THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE

Préparée à l'UMR 1347 Agroécologie

AgroSup Dijon, CNRS, INRA, Univ. Bourgogne Franche-Comté

Ecole doctorale n°554

Environnements - Santé

Doctorat en Instrumentation et Informatique de l'Image

Par

Madame Hania AL SADDIK

"Spectral and Textural Analysis of High-Resolution Data for the Automatic Detection of Grapevine Diseases"

Thèse présentée et soutenue à AgroSup Dijon, le 04/07/2019.

Composition du Jury :

Monsieur, ROGER, Jean-Michel	Directeur de Recherche à l'IRSTEA	Rapporteur
Monsieur, MACAIRE, Ludovic	Professeur à l'Université de Lille	Président
Monsieur, DEBAYLE, Johan	Professeur à l'EMSE	Rapporteur
Monsieur, GERMAIN, Christian	Professeur à Bordeaux Science Agro	Examinateur
Madame, DUTHOIT, Sylvie	Responsable R&D à TerraNis	Examinatrice
Monsieur, BROUSSE, Olivier	Responsable R&D à GST	Examinateur
Monsieur, COINTAULT, Frédéric	MCF-HDR à AgroSup Dijon	Directeur de thèse
Monsieur, SIMON, Jean-Claude	MCF à AgroSup Dijon	Co-directeur de thèse
Madame, MAHE, Héloise	Responsable Technique au BIVB	Invitée
Madame, DUBOIS, Alice	Chargée de mission pour la DRAAF PACA	Invitée

ABSTRACT

'Flavescence dorée' is a contagious and incurable disease present on the vine leaves. In order to contain the infection, the regulations require growers to control each of the vine rows and to remove the suspect vine plants. This monitoring is done on foot during the harvest and mobilizes many people during a strategic period for viticulture.

In order to solve this problem, the DAMAV project (Automatic detection of Vine Diseases) aims to develop a solution for automated detection of vine diseases using a micro-drone. The goal is to offer a turnkey solution for wine growers. This tool will allow the search for potential foci, and then more generally any type of vine diseases detectable on the foliage. To enable this diagnosis, the foliage is proposed to be studied using a dedicated high-resolution multispectral camera.

The objective of this PhD-thesis in the context of DAMAV is to participate in the design and implementation of the Multi-Spectral (MS) image acquisition system and to develop the image pre-processing algorithms, based on the most relevant spectral and textural characteristics related to 'Flavescence dorée'.

Several grapevine varieties were considered such as red-berried and white-berried ones; furthermore, other diseases than 'Flavescence dorée' (FD) such as Esca and 'Bois noir' (BN) were also tested under real production conditions. The PhD work was basically performed at a leaf-level scale and involved an acquisition step followed by a data analysis step.

Most imaging techniques, used to detect diseases in field crops or vineyards, operate in the visible electromagnetic radiation range. It turns out that for disease detection, when the symptoms are already present, the visible may be sufficient, although the colorimetric information is not the only one to consider. In our case, it is advised to detect the disease as early as possible, the information of the visible spectrum does not seem sufficient and it is therefore necessary to investigate broader information. Reflectance responses of plant leaves can be obtained from short to long wavelengths with convenient sensors. These reflectance signatures describe the internal constituents of leaves. This means that the presence of a disease can modify the internal structure of the leaves and hence cause an alteration of its reflectance signature.

A spectro-radiometer is used in our study to characterize reflectance responses of leaves in the field. Several samples at different growth stages were used for the tests. To define optimal reflectance features for grapevine disease detection (FD, Esca, BN), a new methodology that designs Spectral Disease Indices (SDIs) has been developed. It is based on two dimension reduction techniques, coupled with a classifier. The first feature selection technique uses the Genetic Algorithms (GA) and the second one relies on the Successive Projection Algorithm (SPA). The new resulting SDIs outperformed traditional Spectral Vegetation Indices (SVIs) and GA performed, in general, better than SPA. The features finally chosen can then be implemented as filters in the MS sensor.

In general, the reflectance information was satisfying for finding infections (higher than 90% of accuracy for the best method) but wasn't enough. The images acquired with the developed MS device can further be pre-processed by low-level techniques based on the calculation of texture parameters. Several texture processing techniques have been tested but only on colored images. A method that combines many texture features is elaborated, allowing to choose the best ones. We found that the combination of optimal textural information could provide a complementary mean for not only differentiating healthy from infected grapevine leaves (higher than 85% of accuracy), but also for grading the disease severity stages (higher than 73% of accuracy) and for discriminating among diseases (higher than 72% of accuracy). This is in accordance with the hypothesis that a MS camera can enable detection and identification of diseases in grapevine fields.

The first experiments of the whole system "sensor-UAV" will be done during future acquisition campaigns in 2019.

KEYWORDS: grapevine diseases detection, multispectral sensor, vegetation indices, genetic algorithms, successive projections algorithm, texture analysis, feature selection, classification.

RESUME

La Flavescence dorée est une maladie contagieuse et incurable de la vigne détectable sur les feuilles. Pour contenir l'infection, la réglementation impose aux producteurs le contrôle des rangs de vigne pour éliminer les plants suspects. Ce suivi se fait à pied pendant la récolte, et mobilise de nombreuses personnes pendant une période très stratégique en viticulture.

Pour résoudre ce problème, le projet DAMAV (Détection Automatique des MAladies de la Vigne) a été mis en place, avec pour objectif de développer une solution de détection automatisée des maladies de la vigne à l'aide d'un micro-drone. Cet outil doit permettre la recherche des foyers potentiels de la Flavescence dorée, puis plus généralement de toute maladie détectable sur le feuillage à l'aide d'un outil multispectral dédié haute résolution.

Dans le cadre de ce projet, cette thèse a pour objectif de participer à la conception et à l'implémentation du système d'acquisition multispectral et de développer les algorithmes de prétraitement d'images basés sur les caractéristiques spectrales et texturales les plus pertinentes reliées à la Flavescence dorée.

Plusieurs variétés de vigne ont été considérées telles que des variétés rouges et blanches; de plus, d'autres maladies que 'Flavescence dorée' (FD) telles que Esca et 'Bois noir' (BN) ont également été testées dans des conditions de production réelles. Le travail de doctorat a été essentiellement réalisé au niveau feuille et a impliqué une étape d'acquisition suivie d'une étape d'analyse des données.

La plupart des techniques d'imagerie, utilisées pour détecter les maladies dans les grandes cultures ou les vignobles, opèrent dans le domaine du visible. Il s'avère que pour la détection de la maladie lorsque les symptômes sont déjà présents, le visible peut être suffisant, bien que les informations colorimétriques ne soient pas les seules à devoir être prises en compte. Dans DAMAV, il est conseillé que la maladie soit détectée le plus tôt possible, l'information du spectre électromagnétique visible semble donc insuffisante. Des informations spectrales plus larges sont nécessaires, notamment dans l'infrarouge. Les réflectances des feuilles des plantes peuvent être obtenues sur les longueurs d'onde les plus courtes aux plus longues avec des capteurs convenables. Ces réflectances sont intimement liées aux composants internes des feuilles. Cela signifie que la présence d'une maladie peut modifier la structure interne des feuilles et donc altérer sa signature.

Un spectro-radiomètre a été utilisé sur le terrain pour caractériser les signatures spectrales des feuilles à différents stades de croissance. Afin de déterminer les réflectances optimales pour la détection des maladies (FD, Esca, BN), une nouvelle méthodologie de conception d'indices de maladies (SDIs) basée sur deux techniques de réduction de dimensions, associées à un classifieur, a été mise en place. La première technique de sélection de variables utilise les Algorithmes Génétiques (GA) et la seconde s'appuie sur l'Algorithme de Projections Successives (SPA). Les nouveaux SDIs résultants surpassent les indices de végétation spectrales (SVIs) traditionnels et GA était en général meilleur que SPA. Les variables finalement choisies peuvent ainsi être implémentées en tant que filtres dans le capteur MS.

Les informations de réflectance étaient satisfaisantes pour la recherche d'infections (plus que 90% de précision pour la meilleure méthode) mais n'étaient pas suffisantes. Ainsi, les images acquises par l'appareil MS peuvent être ensuite traitées par des techniques bas-niveau basées sur le calcul de paramètres de texture. Plusieurs techniques de traitement de texture ont été testées mais uniquement sur des images couleur. Une nouvelle méthode combinant plusieurs paramètres texturaux a été élaborée pour en choisir les meilleurs. Nous avons constaté que les informations texturales pouvaient constituer un moyen complémentaire non seulement pour différencier les feuilles de vigne saines des feuilles infectées (plus que 85% de précision), mais également pour classer le degré d'infestation des maladies (plus que 74% de précision) et pour distinguer entre les maladies (plus que 75% de précision). Ceci conforte l'hypothèse qu'une caméra multispectrale permet la détection et l'identification de maladies de la vigne en plein champ.

Les premiers essais du dispositif « capteurs-drone » se dérouleront lors de la future campagne d'acquisition de 2019.

ACKNOWLEDGMENTS

CONTENTS

GLOSSARY

ANN	ARTIFICIAL NEURAL NETWORKS	**HIST**	HISTOGRAM MOMENTS
ARI	ANTHOCYANIN REFLECTANCE INDEX	**HS**	HYPER-SPECTRAL
AUC	AREA UNDER ROC CURVE	**HSI**	HYPER-SPECTRAL IMAGING
BN	'BOIS-NOIR'	**LBP**	LOCAL BINARY PATTERN
BP	BIOPHYSICAL PARAMETERS	**LOG**	LAPLACIAN OF GAUSSIAN
CN	'COURT-NOUEE'	**MCAI**	MODIFIED CHLOROPHYLL ABSORPTION INTEGRAL
DAMAV	AUTOMATIC DETECTION OF GRAPEVINE DISEASES	**MS**	MULTI-SPECTRAL
DET	DETRENDING	**MSC**	MULTIPLICATIVE SCATTER CORRECTION
DWT	DISCRETE WAVELET TRANSFORM	**MSI**	MULTI-SPECTRAL IMAGING
ELISA	ENZYME-LINKED IMMUNOSORBENT ASSAY	**NDVI**	NORMALIZED DIFFERENCE VEGETATION INDEX
FD	'FLAVESCENCE DOREE'	**PACA**	'PROVENCE-ALPES-COTES D'AZUR'
FI	FLUORESCENCE IMAGING	**PCA**	PRINCIPAL COMPONENT ANALYSIS
FNR	FALSE NEGATIVE RATE	**PCR**	POLYMERASE CHAIN REACTION
FPR	FALSE POSITIVE RATE	**PRI**	PHOTOCHEMICAL REFLECTANCE INDEX
FRAC	FRACTAL DIMENSION	**PSSRA**	PIGMENT SPECIFIC SIMPLE RATIO CHLOROPHYLL INDEX A
FS	FLUORESCENCE SPECTROSCOPY	**PSSRB**	PIGMENT SPECIFIC SIMPLE RATIO CHLOROPHYLL INDEX B
GA	GENETIC ALGORITHMS	**PSSRC**	PIGMENT SPECIFIC SIMPLE RATIO CAROTENOIDS
GLCM	GRAY-LEVEL CO-OCCURRENCE MATRIX	**ROC**	RECEIVER OPERATING CURVE
GLD	GRAPEVINE LEAF-ROLL DISEASE	**RS**	REMOTE SENSING
GLRM	GRAY-LEVEL RUN-LENGTH	**SDI**	SPECTRAL DISEASE INDEX
GM1	GITELSON AND MERZYLAK 1	**SIPI**	STRUCTURE INSENSITIVE PIGMENT INDEX
GM2	GITELSON AND MERZYLAK 2	**SPA**	SUCCESSIVE PROJECTION ALGORITHMS
SVM	SUPPORT VECTOR MACHINES	**UAV**	UNMANNED AERIAL VEHICLE

TCARI/OSAVI	RATIO OF THE TRASNFORMED CHLOROPHYLL ABSORPTION IN REFLECTANCE INDEX AND OPTIMIZED SOIL-ADJUSTED VEGETATION INDEX	**WAV**	WAVELET
TI	THERMAL IMAGING	**WI**	WATER INDEX
TS	THERMAL SPECTROSCOPY	**ZTM**	ZARCO-TEJADA MILLER

CHAPTER 1

1 INTRODUCTION

France has a reputation for producing wine of high quality. Around 790,000 hectares of vineyards can be found in over seventy departments and thirteen regions within France [1]. The economic and territorial development of the country highly relies on the wine and vine sector. In fact, the sector provides employment to around 100,000 people directly and 500,000 people indirectly. Furthermore, among the seven billion bottles of wine produced every year, more than 147 million are exported with a total value of 7 billion euros [2].

Since the main losses in this sector are due to diseases, a continuous protection approach is required, which means that fungicides/pesticides are uniformly sprayed in vineyards at regular and frequent intervals. More than ten treatments are carried out per season in several of the main wine-producing regions. Thus, it is crucial to detect initial symptoms of diseases in order to target their treatment selectively, preventing and controlling the formation of infection and its epidemic spread to other patches or even to the whole vineyard.

Currently there are two widely-used techniques for in-field disease detection: naked eye observations and biological approaches. Visual inspection is the most used technique for grapevine monitoring. It is a relatively easy approach when visual symptoms characterizing a certain disease are visible. Nevertheless, the accuracy of the diagnosis is always subjective. Although symptoms provide important information on the nature of diseases, one cannot make a decision upon appearance only. Usually, other confirmatory tests are performed after visual inspection to ensure an accurate diagnosis. The additional tests are usually biological ones. They have revolutionized the identification and quantification of pathogens and diseases. Under this category, we can find serological and molecular tests. The Enzyme-Linked Immunosorbent Assay (ELISA) is a serological method that relies on the proteins produced by a pathogen. The Polymerase Chain Reaction (PCR) is a molecular technique that depends on the specific DNA (Deoxyribo Nucleic Acid) sequences of a pathogen. There are many variants of the polymerase chain reaction and the enzyme-linked immunosorbent assay techniques that can be used for detecting different pathogenic viruses and bacteria in plants.

The above-mentioned biological tests have their limitations. To obtain reliable and accurate results, a definite procedure must be used, especially during sample preparation (collection and extraction), which implies more labor and more time. In order to be more effective, the biology-based techniques must directly be used in the field.

These mentioned traditional techniques are not capable of detecting early stages of disease development. An ideal system should localize the initial symptoms and manage the treatments to prevent the infection from growing in the vineyards. The control of grapevine diseases would greatly benefit from an innovative, rapid, non-invasive sensing of vegetation status. Remote Sensing enables the gathering of information about an object, area or phenomenon at a distance. It can be adapted for the identification of crop diseases based on the assumption that the stress induced by the pest modifies the physiological structure of the plant and affects the absorption of light energy and reflectance spectrum.

In this context, a new French collaborative project called DAMAV (Détection Automatique des MAladies de la Vigne - Automatic Detection of Grapevine Diseases) was initiated between many

industrial and academic institutions. It aims to detect in-field grapevine diseases, especially 'Flavescence dorée' (FD) by Unmanned Aerial Vehicle (UAV) imagery *(Figure 1.1)*. The DAMAV project combines the advantages of using an aerial platform with a multispectral imaging tool.

Figure 1.1 UAV equipped with a camera for grapevine disease detection.

FD is propagating in French vineyards inducing economic impact. There is no efficient, fast and in-situ solution and the monitoring takes a lot of time and mobilizes many human resources. The period of inspection is short and takes place during harvest; observations are made vine per vine and it is not certain that the entire vine will be checked. Finding a method that automatically inspects the vineyard instead of the winegrowers is an interesting challenge and might be of great interest for the viticulture domain. As the majority of infections appear during the harvest period, a control solution based on the use of no-contact sensor becomes even more interesting. Among the proxy-detection aerial image acquisition techniques, UAV-imaging is a booming technology that offers interesting advantages over remote sensing by satellite or manned aircraft. To control the FD spread, which is highly contagious and incurable, it is essential to detect any outbreaks as early as possible. Instead of using a regular camera that could only detect visible information, a MS sensor is employed. This sensor is able to acquire not only visible but also near and short-wave infrared information which may offer an advantage in detecting the occurrence of diseases early in time.

The DAMAV project involved many partners:

- **Novadem**, project leader, has UAV specialists while Airbus Defense & Space (D&S) offers expertise in navigation, both of them will make sure that the UAV can navigate correctly to capture the images.
- **Global Sensing Technologies**, can acquire and process images; **UMR Agroecology** provides researchers in agricultural imaging, together they will build the image acquisition and processing chain.
- **Fredon Provence-Alpes-Côte-d'Azur**, **office of vines and wines of the Burgundy region** and **Bernard Magrez** domain from Bordeaux have expertise in grapevine diseases and will offer the testing fields.

The results and products of this project will reduce costs with respect to the use of traditional diagnosis techniques and will enhance the yield that dropped due to FD. The project also has an environmental impact since it tends to reduce the use of phytosanitary products: instead of being applied uniformly for prevention, they will only be used on detected foci.

The present PhD-thesis was part of the DAMAV project. Its objective was to participate in the design and implementation of a MS image acquisition system. This is done by defining the most relevant reflectance and image characteristics related to FD and possibly other grapevine diseases.

In the scientific literature, reflectance and image data were found capable of detecting some grapevine diseases like Mildew and Grapevine Leaf Roll on one grapevine variety usually, is it the case for other infestations? How about testing other infestations on many varieties? To answer this question, we consider in this study reflectance and image data of healthy and diseased grapevine samples affected by three different infections (FD, BN, Esca) on multiple red and white varieties. The majority of the studies analyzing reflectance data for crop disease detection were done inside a laboratory under controlled conditions. A specific disease is inoculated into a specific plant and the data is acquired and analyzed with time. However, some diseases such as FD for vines are contagious and very difficult to cultivate in a laboratory so they need to be studied under production conditions immediately. We collected hyperspectral reflectance data and color images from grapevine fields during our campaigns but then **how can we insure the reliability of reflectance and image data acquired under real field conditions?** To answer this question, we defined repeatable protocols to ensure a correct acquisition of the data. From this, another question arised, **is our reflectance data acquired under field conditions capable of detecting diseases? How about discriminating among these diseases?** For this question, a visual and statistical analysis was done to see at what extent the reflectance data could identify grapevine diseases. If reflectance data is really beneficial, then **how can we define general optimal reflectance features capable of detecting a specific disease?** For this question, some traditional and new methods for variable selection are proposed and compared in order to assess the potential of few reflectance features in grapevine disease detection. **For a MS imagery practical case, having the minimal number of permitted bands (four), how can we exactly define these four optimal reflectance features?** For this question, an extension of the variable selection methods was put in place. It consisted on designing new indices that are disease specific (SDIs) and suggested better results than the traiditonal vegetation indices (SVIs) found in the literature. **After detecting grapevine diseases with reflectance features, what other image related features could be adopted for discriminating these diseases?** Here, texture features are extracted from images and are investigated in grapevine disease identification. In addition, a method for selecting the best texture features is presented.

In general, the imaging technology isn't much used in grapevine crop lands; the amount of work that has been employed to automate the visual detection of vine infections is small with respect to the major impact of those diseases on wine quality. There is still much space for improvement: firstly, the most successful work to date has not been done under production conditions. Secondly, we should be able to detect the disease early when symptoms are not yet clear. Thirdly, a vine disease must be studied in many cultivars including white and red-berried ones. Fourthly, the work that has been undergone mainly targeted the detection of powdery and downy mildew and GLD in grapes so we

should explore the potential to identify other diseases and stresses. Our PhD work actually tries to cover those mentioned improvements.

Our work contains a data acquisition phase and a data analysis phase. In our acquisition campaigns, we performed the complete reflectance measurements at a leaf level scale. In practice, the first campaigns started in the summer of 2016. The reflectance acquisition was dedicated at that time to FD and was conducted in the French Provence-Alpes-Côte-d'Azur (PACA) and Burgundy regions. Reflectance is measured with a suitable instrument and color images were also recorded in addition to reflectance to keep track of the spatial information of leaves. After these first measurements, we focused on reflectance analysis and identified discriminating reflectance features. In the summer of 2017, the reflectance and color image acquisitions were performed in new regions, Nouvelle-Aquitaine and Burgundy in addition to PACA. We also considered other diseases, 'Bois Noir' and Esca, in addition to FD. We found that reflectance features succeeded in disease detection but weren't enough for disease identification, so we started considering other features. Texture computation showed great potential in the scientific literature in disease discrimination so we employed it to enhance our analysis. In 2018, a texture analysis was put to place and was tested on RGB images. Texture features gave promising results.

Our work serves the DAMAV project; in fact, the reflectance features found will be implemented as filters. Furthermore, texture analysis correctly identified diseases on colored images. Hence, they are expected to be re-tested on MS images. Necessary corrections will be later needed in order to take into account passing from the leaf-level to the whole grapevine level or even to the canopy level with the MS sensor and the UAV.

1.1 ORGANIZATION

Our work involves a detailed reflectance analysis along with an assessment of the potential of texture analysis on RGB images that is expected to be later examined on MS images. This manuscript contains seven Chapters:

CHAPTER 1 explains the aim, the organization and the main contributions of the PhD work.

CHAPTER 2 shows how remote sensing was applied in the agricultural field for disease detection directly under production conditions and then a special focus is given to only one crop, grapevines. The grapevine diseases considered in this thesis are also presented with some state-of-the-art studies.

CHAPTER 3 presents how the reflectance and image data were measured and details the corresponding protocols.

CHAPTER 4 describes the reflectance data analysis. Reflectance data alone was capable of detecting the presence of a disease but was uncapable of grading the stage of an infection or differentiating among various infections. Since the reflectance is of high dimensionality, a dimension reduction technique, and in particular, a variable selection technique, is needed. A traditional method of variable selection used in the scientific literature and based on Spectral Vegetation Indices (SVIs) is first explained. Then, two methodologies that couple two different feature selection techniques (genetic algorithms and successive projection algorithm) with a K-Nearest Neighbors classifier are proposed. For the practical case of MS imagery, we introduced a novel methodology for designing indices that are disease-specific (SDIs). To our knowledge, this is the first time the successive projection algorithm is coupled to KNN for the purpose of feature selection. Although the genetic algorithm was already coupled to KNN for feature selection in previous studies, it is the first time that this coupling

is employed for an agricultural application. To our knowledge also, this is the first time a spectral disease index is created having our proposed particular form, and this is also the first time a spectral disease index is ever created based on genetic algorithms or successive projection algorithms. The new spectral disease indices based on the two feature selection methods demonstrated better results than using the complete spectral data or traditional SVIs. The selected features that are the result of the practical reflectance analysis are of great value in the design of the multispectral sensor, in fact the chosen features can be implemented as filters in the final sensor.

CHAPTER 5 exposed the texture analysis to further describe the infection appearance; this helped us in not only identifying a contamination but also grading its severity stage. The ReliefF algorithm was previously used to select the optimal textural features for disease identification in the medical field. To our knowledge, this is the first time it is actually applied on texture data in the agricultural field. Texture processing was only tested on colored images but is expected to be later used on MS images.

CHAPTER 6 presented the outcome of both reflectance and texture data analysis. A cross-test between different varieties (red and white berried vines) was done for reflectance data and another one between different regions (or diseases since in each region we studied one disease) was done for texture data.

CHAPTER 7 contained the conclusions and the perspectives of this work.

1.2 CONTRIBUTIONS

The multidisciplinary work done in this thesis allowed establishing the following main contributions:

- Definition of reproducible protocols for reflectance and image data acquisition directly under field conditions.
- Visual demonstration of how the reflectance pattern changes with respect to the grapevine variety and to the occurrence of a disease. Generally, a different pattern is seen for healthy and contaminated reflectance signatures. However, two diseases on similar varieties expressed similar reflectance patterns.
- Revelation of the weakness point of our sensing technique; in fact, reflectance information can detect the presence of a disease but may be unable to identify it or grade it.
- Assessment of the potential of traditional SVIs in detecting different diseases on various grapevine varieties. Definition of the best SVIs for grapevine disease detection.
- Proposition of a methodology for designing new disease-dedicated indices (SDIs). Two feature selection techniques were tested and expressed better results than traditional indices in detecting some grapevine infections on many varieties. This is the principal contribution of our work.
- Assessment of the potential of traditional textural features in discriminating diseases and grading their severity level. Definition of the best texture features for grapevine diseases and their optimal combination. The importance of texture as complementary information to reflectance features was hence demonstrated.

The main contribution of this PhD is providing a methodology that could be generalized to automatically detect and identify any disease or stress on any type of vegetation.

CHAPTER 2

2 REVIEW OF REMOTE SENSING TECHNIQUES FOR IN-FIELD FOLIAR DISEASE DETECTION: APPLICATION TO VINEYARD

2.1 INTRODUCTION

Generally, the spread of a disease in a field is not homogeneous and has a patchy behavior; therefore, sometimes pesticides are applied to some disease-free areas. In the ideal case the areas where the disease is present are defined and sprays are only applied to those infected parts of the field.

Plant disease monitoring is often defined as the detection of a disease early in time, the distinction among different diseases, the separation of diseases from abiotic stresses and finally the quantification of disease severity.

When a specialist determines the onset of symptoms, the presence of the infection is verified using some disease detection techniques. Presently, the plant disease detection methods available are Enzyme-Linked ImmunoSorbent Assay (ELISA) and Polymerase Chain Reaction (PCR). Despite the availability of these techniques, a fast, accurate and selective method is needed not only to replace, to some extent, the inspectors but also to target the infection in its early stages and control its spread.

Remote Sensing (RS) technologies have progressed tremendously since several decades. Consisting of a platform holding a sensor, RS technologies have the potential, for instance, to enable direct detection of foliar diseases under field conditions. However, many of the studies found in the literature are performed under controlled conditions, not taking into account direct in situ measurements under real production conditions; therefore they will not be considered here.

In this Chapter, we begin by assessing the possible in-field foliar disease detection techniques for different crops with emphasis on the RS methods (analysis techniques, platforms and sensors), the advantages and limitations of various sensors technologies for crop protection in precision pest management.

