

Felzenszwalb, P. F. and Huttenlocher, D. P. (2004). Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision*, 59(2):167–181.

Forbes, G., Perez, W., and Andrade-Piedra, J. (2014). *Field assessment of resistance in potato to Phytophthora infestans, International Cooperators Guide*.

Franceschini, M. H. D., Bartholomeus, H., van Apeldoorn, D. F., Suomalainen, J., and Kooistra, L. (2019). Feasibility of unmanned aerial vehicle optical imagery for early detection and severity assessment of late blight in Potato. *Remote Sensing*, 11(3).

Franke, J. and Menz, G. (2007). Multi-temporal wheat disease detection by multi-spectral remote sensing. *Precision Agriculture*, 8(3):161–172.

Gago, J., Douthe, C., Coopman, R. E., Gallego, P. P., Ribas-Carbo, M., Flexas, J., Escalona, J., and Medrano, H. (2015). UAVs challenge to assess water stress for sustainable agriculture. *Agricultural Water Management*, 153:9–19.

Genuer, R., Poggi, J. M., and Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14):2225–2236.

Ghanem, M. E., Marrou, H., and Sinclair, T. R. (2015). Physiological phenotyping of plants for crop improvement. *Trends in Plant Science*, 20(3):139–144.

Gibson-Poole, S., Humphris, S., Toth, I., and Hamilton, A. (2017). Identification of the onset of disease within a potato crop using a UAV equipped with un-modified and modified commercial off-the-shelf digital cameras. *Advances in Animal Biosciences*, 8(2):812–816.

Gitelson, A. and Merzlyak, M. N. (1994). Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of Plant Physiology*, 143(3):286–292.

Gitelson, A. A., Gritz †, Y., and Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160(3):271–282.

Glasbey, C. A. (1993). An Analysis of Histogram-Based Thresholding Algorithms.

Granier, C. and Vile, D. (2014). Phenotyping and beyond: Modelling the relationships between traits. *Current Opinion in Plant Biology*, 18(1):96–102.

Hamuda, E., Glavin, M., and Jones, E. (2016a). A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture*, 125:184–199.

Hamuda, E., Glavin, M., and Jones, E. (2016b). A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture*, 125:184–199.

Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., Xiao, Y., and He, Z. (2018). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*, (October 2017):1–9.

Hsu, C. W. and Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2):415–425.

Hu, P., Chapman, S. C., Wang, X., Potgieter, A., Duan, T., Jordan, D., Guo, Y., and Zheng, B. (2018). Estimation of plant height using a high throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: Example for sorghum breeding. *European Journal of Agronomy*, 95(November 2017):24–32.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., and Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, (83):195–213.

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3):295–309.

Hussain, T., Singh, B. P., and Anwar, F. (2014). A quantitative real time PCR based method for the detection of Phytophthora infestans causing late blight of potato, in infested soil. *Saudi Journal of Biological Sciences*, 21(4):380–386.

Jiang, Y., Zebarth, B. J., Somers, G. H., Macleod, J. A., and Savard, M. M. (2012). *Sustainable potato production: Global case studies.*

Jiang, Z., Huete, A. R., Didan, K., and Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10):3833–3845.

John Reader (2009). *POTATO. A history of the propitious esculent.*

Jones, H. G. and Vaughan, R. A. (2010). *Remote sensing of vegetation - principles, techniques, and applications*. University Press Oxford, United Kingdom, 1 edition.

Kamberova, G. and Bajcsy, R. (2006). Sensor Errors and the Uncertainties in Stereo Reconstruction. Technical report, University of Pennsylvania.

Kaur, D. and Kaur, Y. (2014). Various Image Segmentation Techniques: A Review. *International Journal of Computer Science and Mobile Computing*, 3(5):809–814.

Klemenjak, C., Faustine, A., Makonin, S., and Elmenreich, W. (2019). On metrics to assess the transferability of machine learning models in non-intrusive load monitoring. *arXiv*.

Kuhl, J. C., Zarka, K., Coombs, J., Kirk, W. W., and Douches, D. S. (2007). Late blight resistance of RB Transgenic potato lines. *Journal of the American Society for Horticultural Science*, 132(6):783–789.

Kumar, J., Kumar, S., and Pratap, A. (2015). *Phenomics in crop plants: Trends, options and limitations*.

Küng, O. ., Strecha, C., Beyeler, A., Zufferey, J.-C., Floreano, D., Fua, P., and Gervaix, F. (2011). The Accuracy of Automatic Photogrammetric Techniques on Ultra-Light UAV Imagery. *Paper presented at: UAV-g 2011 - Unmanned Aerial Vehicle in Geomatics*.

Lawrence, R., Bunn, A., Powell, S., and Zambon, M. (2004). Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment*, 90(3):331–336.