The *Figure 2.1* summarizes some of the methods that might be used for in-field disease detection.
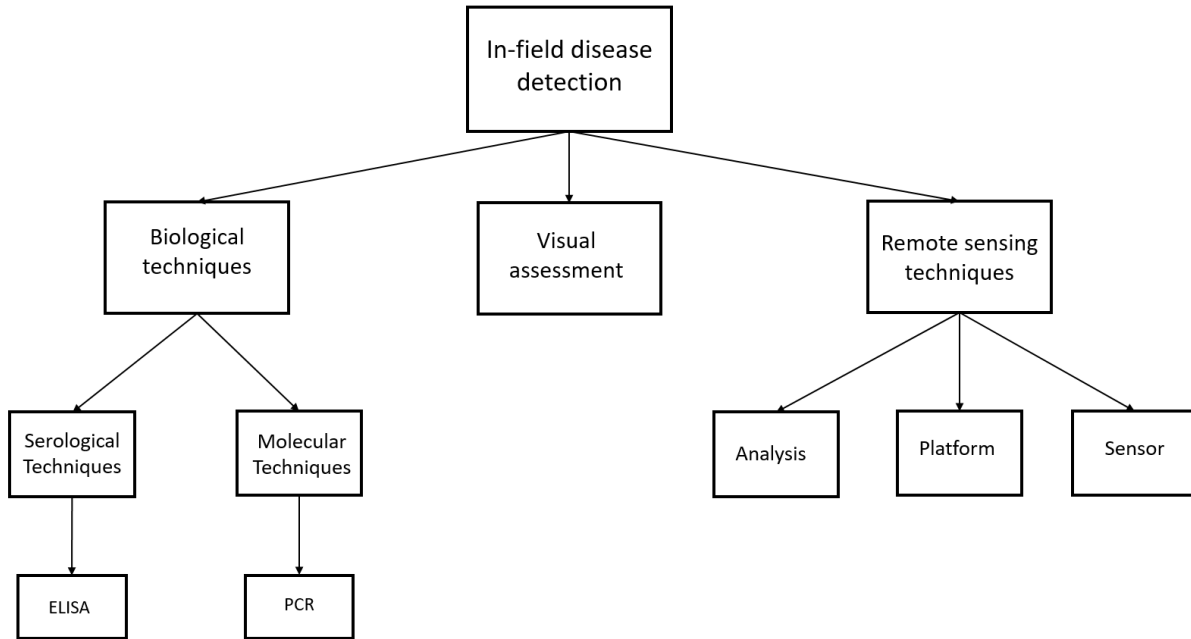
Figure 2.1 Different in-field disease detection techniques.

We also present in this Chapter, some recent studies that employed RS techniques to detect crop diseases directly under field conditions. A special emphasis is given afterwards to the vineyard sector. We discuss, in details, the grapevine diseases considered in the context of the thesis and, later, we study the potential of RS in detecting infections (early or not) in vineyards.

2.2 CURRENT REMOTE SENSING TECHNIQUES USED FOR DISEASE DETECTION IN THE FIELD

Visual inspection is a subjective procedure since it depends on the pathologist's experience and it is affected by temporal change. Moreover, the expert should be present for permanent monitoring which means that visual diagnosis is also expensive mainly in the case of wide canopies. Symptoms are generally a good indicator of a disease; however, it is not advised to make a decision only based on visual assessment. In addition to symptomatology, more tests are done to confirm the presence of an infection [3].

For instance, biology-based techniques have revolutionized the identification and quantification of pathogens and diseases. In the pool of serological methods, the ELISA technique is often used.

The antigen specific to a pathogen is first identified and purified. It is then injected to immune animals. Antibodies are formed in response and can be extracted from the animals. Enzymes have the capacity to associate with antibodies; the resulting molecule is serologically active and can be recognized even at low concentrations. Enzymes provide effective labels and act on some substrates by changing their colors. Once the sample fluoresce this means that the antibody is present and the rate of color change gives an indication about the amount of existing antibodies. This confirms the presence of a particular plant disease. Note that the enzyme label may be attached directly or indirectly to the antibody specific to a certain antigen. ELISA has the advantage of being stable, inexpensive

and safe to use and can be used successfully without sophisticated equipment. In fact, ELISA kits are widely available because of their low costs compared with other detection methods. However, ELISA is somewhat time consuming, labor-intensive and has low potential for spatialization.

One of the most popular molecular techniques is probably the PCR technique. This technique is adopted to create many copies of a Deoxyribo Nucleic Acid (DNA) target. The target is placed in a tube containing primers, free nucleotides and a DNA polymerase enzyme. The primers are small DNA sequences that stick to the beginning and the end of the DNA target. The sample is then placed in a PCR machine. The machine automatically increases and decreases the temperature of the mixture. The DNA polymerase begins to synthesize new segments of DNA starting from the primers. The newly formed DNA copies serve as templates in following cycles, which allows the PCR technique to amplify the DNA target millions of times. The DNA of a microorganism causing a plant disease is extracted, purified, and amplified exponentially. A gel electrophoresis is performed; when a specific band in gel electrophoresis is found, this confirms the presence of the disease.

The affordable kits, reagents and equipment make PCR available for easy use. The technique is effective but is usually applied to confirm diagnosis after symptoms appear. This means that this method is unable to detect early infections and is actually quite unreliable before symptoms appear. Also, similarly to ELISA, PCR is time consuming, labor-intensive and spatialization is difficult. Moreover, for the two previous most common biological techniques, there is a procedure that should be well-respected during the preparation of the samples and the analysis, allowing the results to be accurate and reliable.

More information on the various approaches applied for identifying virus and bacteria in crops, with special emphasize on PCR and ELISA based techniques and their variants can be found in [4].

Thus, innovative tests that are capable of directly detecting pathogens under natural conditions and not only in the laboratory are needed.

Remote sensing gathers data about a remote phenomenon without direct physical contact. It involves registering reflected or emitted energy from an object and then process, analyze and apply that data. Remote sensing tools can not only acquire electromagnetic energy such as visible and infrared light, thermal radiation and microwaves. Remote sensing is also able to monitor the vegetation status remotely, evaluate the crop spatial features, vigor and health demands. Novel developments using different types of platforms, highly sensitive sensors and multiple data analysis techniques have been published.

2.3 ANALYSIS TECHNIQUES

Data produced by RS tools are sometimes huge, implying a computational burden. As a consequence, in order to successfully assess plant-diseases, an adequate handling of high dimensional data is needed. Many data processing and analysis techniques are proven to be effective particularly for dealing with hyperspectral data.

The scientific literature of remote sensing for plant disease detection can be organized under four categories:

- Correlation and regressions analysis of the infestation scales with respect to spectral signatures.

- Use of general spectral vegetation indices or derivation of specific spectral indices characteristic of a certain disease.
- Dimension reduction techniques (feature selection/extraction) adopted for spectral data processing
- Image analysis and classification techniques (supervised/unsupervised) to detect plant diseases using digital images, to indicate the presence or the absence of a disease, or to distinguish between several types of infections.

The use of a specific analysis technique remains, usually, application-dependent.

2.4 PLATFORMS

It is used to install the sensor. The platform can usually be held in strategic points or can be movable.

Ground-based RS methods such as on-board systems on land vehicles (tractors, quads ...) present some drawbacks, they are costly and can affect the crops (injuries, soil compaction), they also risk promoting the propagation of the disease and aerial imaging proves to be more appropriate in this context.

Satellite imaging is clearly, at present time, not suitable for disease detection due to the limited spatial resolution in the images. Satellite sensing can, hence, only highlight relatively large infected areas. Moreover, satellites depend on cloud cover which makes them non-flexible. Also, their revisit time makes them unpractical because the information may not be available when needed. On the other hand, Aircrafts and Unmanned Aerial Vehicles don't have these constraints and could be used whenever needed. Acquisition equipment will have to be, however, more complex and more expensive in the case of an air-based system than a ground-based one.

The air-based RS technology and, in particular, the use of Unmanned Aerial Vehicle (UAV), has become a trend. Due to their flexibility and speed, UAV seem to be the best choice for monitoring diseases directly in the field.

2.5 SENSORS

RS methods are closely linked to the device used for data collection. RS sensors, are divided into two categories (*Figure 2.2*) depending on the application: spectroscopic (non-imaging) or imaging sensors.
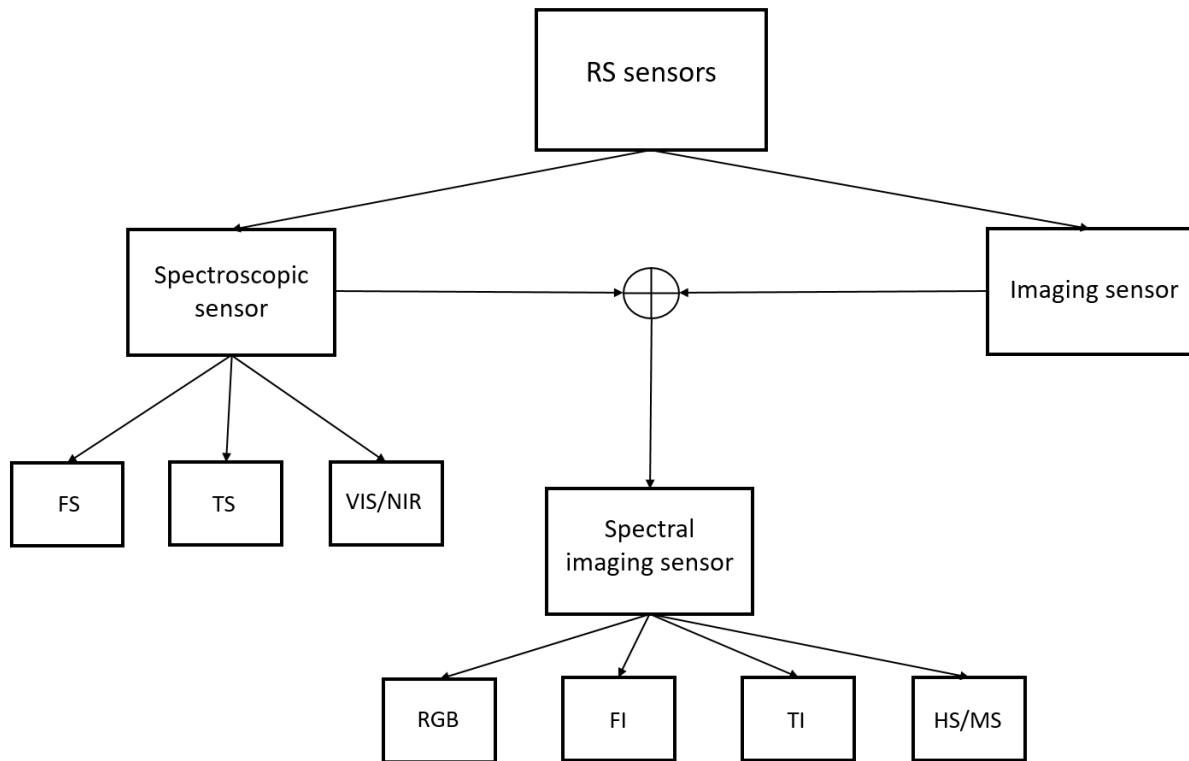
Figure 2.2 Different remote sensing sensors for in-field disease detection.

FS	Fluorescence Spectroscopy	RGB I	Red/Green/Blue Imaging
TS	Thermal Spectroscopy	TI	Thermal Imaging
VIS-NIR S	Visible-Near Infrared Spectroscopy	FI	Fluorescence Imaging
		HS/MS I	Hyperspectral / Multispectral Imaging

2.5.1 SPECTROSCOPIC SENSORS

They provide only spectral data. A reflectance signature is the percentage of reflectance emitted by a body with respect to the wavelengths, and can be used to characterize, identify and discriminate between objects. Indeed, each object has a special spectrum that informs about its properties and chemical composition, peaks at particular wavelengths indicating specific components. It is obvious that objects having similarities in their chemical composition also present similarities in their reflectance patterns. Add to that, different concentrations of the important compounds in the tested sample make the difference in reflectance values at some key wavelengths.

Among spectroscopic sensors, we mention the visible/near infrared sensors. Other technologies may include Fluorescence Spectroscopy (FS) and Thermal Spectroscopy (TS), but thoses weren't in the scope of our work.

2.5.1.1 *VISIBLE-NEAR INFRARED SPECTROSCOPY (VIS-NIR S)*

Figure 2.3 shows the typical reflectance signature of a leaf. It was analyzed by [5] as follows:

In the visible spectral region, the high absorption of radiation energy is due to leaf pigments, primarily the chlorophylls. Particularly, in the green region, the reflectance is low with a peak at around 550 nm. This peak accounts for the green color of plants perceived by the human eye.
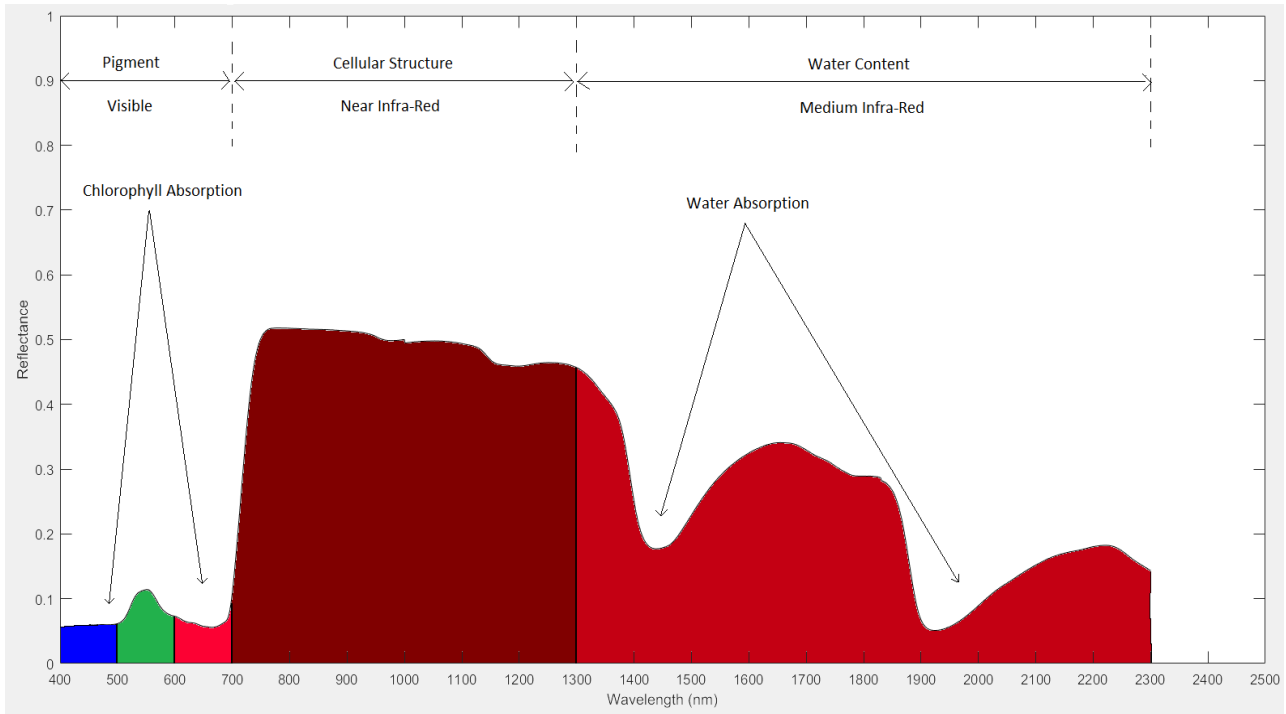
Figure 2.3 Typical reflectance signature of a green leaf (Source: [6]).

In the infrared spectral region, the reflectance increases over the range 700 to 1300 nm, but gradually decreases to a low value at about 2700 nm. The strong absorption by a leaf in the infrared beyond 1300 nm is due to water. The high infrared reflectivity of leaves appears to be caused by their internal cellular structure.

Outside the spectral range (400 – 2700 nm), near the short wavelengths in the ultraviolet domain or the long wavelengths in the far infrared domain, the leaf reflectance is small and approximately constant.

The latest studies, applying VIS/NIR spectroscopy in the field for disease detection, have used VIS/NIR spectroscopic sensors (mainly radiometers or spectro-radiometers), hand-held, meaning that no platform was needed. Authors in [7] studied the effect of an infestation caused by leafhopper on leaf chlorophyll concentration and relative water content for cotton leaves. They used in their study spectral vegetation indices and regression analysis. Similarly, in the work [8], reflectance data was applied to assess the infestation level caused by rice leaf folder diseases. Wheat diseases were studied in [9] and [10] with different approaches: the former used correlation analysis on vegetation indices

to predict the virus presence and the latter applied discriminant and regression analysis to discriminate between three wheat disease conditions.

In conclusion, spectroscopy is a simple, precise, time-efficient tool necessitating very few sample preparation. It has the advantage of analyzing material contents, control the quality of various agro-food products and monitor industrial operations. However, when a measurement is taken, spectroscopy performs average calculation of the composition over only a small area of the tested sample, which means that spatial heterogeneity is not taken into consideration.

2.5.2 SPECTRAL-IMAGING SENSORS

Imaging sensors, unlike spectral ones, provide spatial information; they enable the interpretation of external characteristics such as size, shape, color, texture and permit the evaluation of external visible disease symptoms. Nevertheless, the chemical composition of the sample (including moisture, fat and protein concentrations) can't be calculated with simple images.

When spectroscopic and imaging techniques are combined, we obtain what we call spectral imaging sensors. These sensors collect the spectral information over contiguous wavelengths for each pixel in an image and then produce a three-dimensional data cube (x and y represent the spatial datasets and z represents the spectral dataset).

Among spectral imaging sensors, we mention color imaging and multispectral or hyperspectral imaging. Other technologies may also include Fluorescence Imaging (FI) and Thermal Imaging (TI), but thoses weren't in the scope of our work.

2.5.2.1 *RED GREEN BLUE IMAGING (RGB I)*

A color image is the association of three wide spectral bands: red, green and blue, in this sense, a color image can be seen as a three-dimensional data cube (x, y, λ), the first two dimensions (x, y) are the spatial dimensions (the spatial location of every pixel) and the third dimension (λ) is a spectral one. A pixel holds the captured data in the three 100-nm-wide spectral bands. These bands generally capture the light in the blue (400 nm, 500 nm), green (500 nm, 600 nm), and red (600 nm, 700 nm) portion of the visible optical spectrum [11]. Every pixel of a common digital color photograph will, therefore, contain, at the same spatial location, three digital numbers to represent the amount of radiation captured in the three broad spectral bands.

RGB images are important tools for judging plant's vigor; they form an acknowledged technology for disease detection, identification and quantification. A progress in the technical specifications of the sensor was achieved over time; in particular, the light sensitivity, the spatial resolution or the optical and digital focus are enhanced [12].

Color, grey levels, texture, dispersion, connectivity, shape parameters can be defined as features in plant disease assessment and the selection of the most relevant features, among these, often increases classification accuracies. Many research works are employing pattern recognition methods and machine learning approaches to diagnose plant diseases by processing digital images.

RGB images of five common infections of beet leaves were studied in [13] with a simple cellphone camera. In [14] a methodology was established to identify three diseases that appear on cotton leaves with a usual RGB camera. Both studies applied feature extraction and image analysis techniques but used different types of classifiers.

Deep neural networks were adopted to identify 26 diseases from 14 plant varieties in the study conducted in [15]. Authors used an available dataset containing images of contaminated and healthy leaves and analyzed two architectures: the AlexNet and the GoogLeNet. Transfer learning gave in general better accuracy (up to 99%) than training from scratch for both architectures. The approaches yielded excellent results on the PlantVillage dataset which was collected in a controlled environment. However, when the network is tested on images sampled from trusted online sources and acquired at different conditions than those used for training, the accuracy is considerably reduced (31%). The study concluded that a more diverse set of training data is needed to improve the general accuracy.

Hence, if we aim to automatically process digital images for crop damage examination, a standard image capture procedure should be respected. In fact, focus, sharpness, illumination, imaging angle, distance between the sensor and the object of interest are all parameters that affect the image quality and should be constant in order to obtain repeatable and accurate results.

The problem with using RGB imaging is that we are not capable of detecting diseases prematurely. This is due to the fact that three spectral bands with limited range are not sufficient to identify the disease before its symptoms become visual.

2.5.2.2 *MULTI-SPECTRAL/HYPER-SPECTRAL IMAGING (MS/HS I)*

RGB devices use three spectral bands with limited range; it is an example of a 3-band multispectral image, where each band records one of the colors red, green and blue. It is common to have more bands in a multispectral image and even more for hyperspectral image. A detailed review was presented in [16] including the principles of HSI and its applications in agro-food products. It was mentioned that a spectral image is a pack of successive images of the same phenomenon each in a different spectral narrow band. The first difference between MS and HS imaging systems is the number of spectral bands considered; for MS, it is normally less than 10 bands while for HS, it can be up to hundreds of adjacent and systematically spaced bands. The second difference is that MSI offers spectra values in only few wavelengths for each image pixel, while in the HSI each pixel has a complete spectrum. Thus, HSI stores more detailed data and therefore is more sensitive to subtle variations in reflected energy but requires specific processing techniques capable of dealing with such amount of information. The third difference is the spectral overlapping of bands; indeed, the spectral bands in MSI are usually between 20 nm and 100 nm wide and can display a small spectral overlap comparable to digital cameras. MS imagers often have adjacent bands with a little spectral gap inbetween them (e.g., (450 nm, 500 nm), (550 nm, 600 nm)). The HS imagers, on the other hand, have a higher spectral resolution (often a bandwidth of about 10 nm) and capture more and nearly contiguous adjacent and not overlapping spectra [11].

A MS camera was fixed on an airborne to determine the Rhizoctonia crown and root rot infestation scale in sugar beet lands in [17]; for that study, spectral vegetation indices and statistical analysis were applied. On the other hand, in [18] the use of the Landsat 5 Thematic Mapper images for surveying wheat canopies affected by streak mosaic virus was assessed. The analysis relied on using spectral mixture analysis techniques for subpixel classification.

A hyperspectral HyMap sensor held on an airborne made the essence of two studies: First, authors in [19], discriminated soil-borne diseases generating stress in sugar beet planted regions using spectral vegetation indices and Spectral Angle Mapper (SAM) classifier. Second, article [20] also used the SAM classifier in addition to Support Vector Vachines (SVM) to detect wheat disease(s).

A combination of the two technologies MS and HS study was conducted in [21], a HS and a MS camera were held on an airborne to survey greening disease in Citrus crops. Spectral un-mixing and SAM classification were applied.

New sensor design and improvement are constantly suggested to make vegetation monitoring more pleasant. For example, work [22] presented a new multispectral instrument, MAIA, for calculating many vegetation indices in the context of agricultural multispectral analysis. MAIA has a high resolution and frame rate frequency; it holds 9 sensors allowing a high and well-distributed number of wavelength bands covering a range of 390 to 950 nm. On another hand, a hyperspectral system, having the advantage of being light-weight, was introduced in [23] for rotor-based unmanned aerial vehicle. The improvement made in the processing chain is explained and shows the potential of the instrument in agricultural applications.

HSI/MSI are also capable of fusing spectroscopy and imaging methods, hence, they are apt to define, on one side, the chemical contents, and on the other side, the spatial distribution of those contents in the object imaged. Nevertheless, under real field conditions and natural illumination, many factors can modify the spectral signatures such as: absorption and scattering of light, cloud shadows, specular reflection, surrounding objects that can reflect illumination onto the area of interest [24]. Although the use of a calibration strategy whenever illumination conditions change, can be efficient, but this might be required too often in a natural environment.

The advantages and the limitations of hyperspectral images were assessed in [25]. It was mentioned that HSI implies simplicity in acquiring data, cheap price per analysis, fast inspection, simultaneous investigation of several compounds; it's a non-destructive and reliable technique that doesn't require a sample preparation. It has the potential to detect diseases in crops, agricultural production and food, but its application is restrained due to the price of the devices.

In addition, for rapid image capture and analysis, HSI demands a very high hardware speed; also, calibration models are important and the use of chemometric models is often required. Another drawback of the HSI includes the overlap of bands which makes it hard to assign pixels to specific chemical groups (pixels are assigned as anomalous). Finally, some of the major challenges in HSI based plant disease detection are the selection of disease-specific spectral bands and the selection of a classification algorithm for a particular application.

HS/MS systems are probably the most interesting for detecting disease under field conditions early in time. Each and every one of the technologies mentioned above has its own limitations. A solution might be to integrate many sensor technologies into one more reliable system. For example, the work in [26] combined a fluorescence imaging and a thermal imaging technique under laboratory conditions. The concept can be extended to distinguish different kinds of pathogens damaging a given crop, and after validation, the approach can be adopted in a greenhouse or in a cultivation surface.

Authors in [27] indicated that a hyperspectral spectrometer and a digital camera demonstrated good correlations with damage by greenbugs in wheat growing under field conditions.

2.6 GRAPEVINE DISEASES

France is the world's number two wine producer. Wine growing makes up 15% of France's agricultural revenues while accounting for only 3% of the land area used [28]. France is the global leader for wine exports, 30% of French wines are sold in export markets. One of the main causes that could compromise this sector is the occurrence of diseases.

Disease recognition in grapevines, like other kinds of crops, is a complicated task [29]: (1) different viruses can produce similar symptoms, (2) different strains of a virus cause distinct symptoms in the same host, (3) a disease can have different signs and symptoms sometimes depending on the grape variety, (4) several diseases can be present at the same time, (5) plants can exhibit disease-like symptoms as a response to unfavorable weather conditions, soil mineral/nutrient imbalances, damage caused by air pollution, and pesticides), (6) factors such as time of infection, and the environment can influence the symptoms exhibited, (7) some diseases may induce no apparent symptoms or cause symptomless infection.

Many grapevine diseases exist with respect to other crops. Powdery and Downy Mildew, Grey-mold, Black-rot, Leafroll are well-known grapevine infections, however, in this thesis, a special emphasis is given to grapevine yellows ('Flavescence dorée', 'Bois Noir') and to Esca.

2.6.1 GRAPEVINE YELLOWS

The yellowing of the grapevine has an economic impact on the vineyard. It can cause crop losses and affect the quality of the harvest. Its epidemic character makes the fight against it essential and obligatory. Yellowing of the grapevine alters the development of plants. It affects many plant parts, from fruit, to growing points, leaves and canes. "Bois Noir" (BN) and "Flavescence Dorée" (FD) are the main grapevine yellows.

All "vitis vinifera" varieties are susceptible to phytoplasmas. Some of the most sensible grapevine species in France are Alicante, Cabernet, white Ugni, Carignan, Grenache and Chardonnay, whereas Merlot appears to be relatively tolerant [30]. Plants infected with FD and BN may recover, but symptoms may reoccur, sometimes years later.

FD is a highly contagious quarantine disease that is incurable in grapevines. Southern Europe is facing a growing threat of FD outbreaks causing significant yield losses and compromising the sustainability of viticulture. Available treatment measures consider monitoring grapevines and uprooting infested ones in addition to using insecticides for vector containment. These solutions are costly and polluting; they pose a real problem for organic winegrowers, violate organic viticulture and all actions aiming to reduce pesticide use.

Phytoplasmas are bacteria in the Mollicutes class that lack cell walls. Phytoplasmas are obligate parasites that cannot complete their life-cycle without exploiting a suitable host. In fact, in order to survive and multiply, they enter the host cells to divert their metabolic activity. Phytoplasma strains responsible for FD are intracellular parasites that reproduce in plant phloem and in sucking insects. Genomic studies have demonstrated that they originated in Europe and already existed in wild plants before being introduced into grapevines. The insect vector comes from the United States and was likely introduced in France when American rootstocks were imported as a part of the fight against

downy mildew and phylloxera in the early twentieth century. It is difficult to cultivate phytoplasma under laboratory conditions, conventional techniques for their isolation, extensive cultivation, and genome sequencing could not be exploited.

FD is difficult to detect because the three characteristic symptoms do not always appear every year or on every shoot. To conclude the presence of grapevine yellows, three symptoms need to be displayed on a single shoot: modification in leaf color, decline in branch lignification and death of grapes/inflorescences. Leaf discoloration depends on the varietal, leaves in white varietals become creamy-yellow (*Figure 2.4 (a)*), red varietals become reddish (*Figure 2.4 (b)*). One of the difficulties in detecting the disease is that symptoms don't always appear every year, and may only be present on one shoot or on a small number of shoots. In addition, "vitis vinifera" varietals are not equally susceptible to FD and may not present symptoms with equal intensity.

(a)

(b)
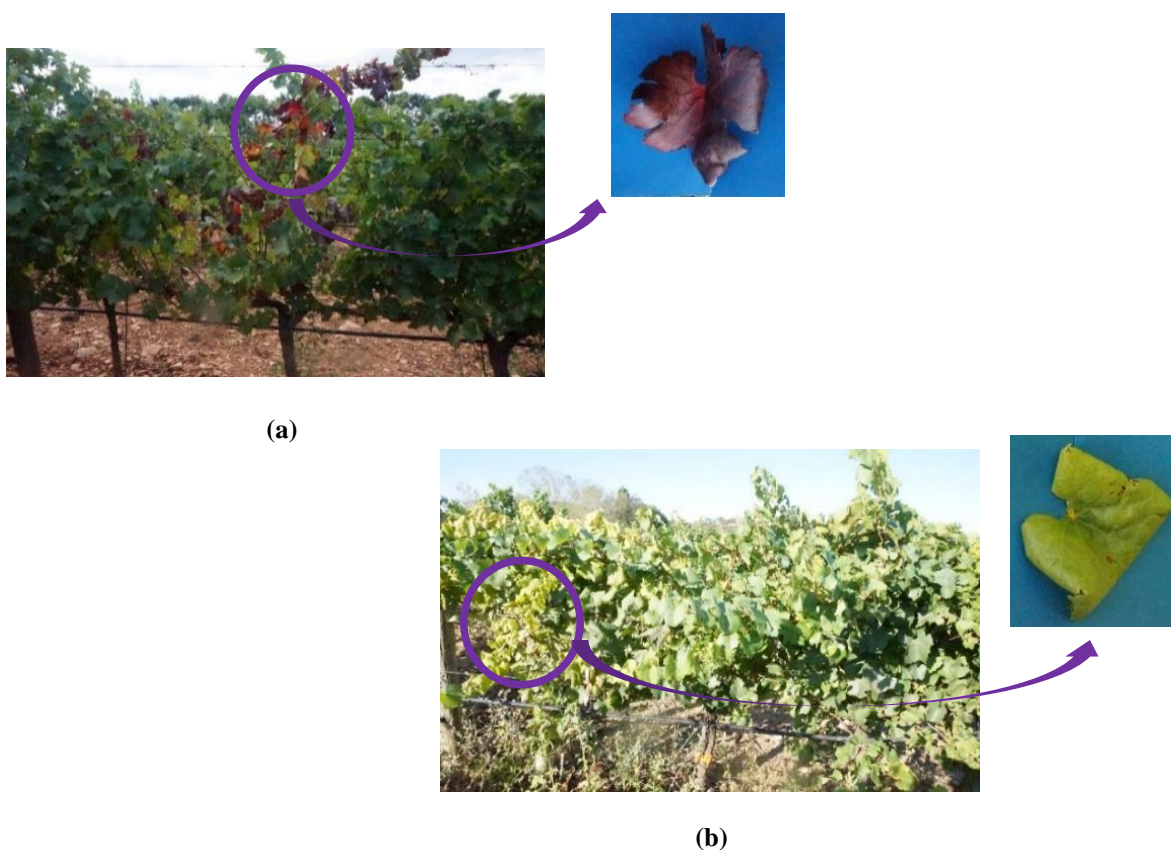
Figure 2.4 FD foliar symptoms on Marselan (red) **(a)** and Chardonnay (white) grapevine varieties **(b)**.

The color of leaves changes; leaves rolls downward resulting in an angular form, their blade become fragile. At the beginning, some spots appear on the surfaces of leaves exposed to the sun, then they increase and become bands until they cover the whole surface of leaves. As mentioned before, the color of the spots and bands is function of the cultivar.

Flowers de-hydrate and fruits won't appear on grapevines affected since the start of the season. Fuit sets will become brown and shrivelled on grapevines that are infested later in the season. Fruits on contaminated plants can fall quickly on some varieties.

Canes of affected plants don't ripen completely; they display instead irregular green and brown bands. These canes will look spongy and exhibit a weeping appearance. Diseased branches will show narrow pustules and long fissures will appear at the base of weak branches. During winter, symptomatic canes will die.

The main phytoplasma vector for "Flavescence Dorée" (FD) is the American leafhopper "Scaphoideus titanus". The sucking insect acquires FD phytoplasma when feeding on an infected vine. The phytoplasma then colonises the insect's body and multiplies. One month later, the leafhopper becomes contaminated and can transmit the phytoplasma to new hosts. It injects phytoplasma to the new plant with infected saliva prior to withdrawing nutrients; this happens each time it feeds. Insects remain infectious till death but phytoplasma is not transmitted to offspring. "S. titanus" is largely responsible for the rapid spread of the "Flavescence Dorée" in France and in Europe.

A single generation is produced each year. Eggs are settled at the end of the season in the bark of branches that are one year old. The insect then enters a period of suspended development due to unfavorable weather conditions. Young insects emerge from the eggs at the end of the spring. There are five larval developmental stages for the insect that extend over a period of 35 to 55 days (*Figure 2.5*) some of them are shown in *Figure 2.6*. The first adults usually appear at the end of July.
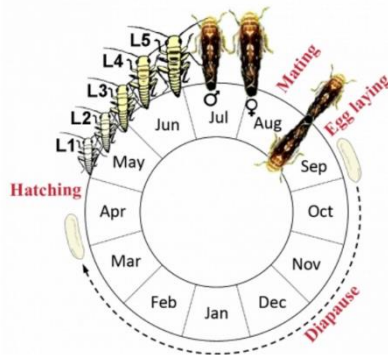
Figure 2.5 Cycle of the leafhopper S.titanus of FD [31].

Figure 2.6 Development stages of S. titanus, principal vector of the FD disease. From left to right: 3rd lavar stage, 5th lavar stage, adult. (Source: [32], modified).

Leafhoppers spread the disease at short distances to the adjacent fields; nevertheless, FD can be transmitted by the human at long distances through transportation of vegetative material. When a damaged plant material is used for grafting for example, this will spread FD in the new field. We

should note that the FD doesn't exist in all canes and shoots corresponding to contaminated grapevines, hence, not all cuttings will hold phytoplasma.

At present, compulsory control measures are applied over more than half of the France's total vineyard area. FD is also detected across other countries in southern Europe such as Italy, Portugal, Serbia, and also in Switzerland. *Figure 2.7* shows the distribution of the FD vector in Europe and in France respectively.

Grapevine Fields

Presence of S. titanus

(a) **(b)**
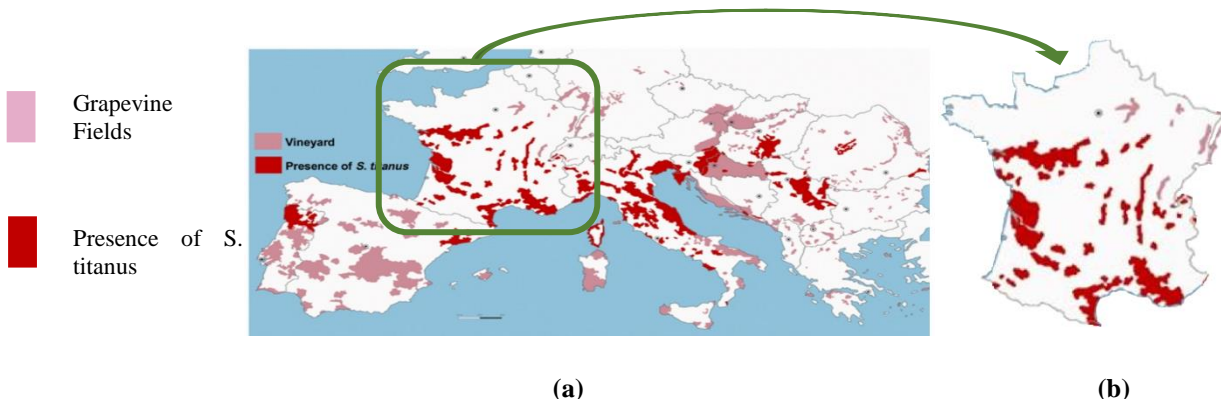
Figure 2.7 Presence of FD in Europe **(a)** and in France **(b)**. (Source: [31], modified).

It was declared as a quarantine pest by the European Union in 1993. A control of the leafhopper vector in all graft and rootstock nurseries across France has been made obligatory by a ministerial decree; the production of vines in wine growing regions infected with FD became closely controlled.

The disease gradually surrounds the initial foci and is transmitted by a specific leafhopper. However, the BN doesn't progress that fast and it is spread by another vector called "*Hyalesthes obsoletus*". Symptoms of BN are shown in *Figure 2.8*.

Figure 2.8 Symptoms of Bois Noir on Chardonnay grapevines in Burgundy region, France.

The presence of symptoms alone is often not enough to identify the Yellowing. In general, additional laboratory tests are needed in order to confirm the presence of the disease.

Current used strategies consist on removing suspected grapevines showing symptoms of FD, supervising the nurseries to insure a production of healthy woods and apply insecticide control to eradicate the FD vector and protect the vineyard. These treatment measures seem to be expensive and can cause pollution. They also became a problem for winegrowers and wineries that are willing to invest in organic viticulture.

Besides being polluting and costly, phytosanitary protection of vineyards can imply a resistance in the long term and badly effect the environment. That's why, on-going research is exploring new control procedures that will decrease the use of insecticides. On one hand, these can be based on studying the insect's behavior. On another hands, hunting for resistant vines might offer an environment-friendly alternative [33].

- Leafhoppers send vibrating signals to attract mates. A possible procedure would be to artificially create other vibrating signals with the aim of disturbing the communication between males and females. That way, the insect can no longer reproduce or will have a very limited chance to do it.
- Use of strategies based on attractive and repulsive stimulation to influence the insect's behavior and then attract them to a definite area where they will be destroyed.
- Seek of grapevine varieties that are naturally immune to phytoplasma. Look for graft combinations that are resistant to FD and determine the genetic principle of this resistance.

2.6.2   ESCA OF THE GRAPEVINE

In the literature, the term 'grapevine trunk diseases' refers to a number of different diseases. The most destructive among these diseases are Esca. Esca, also known as black measles, is a complex of fungal infections. It is a complicated dieback infection associated with pathogens and fungus that degrade the woody part and induce foliar symptoms. The disease is, till now, poorly understood as it has not been fully studied.

Esca usually occurs in adult plants aged 10 years or more. Esca has gained importance due to the damage in vineyards and the economic losses that it induced worldwide. Grapevine trunk diseases are considered to be the most fatal diseases of grapevine of the past 30 years and are quickly growing concern in all wine producing countries. The price for the replacement of dead grapevine plants alone is approximately predicted to be more than 1.5 billion dollars per year worldwide [34]. In the last years, several European wine countries, such as France, Italy and Spain have noted an increase of occurrence of the disease. Affected countries extended their research about the disease because there is, at the moment, no definitive cure for it.

A study was conducted to assess the nature of Esca spread within commercial vineyards Nouvelle-Aquitaine regions in France for a long period of time [35]. Another work was carried out from 2004 to 2006 in five mature vineyards in Aquitaine, France to monitor symptoms severity over time [36]. It was stated in [37] that Phaeomoniella chlamydospora is usually associated with grapevine Esca [38].

Leaf symptoms principally appear as a stripe pattern [39], they can begin to emerge any time during the growing season, but are mostly seen during July and August. When looking inside the trunks and main branches, a cross-sectional cut of symptomatic shoots uncovers concentric rings of dark spots [40]. Symptoms are often limited to an individual shoot or to shoots originating from the same spur or cane. Symptomatic leaves can become dehydrated completely and drop too early.

There are two forms of the Esca disease which can be distinguished depending on their severity and their rapidity of appearance. The first is a chronic form and the second is an acute form (also called apoplexy): it is a severe form of Esca and grapevines rarely recover from it. It results in a sudden dieback of one or many shoots, rather than a gradual development of foliar symptoms. The apoplectic form that kills the plants within a few days is more frequently present in Europe.

The chronic form of Esca produces symptoms inside the trunk and large branches, on the shoots and small branches, on the leaves and on the berries.

A red striping between the veins appears for red cultivars, where, discolorations surround some necrotic red-brown tissue (*Figure 2.9*). For white cultivars, a yellow color alteration may start as small, bright-yellow spots, gradually enlarging and becoming necrotic at the center later in the season. The spots start small then increasingly spread and fuse and eventually leave only a thin stripe of unaffected green tissue along the principal veins.

Figure 2.9 Symptoms of Esca on Cabernet Sauvignon grapevines in Aquitaine region, France.

The main symptom is the white rot that gives to the wood a rubbery appearance. Deteriorated tissue becomes whitish or creamy-yellow. The inside of the trunk shows a dark line that separates the decayed wood from the wood that hasn't rotted yet. When the mold arrives to the surface, the trunk begins to show fissures.

Symptoms begin to appear in spring disabling the correct growth of canes as these don't rippen properly. During summer, some shoots or branches may collpase and cannot recover.

Brown or violet marks can appear frequently on the berries. Similarly to foliar symptoms, the symptoms on the fruit can develop on only few berrries in a cluster, or all clusters belonging to a single branch or several branches.

The detection of Esca is a difficult task because symptoms don't necessarily appear every year. For more details on the syndrome development, refer to [41].

Identifying the creatures causing the illness has brought researchers and growers one step closer to recognizing the disease cycle and determining a management program. However, the actual treatments relying on the use of chemical products can't cure the grapevine. Preventive procedures are usually applied in vineyards. Measures start with the necessity of having a healthy vine since it is the first in the production chain. In fact, many actions are considered in the nursery before planting,

for instance, hot water treatments is often adopted to disinfect vegetative material and produce healthy clean plants. Other control measures are important to stop the spread of the disease. For example, removing and pruning residues, dead branches or vines are necessary preventive techniques. It is also advised to protect pruning wounds and to fix dead branches or shoots; if this is not possible then the plant should be completely replaced.

NaAsO2 was an effective product used in the past to eradicate the trunk diseases through the xylem. It was discovered that this product can be toxic for wine growers so the chemical product was forbidden since 2003. Today, there is no fungicide that is authorized and that can be used against trunk diseases.

Current research aims to focus again on the Sodium arsenate. But this time the idea is to understand the mode of action of the product on the infection and try to find an alternative that isn't toxic and that would have the same effect. Another area of research involves the hunt for resistant grapevines to trunk diseases. For instance, Cabernet Sauvignon, Sauvignon blanc, Mourvedre, Ugni Blanc, Cinsault and Trousseau are found to be more susceptible to Esca that Merlot, Carignan, Roussane and Pinot. The genetic potential immunity of vitis vinifera with respect to Esca should be considered more seriously.

Symptoms of grapevine yellows like wilting and leaf discoloration may be confused with esca disease, however in the case of Esca, the plant will collapse suddenly in July or August. Hence, detecting before to planting is crucial to assure durability of newly established vineyards.

2.7 REMOTE SENSING FOR GRAPEVINE DISEASE DETECTION

Many grapevine diseases exist with respect to other crops; the most important ones are probably Powdery and Downy Mildew, Esca, Grey-mold, Black-rot, Leafroll and grapevine yellows.

New technologies are aiming to improve precision viticulture and promising better wine and grape quality, so interest in the application of these technologies is growing. Pesticides are applied uniformly in the vineyard and according to a specific calendar in order to ensure a permanent protection of the vegetation against pathogenic attacks. Since pests are the principal reason behind losses in the vineyard sector, more than ten fungicide treatments are accomplished each season in major grapevine planted regions around the world. There is, hence, an interest in exposing fresh symptoms of diseases to selectively target their treatment, preventing and controlling the formation of the infection and its epidemic propagation to broader patches or to the whole vineyard.

Remote sensing is probably one the most applied methodologies in precision viticulture due to its robustness. It has the potential to produce a mapping of the vigor of wine-growing regions.

Some studies were found in the literature discussing the use of RS in viticulture. First study [50] considered RS as a tool for precision viticulture and they introduced the reader to RS in general and then reviewed its potential, applications not only on vines but also on soils. Second study [51] found that portable near infra-red devices provide an efficient tool for product quality control analysis, so they outlined the contemporary applications in Viticulture and other agro-food sectors. Third study [52] presented an overview of the studies made till now for the management of vineyards by using image processing and focused on the fundamental challenges for in-field ground-based image acquisition and analysis.

Current research axis of grapevine imaging include: yield estimation, quality evaluation, phenology prediction and finally disease detection. We will focus next on the in-field grapevine disease detection and state the relative existing studies.

2.7.1 SPECTROSCOPY

The spectroscopic sensors used for grapevine disease detection reported in scientific literature were mainly operative in the visible and near infra-red ranges.

2.7.1.1 *VIS-NIR S*

The two commercial proximal optical sensors GreenSeeker and Crop Circle were compared in [45] for detection of grapevine downy mildew symptoms under laboratory conditions, but these sensors have been conceived to work in the field. The experimentation used regression analysis on NDVI (Normalized Difference Vegetation Index) values which were correlated with growing proportions of infected areas on leaves. The results revealed that the green-seeker is more accurate in predicting the disease's stage. A limitation of the study involved the incapacity to conduct surveys under windy conditions because wind can change the orientation of the leaves and NDVI had smaller values on the abaxial surface of the leaves than on the adaxial surface.

Specific differences in vegetation indices and wavelength intervals were observed in [46] between leaf-roll GLD (Grape Leaf-roll Disease) infected and normal grapevine leaves. A portable spectrometer Field Spec acquired leaf reflectance under field conditions; those measurements enabled the calculation of some well-known vegetation indices. Discriminant analysis was then employed, providing specific parameters for detection of GLD during both non-symptomatic and symptomatic stages of the disease in red-berried cultivars. Similarly, authors in [47] used a portable spectrometer Field Spec to acquire reflectance measurements from GLD infected Merlot and Cabernet Sauvignon grapevine fields in China. The disease was classified into three stages according to its infection severity. To identify the GLD disease, a feature vector made of eleven parameters in conjunction with a clustering algorithm called "Ant colony clustering algorithm" are employed in the study. The feature vector was generated based on some vegetation indices to enhance the spectral differences and spectral discrimination between diseased and healthy grapevines. The clustering algorithm is then arranged to solve the fuzziness of the multi-spectral images and successfully identify GLD from normal grapevine leaves.

The two last studies were capable of differentiating GLD-infected from healthy grapevines; however, both methodologies were only tested on red-berried cultivars. As a matter of fact, white cultivars do not express visual symptoms similar to those observed in red ones, meaning that additional research is required before practical implementation. Moreover, the used approaches must be validated with other stress factors and diseases than GLD to prove its general application.

2.7.2 SPECTRAL IMAGING

This section introduces the different imaging sensors used for the purpose of grapevine disease detection in the scientific literature.

2.7.2.1 *RGB I*

Unlike studies where environments of systems are usually controlled in many aspects, no complex background is present and only a single leaf is examined, work [48] presented a grape leaf disease

investigation system that uses RGB images taken in a complex environment. Multiple grape leaves are considered with different sizes, orientations and light conditions, using a regular digital camera. The system was able to first segment the leaf from the background, then to extract the diseased area from the leaf and finally to classify the disease. Even though, the system demonstrated automatic diagnosis capability with promising performance, some difficulties were faced in extracting ambiguous color pixels from the background of the image.

Convolutional deep neural networks were used to classify thirteen diseases occurring in many crop species. Grapevine infections such as: wilt, mites, powdery mildew and downy mildew were considered. Authors in [49] extended an existing database and trained their network using Caffe framework to achieve an accuracy of 96.3%.

2.7.2.2 *HS/MS I*

Researchers examined the effect of the viewing angle on the sensitivity of the recognition of powdery mildew in grapevine leaves [50]. This study is conducted by applying a multispectral imaging approach: two appropriate combinations of spectra bands (vegetation indices) were tested according to different configurations of viewing angles for disease detection. Global results illustrated that when the viewing angle increases, the sensitivity of powdery mildew also increases. This finding is supposed to help in the design of an optimal automatic system capable of detecting powdery mildew in grapevine crop areas or in other vegetation covers susceptible to this disease.

In order to scale down economic impacts related to grapevine Leafroll disease, scientists in [51], studied the hyperspectral aerial imagery to recognize and estimate the disease severity at the vineyard scale. The goal of the study was to specify some of the best parameters to get a better sensitivity of disease identification. Those parameters included the radius for the pixel zone around the grapevine, the period of seasonal changes and the factors affecting image quality.

Recently, article [52] estimated the ability of a UAV multispectral imagery to remotely separate Flavescence Dorée symptomatic from asymptomatic hosts in four red cultivars. A set of eleven vegetation indices and four biophysical parameters are selected as interesting for the identification of Flavescence Dorée.

A multispectral UAV-based approach proposed in [53] and was tested on one red grapevine variety to evaluate the presence of grapevine leaf stripe disease foliar symptoms. The Normalized Differential Vegetation Index showed high correlation and good potential for identifying symptomatic hosts.

Finally, a merging of different kinds of imaging technologies was also studied in the context of grapevines. Scientists in [54] explored the design of a predictive model for pest detection in grapevine fields by processing airborne-collected Red-Green-Blue, multispectral and hyperspectral data at two different levels of phylloxera infection in vineyards. The collection, analysis and integration of different types of data is detailed in order to evaluate both old and new vegetation index.

The availability of a reproducible, efficient method that remotely detects and quantifies infections at the vineyard scale has great potential for bringing automatic disease management one step closer to become a reality. The technologies that were proposed in the above-mentioned studies give promising results but they must be refined, because, as we can see, each time a specific disease is tested in some specific cultivars only.

2.8 CONCLUSION

Highly authentic and non-contact systems are required to automatically examine plant diseases. RS devices provide the possibility to spatially locate the results; they make a valid tool capable of identifying infections early in time [55]. RS means were found, in many studies, able to quickly analyze crop canopies, to detect initial infections early, avoiding the spread of the contamination. RS tools are environment friendly and provide non-destructive analysis; however, they are affected usually by surrounding conditions and a calibration protocol is often required to compensate interfering sources. The number of published studies in this domain has been increasing in the last decade; however, there is still work to be done in the efforts to overcome the above-mentioned limitations.

It is also clear that the vine is a crop of great importance in France and detecting the occurrence of diseases by automatic RS means was highlighted in this Chapter. There are many grapevine varieties in France, some of them are more sensible to diseases than others. The targeted diseases were 'Flavescence dorée', Esca and 'Bois Noir'. The grapevine yellows include FD and BN, which have similar symptoms and can't be distinguished by the naked eye unless a laboratory test is undertaken. The FD is contagious and is spread via an insect called "Scaphoideus titanus". The Esca disease is a complex infection that has a chronic, soft and an acute, deadly form. Current research is looking for easy non-polluting means to reduce the negative effect of those diseases.

The tests, in this thesis, were done directly in the field in many French regions and different colored varieties were also considered. In the next Chapter, we will demonstrate how the measurements were acquired in the context of our study.

CHAPTER 3

3   DATA ACQUISITION

3.1   INTRODUCTION

In the previous Chapter, we reviewed some RS techniques for in-field disease detection. Moreover, we presented the grapevine diseases considered in the thesis and explained their symptoms, causes and treatments.

Our study consists on gathering data from sample leaves directly from the field. In fact, among the diseases studied, FD is a quarantine and contagious one, needing a suitable laboratory to cultivate it. This kind of laboratory wasn't available for our study. Furthermore, since the thesis was part of a project that aims to detect diseases using a UAV directly in the vineyard, we decided to get closer to the practical reality and conduct our study under uncontrolled conditions. The data were acquired at a leaf-level scale but corrections will be made later when going to the canopy scale with a UAV.

In this Chapter, we begin by introducing some reminders about the nature of our data, then, we present the materials needed for acquiring the data. Next, we introduce the protocols for reflectance and image data acquisition and we finally explain the parameters that may affect the accuracy of our measurements. The whole aim of these acquisition campaigns was to have a database to analyze, in order to define the best discriminating reflectance bands for the MS sensor design and to check the potential of texture analysis in grapevine disease detection. Reflectance and image data acquisition campaigns were performed in three different French regions, considering different grapevine varieties, different soil and climatic conditions along with multiple diseases. The protocols explained in this Chapter and all the tuned parameters are the same for all acquisition campaigns, all varieties and all diseases.

The main contribution behind this Chapter was the definition of reproducible protocols for a correct acquisition of reflectance and image data under uncontrolled conditions.

3.2   NATURE OF THE ACQUIRED DATA

Before describing the data acquisition procedure, it is important to understand the basic principles of radiance and how the surface materials interact with photons. The sun's energy propagates through space isotropically, so that at a given distance, the sun's emission can be measured as Watts per square meter [56]. Light interacts with matter in four unique ways. Transmission is when there is no perceived attenuation of the incident radiation as it passes through matter. Reflection is when incident radiation bounces off matter in a single expected direction. Scattering is when incident radiation is spread out in many unexpected directions. Absorption is when incident radiation is captivated by the matter and transformed into internal heat energy (*Figure 3.1*).
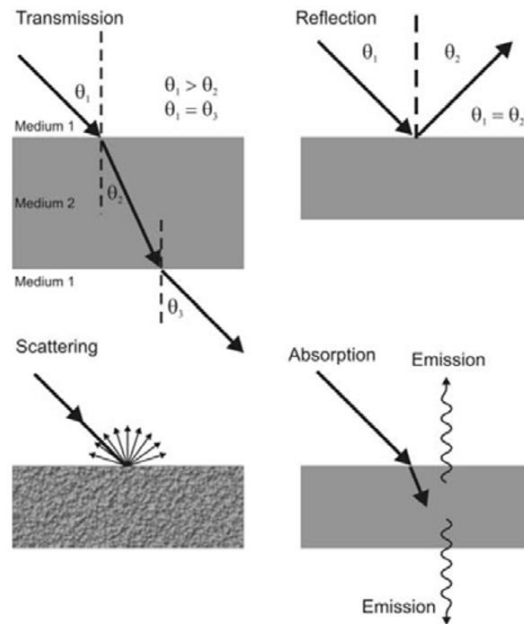
Figure 3.1 Behaviors of light when interacting with distinctive surfaces (Source: [57]).