Lee, Y. H., Wei, C. P., Cheng, T. H., and Yang, C. T. (2012). Nearest-neighbor-based approach to time-series classification. *Decision Support Systems*, 53(1):207–217.

Li, B., Xu, X., Han, J., Zhang, L., Bian, C., Jin, L., and Liu, J. (2019). The estimation of crop emergence in potatoes by UAV RGB imagery. *Plant Methods*, 15(1):1–13.

Liu, J., Pattey, E., and Jégo, G. (2012). Assessment of vegetation indices for regional crop green LAI estimation from Landsat images over multiple growing seasons. *Remote Sensing of Environment*, 123:347–358.

Liu, S., Li, L., Gao, W., Zhang, Y., Liu, Y., Wang, S., and Lu, J. (2018a). Diagnosis of nitrogen status in winter oilseed rape (Brassica napus L.) using in-situ hyperspectral data and unmanned aerial vehicle (UAV) multispectral images. *Computers and Electronics in Agriculture*, 151(June):185–195.

Liu, T., Li, R., Zhong, X., Jiang, M., Jin, X., Zhou, P., Liu, S., Sun, C., and Guo, W. (2018b). Estimates of rice lodging using indices derived from UAV visible and thermal infrared images. *Agricultural and Forest Meteorology*, 252(January):144–154.

Mahlein, A.-K. (2016a). Present and Future Trends in Plant Disease Detection. *Plant Disease*, 100(2):1–11.

Mahlein, A.-K. (2016b). Present and Future Trends in Plant Disease Detection. *Plant Disease*, 100(2):1–11.

Mamaghani, B. and Salvaggio, C. (2019). Multispectral sensor calibration and characterization for sUAS remote sensing. *Sensors (Switzerland)*, 19(20).

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information.*

Micasense (2018). Rededge Integration Guide.

Minervini, M., Fischbach, A., Scharr, H., and Tsaftaris, S. A. (2016). Finely-grained annotated datasets for image-based plant phenotyping. *Pattern Recognition Letters*, 81:80–89.

Mishra, Y., Johansson Jänkänpää, H., Kiss, A. Z., Funk, C., Schröder, W. P., and Jansson, S. (2012). Arabidopsis plants grown in the field and climate chambers significantly differ in leaf morphology and photosystem components. *BMC Plant Biology*, 12.

Moore, A. W. (1991). An introductory tutorial on kd-trees, Robotics Institute. *Efficient Memory-based Learning for Robot Control*, (209):1 – 20.

Namugga, P., Sibiya, J., Melis, R., and Barekye, A. (2017). Phenotypic characterisation of potato (Solanum tuberosum) genotypes in Uganda. *South African Journal of Plant and Soil*, 1862:1–8.

OECD (1997). Organization for Economic Co-operation and Development. Consensus document on the biology of solanum tubersoum subsp. Tuberosum (potato). *Http://Www.Oecd.Org/Science/Biotrack/46815598.Pdf*, (8):1–38.

Omar, L. and Ivrissimtzis, I. (2019). Using theoretical ROC curves for analysing machine learning binary classifiers. *Pattern Recognition Letters*, 128:447–451.

Omohundro, S. M. (1989). Five balltree construction algorithms. *Science*, 51(1):1–22.

Ose, K., Corpetti, T., and Demagistri, L. (2016). Multispectral Satellite Image Processing. *Optical Remote Sensing of Land Surface: Techniques and Methods*, pages 58–124.

Otsu, N. (1979). A Tlreshold Selection Method from Gray-Level Histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66.

Pacilly, F. C., Groot, J. C., Hofstede, G. J., Schaap, B. F., and van Bueren, E. T. (2016). Analysing potato late blight control as a social-ecological system using fuzzy cognitive mapping. *Agronomy for Sustainable Development*, 36(2).

Pepe, M., Fregonese, L., and Scaioni, M. (2018). Planning airborne photogrammetry and remote-sensing missions with modern platforms and sensors. *European Journal of Remote Sensing*, 51(1):412–435.

Pradal, C., Artzet, S., Chopard, J., Dupuis, D., Fournier, C., Mielewczik, M., Nègre, V., Neveu, P., Parigot, D., Valduriez, P., and Cohen-Boulakia, S. (2017). InfraPhenoGrid: A scientific workflow infrastructure for plant phenomics on the Grid. *Future Generation Computer Systems*, 67:341–353.

Puissant, A., Rougiera, S., and Stumpf, A. (2014). Object-oriented mapping of urban trees using random forestclassifiers. *International Journal of Applied Earth Observation and Geoinformation*, 26(1):235–245.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., and Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67(1):93–104.