Authors in [58] define imaging spectrometry as "the acquisition of images in many contiguous spectral bands such that for each pixel a radiance spectrum can be derived". Raw measurements of radiance are typically acquired as digital numbers, which characterize the intensity of a pixel for a given wavelength. These are converted through sensor calibration information to units of radiance. Because RS systems do not directly measure reflectance, a further conversion of radiance into reflectance is needed to allow comparison to field and laboratory reflectance measurements. This is done by dividing the total radiance by the radiance of a near perfect reflector. For laboratory and field spectrometry applications, this is accomplished by collecting radiance from a material that is a near-perfect diffuse reflector throughout the entire spectral range that is being utilized. Spectralon is the product most generally used as a reference for reflectance target.

3.3 MATERIALS FOR DATA ACQUISITON

This section deals with the hardware and software needed for properly acquiring reflectance and image data from leaf samples directly under field conditions.

3.3.1 SPECTRAL DATA ACQUISITION MATERIALS

- Spectro-radiometer: Spectral reflectance measurements from leaf surfaces were acquired using a portable Spectro-radiometer (FieldSpec 3, Analytical Spectral Devices, Boulder, CO, USA), shown in *Figure 3.2*.

Figure 3.2 Spectro-radiometer FieldSpec *3.*

The model number was FS3 350-2500 and the serial number was 16576. Each sample data was taken every 1 nm from 350 nm to 2500 nm using three different sensors. This was the result of an interpolation performed by the software because the true spectral resolution of the instrument is about 3 nm at 700 nm wavelengths and about 10 nm at 1400 or longer wavelengths. More details about its specifications can be found in *Table 3.1*.

Table 3.1 Specifications of the spectro-radiometer FieldSpec 3.

Spectral range	350-2500 nm
Spectral resolution	3 nm at 700 nm 10 nm at 1400-2100 nm
Sampling interval	1.4 nm at 350-1050 nm 2 nm 1000-2500 nm
Scanning time	100 ms
Detectors	One 512 element Si photodiode array 350-1000 nm Two separate, TE cooled, graded index, InGaAs photodiodes 1000-2500 nm
Input	1.5 m fiber optic (25" field of view) optional foreoptics available
Noise equivalent radiance (NEdl)	UV/VNIR 1.1 x 10-9 W/cm2/nm/sr at 700 nm NIR 2.4 x 10-9 W/cm2/nm/sr at 1400 nm NIR 4.7 x 10-9 W/cm2/nm/sr at 2100 nm
Weight	12 lbs (5.2 kg)
Calibrations	Wavelength, reflectance, radiance*, irradiance* All calibrations are NIST traceable (* radiometric calibrations are optional)

- Contact probe for vegetation: A contact probe has the general advantage of reducing the effect of environmental light scattering to insure better measurement accuracy (*Figure 3.3*). Measurements should be made on each leaf using a plant probe, specially designed with a low power source, for sensible vegetation surfaces, leaving no observable damage. It was employed instead of the traditional contact probe that could burn leaves due to a much more powerful Halogen bulb.

Figure 3.3 Contact Plant probe.

The serial number was 16576, more details about the specifications of the plant probe can be found in *Table 3.2*.

Table 3.2 Specifications of the plant probe to be used with FieldSpec3.

Length	10" (25.4 cm)
Weight	1.5 lbs (0.7 kg)
Power requirements	12-18 VDC, 6.5 W
Light source type/life (approx.)	Halogen bulb/1500 hours
Halogen bulb color temperature	2901 +/- 10% K
Spot size	10 mm

- Spectro-radiometer batteries for field usage: Two FieldSpec3 batteries must be previously charged and taken to the field to supply necessary power for the spectro-radiometer.
- White reference (*Figure 3.4*): The spectral response of a reference white surface is required because the spectral of a target is computed by dividing its spectral response by that of the reference sample. Applying this method cancels the parameters that are multiplicative in nature and present in both the spectral response of a reference sample and of the target. An inherent assumption is that the characteristics of the illumination are the same for the reference and target materials.

Figure 3.4 White reference used for calibration.

- Black support (*Figure 3.5*): A black surface is placed behind each sample leaf prior to reflectance measurement acquisition. This approach is used to ensure that the radiation

reflected is uniquely due to the leaf since the black support is supposed to absorb the transmitted radiation by the leaf and not reflect any.

Figure 3.5 Black support used behind sample leaves during reflectance data acquisition.

- Software RS3: The software is provided with FieldSpec3, and allows the user to adjust the measurement configuration, enables the spectral acquisition with the spectro-radiometer and saves the corresponding spectral data into memory. We used the version 6.0.10 of RS3.
- Software ViewSpec Pro: the software enables the visualisation of spectral data; we used the version 6.0.15 of Viewspec Pro.

3.3.2 DIGITAL DATA ACQUISITION MATERIALS

- Red-Green-Blue (RGB) camera: A digital camera Sony Alpha 5000 with 20 Mpixels resolution (*Figure 3.6*). More details about its specifications can be found in *Table 3.3*.

Figure 3.6 Sony Alpa 500 camera.

Table 3.3 Specifications of the RGB camera.

Maximum resolution	5456 x 3632
Effective pixels	20 megapixels
Sensor size	APS-C (23.2 x 15.4 mm)
Sensor type	CMOS
Lens Mount	Sony E
Max shutter speed	1/40000 s
Custom white balance, focus and exposure	Yes
Format	Uncompressed: raw, JPEG standard quality
USB	USB 2.0 (480 Mbit/sec)

- Color checker: A color checker is a collection of 24 accurately prepared, pure, flat, rich colors (*Figure 3.7*). Each colored squared area is prepared individually to produce natural, chromatic and primary colors in a wide range of shades (using the CIE). This is used to perform color calibration of the RGB digital images since the natural illumination can vary during the day.
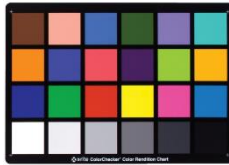
Figure 3.7 Color checker for color calibration.

- A uniform background having a color different than healthy or contaminated leaves is needed (not green, nor yellow or red). It is placed behind each sample leaf prior to image acquisition to facilitate leaf segmentation later.

3.3.3 EXTRA MATERIALS

- GPS to mark the tested grapevines in the field and to enable timely follow-up.
- Colored tapes for marking the grapevines and the leaves to facilitate the timely follow-up.
- A paper and a pen to keep track of some meta-data at each acquisition campaign. Error of measurements may have different sources: atmospheric conditions, operator, instrument and vegetation itself. Since it is difficult to control all parameters, it seems important to save them in a metadata file. More details about metadata will be presented in the rest of this Chapter.

3.4 PROTOCOLS FOR DATA ACQUISITION

In the second part of this Chapter, we discuss the protocols related to sample leaves selection and we state in order the protocols needed for a correct acquisition of reflectance and image data from grapevine fields.

3.4.1 LOCATION AND WEATHER REQUIREMENTS

1. Do not limit the measurements to one field. Try testing several locations to verify the reliability of the results.
2. Consider many varieties of the vegetation in order to determine if the impact of the disease is the same or if it is variety-dependent.
3. Perform measurements several times during the season to see the effect of the severity stage on the results. For the case of FD disease on grapevines, it is important to perform at least two acquisition campaigns. The first one must take place in July when the symptoms affect only some parts of the leaves. Then a second one could be performed in September or in the beginning of October, when FD symptoms appeared clearly on the entire leaves.
4. Do not perform measurements when it is raining or directly after also do not take measurements in case of highly smoky and haze conditions because the water left on leaves can highly impact the spectral reflectance.
5. Do not perform measurements in case of extreme windy conditions because the digital images might be blurred if the leaves are moving fast.
6. Do not perform an acquisition campaign in extreme high temperatures (above 40 degrees) because the DC noise is sensitive to temperature.

3.4.2 SAMPLE LEAVES SELECTION PROTOCOL

1. Choose carefully symptomatic leaves. Let an expert be present during the acquisition campaign to confirm the occurrence of the disease and its severity stage. The distinction between different crop diseases is a complicated task. Indeed, various infestations can cause almost similar

symptoms. Furthermore, marks can be the result of a fusion of many infections affecting the plant at the same time. Some factors such as nutrient deficiencies, pollution and pesticides can produce expressions indistinguishable from those of diseases. That's why it is advised to have a plant pathologist while selecting samples.

2. Choose leaves from three levels (bottom, medium, top), defined on the grapevine (*Figure 3.8*).

3. Select a range of healthy leaves of different ages. Choose a set of infected leaves in order to get a complete and representative range of the damage symptoms.

4. Examine at least a total of 30 infested and healthy samples for each variety; this is important to keep the study statistically significant. In fact, a sampling size that is statistically representative includes a number of sample sizes between 30-40 measurements with a minimum of 15.

5. Test the same number of healthy and diseased samples for each variety to avoid the bias in the analysis.

6. Perform laboratory analysis on tested symptomatic leaves after the end of the acquisition campaigns. The presence of symptoms alone is often not enough to identify the Yellowing in general; additional laboratory tests need to be done in order to confirm the presence of the disease. PCR analysis for example should be done in the closest laboratory to support the inspector's claim [59]. For multiplex-PCR protocol enabling the simultaneous detection of FD and BN please refer to [60].

7. Locate the grapevines using a GPS and label all tested leaves in order to ensure timely follow-up.

8. Choose another neighboring leaf when the old one is not available anymore. After testing some leaves in a campaign, they may not be present in the following one. Some of the leaves may either naturally fall or might be cut by the winegrower. Thus, when the leaf is not found, another candidate located on the same branch must be considered.
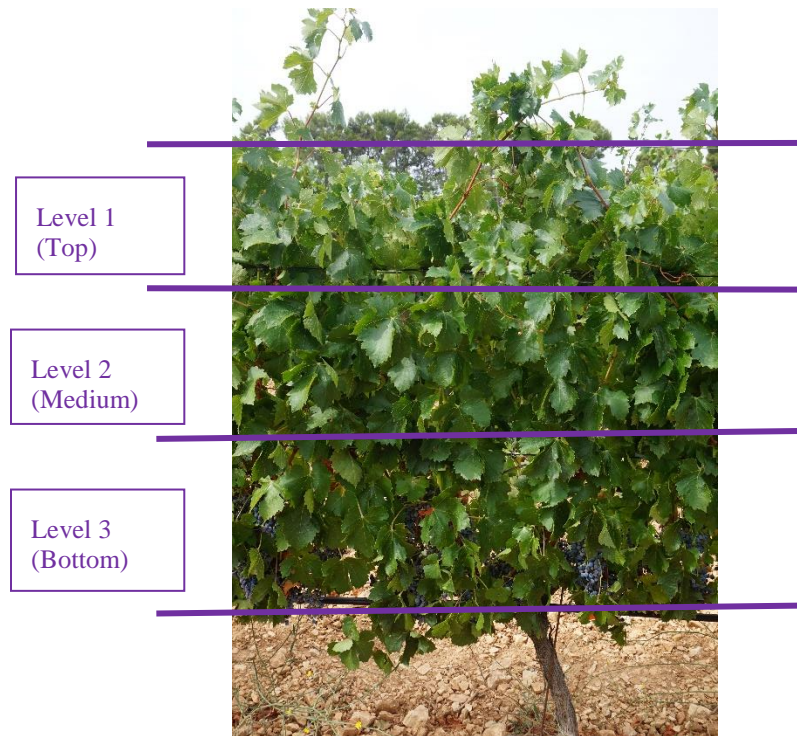
Figure 3.8 Choice of leaf sampling positions.

3.4.3   METADATA ACQUISITION PROTOCOL

Metadata must be noted for each acquisition, including factors that affect standardized measurements. Some of them are presented next and summarized in *Figure 3.9*.
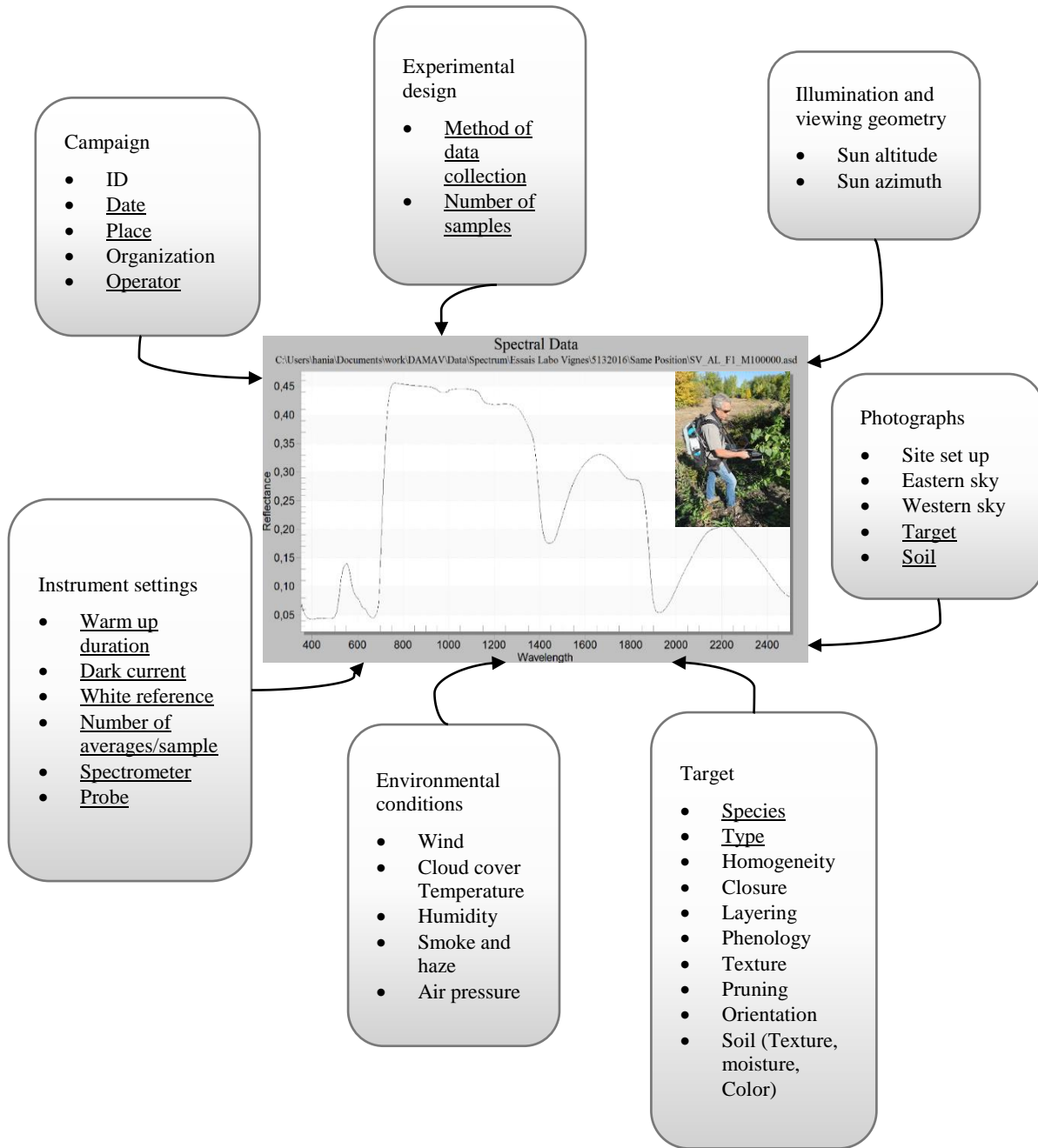
Figure 3.9 Meta-data considering uncontrollable parameters during reflectance data acquisition.

1. Start by noting the experimental design, the date and time of the acquisition, the method of data collection including geometry, scale and the number of samples

2. Check the illumination and viewing geometry: date, time, position, sun altitude, sun azimuth Fore optic degree, fore optic height above target, fore optic height above ground, IFOV (not used in case of contact probe)

3. Memorize the measurement information: number of averages/sample, dark current integrations, white reference integrations.

4. Observe the environmental conditions: wind speed and direction, cloud cover and type, temperature and humidity, aerosols/smoke/haze, air pressure

5. Take photographs for the site set up, eastern sky, western sky, nadir azimuth, target and soil

6. Perceive the target and note the species name, its health status and its texture (specular, diffuse)

7. Check the pattern of distribution, an even distribution describes a uniform cover of vegetation over the ground and a clumped one describes vegetation that presents as distinct clumps across plot.

8. Estimate and record the amount of layering within the vegetation plot where single describes a layer of vegetation where all plant components are at the same level and little scattering would occur, and multiple describes the vegetation that grow in layers as either different components of the plants or as different growth heights of individual plants.

9. Estimate and record the cover homogeneity: It is a percentage of cover of the target vegetation; ideally all plots must have 100% of the target species, at times, a percentage component of exposed soil interspace, leaf litter or an alien species.

10. Measure and record the maximum plant height mean density, this refers to the height at which most biomass occurs.

11. Describe the phenology, the vegetation can be in a green growth stage, flowering, seeding and senescing or dying; record any disturbances visualized.

The meta-data that are underlined are the most important ones to keep. Whether these changes are a result of biophysical changes of the target or attributable to the illumination conditions can only be assessed by an increased length of the sampling record, it is only with accurate reflectance and metadata collection that both averaged reference spectra and any significant temporal changes in reflectance response can be identified.

3.4.4 REFLECTANCE DATA ACQUISITION PROTOCOL

Spectral reflectance is the ratio of incident to reflected radiant flux measured from a surface over a defined range of wavelengths. Based on [61] and [62], the following reflectance in-field measurements protocol was established for all our acquisition campaigns:

1. Let the spectrometer warm up for at least 15 min before starting the measurements.

2. During the warm up time, use some colored patches to determine the locations of the grapevines and the corresponding leaf samples that will be measured in order to enable a temporal tracking.

3. Connect the PC to the spectrometer and attach the plant probe.

4. Start the RS3 instrument software.

5. Adjust the measurement configuration: fore optic selection (not used); spectral averaging selection spectrum averaging is the number of samples taken per observation: spectrum (30), dark current (100) and white reference (100). It is advised to use the same instrument settings for each acquisition campaign.

6. Place the probe on the white panel and perform an optimization, the profile will change while the instrument is adjusting; the different regions of the 3 detectors will be visible.

7. Perform white reference calibration by maintaining the probe on the white panel, a reflectance curve with a near horizontal line at a value of 1 should appear if the set-up is correct.

8. Place a black support under each leaf before acquiring the reflectance data to avoid interference.

9. Reflectance measurements are taken from four locations on the leaves (shown in *Figure 3.10*), according to expert recommandations and one measurement is taken per location.

The locations were chosen according to the disease that has tendency to start growing between the veins first. A range of two-four measurements is considered depending on the leaf surface with respect to the probe diameter. When the leaf is small, only two reflectance spectra from two locations are acquired, and when the leaf is wide enough, four reflectance tests are taken from all the four locations. When ready, press the spacebar to save the measurement.

10. Repeat the white reference calibration, whenever there are changes in illumination conditions.
11. At the end of the reflectance data acquisition, two, three or four reflectance measurements are acquired from one sample leaf.
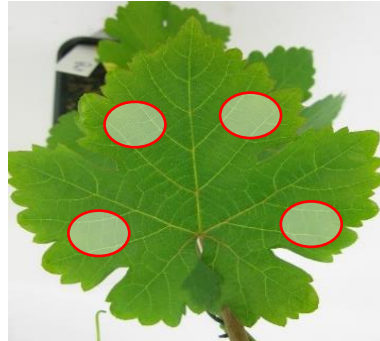
Figure 3.10 Locations of the reflectance measurements on a sample leaf.

3.4.5 RGB DIGITAL IMAGE ACQUISITION PROTOCOL

In order to keep track of the spatial appearance of the disease on the leaf, its location and how it is spreading on the leaf, we took an RGB image of the sample. The protocol for image acquisition for all campaigns is as follows:

1. Prepare a uniform color cardboard and place a color chart on it. This will facilitate the leaf extraction task and the color calibration.
2. Keep the same manual settings during all acquisition campaigns.
3. Place the cardboard behind each sample leaf.
4. Take one RGB image for each sample leaf (*Figure 3.11*).

Figure 3.11 Example of RGB image taken on a healthy Grenache leaf.

At the end of the campaign make sure to do the correspondence between the reflectance measurements saved and the color images of the leaves.

3.5 PARAMETERS AFFECTING THE DATA

Many factors can compromise the accuracy of measured reflectance data in the field as mentioned in [62]. Some factors related to the instrument of measurement, the operator, the surrounding conditions

and the vegetation could alter the reflectance response inducing errors. These are detailed in ANNEX A. Some propositions were given to prevent or reduce some of them (the other ones seem to be uncontrollable).

The above-mentioned factors are not to be considered in the case we measure reflectance data at a leaf-scale using a spectro-radiometer equipped with a contact probe. Actually, the contact probe is the most precise and consistent method for collecting spectra because most of the lighting and imaging geometry are controlled. Therefore, spectra will vary much less given any lighting condition or location. However, surrounding factors might modify the internal contents of sample leaves and hence alternate the reflectance. For example, stomata opening or closing to respectively release or detain water vapor is function of the temperature. Hence, it is desirable to repeat the instrument calibration process when we notice any modification in the environmental conditions. The operator should pay attention during the acquisitions for any sunlight illumination changes due to clouds passing for example. Furthermore, it seems interesting to also gather meta-data, if we analyzed the data collected from two acquisition campaigns and had strange results, we could refer to the metadata file and attribute the difference in the result to the modified meta-data parameters between the two acquisition campaigns.

Imaging systems on the other hand, keep evolving and have reached great performances recently. The perceived quality of a digital image, and hence its utility for research purposes, depend upon many interrelated factors. These include the operator, the imaging device, the object under study and the environmental conditions. More details about those factors are also given in ANNEX A.

3.6 SUMMARY OF ACQUISITION CAMPAIGNS

We have performed acquisition campaigns in the summer seasons of 2016 and 2017. We considered in 2016 the Provence-Alpes-Côte-d'Azur (PACA) and the Burgundy region; however, in 2017 we studied in addition, the Nouvelle-Aquitaine region. These regions can be visualized in the *Figure 3.12*.
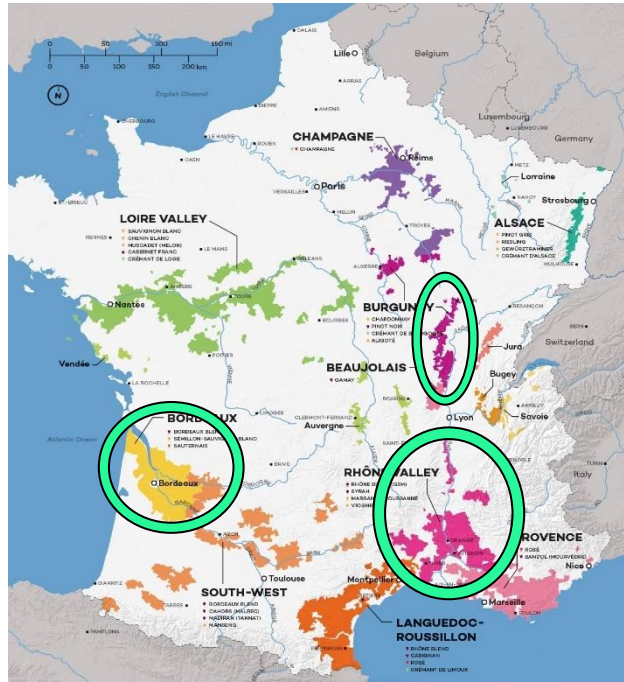
Figure 3.12 Map showing French wine growing regions. Studied regions are PACA, Burgundy and Nouvelle-Aquitaine (green ellipses).

The same four grapevine varieties infested with FD were tested in PACA in 2016 and 2017; two red-berried (Marselan, Grenache) and two white-berried (Vermentino, Chardonnay). Red-berried fields were measured first in the morning from (10:00 to 12:00) and white-berried fields were measured next in the afternoon (14:00 to 16:00).

In the Burgundy region, we examined in 2016 and 2017 a white berried variety of Chardonnay grapevines that were infested with BN and it turned out that there were also other diseases present. In the Nouvelle Aquitaine region in 2017, we investigated a red-berried variety of Cabernet Sauvignon grapevines infested with Esca first in the morning and another red variety of Merlot attained with Mildew. Traditional cultural practices are applied in the tested fields.

In the PACA region, four acquisitions were executed in the 2016 season and three others were done during the summer of 2017. The relative number of samples for each measurement campaign can be found in *Table 3.4*.

Table 3.4 Number of samples in PACA

Year	Date	Grapevine variety	Number of diseased samples	Number of healthy samples	Effective disease
	12/07	Noisy data acquired			
	09/08	Marselan	15	15	FD
		Grenache	15	15	FD
		Vermentino	15	15	FD

50

2016		Chardonnay	15	15	FD
		Red	30	30	FD
		White	30	30	FD
		All	**60**	**60**	FD
	06/09	No data acquired			
	27/09	Marselan	15	15	FD
		Grenache	15	15	FD
		Vermentino	15	15	FD
		Chardonnay	15	15	FD
		Red	30	30	FD
		White	30	30	FD
		All	**60**	**60**	FD
2017	27/07	Marselan	15	15	FD
		Grenache	15	15	FD
		Vermentino	0	0	FD
		Chardonnay	15	15	FD
		Red	30	30	FD
		White	15	15	FD
		All	**45**	**45**	FD
	11/09	Marselan	15	15	FD
		Grenache	15	15	FD
		Vermentino	15	15	FD
		Chardonnay	15	15	FD
		Red	30	30	FD
		White	30	30	FD
		All	**60**	**60**	FD
	04/10	Marselan	15	15	FD
		Grenache	15	15	FD

		Vermentino	15	15	FD
		Chardonnay	15	15	FD
		Red	30	30	FD
		White	30	30	FD
		All	**60**	**60**	FD
Total			285	285	

On the 12th of July 2016, the data we acquired were noisy due to an operator error. When placing the probe on the leaf, we shouldn't press save until the screen is refreshed twice and the spectra is stabilized. The spectra were hence not exploitable and not taken into account in this thesis. On the 6th of September 2016, we couldn't acquire any reflectance data due to a hardware error. Indeed, the Ethernet connection cable wasn't correctly functioning, only few reflectance data were taken before the connection was completely lost with the spectro-radiometer. Only few data were successfully acquired so they weren't considered in the rest of the study.

In Burgundy, three acquisitions were executed in 2016 and two others were done during the summer of 2017. The relative number of samples for each measurement campaign can be found in *Table 3.5*.

During 2016 in Burgundy, the expert told us that the grapevines seem to be infested with BN, however after the PCR analysis at the end of the season, it was shown that there were two other diseases: FD and court-noué that were present also among the tested grapevines. The CN, BN and FD samples from the Burgundy region in 2016 were ignored in the following analysis and discarded because we have less than 15 samples for each of the diseases (CN, BN and FD) which are not statistically significative to build a classification model.

Table 3.5 Number of samples in Burgundy acquisition campaigns.

Year	Date	Grapevine variety	Number of diseased samples	Number of healthy samples	Effective disease
2016	20/07	Chardonnay	15	15	BN,FD,CN
	14/09	Chardonnay	15	15	BN,FD,CN
	21/09	Chardonnay	15	15	BN,FD,CN
2017	17/08	Chardonnay	15	15	BN
	25/09	Chardonnay	15	15	BN
Total			**75**	**75**	

In the Nouvelle-Aquitaine region, two acquisitions were executed in the summer of 2017. The relative number of samples for each measurement campaign can be found and *Table 3.6*.

In Nouvelle-Aquitaine region during 2017, we acquired data from the Merlot grapevines type infected with Mildew according to the expert but it turns out eventually that there were two different types of Mildew present: Powdery and Downy. The DM and PM Merlot samples from Nouvelle-Aquitaine in 2017 weren't considered in the following analysis and were discarded because we have less than 15 samples for each of the diseases (DM, PM) which are not statistically significative to build a classification model.

Table 3.6 Number of samples in Nouvelle-Aquitaine acquisition campaigns.

Year	Date	Grapevine variety	Number of diseased samples	Number of healthy samples	Effective disease
2017	12/07	Cabernet Sauvignon	15	15	Esca
		Merlot	15	15	DM, PM
	19/09	Cabernet Sauvignon	15	15	Esca
		Merlot	15	15	DM, PM
Total			**60**	**60**	

3.7 CONCLUSION

In the present chapter, we explained in detail how the measurements were acquired, using which hardware and software. Our data were made of reflectance acquired by a spectro-radiometer and of RGB images acquired by a digital camera. The reflectance data neglect the spatial properties whereas the image data is a spatial representation of radiances. The higher the spatial resolution (case of images), the lower the spectral resolution is. The higher the spectral resolution (case of reflectances), the lower the spatial resolution is.

We also pointed out that there are some factors that seem to be uncontrollable and that can alter the data acquired and induce errors. We hence advise to register some metadata to take into account all these parameters. Next in this Chapter, we detailed the number of measurements taken in each campaign but not all of them were actually considered in the analysis. Accordingly, for the rest of the thesis, four varities (Marselan, Grenache, Vermentino and Chardonnay) are considered in PACA region and we focused on the FD disease. One red variety (Cabernet Sauvignon) is retained in Nouvelle-Aquitaine region and we focused only on Esca. One white variety (Chardonnay) is tested in Burgundy region and we focused on BN only.

In the next Chapter, we will start by finely assessing and analyzing our reflectance data.

CHAPTER 4

4 REFLECTANCE DATA ANALYSIS

4.1 INTRODUCTION

In the previous Chapter, we discussed how the data was acquired. The next step is to check the robustness of our reflectance measurements in the distinction between infected and healthy leaf samples.

A visual investigation of the shape of reflectance data and how the pattern changes depending on the variety type and the nature of the infection is given in ANNEX B. This visual analysis of reflectance data enabled the elaboration of some conclusions. Changes perceived suggest that the reflectance signature depends on the pathogen-host interaction, in fact, the differences depicted in the reflectance signature of diseased leaves were possible since the pathogen interacts with the host and modifies the internal structure of leaves. In other words, healthy reflectance patterns from grapevine leaves were somehow different from FD, Esca or BN affected ones. Moreover, the modification of the reflectance pattern produced by the occurrence of a particular disease is function of the grapevine variety. This means that a leaf from a red grapevine variety affected by a certain disease doesn't display the same reflectance pattern as a leaf from a white grapevine variety affected by the same disease. Furthermore, FD and Esca diseased reflectance patterns on one hand, BN and FD diseased reflectance patterns on another hand, expressed similar signatures on red and white varieties respectively.

A statistical investigation of the reflectance data was also done using a simple Principal Component Analysis (PCA) in ANNEX B. The aim was to check the potential of reflectance data in detecting FD, BN and Esca diseases, grading their severity stage and discriminating among them on different grapevine varieties. We applied PCA and calculated intra-class and inter-class distances on the obtained clusters. When inter-class distance (distance between two classes) is higher than the intra-class distances (distances inside each class), we can suppose that these classes are easily distinguishable. We concluded from this analysis that reflectance data could be employed with confidence to detect diseases but not to distinguish among them or to define their infestation level.

This result was also in accordance with what we found on the effect of diseases on the contents of leaves in the biological field. In fact, many biologists focused on the effect of phytoplasma (causing BN and FD) on leaves and found that besides the morphological changes that occur due to the presence of the disease, other functional changes may appear in infested leaves as a result of modifications in their contents. First, changes in pigment concentrations are noticed; the chlorophyll content is greatly reduced according to [63], 30% and 52% reduction for moderately and highly infected leaves respectively. Similarly, the Carotenoids (Car) content is decreased but less rapidly. In fact, authors in [64] found that the Car/Chl ratio was bigger that the Chl a/b ratio in phytoplasma affected leaves. Anthocyanin in phytoplasma infected red-berried grapevine leaves is accumulated and induces the purple symptoms. This purple discoloration of leaves seem to be appreciated by sucking insects that spread the disease. Sucrose is also accumulated which helps attracting many insect vectors and helps propagating the disease faster [65]. Serious variations in midday relative water content in September were observed for symptomatic leaves [66]. Also, the infested leaves displayed a reduction in stomatal conductance [67]. A low concentration of total soluble proteins was found after phytoplasma infection, this decrease might be the result of shrinkage in the production of

the main protein of leaf (RuBPC). A decline in protein synthesis can partially be responsible for chloroplast destruction as stated by [64].

Studies on grapevines and apples showed that the presence of phytoplasma causes a decline in photochemical efficiency, active reaction centers and electron transport rate inducing an inactive and inefficient photosynthetic activity. Researchers in [68] stated that once the Photosystem 2 (PS2) is influenced, the soluble carbohydrates and starch are accumulated in infected leaves. This phenomenon demonstrates that phytoplasma transforms leaves from source organs into strong sink organs [69]. Some hormonal changes are also noted after contamination and hormones play an important role in adjusting the growth and allocating nutrients in the plant. Apparently phytoplasma infection encourages the CA^{2+} flow into sieve tubes leading to blockage. Several studies like the one conducted in [70] demonstrated that phytoplasma originates modification in the phloem of host plants critically disturbing phloem transport.

Reduction in pigment concentration was also reported for Esca infected hosts. The content of chlorophyll badly decreased causing degradation in photosynthesis. However, in contrast to the phytoplasma case, the carotenoid content was not affected by esca since their concentration was similar in healthy and symptomatic leaves [71]. Similarly to phytoplasma, the PS2 inactivation is attributed to an imbalance in electron transport reflecting a reduction in the fraction of open reaction centers. Furthermore, Esca caused stomatal conductance, and transpiration to decline in symptomatic leaves. The intensity of gas exchange was in general found to be associated to the disease developmental level. A decrease in carbohydrate storage and starch were recorded which may affect the plant vigor the year after.

Some studies like [72] investigated the chemical composition and phenolic composition of Cabernet-Sauvignon grapes from Nouvelle-Aquitaine and suffering from Esca, since such composition may highly influence the wine quality. Changes in the internal structure and composition of berries, musts and wines were assessed in three groups of vines: vines showing esca symptoms, vines with asymptomatic Esca symptoms and healthy vines [73].

In general, the effects observed at a biological level for the diseases under investigation are very common. Hence, Phytoplasma and Esca cause non-specific disease responses in leaves and the changes in contents and activities connected with photosynthesis of the infected leaf tissues are comparable. This finding may explain why the reflectance signatures of yellowing and Esca cases were very much alike for the same grapevine species and why the hyperspectral data might seem insufficient for disease identification.

Our reflectance data could then be trusted only to detect the presence of a disease. For analyzing those data, one must notice that each a reflectance holds 2150 values or features, implying a high dimensionality problem. In the rest of this Chapter, our reflectance data are processed. We could either consider them as an input directly to the classification model, but this method isn't feasible due to the curse of dimensionality. Or we could perform a dimension reduction phase.

In dimension reduction problems, the data might be first pre-processed and prepared for further analysis. Then, two things could be considered, the first is the type of learning algorithm, supervised or unsupervised; the second is the choice of features, it can be a feature selection or a feature transformation technique. For the rest of this Chapter, when talking about a feature, it actually refers to a reflectance at a particular wavelength. The general flow of a dimension reduction is shown in *Figure 4.1*.
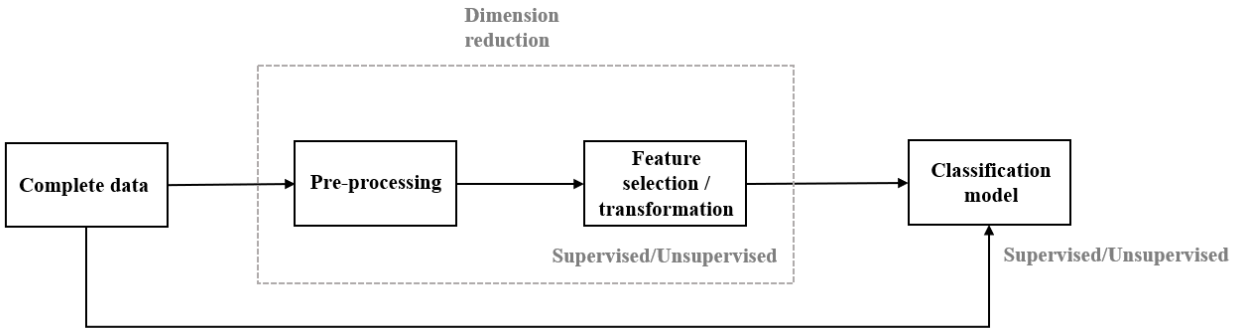
Figure 4.1 General schema for high-dimensional data classification, use of feature transformation or feature selection for dimension reduction.

Feature transformation consists on transforming the set of features to another new reduced one. It creates a substitue more compact set of features that holds the essential information. The new features formed don't necessarily have a physical meaning to the expert; we can find under this category feature extraction and feature generation. Principal component analysis is an example of feature transformation that is unsupervised. It aims to maintain the majority of the variance in the data and doesn't require any class labels to operate.

Feature selection spots the optimal minimal predictors from the orginal data space. In other words, it selects only some of the original features This technique has the advantage of keeping the physical meaning of the features. In fact, the selected features can be directly interpreted afterwards. It is indeed a challenge to determine a small number of variables among a huge number of variables particularly when the sample size is small. For instance, it can enhance the prediction and estimation accuracy, facilitate the interpretation and reduce the complexity of the model. Many unbiased methods for selecting important predictors have been proposed in the scientific literature. Similarly to feature transformation, a feature selection technique could be supervised or unsupervised.

Feature transformation techniques consist on transforming features which isn't effcient with respect to the goal of this thesis. In fact, the aim of the work was to define a subset of features to be implemented as filters in the MS sensor, those features need to be physically real and they should also be able to discriminate healthy from disease leaves. Hence, feature transformation techniques were discarded.

An example of traditional feature selection technique is considered and then two new selection methods were proposed and discussed. For the feature selection methods, the features are preserved and not transformed. They were tested for each variety, each disease and each level of infestation. All dimension reduction techniques take as input the hyperspectral reflectance data acquired in the campaigns.

In resume, the main contributions from this Chapter were: first, Spectral Vegetation Indices (SVIs), which are one of the main techniques for spectral analysis, were tested. Second, two feature selection techniques were proposed in the process of creating new spectral indices that are disease-dependant (SDIs) and could outperform common SVIs.

In this work and after PCR analysis, the data at disposal were labeled, which means that we had the classes of the samples (we know if the sample is infected and what is the infection in particular or if it is healthy), so a supervised classification model is employed.

4.2 TRADITIONAL REFLECTANCE ANALYSIS

4.2.1 INTRODUCTION

One straightforward way for analyzing our reflectance data is to use them entirely (*Figure 4.2*). This means that for each sample we will have 2150 reflectance values (from 350 nm to 2500 nm). The problem with this kind of data is that it contains noise and correlated features that might falsify the results; in addition to this, analyzing them is extensive. In fact, using high dimensional data shows its limitation from a computational point of view so in order to obtain a proper classification model able to differentiate symptomatic leaves from asymptomatic ones, a dimension reduction step seems mandatory.

Figure 4.2 Classification of high-dimensional data with no dimension reduction step.

4.2.2 TRADITIONAL SPECTRAL VEGETATION INDICES (SVIs)

An example of variable selection is the use of SVIs, here no particular pre-processing was needed. From the reflectance data, we can calculate multiple SVIs and each one of them could be used as an input into a classification model (*Figure 4.3*).
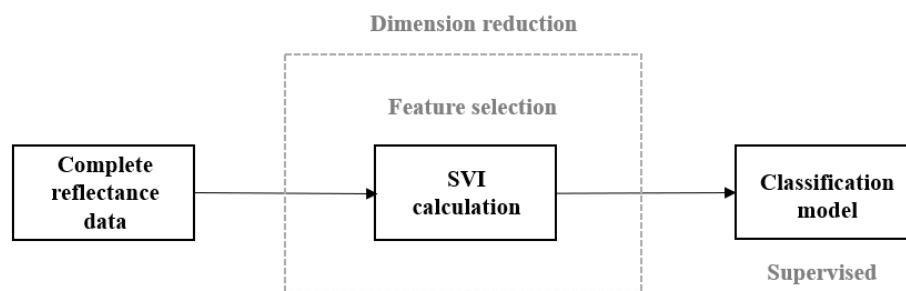
Figure 4.3 Classification of high-dimensional data using traditional spectral vegetation indices, a feature selection technique.

SVIs enable the analysis of vegetative spectral modifications due to the appearance of a symptom by calculating ratios of reflectance at different wavebands. SVIs are simple and effective algorithms for quantitative and qualitative assessment of vegetation cover, strength, and development dynamics and

can be applied for any other applications that include crop management. They are designed to improve the contribution of vegetation characteristics and allow true spatial and temporal inter-comparisons of terrestrial photosynthetic activity and canopy structural diversity.

Many researchers investigated the potential of SVIs in detecting diseases. For example, article [74] studied some sugar beet fungal diseases; article [75] on the other hand, analyzed basal stem rot infection in oil palm trees. Focusing mainly on grapevines, authors in [46], analyzed the spectral reflectance of red-berried leaves infected by Grapevine Leaf Roll disease (GLD) and found a set of variables capable of detecting pathogen presence. Another study was conducted in [47], it identified a feature vector made of 11 indices, able to spectrally differentiate healthy from GLD infected data.

SVIs enable the dimension reduction of the hyperspectral data. More than 150 SVIs have been published in the scientific literature. Some SVIs correlate with the biochemical constituent concentrations of vegetation (chlorophyll, carotenoids, water, cellulose, lignin, dry matter…), thus, associating the physiological status of crops to other hyperspectral data. Other SVIs have not been systematically tested or do not have a biophysical base. In our study, we only selected 13 of the most commonly used SVIs in the scientific literature. The NDVI, the PRI, the ARI, the SIPI, the mCAI, the PSSRa, PSSRb and PSSRc, the GM1 and GM2, the WI, the ZTM and the TCARI/OSAVI indices used in this work are displayed in *Table 4.1* below.

Table 4.1 Typical vegetation indices found in literature and used in the context of this thesis.

Index Name	Formula	Association with Relevant Plant Pigment	Reference Example
Normalized Difference Vegetation Index (NDVI)	$NDVI_{705} = \dfrac{R_{750} - R_{705}}{R_{750} + R_{705}}$	NDVI is a very typical index. Positive values suggest vegetated areas.	[76], [77]
Photochemical Reflectance index (PRI)	$PRI = \dfrac{R_{570} - R_{531}}{R_{570} + R_{531}}$	PRI index is a function of the reflectance at the 531 nm, this reflectance is related to xanthophyll. When the xanthophyll activity is high, the light use efficiency is low, meaning a possible stress occurred.	[78], [79], [80]
Anthocyanin Reflectance Index (ARI)	$ARI = \dfrac{1}{R_{550}} - \dfrac{1}{R_{700}}$	ARI index is designed to estimate the stack of anthocyanin in senescing and stressed leaves.	[81]
The structure insensitive pigment index (SIPI)	$SIPI = \dfrac{R_{800} - R_{445}}{R_{800} + R_{680}}$	The SIPI index is responsive to the ratio of carotenoids to chlorophyll. It is very practical to use when the canopy structure or leaf area index are inconsistent.	[82], [83]
Modified chlorophyll absorption integral (mCAI)	$mCAI = \left(\dfrac{R_{545} + R_{752}}{2}\right)(752 - 545) - (\sum\limits_{R_{545}}^{R_{752}} 545 - 752(1.158\, R_{\lambda})$	The mCAI is sensitive to the chlorophyll content. It calculates the area between a straight line connecting two points (the green peak at 545 nm and 752 nm) and the curve itself	[84]

Pigment specific simple ratio chlorophyll a (PSSRa)	$PSSRa = \dfrac{R_{800}}{R_{680}}$	The pigment specific ratio indices were suggested to estimate the pigment's content at the leaf level. Samples from trees at different senescence stages were studied aiming to empirically determine the best individual wavebands for pigment assessment (680nm for chlorophyll a, 635nm for chlorophyll b, 470 nm for the carotenoids).	[82], [85], [86]
Pigment specific simple ratio chlorophyll b (PSSRb)	$PSSRb = \dfrac{R_{800}}{R_{635}}$		[82], [87], [88]
Pigment specific simple ratio carotenoids (PSSRc)	$PSSRc = \dfrac{R_{800}}{R_{470}}$		[82], [88]
Gitelson and Merzlyak 1 (GM1)	$GM1 = \dfrac{R_{750}}{R_{550}}$	GM1 and GM2 were created to measure the chlorophyll content in vegetation leaves.	[89], [90]
Gitelson and Merzlyak 2 (GM2)	$GM2 = \dfrac{R_{750}}{R_{700}}$		[91]
Water index (WI)	$WI = \dfrac{R_{1300}}{R_{1450}}$	WI is related to water deficit	[92]
Zarco-Tejada Miller (ZTM)	$ZTM = \dfrac{R_{750}}{R_{710}}$	ZTM is a Red edge index highly correlated to chlorophyll content. At the canopy level, it has the advantage of minimizing shadow effects.	[93], [94]
Ratio of the Transformed Chlorophyll Absorption in Reflectance Index and Optimized Soil-Adjusted Vegetation Index (TCARI/OSAVI)	$TCARI = 3((R_{700} - R_{670}) - 0.2(R_{700} - R_{550})\dfrac{R_{700}}{R_{550}})$ $OSAVI = (1 + 0.16)\dfrac{R_{800} - R_{670}}{R_{800} + R_{670} + 0.16}$	A combination of the Transformed Chlorophyll Absorption in Reflectance Index (TCARI) and the Optimized Soil-Adjusted Vegetation Index (OSAVI). It is sensitive to chlorophyll content variations and resistant to variations in Leaf Area Index (LAI) and underlying soil background effect.	[95], [96]

After calculating the different values of vegetation indices from the reflectance data, these are injected one at a time into the classification model.

4.3 PROPOSED REFLECTANCE ANALYSIS

4.3.1 INTRODUCTION

The SVIs assessed are common ones; however, the impact of plant diseases on the physiology and phenology of plants varies with the host-pathogen interaction and each disease may influence the reflectance signature in a different way. Common SVIs are not disease specific (or disease-dependent); hence, it seems beneficial to design special indices for each infection (SDIs), as these might simplify the disease detection by spectral sensors.

In the rest of the Chapter, we propose a technique based on designing new SDIs.

4.3.2 THEORETICAL DEFINITION OF OPTIMAL REFLECTANCE FEATURES

We investigated the efficiency of two feature selection techniques: GA and SPA techniques. These two methods were used extensively in the scientific literature to select feature from high dimensional and highly correlated data such as reflectance. The overall flow is presented in *Figure 4.4* next.
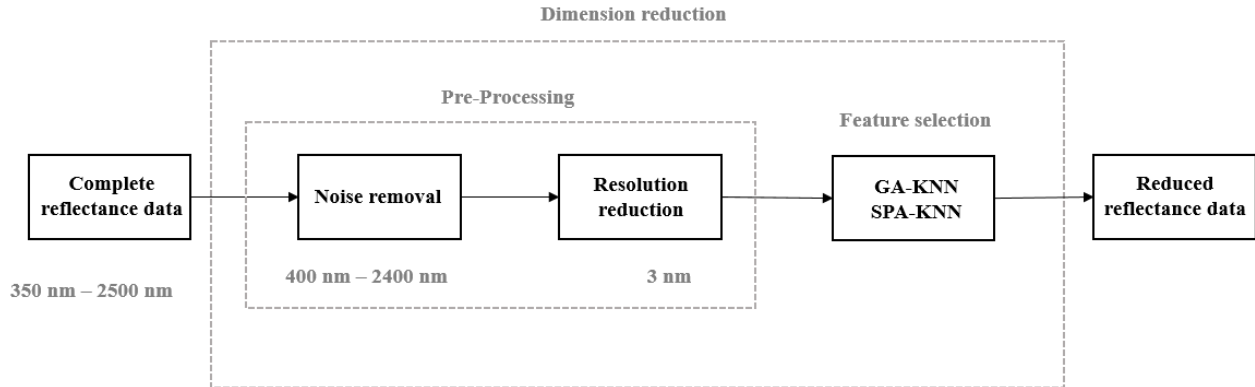
Figure 4.4 Classification of high-dimensional data by designing new SDIs, a feature selection technique.

4.3.2.1 *SPECTRAL NOISE REMOVAL*

After the acquisition of the reflectance signatures of leaves from the field, we obtained a set of healthy and diseased observations ranging from 350 to 2500 nm, with a total of 2151 features or wavelengths per sample. Since the reflectance data were noisy at the extremities and just like other similar studies, we only considered values between 400 nm and 2400 nm, giving a total of 2000 wavelengths.

4.3.2.2 *SPECTRAL RESOLUTION REDUCTION*

Following this, the reflectance resolution was reduced by a factor of three due to high correlation between adjacent wavelengths. This factor was a compromise between scaling down data redundancy while not eliminating important information. This is logical since the true spectral resolution of the spectro-radiometer is about 3 nm. In consequence, only 665 wavelengths were considered for the rest of the analysis (*Figure 4.5*).
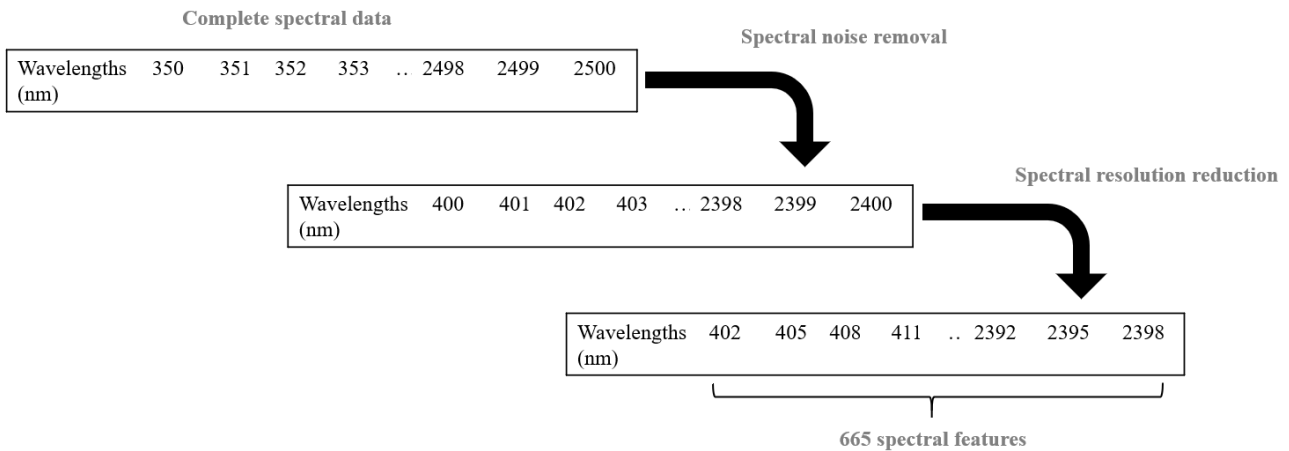
Figure 4.5 Noise removal and resolution reduction of reflectance data.

4.3.2.3 *FEATURE SELECTION*

SUCCESSIVE PROJECTIONS ALGORITHM-KNN

The objective of the reflectance analysis presented here is to select a subset of wavelengths with the most useful reflectance information; features were included or excluded based on good discrimination and classification efficiency. Once the optimum features are selected, a proper model can be designed using them instead of the whole reflectance data.

In this section, we tested the SPA algorithm. Although the algorithm itself produces chains of independant variables in an unsupervised way, we employed a supervised method for selecting the best chain among them, which means that the a priori knowledge of the spectral response (healthy, diseased) was eventually involved in selecting the features.

The SPA algorithm is introduced in [97]. The algorithm has been successfully employed in several analytical applications over the past 10 years; it often leads to better results compared to the use of full spectra. The technique was also investigated in many other studies such as [98] or [99].

Three main steps are needed for SPA. Firstly, the reflectance data are used to design a set of chains of variables according to a sequence of operations engaging vector projection. These operations are designed to minimize collinearity among the variables of the same chain. Secondly, for each candidate chain holding a subset of variables, a classification model is created. Finally, each model is evaluated and the optimal subset is chosen according to the best performance reached.

Next, matrices and linear operators are represented in bold capital letters, column vectors by bold lower-case letters and scalars by italic characters. Elements of a sequence are denoted by italic characters with an index between parenthesis. The superscript T means transposed.

SPA is a forward selection method: it starts with one variable then incorporates a new one at each iteration until a specified number N of variables is reached. Its purpose is to select features whose information content is minimally redundant in order to solve collinearity problems. SPA steps are described below:

Let $k(n)$ be the feature selected at the nth iteration of SPA. Let $\mathbf{X}_{cal}$ be the matrix ($M_{cal}\mathrm{x}J$) of instrumental response data, J are wavelengths and M_{cal} is the number of observations.

Assume that the first variable is $k(0)$ and the number N are given.