Rouse J.˜W., J., Haas, R. H., Schell, J. A., and Deering, D. W. (1974). Monitoring Vegetation Systems in the Great Plains with Erts. In *NASA Special Publication*, volume 351, page 309.

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., Miklas, P. N., Carter, A. H., Pumphrey, M. O., Knowles, R. R., and Pavek, M. J. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review.

Sankaran, S., Zhou, J., Khot, L. R., Trapp, J. J., Mndolwa, E., and Miklas, P. N. (2018). High-throughput field phenotyping in dry bean using small unmanned aerial vehicle based multispectral imagery. *Computers and Electronics in Agriculture*, 151(June):84–92.

Schumann, G. L., D'Arcy, C. J., and Ristaino, J. (2000). Late blight of potato and tomato.

Shakoor, N., Lee, S., and Mockler, T. C. (2017). High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Current Opinion in Plant Biology*, 38:184–192.

Singh, A., Ganapathysubramanian, B., Singh, A. K., and Sarkar, S. (2016). Machine Learning for High-Throughput Stress Phenotyping in Plants.

Śliwka, J., Jakuczun, H., Chmielarz, M., Hara-Skrzypiec, A., Tomczyńska, I., Kilian, A., and Zimnoch-Guzowska, E. (2012). A resistance gene against potato late blight originating from Solanum x michoacanum maps to potato chromosome VII. *Theoretical and Applied Genetics*, 124(2):397406.

Speiser, J. L., Miller, M. E., Tooze, J., and Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134:93–101.

Srbinovska, M., Gavrovski, C., Dimcev, V., Krkoleva, A., and Borozan, V. (2015). Environmental parameters monitoring in precision agriculture using wireless sensor networks. *Journal of Cleaner Production*, 88:297–307.

Strobl, C., Boulesteix, A. L., Zeileis, A., and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8.

Struik, P. C. (2010). Can Physiology Help Us to Combat Late Blight in Potato? *Potato Research*, 53(4):277–287.

Sugiura, R., Tsuda, S., Tamiya, S., Itoh, A., Nishiwaki, K., Murakami, N., Shibuya, Y., Hirafuji, M., and Nuske, S. (2016). Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle. *Biosystems Engineering*, 148:1–10.

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., and Bennett, M. (2017). Plant Phenomics, From Sensors to Knowledge. *Current Biology*, 27(15):R770–R783.

Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*.

Tokekar, P., Hook, J. V., Mulla, D., and Isler, V. (2016). Sensor Planning for a Symbiotic UAV and UGV System for Precision Agriculture. *IEEE Transactions on Robotics*, 32(6):1498–1511.

Van Der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T. (2014). Scikit-image: Image processing in python. *PeerJ*, 2014(1):1–18.

Varoquaux, G., Buitinck, L., Louppe, G., Grisel, O., Pedregosa, F., and Mueller, A. (2015). Scikit-learn. *GetMobile: Mobile Computing and Communications*, 19(1):29–33.

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., Feldmann, K. A., French, A. N., Heun, J. T., Hunsaker, D. J., Jenks, M. A., Kimball, B. A., Roth, R. L., Strand, R. J., Thorp, K. R., Wall, G. W., and Wang, G. (2012). Field-based phenomics for plant genetics research. *Field Crops Research*, 133:101–112.

Wiik, L. (2014). Potato Late Blight and Tuber Yield: Results from 30 Years of Field Trials. *Potato Research*, 57(1):77–98.

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., and Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the American Society of Agricultural Engineers*, 38(1):259–269.

Xu, H., Zhang, H., He, W., and Zhang, L. (2019). Superpixel-based spatial-spectral dimension reduction for hyperspectral imagery classification. *Neurocomputing*, 360:138–150.

Zhang, C., Chen, W., and Sankaran, S. (2019). High-throughput field phenotyping of Ascochyta blight disease severity in chickpea. *Crop Protection*, 125:104885.

Zhang, F. and Yang, X. (2020). Improving land cover classification in an urbanized coastal area by random forests: The role of variable selection. *Remote Sensing of Environment*, 251(January):112105.

Zhao, C. J. and Jiang, G. Q. (2010). Baseline detection and matching to vision-based navigation of agricultural robot. *2010 International Conference on Wavelet Analysis and Pattern Recognition, ICWAPR 2010*, (July):44–48.

Zhao, H., Yang, C., Guo, W., Zhang, L., and Zhang, D. (2020). Automatic estimation of crop disease severity levels based on vegetation index normalization. *Remote Sensing*, 12(12):1–17.

Zhou, J., Pavek, M. J., Shelton, S. C., Holden, Z. J., and Sankaran, S. (2016). Aerial multispectral imaging for crop hail damage assessment in potato. *Computers and Electronics in Agriculture*, 127:406–412.