Before the first iteration ($n=1$), let $\mathbf{x}_j=j$th colum, of $\mathbf{X}_{cal}$, $j=1,..,J$

Let S be the set of wavelengths which have not been selected yet. That is S={j such that $1\leq j\leq J$ and j doesn't belong to $\{k(0),..,k(n-1)\}$

Calculate the projection of $\mathbf{x}_j$ on the subspace orthogonal to $\mathbf{x}_{k(n-1)}$:

$$\mathbf{Px}_j = \mathbf{x}_j - (\mathbf{x^T}_j\, \mathbf{x}_{k(n-1)})\, \mathbf{x}_{k(n-1)}\, (\mathbf{x^T}_{k(n-1)}\, \mathbf{x}_{k(n-1)})^{-1}$$

For all j belongs to S where $\mathbf{P}$ is the projection operator

       Let $k(n)$ =argmax($\|\mathbf{Pxj}\|$, j belongs to S)

       Let $\mathbf{xj} = \mathbf{Pxj}$ j belongs to S

       Let $n=n+1$ if $n<N$ go back to Step 1

End

The resulting variables are $\{k(n), n=0,..,N-1\}$

The number of projection operations performed in the selection process can be shown to be $(N-1)(J-N/2)$. An example of SPA is given next in *Figure 4.6*.
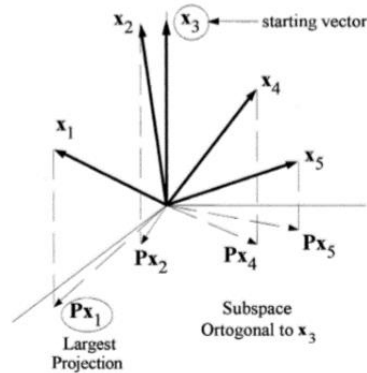
Figure 4.6 Example of SPA with $J=5$, Mcal=3 and $k(0)=3$. Result of first iteration: $k(1)=1$. (Source: [97]).

Each candidate subset of variables made of the stored chains is considered and evaluated. The essence of our contribution relies on choosing the best subset of variables. The subset is in fact chosen on the basis of its ability to discriminate healthy from diseased spectra with a KNN-cross validated classification model. A KNN non-parametric classifier was picked to avoid the burden of parameter tuning that may influence the feature selection result. The maximal accuracy was elected as a criterion for assessing the performance of a defined SPA subset; the subset displaying the higher accuracy is kept.

GENETIC ALGORITHMS-KNN

From the modified set of observations obtained, the best wavelengths were chosen by applying a Genetic Algorithm (GA) feature selection tool.

Genetic algorithms provide a valid tool for solving optimization and searching problems; it imitates the natural human evolution process [100]. GA manipulates one population to produce a new one based on some genetic operators.

The five important steps in GA [101] are: (1) chromosome encoding, (2) fitness evaluation, (3) selection mechanisms, (4) genetic operators and (5) criteria to stop the GA (*Figure 4.7*).

Human genetics vocabulary is often used in GA; chromosomes are the bit strings (individuals that form the population), gene is the feature [102]. In our study, a binary space is assumed: a gene value "1" indicates that the feature indexed by the "1" is chosen. Contrarily, (i.e., if it is 0), the feature is not chosen for evaluation. At the beginning, a matrix of dimension (Population size (300 samples) x Number of wavelengths (633 spectral features)) containing random binary digits is created, which forms the initial population.

```
                    ┌─────────────┐
              1     │ Chromosome  │
                    │  encoding   │
                    └──────┬──────┘
                           │
                    ┌──────┴──────┐
                    │   Initial   │
                    │   binary    │
                    │ population  │
                    └──────┬──────┘
                           │
                    ┌──────┴──────┐
              2     │   Fitness   │
                    │ Evaluation  │
                    └──────┬──────┘
                           │
                    ┌──────┴──────┐
              3     │  Selection  │◄────────────────┐
                    └──────┬──────┘                 │
                           │                        │
                    ┌──────┴──────┐                 │
                    │   Genetic   │                 │
              4     │  Operators  │                 │
                    └──────┬──────┘                 │
                           │                     No │
                    ┌──────┴──────┐                 │
                    │     New     │                 │
                    │   binary    │                 │
                    │ generation  │                 │
                    └──────┬──────┘                 │
                           │            5           │
                    ┌──────┴──────┐   ╱─────────╲   ┌──────────┐
                    │   Fitness   │──▶ Termination   │   Best   │
                    │ Evaluation  │  ╲ Condition ╱──▶│individual│
                    └─────────────┘   ╲─────────╱ Yes└──────────┘
```
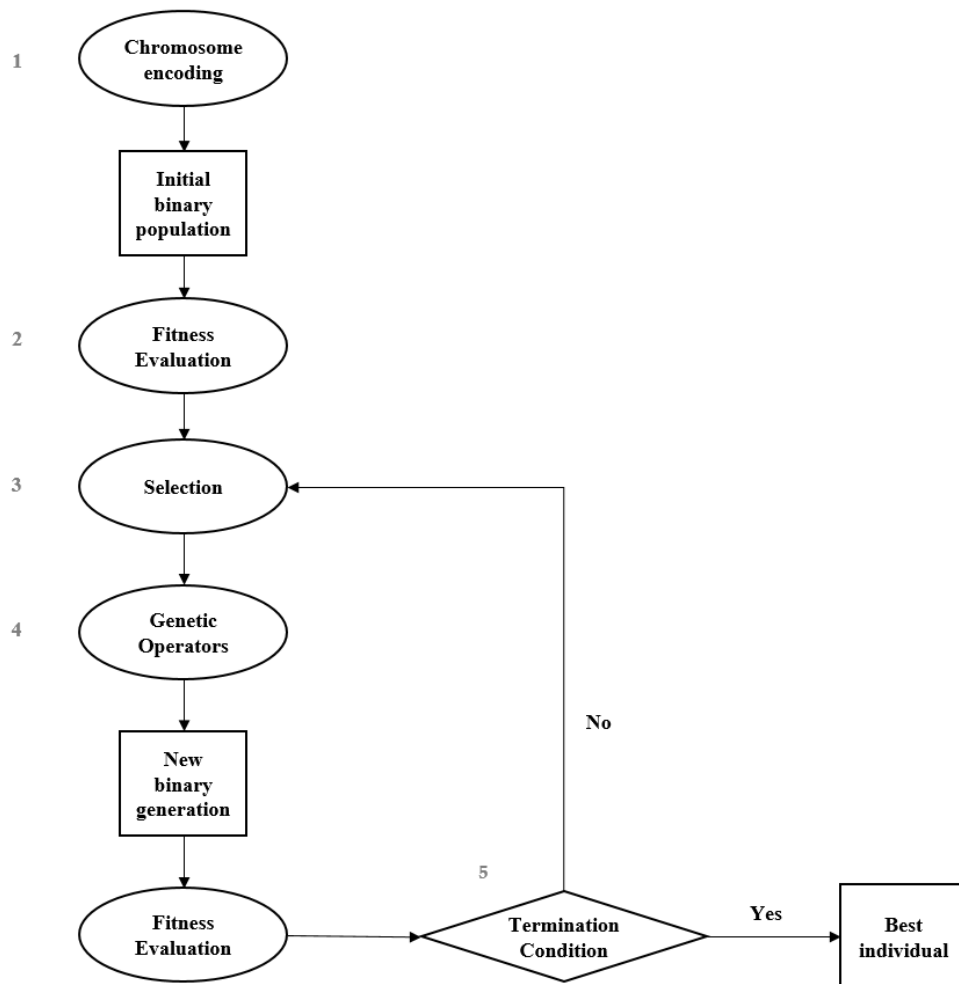
Figure 4.7 GA-Based feature selection.

A fitness function evaluates the discriminative capacity of the population, made of chromosomes, each selecting a subset of features. The GA-based technique was designed in a supervised way, which means that the spectral response (*a priori* knowledge of the diseased and healthy samples) affects the choice of features. We chose the loss obtained by a cross-validated K-Nearest-Neighborhood (KNN) classification model as a fitness function that guides feature selection. This seemed very logical because we wanted to choose reflectance features based on their discrimination potential. A KNN non-parametric classifier was picked to avoid the burden of parameter tuning that may influence the feature selection result.

Individuals are ranked, based on the values reported by the fitness function; then, the Elite kids with the best fitness values are selected to survive and are, hence, transferred to the next generation.

The selection operation provides individuals for genetic cross-over and mutation; it ensures that the population is being constantly improved. Tournament Selection was used here due to its simplicity, speed and efficiency. For tournament selection, two functions are needed; the first one generates the

parents and the second one produces the chromosome that succeeded in the tournament. The winner is defined by calculating the fitness of all selected chromosomes and choosing the best one. As an example, in tournament selection of size 2, two individuals are selected from the current population, excluding the Elite kids, then the best of the two chromosomes, using fitness ranking, is chosen. Tournament selection is done iteratively until the new generation is created.

Cross-over consists on combining two parent individuals to form children in the new generation. Two parent chromosomes are required to accomplish a crossover operation. In our case, the two chromosomes are captured from tournament selection. XOR operation is performed in this case since parent chromosomes are binary [103]. The number of new children produced due to the cross-over operator, is defined based on the cross-over fraction.

Mutation is another genetic operator and induces a perturbation of chromosomes by applying a bit flipping procedure depending on the mutation probability. Mutation ensures genetic diversity, eliminating premature convergence. A uniform mutation is applied in this study and a set of 633 random numbers following a uniform distribution is created. Each number is compared with the mutation probability: if the number is bigger than the probability, the bit at the position of the number remains; if not, the bit is flipped. The number of new children produced by the mutation operator is defined by subtracting the population size from the number of elite children and the number of children obtained by cross-over.

Each new generation, formed by GA, contains individuals from Elite kids, crossover kids and mutation kids [104]. The new population is evaluated again and the GA continues to evolve until the stopping condition is met. Two stopping conditions are applied in this study: Maximum Number of Generations and Stall Generation Limit. The choice of the number of generations is essential because GA can stop too soon if the number is not correctly set. GA terminates if the average changes in the fitness values among the chromosomes over Stall Generation Limit generations is less than or equal to tolerance function. The goal is to insure genetic homogeneity All the GA parameters used in our study are described in *Table 4.2*. Population size in evolutionary algorithms needs to be large enough to initialize with a rich set of solutions. We need to modulate the minimum size to cope with premature convergence and other problems; here a population size of 300 was selected.

Table 4.2 Numerical values of different GA parameters considered in this study.

GA parameter	Value
Population size (Number of Chromosomes)	300
Genome length (Number of genes/features)	300
Population type	Bit strings
Fitness Function	SVM-Based Classification Error
Number of generations	300
Stall generation limit	50
Crossover	Arithmetic
Crossover Probability	0.8
Mutation	Uniform Mutation

Mutation Probability	0.2
Selection scheme	Tournament of size 2
Elite Count	2

Figure 4.8 displays an example of the fitness values with respect to the generation's number on a red-grapevine variety using the GA selection tool.
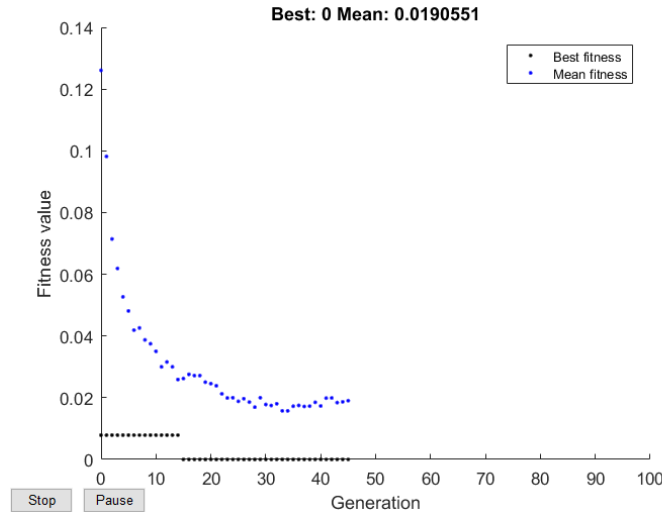
Figure 4.8 Simulation of the GA on Marselan data.

When GA terminates, one individual is chosen providing the convergence. This individual contains the optimal features, which is a binary set with "1" meaning that the feature at this specific index is considered.

Since the initial population is randomly created in GA, the number of selected features by the GA tool cannot be predicted and is function of the data. In fact, the GA keeps evolving until convergence and the number of features isn't known in advance.

4.3.3 PRACTICAL DEFINITION OF OPTIMAL REFLECTANCE FEATURES FOR MS SENSOR BY SPECTRAL DISEASE INDEX (SDI) DESIGN

In theory, the reflectance features selected by GA-KNN and SPA-KNN are enough to detect healthy from infected reflectance data. The number of selected features can't be chosen in advance (at least not for GA). We wanted to check if with the minimum number of bands required to have a MS imaging system (four bands), we can classify with good precision.

In practice, a MS sensor can be made of 4-10 bands but under price and weight constraints, an exact number might be imposed. The question here is what top bands should be selected among the ones chosen by GA-KNN and SPA-KNN. To solve this, we propose to design special spectral indices that are disease-dedicated. SDIs, unlike general SVIs, are designed to identify a specific damage in plants. They couple information from parts of the electromagnetic spectra which are characteristic of an infection and not visible to the human eye. Therefore, they may have the potential to automate the

65

disease detection procedure, not only replacing to some extent the pathologist but also predicting the presence of an infection before its symptoms become obvious.

Special spectral disease indices sensitive to the FD, Esca and BN diseases were developed next. *Figure 4.9* shows the detailed approach that was adapted to compute and to evaluate SDIs. The two acted as a tool for reducing the complexity and the number of our data and aided in the forming of new SDIs with respect to the common SVIs, which are frequently used in remote sensing applied to agricultural fields. The new candidate SDIs are injected each as an input to a classification model.
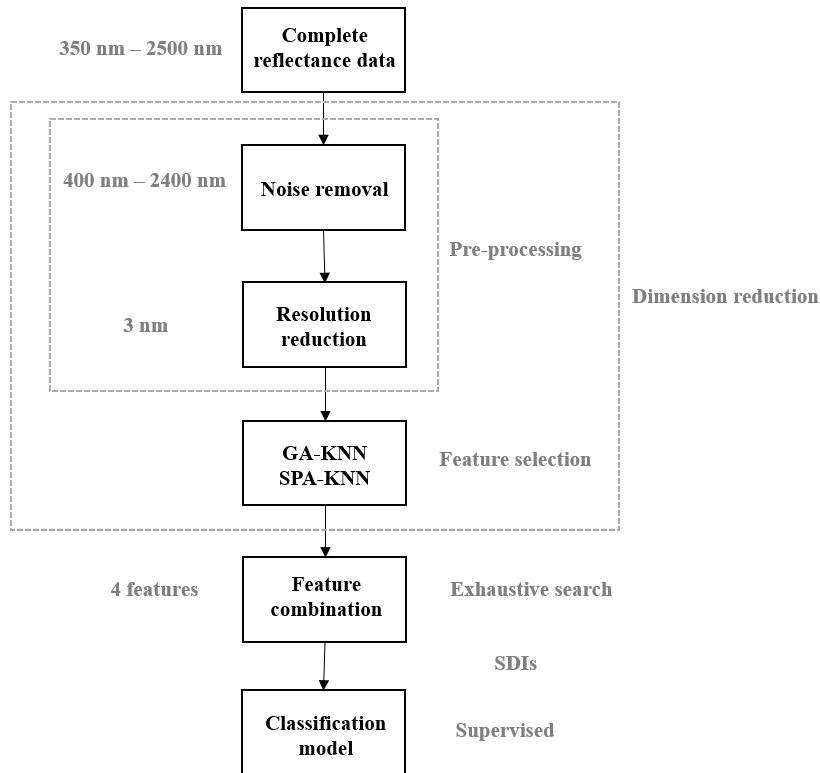
Figure 4.9 Detailed approach proposed for the development of SDIs from hyperspectral reflectance data.

4.3.3.1 *FEATURE COMBINATION*

GA-KNN and SPA-KNN provide a chain of optimal variables. For each technique, we can combine resulting features in order to design the best SDI.

In the context of DAMAV project, only four cameras could be supported to form the MS sensor. Filters placed on the four cameras are, on the other hand, interchangeable depending on the application.

The indices developed aim at detecting a specific plant disease. Next, we will detail the case of an SDI holding four spectral bands. Our choice of the form of the SDI was inspired from [105]; authors used a single weighted reflectance added to a normalized reflectance difference. We, instead of using a single reflectance, adopted a ratio of reflectances. In fact, the use of a single reflectance poses a problem because a single reflectance taken on the raw spectrum is subject to uncontrolled variations. Thus, a combination of a ratio of two reflectances and a normalized reflectance difference seemed suitable. A weighting factor for the ratio of reflectances was determined and the possible weights were: −1, −0.5, 0.5 and 1. An exhaustive search of the best SDI is undertaken; combinations of a

weighted ratio of two reflectances and a normalized difference of two reflectances are tested. Each combination of four wavelengths and a weighting factor forms an index. The SDI design is shown in the next Equation *(4.1)*:

$$SDI = \frac{R_{\lambda_2} - R_{\lambda_1}}{R_{\lambda_2} + R_{\lambda_1}} + w \frac{R_{\lambda_3}}{R_{\lambda_4}} \quad (4.1)$$

where R_{λ_1}, R_{λ_2}, R_{λ_3} and R_{λ_4} are reflectances at optimal wavelengths chosen from the pool of variables selected by GA-KNN or SPA-KNN ($\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_4$) and w is the weighting factor.

The set of indices defined are assessed for their classification ability using a classification model and the configuration providing the best classification accuracy is retained. This configuration corresponds to the best SDI.

λ_1 , λ_2 , λ_3 , λ_4 could be implemented later as filters on the MS sensor equipped with four cameras. For example, among the Narrow-band filters from the MidOpt catalogue, we can choose the ones having the central wavelengths that correspond to or is the closest to λ_1 , λ_2 , λ_3 , λ_4. The computation of the SDI could be calculated later on, using the corrected reflectance values gathered form the four cameras.

The computational load is higher when the number of wavelengths that have been selected by GA-KNN or SPA-KNN is higher because the exhaustive search will look for the best combination of four reflectances and weights among the pool of available reflectances. Hence, the number of combinations to be tested varies depending on the number of features given for each case.

4.3.3.2 *CLASSIFICATION MODEL*

There are hundreds of classifiers in the literature (support vector machines, discriminant analysis, neural networks, classification trees, naïve bayes, adaboost ...) and it is often difficult for researchers to choose an appropriate classifier for a specific application. The easiest approach that is used to address this issue is to try empirically several classifiers and select the one having the highest accuracy.

The Support Vector Machines is a supervised machine learning algorithm, mostly used to solve classification problems. It consists on defining a boundary (line/hyperplane) that best separates two classes [106]. The closest points to the boundary are called support vectors; the margin is the perpendicular distance calculated from the boundary to the support vectors. A maximal-margin classifier defines a hyperplane separating two classes and having the largest margin. However, a soft-margin classifier allows points to lie between the margins or on the wrong side of the plane. It is usually used when classes are not fully separable. A cost parameter, C, defining the trade-off for misclassifying objects, is used and adjusts the importance of the separation error in the creation of the separation surface.

In practice, SVM are implemented using kernels. When applying non-linear Kernels (polynomial or radial), non-linear boundaries are created and the accuracy improves. Due to its flexibility, the Radial Basis Function (RBF) kernel is however the most used, so we decided to employ it in our study (Equation *(4.2)*).

$$K(x_i, x_j) = exp(-\gamma \|x_i - x_j\|^2) \quad (4.2)$$

where x_i and x_j are objects i and j, respectively. Υ defines the width of the kernel function; in other words, it is a shape parameter determining the smoothness of the boundary between the groups in the original object space. This kernel is recommended as a first choice when selecting SVM kernel, due to its flexibility.

One of the most known methods for fitting SVM is the sequential minimal optimization method. The concept and the applications of SVM are discussed in detail in [107].

Each feature variable was first scaled to mean "zero" and standard deviation "one". The shape parameter Υ and the cost parameter C were found using a grid search. The performance of the classification for a given value (γ, C) is evaluated by computing the accuracy across all subjects.

In [108], hyperspectral data were recorded from healthy leaves and leaves inoculated with different pathogens causing Cercospora leaf spot, sugar beet rust and powdery mildew. Nine spectral vegetation indices, were used as features for an automatic classification. Early differentiation between healthy and inoculated plants as well as among specific diseases was achieved by a Support Vector Machine with a radial basis function as kernel. The discrimination between healthy sugar beet leaves and diseased leaves resulted in classification accuracies up to 97%. In [109], authors presented a non-destructive method to discriminate P. aeruginosa strains from simple NIR spectroscopy measurements aided by variable selection methods. Variable selection techniques were used to select an appropriate subset of features and are given to LDA for a classification task. The variables selected were then used for discriminating the strains.

Inspired by the two previous works, we employed the SVM model to identify the potential of the combination of the variables selected by GA-KNN and SPA-KNN (or more generally SDIs) in discriminating infected from healthy leaves. In particular, a SVM-based cross-validated classification model was used in this work; it takes N training samples, trains the classifier on N-1 samples, then uses the remaining one sample to test. This procedure is repeated until all N samples have been used as the test sample.

4.4 CONCLUSION

A statistical analysis of reflectance data revealed the weakness point of our sensing technique; in fact, it can detect the presence of a disease but may be unable to identify it or grade its severity stage. This conclusion was in accordance with the results obtained by biologists that tried to define the effect of diseases on leaf contents. In fact, the laboratory experiments showed that modifications identified at a leaf-level were very similar for the three considered diseases. At this point, we have an idea about potential of reflectance data. To reach more than a simple disease detection task, other features than reflectance seem necessary. A texture analysis will be presented in the next Chapter.

As we mentioned before, each reflectance measurement acquired, holds a large number of values on contiguous narrow bands. Analyzing such high dimensional data is a complex and time-consuming task; therefore, reducing the dimensionality of the data seems crucial. SVIs were introduced in this Chapter as an efficient traditional way of decreasing data dimensionality. It can be applied for the purpose of variable selection. SVIs are widely used to detect leaf contents (pigments, water ...); nevertheless, these standard indices are not disease-specific.

The difference between reflectance signatures of healthy and diseased plants can be the key to identify efficient variables correlated with a specific disease. Two new dimension reduction techniques that

are disease-dedicated were suggested. A data analysis procedure to design specific grapevine disease indices, that challenge common vegetation indices for a MS imaging application, is exposed. Two feature selection tools: the GA-KNN and SPA-KNN based approaches, when compared to using complete hyperspectral data, gave better results. This is somewhat predictable since the collinearity in the full spectral data will highly degrade results.

The SPA algorithm uses straightforward operations in a vector space to obtain subsets of variables having reduced collinearity. The overall aim is to remove redundant and unnecessary variables in order to solve collinearity problems. SPA doesn't modify the original data vectors, in fact, projections are adopted only for selection purposes thus, the relation between spectral variables and data vectors is maintained. GA is one of the most advanced techniques used in the field of predictive analysis, it is computationally expensive but it performs better than common selection techniques and has the advantage of manipulating large data sets.

Reflectance features on specific wavelengths obtained with SPA-KNN and GA-KNN will be compared in terms of accuracy in CHAPTER 6. After being selected by SPA-KNN or GA-KNN, reflectance features are then combined to design the SDIs. SDIs coming from both methods will also be tested in CHAPTER 6. Depending on the disease, on one hand and on the grapevine variety on the other hand, a combination of ratio and normalized reflectances is required each time to correctly detect the disease. For each disease and after a specific learning task, the selection algorithm coupled with a classifier could define a certain combination of reflectances at sub-bands that seem interesting for detecting a specific infection.

CHAPTER 5

5 TEXTURAL ANALYSIS

5.1 INTRODUCTION

Numerous non-destructive sensing technologies have been implemented to assess the physiological status of vegetation and to respond to the need for the automatic identification of diseases. As previously indicated, when an infection occurs, the foliar concentration of pigments and biochemical components is altered ([110], [111]). As a result, the infected leaf reflectance in the visible and near-infrared regions of the electromagnetic spectrum is no longer perfect and similar to a healthy reflectance. This assumption was confirmed in this thesis; indeed, we found that the reflectance data can be considered as a tool for damage detection in grapevines. However, reflectance data alone might not be capable of identifying the infection that's why other features need to be tested.

The identification of the symptoms of plant diseases by means of an automatic system may provide solutions for farmers in their struggle against disease outbreaks. In the last Chapter, we proposed some techniques for determining a set of MS bands. However, the tests in the context of this thesis are done at a leaf-level scale. When going from the leaf scale to the field scale with a UAV, the sources of noise will multiply. The reflectance data obtained require corrections and might not be enough, that's why other parameters are needed to enforce the analysis, such as color, texture or shape of the disease.

The color modification is an interesting factor to detect a disease but many diseases can cause the same discoloration on leaves. Disease identification based on color only, seems insufficient. How the disease spreads spatially and its shape are on the other hand interesting factors to identify it. However, this would mean that the symptoms are already present. Contrarily, the texture may change even before the onset of symptoms, which may be an advantage for early detection. In addition to this, depending on the application, texture features can be more advantageous than shape features for detecting diseases. In [112], diseased regions shown in digital pictures of cotton crops were enhanced, segmented, and a set of features were extracted from each of them. Features were then used as inputs to a SVM classifier and tests were performed to identify the best classification model. With this procedure we found that features composing the group texture achieved the highest classification rate (83%) while features composing the group shape achieved only (55%).

In general, texture from colored images provide a valuable mean for assessing grapevine contamination. Employing texture features might not only optimize the infection detection but could also enable disease discrimination. Actually, symptoms of diseases appear differently, for example Esca disease induces stripes while FD and BN cause uniform discoloration. Hence, by studying the distribution of gray levels on infected leaves, we may be able to differentiate among grapevine infestations.

A statistical investigation of the texture data calculated from RGB images was done using a simple Principal Component Analysis (PCA) but wasn't shown here. This analysis should be later tested on texture data extracted from MS images. The overall aim was to check the potential of texture data in detecting FD, BN and Esca diseases, grading their severity stage and discriminating among them. PCA was adopted and intra-class and inter-class distances on the obtained clusters were calculated.

We concluded from the analysis that texture data could be employed with confidence to detect diseases, to distinguish among them and to define their infestation level.

In the scientific literature, extensive studies investigated the potential of texture features in disease detection and found promising results. In [113], three diseases were investigated: rice bacterial leaf blight, rice sheath blight and rice blast. After segmentation of the infested spots, shape and texture descriptors were defined. The SVM classifier detected and classified these disease spots with an accuracy of 97.2%. Many authors studied the potential of image processing in classifying plant diseases. A review of the image analysis techniques used to identify and classify fungal disease symptoms can be found in [114]. In that study, four different classes of agriculture and horticulture crops: fruit crops, vegetable crops, cereal crops and commercial crops affected by fungal disease were studied. Algorithms for feature extraction and classification based on image processing techniques were also exposed. Similarly, the paper [115] stated various methods used to study leaf disease detection. Finally, the article [116], discussed various techniques to extract the affected part of a plant along with some feature extraction and classification techniques.

The idea of using both reflectance and image data for the sake of classifying crop infections was already inspected in other studies. For example, texture parameters were extracted from hyperspectral images and both spectrum and texture for early blight disease detection on eggplants were investigated [117]. Four effective wavelengths were defined and then the corresponding gray-level images were further processed. Authors extracted eight texture features based on gray level co-occurrence from these gray images. K-Nearest Neighbor (KNN) and AdaBoost classifiers were used for building models that detect healthy and infected samples. Reported results from this study were very good (over 88.46% precision). Another work of Xie et al. [118] employed texture features from hyperspectral images to detect different tomato leaves infections. Five sensitive wavelengths were chosen and the corresponding MS images are extracted. Eight texture features based on gray level co-occurrence matrix are computed from these MS images. Once the features are ready, extreme learning machine (ELM) classifier models were established for reflectance data and for textural data. Spectral detection models yielded an excellent overall classification accuracy of 97% to 100%, textural detection models also carried effective information and scored good results (72%).

In this Chapter, we will test some traditional texture features and try an optimal combination for the sake of grapevine disease identification when symptoms start to become visible. The number of texture parameters calculated on each colored band of the sample leaf images, might rapidly become extensive; hence a feature selection step seems essential (ReliefF). Each reduced set of texture characteristics is then injected as an input to a classification model. Just like the reflectance data analysis, we had the classes of the samples (infected and what infection in particular, healthy), so a supervised classification model is employed also for texture data analysis. The general flow of the procedure is shown in *Figure 5.1*.
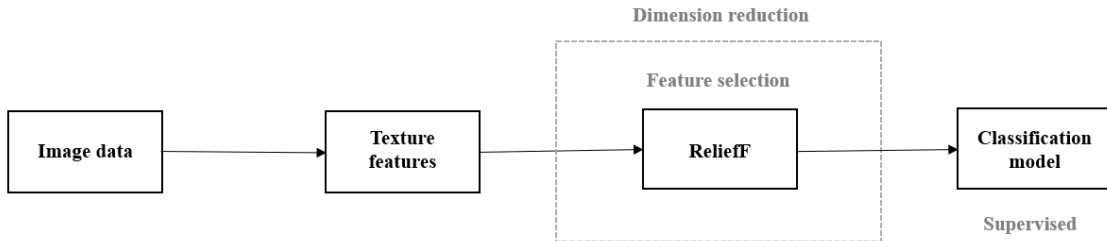
Figure 5.1 Classification of textural data by using ReliefF, a feature selection technique.

5.2 PROPOSED TEXTURE ANALYSIS

RGB leaf images are converted to the HSV space then segmented. Pixels corresponding to the leaf are kept and afterwards texture features are extracted from the Lab space.

Figure 5.2 presents the different steps of the pre-processing strategy. The HSV color space is used for leaf segmentation. Then, texture features, that will be presented next, are extracted from each band of the Lab color space.
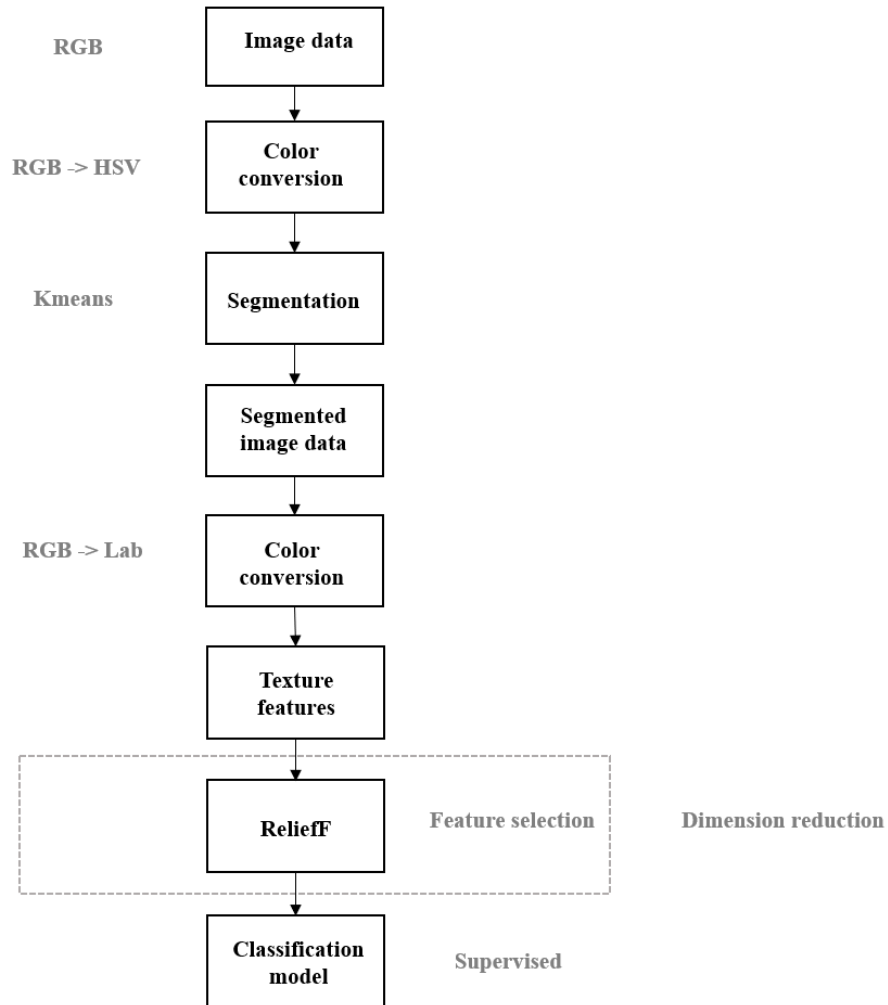
Figure 5.2 Detailed approach proposed for the classification of textural features extracted from RGB images.

5.2.1 RGB TO HSV CONVERSION

First, RGB images are converted to Hue Saturation Value (HSV) space through a non-linear transformation (Equations *(5.1), (5.2), (5.3), (5.4)*). Note that these equations are function of the illuminant. This color space was tested since it is a good descriptor that is close to human perception. According to reference [119], RGB is ideal for color generation but HSV model is ideal for color description. Hue describes the actual color as perceived, saturation is the purity of a color and value means the lightness or darkness of a color.

$$v = max(r, g, b) \qquad (5.1)$$

$$s = \begin{cases} (v - min(r, g, b) * 255)/v & v \neq 0 \\ 0 & v = 0 \end{cases} \qquad (5.2)$$

$$h = \begin{cases} (g - b) * \frac{60}{s} & v = r \\ 180 + (b - r) * \frac{60}{s} & v = g \\ 240 + (r - g) * \frac{60}{s} & v = b \end{cases} \qquad (5.3)$$

$$if \ h < 0 \quad then \quad h = h + 3 = 60 \qquad (5.4)$$

Where r, g and b represent pixel values of red, green, blue (RGB) color of image respectively and h, s and v represent the new color channels in the HSV color space.

5.2.2 LEAF SEGMENTATION

After HSV color space conversion, a background subtraction was performed. The goal is to divide the image into 2 classes (background and leaf). The leaf was segmented from each HSV converted image using the K-means algorithm. K-means is an iterative clustering algorithm that is well-used in many applications. It aims to partition a given set of patterns into different groups. The patterns belonging to the same cluster are similar and are distinguishable from those belonging to another cluster. Many applications studying the image processing potential in classifying diseases in the agricultural field, adopted the K-means algorithm. The work [120] explained that K-means finds a relevant partition in which elements in each cluster are close to each other, and at the same time far from objects in other clusters. Each cluster in the partition is defined by its member objects and by its centroid; K-means minimize the sum of distances from each element to its cluster centroid, over all clusters that have been created. This algorithm moves objects in between clusters until the sum cannot be reduced any further. The result is a set of clusters that are as compacted and are well-separated as far as possible. The minimization can be modified using possible input parameters to K-means algorithm; this involves the initial values of centroid of each cluster and the allowed number of iterations. By default, K-means use the cluster center initialization and the squared Euclidean distance metric to determine distances.

K-means segment the images into two distinct regions, then the background pixels are set to zero and only the pixels relative to the leaf are kept and used in the next steps of the analysis.

5.2.3 RGB TO LAB CONVERSION

The HSV color conversion is applied again on the segmented RGB image using the above-mentioned equations. Another Lab color space conversion is also performed, the D65, CIE standard illuminant was used for this purpose. The Lab system [121] is defined by the International Commission on

Illumination (CIE) to describe colors in agreement with the human vision. To transform an image from the RGB to the Lab color spaces, the RGB image must be converted into CIEXYZ space first using the following (Equations *(5.5), (5.6), (5.7)*) [122]:

$$X = 0.4124 * r + 0.3576 * g + 0.1805 * b \quad (5.5)$$

$$Y = 0.2126 * r + 0.7152 * g + 0.722 * b \quad (5.6)$$

$$Z = 0.0193 * r + 0.1192 * g + 0.9505 * b \quad (5.7)$$

Brightness and color information of Lab color model are independent one another. In Lab color model, 'L' describes color brightness; 'a' describes the color ranging from green to red; 'b' describes the color ranging from blue to yellow. Conversion Formulas for Lab color model are shown in Equations(*(5.8), (5.9), (5.10)*) [123].

$$L = 0.2126 * r + 0.7152 * g + 0.0722 * b \quad (5.8)$$

$$a = 1.4749 * (0.2213 * r - 0.3390 * g + 0.1177 * b) + 128 \quad (5.9)$$

$$b = 0.6245 * (0.1949 * r + 0.6057 * g - 0.8006 * b) + 128 \quad (5.10)$$

The Lab color space is represented as a cube. The L axis refers to luminosity and runs from top to bottom with a maximum of 100 (representing white) and a minimum of 0 (representing black). The 'a' and 'b' axis refer to chromaticity indicating where the color falls along the red-green axis and blue-yellow axis, respectively. These axes have no specific numerical limits.

The Lab is expected to provide a nearly uniform color space, in which the differences between two points in the color space resemble to visual differences between two sketched colors. This color space was adopted in [124]. Authors in [125] also used the Lab color space but chose only the band b as input for their segmentation.

5.2.4 TEXTURE CALCULATION

Extensive research has been conducted to explore various methods for automatic detection of plant diseases. The measurements of texture from digital images can be used as a useful tool for disease outbreaks diagnosis. Many texture parameters are extracted from each band of the Lab sample leaf images. These parameters are traditional ones found in the scientific literature. The intuition behind testing the following texture features is that they performed well in the task of disease detection in the medical field.

5.2.4.1 *HISTOGRAM MOMENTS*

The first step in our texture analysis was to examine the first-order or low-level statistics describing the image gray-level distribution (histogram). By direct use of histogram calculated from the image pixel gray-level values, four first-order statistics were evaluated: mean (Equation *(5.11)*) , standard deviation (Equation *(5.12)*), skewness (Equation *(5.13)*), and kurtosis (Equation *(5.14)*) [126], leading to 12 features (4 features x 3 bands) for each image in a defined color space.

In [127], there are details about the histogram moments. A histogram presents the gray-level intensities on the x-axis and the frequency of occurrence of any particular gray-level in the y-axis. It is an estimate of the probabilistic density function, or, in practice, the discrete probabilities p(i) of grey levels i, with i =0,…, L-1 (usually L=256), in an image.

The statistical average gray level:

$$\mu_1 = \sum_{i=0}^{L-1} i\, p(i) \quad (5.11)$$

The second order moment is the variance:

$$\mu_2 = \sum_{i=0}^{L-1} i - \mu_1)^2 p(i) \quad (5.12)$$

Skewness (moment order 3 μ_3):

$$\mu_3 = \sum_{i=0}^{L-1}(i - \mu_1)^3 p(i) \quad (5.13)$$

Kurtosis:

$$\kappa = \frac{\mu_4}{\mu_2{}^2} \quad\quad (5.14)$$

A study conducted in [128] examined the efficiency of applying histogram features for characterizing coronary plaques in intravascular ultrasound images.

5.2.4.2 *STATISTICS OF GRYAY-LEVEL CO-CCURRENCE MATRIX (HARALICK PARAMETERS)*

Grey Level Co-occurrence Matrix (GLCM) is one of the most frequently used approaches for texture analysis. It can describe essential characteristics of an image related to second order statistics introduced by Haralick [129].

Many scientific studies applied texture parameters from GLCMs for disease classification of various crop applications. The authors in [130] investigated the use of texture features extracted from hue, saturation and intensity color space for the classification of citrus leaves under laboratory conditions. The research conducted in [131] studied diseases from ten species (banana, beans, jackfruit, lemon, mango, potato, tomato, and sapota) and the algorithm could classify diseases with an accuracy of 94%. Focusing more on the grapevine crops, authors in [132], studied downy mildew and powdery mildew in their experiments. Disease diagnosis was possible by applying image processing and artificial intelligence techniques.

GLCMs can be computed by defining the neighbor relationship between pixels, in fact, the distance step length d and the direction θ should be determined to create the co-occurrence matrix G(i, j | d, θ). The element (i,j) of the matrix G is the frequency of appearance of grey-tone j near the reference grey-tone i within the image at a defined distance and direction. The GLCM is a square matrix where the number of rows and columns equals the number of gray levels considered.

In this study, symmetrical GLCMs are constructed for each single channel image and calculated for (θ = 0°, θ = 45°, θ = 90°, θ = 135°) with an offset of 1 pixel.

A set of features derived from four directional symmetrical GLCMs are considered for texture characterization. Based on the co-occurrence matrix, we can extract texture features called second order statistical features.

Haralick proposed 14 different features. Although some of them are redundant, they are entirely used in our analysis, but later on, only some of them will be selected. Amongst all 14 features, contrast, correlation, energy, and homogeneity, are probably the most important ones and they are detailed next ([133], [134]).

The probability of appearance of a grey-tone i near the reference grey-tone j within the image at a defined distance and direction is $p(i,j)$.

Contrast (K) (Equation (5.15)) measures local variations in gray level from a pixel to its neighbor in an image, it shows the texture fineness.

$$K = \sum_{i,j}(i-j)^2 p(i,j) \quad (5.15)$$

Correlation (R) (Equation (5.16)) measures the linear dependence of gray-levels in the co-occurrence matrix or in other words, the correlation intensity between neighboring pixels.

$$R = \sum_{i,j}\frac{(i-m_i)(j-m_j)p(i,j)}{\sigma_i\sigma_j} \quad (5.16)$$

m_i and m_j are the average of row i and column j in the GLCM, respectively. σi and σj are the standard deviations of row i and column j in the GLCM, respectively.

Energy (E) (Equation (5.17)) known as the angular second moment, it's simply the sum of squared elements in the GLCM. It measures uniformity of an image.

$$E = \sum_{i,j}p(i,j)^2 \quad (5.17)$$

Homogeneity (H) (Equation (5.18)) is a measure of the closeness of the distribution of elements in the GLCM to the diagonal. Homogeneity is unity for a diagonal GLCM. This is the case where all pixels in the original image have the same value as their neighbor.

$$H = \sum_{i,j}\frac{p(i,j)}{1+|i-j|} \quad (5.18)$$

5.2.4.3 *STATISTICS OF GRAY-LEVEL-RUN-MATRIX*

In coarse texture, long gray-level runs would often appear, however, in fine texture short runs are mainly present. Based on this observation, Galloway proposed the use of a run-length matrix for texture feature derivation [135]. For a given image, a run-length matrix p(i,j) is defined as the number of runs with pixels of gray level i and run length j. Various texture features can then be derived from this run-length matrix [136]. From an original run-length matrix p(i,j), the first five original features of run-length statistics were derived in [135]. Then work [137] suggested two new features, to extract gray-level knowledge in the matrix. Afterwards, article [138] described other four features adopting the idea of joint statistical measure of grey-level and run-length, their equations are given next.

Short Run Low Gray-Level Emphasis (SLRGE) (Equation (5.19)) is the distribution of the short homogeneous runs with low grey-levels.

$$SRLGE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j)}{i^2 j^2} \quad (5.19)$$

Short Run High Gray-Level Emphasis (SRHGE) (Equation (5.20)) is the distribution of the short homogeneous runs with high grey-levels.

$$SRHGE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j).i^2}{j^2} \quad (5.20)$$

Long Run Low Gray-Level Emphasis (LRLGE) (Equation (5.21)) is the distribution of the long homogeneous runs with low grey-levels.

$$LRLGE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j).j^2}{i^2} \qquad (5.21)$$

Long Run High Gray-Level Emphasis (LRHGE) (Equation *(5.22)*) is the distribution of the long homogeneous runs with high grey-levels.

$$LRHGE = \frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j).i^2.j^2 \qquad (5.22)$$

With n_r corresponds to the number of homogeneous runs in the region of interest.

In [138] all eleven features were tested on the classification of a set of cell images and showed that the last four features are more efficient. These features are all based on an intuitive interpretation, in order to catch some apparent properties of the run-length distribution. The problem is that many of these features are highly correlated with each other. A classification of samples characteristic of the CT appearance of three obstructive lung diseases and normal lung tissue was tested in [139]. Sensitivity and specificity of 73.6% and 91.2% were achieved, respectively. The method gives good accuracy, which suggests that it should be included as one of the main CT feature extractors for the automated detection of obstructive lung diseases.

5.2.4.4 *FRACTAL DIMENSION*

Fractal dimension measurements can be used to estimate and quantify the complexity of the shape or texture of objects ([140], [141]). Fractal geometry involves various approaches to define fractional dimensions; the most familiar one is the Hausdorff's dimension. Considering an object that possesses a Euclidean dimension E, the Hausdorff's fractal dimension (Equation *(5.23)*) D_0 can be computed by the following expression:

$$D_0 = \lim_{\varepsilon \to 0}\frac{\log N(\varepsilon)}{\log \varepsilon^{-1}} \qquad (5.23)$$

Where $N(\varepsilon)$ is the counting of hyper-cubes of dimension E and length ε that fill the object.

If we consider an object represented by a binary image I_b, an approximation D for D_0 can be obtained through the box counting algorithm [142]. Without loss of generality, the algorithm for the 2D case can be described as follows. Initially, the image is divided into a grid composed of squares of size ε x ε. The next step consists in counting the number $N(\varepsilon)$ of squares of size ε x ε that contains at least one pixel of the object. By varying the value of ε x ε, it is possible to create a log $N(\varepsilon)$ vs log ε^{-1} curve. Finally, this curve is approximated by a straight line using a line fitting method (like least squares fitting). The fractal dimension D corresponds to the slope of this line.

We employed in this work, the Segmentation-based Fractal Texture Analysis, or SFTA presented in [143]. The algorithm consists on decomposing the input image into a set of binary images. For this purpose, a two-threshold binary decomposition algorithm is employed. Then, from these binary images, the fractal dimension is calculated for the segmented regions to describe the texture patterns. SFTA was further evaluated in the study for the application of content-based image retrieval and image classification. Its performance was compared to other feature extraction techniques such as the Gabor filter banks which is one of the most widely used in the field. SFTA provided a higher precision and accuracy for both applications. Additionally, Gabor was at least 3.7 times slower than SFTA.

Some studies, like in [140] investigated the use of shape and fractal analysis in classifying breast masses. Results proved a significant difference between malignant and benign tumor masses in terms of fractal dimension. Fractal dimensions also showed a similar accuracy to some shape features.

5.2.4.5 STATISTICS OF LAPLACIAN OF GAUSSIAN

The theoretical framework of the LoG technique is detailed in [144]. The Laplacian of Gaussian (LoG) filter can be viewed as the combination of a Gaussian smoothing operator with a kernel of width sigma (σ) and anisotropic filter, the Laplacian, which measures the second spatial derivative in the image. The LoG is commonly used to detect edges and blobs at various scales. For each segmented region of interest, a LoG filter was applied using σ values of 0.5 (fine texture type), 1.5 (medium texture type) and 2.5 (coarse texture type). The region was then quantified by computing the average, standard deviation, skewness, kurtosis and entropy of its LoG values.

The LoG filter is mathematically defined in [145] as $\Delta G(x,y)$ with Δ the Laplace operator and $G(x,y)$ a two-dimensional Gaussian (Equation *(5.24)*). The gaussian was as follows:

$$G(x,y) = -e^{-\frac{(x^2+y^2)}{2\sigma^2}} \qquad (5.24)$$

This gives a LoG filter described as in Equation *(5.25)*:

$$\Delta G = \nabla . \nabla G = \frac{4}{2\sigma^2}(1 - \frac{x^2+y^2}{2\sigma^2})e^{-\frac{(x^2+y^2)}{2\sigma^2}} \qquad (5.25)$$

where x and y are the variates for a 2D case.

Fine texture features calculated form the LoG filter were associated with a low 5-year overall survival rate in sick patients in [146], meaning that LoG features could be a good indicator of the absence or presence of colorectal cancer.

5.2.4.6 STATISTICS OF WAVELET

The continuous wavelet transform of a signal x(t) (Equations *(5.26), (5.27)*), is defined as $\psi(t)$ in [147]:

$$W_\psi(c,d) = \int_{-\infty}^{+\infty} x(t) * \psi_{c,d}(t)dx \qquad (5.26)$$

Where

$$\psi_{c,d}(t) = \frac{1}{\sqrt{|c|}}\psi(\frac{t-c}{d}) \qquad (5.27)$$

The wavelet $\psi_{a,b}$ is computed from mother wavelet ψ by translation and dilation; where a is the dilation factor and b is the translation parameter (both are real positive numbers). By restraining c and d to a discrete lattice ($c=2^d$) we can obtain the discrete wavelet transform. There are many wavelets which have gained popularity throughout the development of wavelet analysis. One important discrete wavelet is the Haar wavelet due to its simplicity [148]. In this work, however, we found that the Biorthogonal wavelet gave better classification accuracy, so we adopted them.

The Discrete Wavelet Transform (DWT) (Equation *(5.28)*) acts on a data vector converting it into a numerically distinct vector of the same length by using a linear transformation. It is a tool that separates data into different frequency components. By using wavelets, the function can be analyzed at various levels of resolution [149].

DWT can be expressed as:

$$DWT_{x(n)} = \begin{cases} d_{j,k} = \sum x(n) h_j^*(n - 2jk) \\ a_{j,k} = \sum x(n) g_j^*(n - 2jk) \end{cases} \quad (5.28)$$

The coefficients $d_{j,k}$ are the detail components in signal $x(n)$, whereas $a_{j,k}$ are the approximation components in the signal.

The functions $h(n)$ and $g(n)$ in the equation represent the coefficients of the high-pass and low-pass filters, respectively, parameters j and k refer to wavelet scale and translation factors.

The DWT is applied on a set of images and statistical features are extracted from the approximation and detail regions of DWT decomposed images, at the two first scales in [150]. In the case of two-level decomposition, four subbands are created, labeled LH1, HL1 and HH1 presenting detail images while the sub-band LL corresponds to the approximation image. The sub-band LL1 can be further decomposed and sampled to obtain the next level of wavelet coefficients. Parameters such as average, standard deviation, skewness, kurtosis and entropy are calculated from the approximation and detail images with a two-level decomposition case.

An experiment performed in [151] considered 242 cases of breast tumors (including benign from 161 patients and carcinomas from 82 patients). A *10*-fold cross-validation sampling was employed to evaluate the performance. For the proposed system, the sensitivity was 98.77%, the specificity was 81.37%, the positive predictive value is 72.73% and the negative predictive value is 99.24%. Experimental results showed that the model using Wavelet features performed very well for breast tumor diagnosis.

5.2.4.7 *STATISTICS OF GRADIENT FILTER*

The algorithmic approach is first to compute directional gradients, Gx (Equation *(5.29)*) and Gy (Equation *(5.30)*), with respect to the *x*-axis and *y*-axis. The Sobel filter (*Figure 5.3*) was used for this purpose.

Figure 5.3 Sobel filter masks (S_x on the left and S_y on the right).

We ran this filter over the image in order to get:

$$\frac{\partial f}{\partial x} = S_x \otimes f \quad (5.29)$$

$$\frac{\partial f}{\partial y} = S_y \otimes f \quad (5.30)$$

f is the image and the resulting image gradient: $\nabla f = [\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}]$

The gradient magnitude and direction could be then computed from their orthogonal components Gx and Gy.

The Gradient's direction is defined in Equation *(5.31)* :

$$\theta = tan^{-1}(\frac{\partial f}{\partial x} / \frac{\partial f}{\partial y}) \quad (5.31)$$

The Gradient's magnitude is defined in Equation *(5.32)* :

$$\|\nabla f\| = \sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2} \quad \textit{(5.32)}$$

The histograms corresponding to the gradient magnitude are calculated and again four statistics were evaluated: mean, standard deviation, skewness, and kurtosis, giving 12 features for each image. In order to distinguishing between different kinds of rice paper, a study was conducted in [152]. The authors calculated second order statistics on the gradient filtered multispectral images of rice paper and reported good average classification accuracy.

5.2.4.8 *COMBINATION OF TEXTURE FEATURES*

Many researchers had the idea of combining texture features. For example, in the study [112], authors assessed the following parameters: texture, shape, fractal dimension, lacunarity, histogram of frequencies, grey-level dispersion. Features composing the group texture achieved the highest classification rate (83%) and features composing the group shape achieved the lowest (55%). On the contrary, when all features were used as inputs, classification accuracy was 90%. This method did not suggest to remove groups of features, rather it advised the use of the whole set. Hence, in our work, we will test the combination of our whole set of features with and without reducing it.

5.3 FEATURE SELECTION WITH RELIEFF

ReliefF is used for random selection of instances for feature weight calculation [153]. It finds a candidate set by filtering out many unimportant features and hence reducing the computational load. This feature selection technique is able to effectively provide quality estimates of attributes in problems with dependencies between attributes. ReliefF finds the weights of predictors in the case where the response is a multiclass categorical variable. The algorithm penalizes the predictors that give different values to neighbors of the same class, and rewards predictors that give different values to neighbors of different classes. The algorithm was extensively evaluated and many enhanced versions were also proposed in different works, an example of a different implementation of the ReliefF algorithm figures in the study [154].

In the scientific literature, GA and SPA were mainly used for hyper-dimensional data that are similar to our reflectance data. For another type of data, like texture features, ReliefF and its variations were extensively used for the aim of feature selection. ReliefF was also based on the KNN classifier which somehow homogenizes our variable selection procedure in this thesis. The resulting reduced set of features is injected into a classification model.

In paper [155], the authors studied the effects of reducing the number of features and selecting the most effective subset of features in the context of content-based image classification and retrieval of objects. Legendre moments are extracted, then ReliefF algorithm selects the most relevant and non-redundant features. Later, a classifier is used to categorize the images. The experimental results show that by selecting much lower number of features with ReliefF, we can improve retrieval in terms of speed and accuracy. The study in [156] established a fully automated computer-aided diagnosis system for the classification of malignant and benign masses via breastmagnetic resonance imaging. A classifier in conjunction with ReliefF feature selection was used to merge the extracted morphological and texture features into a classification score. The ReliefF algorithm was useful in selecting an optimal subset of breast tumor features. The subset could be used to decrease feature

dimensions and weight minimum distance classifiers. In terms of computerized characterization, the Fisher and SVM methods were used separately to select morphological and texture features. It was demonstrated that the ReliefF-SVM model exhibited the best performance. Article [157] proposed a method to predict and classify cancer. Feature selection techniques are important for detecting the necessary training features. In this work, an improved method of the Relief algorithm called ReliefK was proposed and yielded satisfactory results. The feature selection was used as preprocessing step for selecting the 14 features before training them by Back propagation neural network.

Since ReliefF seemed convenient for analyzing texture data, we applied it in our work. We reduced the set of features by selecting the top texture features chosen by the ReliefF technique for further classification. In fact, ten features were chosen because this is the minimum size of features that gave high classification accuracy in most cases. Adding more features won't change the accuracy much.

5.4 CLASSIFICATION MODEL

Choosing the most appropriate classifier for an application is a tricky task because there are hundreds of classifiers in the scientific literature. In [158] the authors reviewed some of the most used classifiers for disease detection using imaging and introduced their advantages and disadvantages. In [159], some recent crop disease identification and the classifiers employed for this purpose are also discussed.

Artificial Neural Networks (ANNs) are a well-known technique that has shown successful results in many real applications, particularly in the agricultural field. For example, authors in [160] gathered data from Florida grapefruit, orange, and tangerine varieties using a color-based vision system and used multiple neural network classification processes to detect blemish-related features for the citrus fruit. Dheeb Al Bashish, et al [161] also used ANNs and employed color and texture to classify five diseases which can impact plants; they are: early scorch, cottony mold, ashen mold, late scorch, tiny whiteness. Scientists in [132] used ANNs for diagnosing and classifying diseases in grape leaves. Downy Mildew and Powdery Mildew infections are tested and the authors claimed to have a training accuracy of 100%. Researchers in [162] further used ANNs to classify the watermelon leaf diseases of Downy Mildew and Anthracnose. The true classification rate depicted was 75.9%.

While ANNs perform well in most applications, their popularity remains inferior to the support vector machines (SVM), which are computationally cheaper and provide similar or even better performance than ANNs [163]. In [164], the authors demonstrated better performance of the SVM classifier over ANNs for differentiation of benign and malignant brain tumors. SVMs were also applied for brain tumor classification in [165], [166]. Other applications of SVMs include the staging of liver fibrosis [167], detection of prostate [168], assessment of osteoarthritis [169], classification of cervical cancer [133], mammogram lesions [170], and Parkinson disease [171].

SVM for texture data classification was widely used in the agricultural field in particular and gave good performance in several applications. For detecting rice diseases early and accurately, image processing techniques and Support Vector Machine (SVM) were employed in [113]. Rice disease spots were segmented and their shape and texture features were extracted. SVM was employed to classify rice bacterial leaf blight, rice sheath blight and rice blast. The results showed that SVM could effectively detect and classify these disease spots to an accuracy of 97.2%.

Another application can be found in [131]. First a color transformation structure for the input RGB image is created; then the green pixels are masked and removed using specific threshold value followed by segmentation process, the texture statistics are computed for the useful segments; finally the extracted features are passed through the SVM classifier. The proposed algorithm's efficiency can successfully detect and classify the examined diseases with an accuracy of 94%.

We employed for our texture analysis, similarly to our reflectance analysis: a SVM cross validated model.

5.5 CONCLUSIONS

Digital image processing is advantageous because images can be reproduced, stored, and used at a later time as historical documents of vegetation status.

The sample leaf images were segmented using K-means, then they were converted to the Lab color space. In fact, the Lab color space was the best for disease detection with respect to HSV and RGB in terms of correct classification accuracy (results for HSV and RGB color spaces were not shown here). A set of texture features was then calculated for each band of the Lab colored leaf images. We investigated if one type of texture features can be sensitive for detecting all considered diseases in our tested grapevine varieties. We also assessed combining texture features and significantly eliminate redundant ones. In fact, only ten texture features were kept when a concatenation of texture techniques is adopted; the ReliefF was efficient for this purpose. The performance of each texture technique solely and their concatenation was evaluated using a supervised classifier.

The main contribution of this Chapter was the demonstration that color imaging in general and texture feature analysis, in particular, could be used for detecting foliar symptoms of FD, Esca and BN diseases but only when symptoms are already starting to become visible. The use of texture analysis was also demonstrated to be feasible for disease identification and severity grading.

The tests that are presented in this Chapter are supposed to be re-used on MS images when the final sensor is ready.

CHAPTER 6

6 RESULTS AND DISCUSSIONS

6.1 REFLECTANCE DATA ANALYSIS

In this first section, we will present and discuss the results of the reflectance data analysis. The use of complete spectral data is presented in ANNEX C. Here, the use of traditional SVIs is first investigated and afterwards, the potential of the feature selection by SPA-KNN and GA-KNN is assessed. Finally, the particular SDIs deduced from SPA-KNN and GA-KNN are also tested and comapred.

The classifier used for reflectance data analysis is SVM with cross-validation. For reflectance analysis, we only tested the potential of different methods in disease detection, which means that a binary configuration is used (healthy vs diseased), only healthy and infested groups exist. Here, there is no distinction between the infestation levels. One disease is tested at a time but a combination of varieties is investigated.

In the PACA region, the reflectance data were acquired in 2016 and the disease studied was FD. In the Nouvelle-Aquitaine region, the reflectance data were measured in 2017 and the infection tested was Esca. In the Burgundy region, the reflectance data used were recorded in 2017 and the infestation considered was BN.

The model accuracy defines the percentage of correctly classified test set samples; the false negative rate (FNR) is the percentage of negative results that are, in fact, positive; and the, false positive rate (FPR) is the percentage of positive results that are, in fact, negative. Note that, in our case we considered the positive class as the diseased class and the negative class as the healthy one. What can be critical in disease diagnosis is probably the FNR, which means that a diseased case was claimed to be healthy. In general, the lower the FNR, the better the classifier is.

ROC curve is a performance measurement for classification problem at various thresholds settings. ROC is a probability curve. The ROC curve is plotted with True Positive Rate (TPR) against the False Positive Rate (FPR) where TPR is on y-axis and FPR is on the x-axis.

AUC represents degree or measure of separability. It tells how much model is capable of distinguishing between classes. An excellent model has AUC near to the 1 which means it has good measure of separability. A poor model has AUC near to the 0 which means it has worst measure of separability. In fact, it means it is reciprocating the result. It is predicting 0s as 1s and 1s as 0s. And when AUC is 0.5, it means model has no class separation capacity whatsoever.

6.1.1 DATA CONFIGURATION (DISEASE DETECTION, BINARY CLASSIFICATION)

6.1.1.1 *PACA REGION*

The data considered for the reflectance analysis in the PACA region correspond to those acquired in 2016. The disease tested here was FD. Two dimensions can be used in the analysis of these data. The first dimension is the severity of infestation which implies the number of classes used. Here, a binary classification was adopted (shown in *Figure 6.1*); in fact, we studied the healthy group versus the

diseased group as a whole (slightly infested measurements from the August acquisition campaign + highly infested measurements from the September acquisition campaign).

The second dimension is the type of measurements. In the testing campaigns, four varieties were taken into consideration: Marselan, Grenache, Vermentino and Chardonnay. They each could have been studied separately or combined which means that the analysis could be conducted based on the grapevine color; so, two groups were considered: Red (measurements from Marselan + measurements from Grenache) vs White (measurements from Vermentino + measurements from Chardonnay). In the last case, all the grapevine measurements were combined together (measurements from Marselan + measurements from Grenache + measurements from Vermentino + measurements from Chardonnay).

Binary classification	Marselan	Grenache	Vermentino	Chardonnay
Severity of infestation 0 (Healthy)				
Severity of FD infestation 1 (Slightly infested)				
Severity of FD infestation 2 (Highly infested)				

— Marselan ---- Grenache –·– Vermentino –··– Chardonnay
··········· White ------ Red ▬▬ All
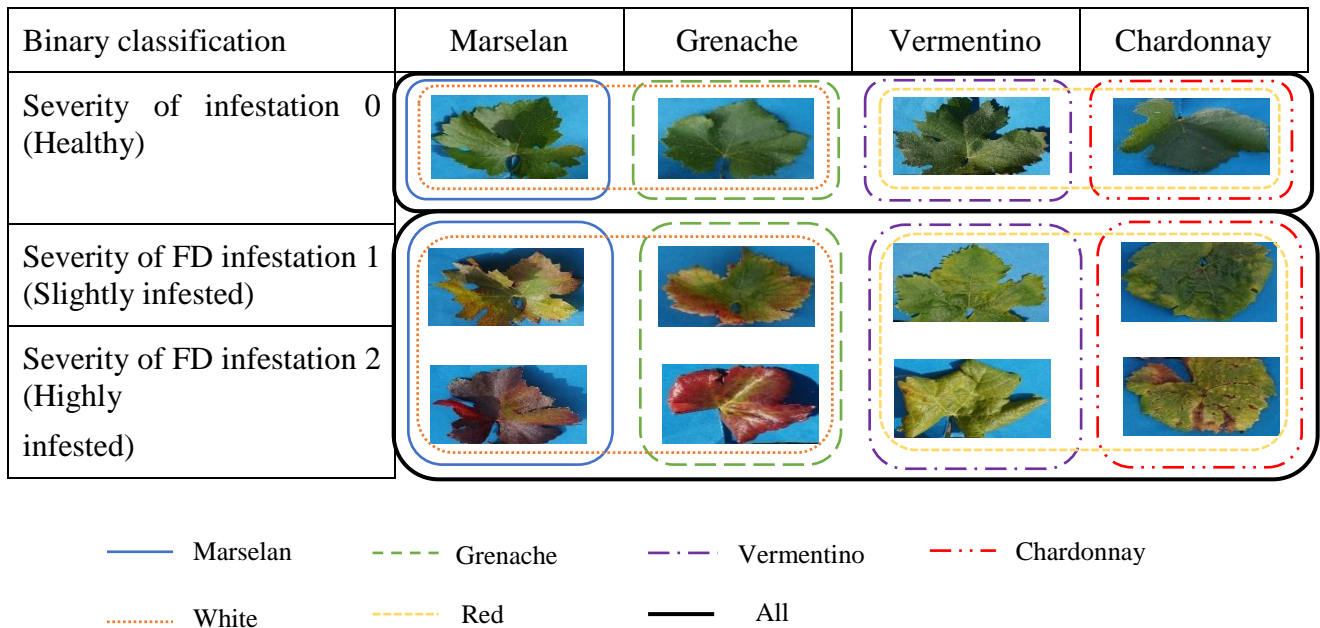
Figure 6.1 PACA data configuration relative to a binary classification.

6.1.1.2 *NOUVELLE-ACQUITAINE REGION*

The data considered for the reflectance analysis in the Nouvelle-Acquitaine region correspond to those acquired in 2017. The disease tested here was Esca. One dimension can be used in the analysis of these data; it is the severity of infestation which implies the number of classes used. Here, a binary classification was employed (shown in *Figure 6.2*); in fact, we studied the healthy group versus the diseased group as a whole (slightly infested measurements from the July acquisition campaign + highly infested measurements from the September acquisition campaign).

Binary classification	Cabernet-Sauvignon
Severity of infestation 0 (Healthy)	
Severity of Esca infestation 1 (Slightly infested)	
Severity of Esca infestation 2 (Highly infested)	

Figure 6.2 Nouvelle-Aquitaine data configuration relative to a binary classification.

6.1.1.3 *BURGUNDY REGION*

The data considered for the reflectance analysis in the Burgundy region correspond to those acquired in 2017. The disease tested here was BN. One dimension can be used in the analysis of these data; it is the severity of infestation which implies the number of classes used. Here, we used a binary classification (shown in *Figure 6.3*); in fact, we studied the healthy group versus the diseased group as a whole (slightly infested measurements from the August acquisition campaign + highly infested measurements from the September acquisition campaign).
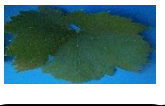
Binary classification	Chardonnay
Severity of infestation 0 (Healthy)	
Severity of BN infestation 1 (Slightly infested)	
Severity of BN infestation 2 (Highly infested)	

Figure 6.3 Burgundy data configuration relative to a binary classification.

6.1.2 TRADITIONAL SVIS

In this section, the results of the classification using the common SVIs are presented. Only the best SVIs will be mentioned next (the ones that scored the highest accuracy for each case).

6.1.2.1 APPLICATION TO PACA 2016 DATA

Table 6.1 presents the result of calculating the best traditional SVI from August measurements (slightly infected leaves). The classification accuracies were satisfying (>91%) and No FNR was reported for Grenache leaves. ARI is convenient for Red-grapevine varieties when considered individually or combined (>90%). ZTM behaved well for Vermentino and Chardonnay (>91%) but when combined, TCARI/OSAVI performed better (92%). This index was also robust when all observations are considered together (92%).

Table 6.1 Results of using the best SVI in classifying different groups of spectral data acquired in the August acquisition campaign for FD detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Marselan	90.32	11.11	7.69	0.94	PRI-ARI
Grenache	96.88	0.00	6.25	1.00	ARI
Vermentino	93.75	5.88	6.67	0.94	ZTM
Chardonnay	90.63	11.11	7.14	0.91	NDVI-ZTM
Red	95.24	5.88	3.45	0.99	ARI
White	92.19	11.11	3.57	0.92	TCARI/OSAVI
All	92.13	13.51	1.89	0.92	TCARI/OSAVI

Table 6.2 presents the result of calculating the best traditional SVIs from September measurements (highly infected leaves). In this case, all accuracies were enhanced with respect to those of August (>94%). The best results were obtained for white-berried signatures and no FNR were found (97–98%). Similar to the first acquisition campaign, the index ARI was interesting for the red-berried signatures (>97%). Moreover, ZTM and GM1 were chosen for Vermentino and Chardonnay (98%) but when combined, GM1 was selected (99%). mCAI was the most robust index when All measurements are mixed together (94.02%).

Table 6.2 Results of using the best SVI in classifying different groups of spectral data acquired in the September acquisition campaign for FD detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Marselan	96.88	4.00	2.17	0.98	ARI
Grenache	97.53	2.70	2.27	0.99	ARI
Vermentino	98.18	0.00	3.23	0.96	NDVI-mCAI-PSSRb-ZTM
Chardonnay	97.73	0.00	3.85	0.96	NDVI-mCAI-PSSRb-GM1-GM2-ZTM-TCARI/OSAVI

Red	98.31	2.33	1.10	0.99	ARI
White	98.99	0.00	1.75	0.99	ARI-GM1
All	94.20	7.52	2.80	0.95	mCAI

Table 6.3 presents the results of calculating the best traditional SVIs from August measurements in addition to those from September (slightly + highly infected leaves). ARI was again chosen to be the best SVI for classifying red grape leaves (>91%) and was also selected when combined white grape leaves reflectance was tested (95%). GM1 was found to be interesting for Vermentino (95%) but ZTM (96%) was more convenient for Chardonnay. mCAI, like the above case, was the most robust index when all measurements are mixed together (88%).

Table 6.3 Results of using the best SVI in classifying different groups of spectral data acquired in the August and the September acquisition campaigns for FD detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Marselan	93.70	8.70	3.45	0.9824	ARI
Grenache	91.13	13.70	1.96	0.9753	ARI
Vermentino	95.40	6.98	2.27	0.9730	GM1
Chardonnay	96.05	5.56	2.50	0.9672	GM2-ZTM
Red	93.23	11.19	0.93	0.9888	ARI
White	95.09	8.64	1.22	0.9433	ARI
All	88.41	16.96	5.98	0.9184	mCAI

From the previous section, we can deduce that the FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA region using SVIs was overall satisfying (accuracy>88%). The SVIs performed in general better than complete spectra (ANNEX C) for early or late season detection. When the severity stages are combined, the SVIs gave better results in few cases only (Marselan and combined white-berried data). The highest detection accuracies for SVIs were at the end of the season (severity stage 2). The most frequently used SVI for the FD detection on red-berried (Marselan, Grenache, Cabernet-Sauvignon) grapevines was ARI. GM1 and ZTM were most useful for white-berried grapevines. For the case where measurements from different varieties are combined the mCAI and TCARI/OSAVI indices performed the best.

6.1.2.2 *APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA*

Table 6.4 presents the result of calculating the best traditional SVI from July measurements. The classification accuracy for slightly infested Cabernet-Sauvignon leaves was very good (98.5%). No FNR was found. The best SVI was ARI like the previous red-berried grapevines in PACA region.

Table 6.4 Results of using the best SVI in classifying different groups of reflectance data acquired in the July acquisition campaign for Esca detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Cabernet-Sauvignon	98.48	0.00	3.33	0.99	ARI

Table 6.5 presents the result of calculating the best traditional SVI from September measurements. The classification accuracy for slightly infested Cabernet-Sauvignon leaves was very good (99 %). No FNR was found. The best SVI was ARI like the previous case of severity stage 1.

Table 6.5 Results of using the best SVI in classifying different groups of reflectance data acquired in the September acquisition campaign for Esca detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Cabernet-Sauvignon	98.89	0.00	2.22	0.98	ARI

Table 6.6 presents the result of calculating the best traditional SVI from July and September measurements combined. The classification accuracy for slightly and highly infested Cabernet-Sauvignon leaves was very good (96%) a bit lower than the previous Cabernet-Sauvignon cases. No FNR was found. The best SVI was PRI here, ARI scored (95.5%) quite similar to PRI.

Table 6.6 Results of using the best SVI in classifying different groups of reflectance data acquired in the July and the September acquisition campaigns for Esca detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Cabernet-Sauvignon	96.15	0.00	1.33	0.98	PRI

Depending on the earlier section, we can say that the Esca detection on a red grapevine variety (Cabernet-Sauvignon) in the Nouvelle-Aquitaine region was nealy perfect when tested early or late in the season. The accuracy was a bit less when the severity stages are combined. The SVI accuracies here were higher than using complete spectral data no matter what the infestation stage (ANNEX C). The best SVIs were ARI and PRI. The detection accuracy of Esca on Cabernet-Sauvignon leaves was higher than the FD detection accuracy on Marselan leaves and on Grenache leaves for all infestation levels.

6.1.2.3 *APPLICATION TO BURGUNDY 2017 DATA*

Table 6.7 presents the result of calculating the best traditional SVI from July measurements. The classification accuracy for slightly infested Chardonnay leaves was very good (98.5%). No FNR was found. The best SVI was GM1 and ZTM like the white varieties in PACA region for the same infestation level 1.

Table 6.7 Results of using the best SVI in classifying different groups of reflectance data acquired in the July acquisition campaign for BN detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Chardonnay	97.78	0.00	4.35	0.99	NDVI, GM1, ZTM, TCARI/OSAVI

Table 6.8 presents the result of calculating the best traditional SVI from September measurements. The classification accuracy for highly infested Chardonnay leaves was very good (96.6%). The best SVI was mCAI like the previous white-berried grapevines in PACA region for the same severity level 2.

Table 6.8 Results of using the best SVI in classifying different groups of reflectance data acquired in the September acquisition campaign for BN detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Chardonnay	96.66	3.28	3.39	0.97	mCAI

Table 6.9 presents the result of calculating the best traditional SVI from August and September measurements. The classification accuracy for highly and slightly infested Chardonnay leaves was very good (94%) a bit less than the previous Chardonnay cases. The best SVI was TCARI/OSAVI.

Table 6.9 Results of using the best SVI in classifying different groups of reflectance data acquired in the August and the September acquisition campaigns for BN detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Best SVI
Chardonnay	93.80	8.11	39.62	0.97	TCARI/OSAVI

As a conclusion from the previous section, we can assume that the BN detection on a white grapevine variety (Chardonnay) in the Burgundy region was nealy perfect when tested early or late in the season. The accuracy was a bit less when the severity stages are combined. The SVI accuracies here were higher than using complete spectral data only for the first the infestation stage (ANNEX C). The best SVIs were ZTM, GM1, mCAI and TCARI/OSAVI. The detection accuracy of BN on Chardonnay leaves was higher than the FD detection accuracy on Vermentino leaves and on Chardonnay leaves in PACA region early in the season and lower late in the season and when infestation stages were combined.

6.1.3 PROPOSED THEORETICAL DEFINITION OF OPTIMAL REFLECTANCE FEATURES WITH SPA-KNN

6.1.3.1 *APPLICATION TO PACA 2016 DATA*

Table 6.10 presents the result of calculating the optimal reflectance features from August measurements (slightly infected leaves) using SPA-KNN. The results are in general satisfying (>85%). The highest accuracy was for the Vermentino variety (97%). There was no FNR for Marselan and red varieties.

Table 6.10 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the August acquisition campaign for FD detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Marselan	87,09	0.00	22.73	0,88	29
Grenache	90,62	7.699	15.79	0,96	30
Vermentino	96,87	8.33	20.00	0,93	30
Chardonnay	90,62	15.38	21.05	0,90	30
Red	88,89	0.00	28.26	0,93	61
White	89,06	10.00	11.76	0,91	62
All	85,04	38.47	19.29	0,81	125

Table 6.11 presents the result of calculating the optimal reflectance features from September measurements (highly infected leaves) using SPA-KNN. The results are generally satisfying (>89.5%) and perfect ones were achieved for white varieties (100%).

Table 6.11 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the September acquisition campaign for FD detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Marselan	96,87	2.32	9.43	0,99	94
Grenache	95,06	2.50	12.19	0,96	79
Vermentino	100.00	3.33	4.00	0,96	53
Chardonnay	100.00	4.00	5.26	0,96	42
Red	93,78	3.57	11.83	0,96	175
White	95,96	5.17	2.44	0,98	97
All	89,49	8.23	7.33	0,96	274

Table 6.12 presents the result of calculating the optimal reflectance features from August and September measurements (slightly & highly infected leaves) using SPA-KNN. Results were good in general but the best ones were also for white varieties (Chardonnay: 96%).

Table 6.12 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the August and September acquisition campaigns for FD detection (Severity of infestation = 1&2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Marselan	92,91	8.62	13.04	0,95	125
Grenache	91,93	2.13	18.18	0,96	122
Vermentino	94,25	6.98	13.64	0,95	85
Chardonnay	96,05	7.14	5.88	0,96	74
Red	91,23	6.79	18.24	0,95	249
White	90,79	8.04	10.53	0,97	161
All	87,43	7.53	11.53	0,94	412

From the last section, we conclude that the FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA region using SPA-KNN was overall satisfying. However, the feature selection method didn't performed a lot better than complete spectral data specially for the case when severity stages or when varieties were combined. The SPA-KNN also didn't perform a lot better than SVIs specially when varieties are mixed. The highest detection accuracies for SPA-KNN were at the end of the season (severity stage 2).

The number of features selected is considered high (>29 features).

6.1.3.2 *APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA*

Table 6.13 presents the result of calculating the optimal reflectance features from July measurements (slightly infected leaves) and the result is satisfying (95.5%).

Table 6.13 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the July acquisition campaign for Esca detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Cabernet Sauvignon	95,45	4.00	12.19	0,95	64

Table 6.14 presents the result of calculating the optimal reflectance features from September measurements (highly infected leaves) and the result is satisfying (91%).

Table 6.14 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the September acquisition campaigns for Esca detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features

91

Cabernet Sauvignon	91,11	2.94	19.64	0,97	88

Table 6.15 presents the result of calculating the optimal reflectance features from July and September measurements (slightly & highly infected leaves) and the result is satisfying (92%).

Table 6.15 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the July and September acquisition campaigns for Esca detection (Severity of infestation = 1&2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Cabernet Sauvignon	92,31	9.37	17.39	0,95	154

We conclude that the Esca detection on red (Cabernet-Sauvignon) grapvines in the Nouvelle-Aquitaine region using SPA-KNN was overall satisfying. However, the features from SPA-KNN had less accuracy than the complete spectra data except for the first severity of infestation. The SVIs behaved better than the chosen features for all cases. The chosen features worked slightly better for the beginning of the season than the end of the season. The Esca detection on Cabernet-Sauvignon was in general better or similar to the FD detection accuracy on Marselan and Grenache but less for an advanced severity stage.

The number of features selected was considered high (>64 features).

6.1.3.3 *APPLICATION TO BURGUNDY 2017 DATA*

Table 6.16 presents the result of calculating the optimal reflectance features from August measurements (slightly infected leaves) and the result is satisfying (96%).

Table 6.16 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the August acquisition campaign for BN detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Chardonnay	96,19	1.07	10.26	0,99	208

Table 6.17 presents the result of calculating the optimal reflectance features from September measurements (highly infected leaves) and the result is satisfying (98%).

Table 6.17 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the September acquisition campaign for BN detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Chardonnay	98,33	0.88	4.76	0,99	238

Table 6.18 presents the result of calculating the optimal reflectance features from August and September measurements (slightly infected leaves) and the result is satisfying (96%).

Table 6.18 Results of using SPA-KNN in classifying different groups of reflectance data acquired in the August and September acquisition campaigns for BN detection (Severity of infestation = 1&2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Chardonnay	96,1	1.07	10.26	0,99	208

In conclusion, the BN detection on white (Chardonnay) grapvines in the Burgundy region using SPA-KNN was good at all moments. The feature selection didn't achieve better results than using complete spectral data. However, compared to SVIs, the selected features achieved similar or slightly better accuracies. The BN detection accuracy on Chardonnay was comparable to the FD detection accuracy on Vermentino and Chardonnay leaves at all moments.

The number of features selected was considered high (>208 features).

6.1.4 PROPOSED THEORETICAL DEFINITION OF OPTIMAL REFLECTANCE FEATURES WITH GA-KNN

6.1.4.1 *APPLICATION TO PACA 2016 DATA*

Table 6.19 presents the result of calculating the optimal reflectance features from August measurements (slightly infected leaves). Results were satisfying in general (>85%) and perfect for Grenache (100%). There was no FNR reported except for white and all mixed varieties. There was also no FPR reported for Grenache.

Table 6.19 Results of using GA-KNN in classifying different groups of reflectance data acquired in the August acquisition campaign for FD detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Marselan	93,55	0.00	11.11	0,96	4
Grenache	100.00	0.00	0.00	1.00	12
Vermentino	96,87	0.00	5.88	0,98	15
Chardonnay	96,87	0.00	5.88	0,99	4
Red	95,24	0.00	8.33	0,96	15
White	95,31	6.25	3.12	0,96	36
All	85,04	21.43	16.81	0,85	11

Table 6.20 presents the result of calculating the optimal reflectance features from September measurements (highly infected leaves). Perfect classifications were achieved for almost all cases. Here there was no FPR and no FNR except for Marselan and all mixed varieties.

Table 6.20 Results of using GA-KNN in classifying different groups of reflectance data acquired in the September acquisition campaign for FD detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Marselan	98,96	0.00	2.04	1.00	9
Grenache	100.00	0.00	0.00	1.00	10
Vermentino	100.00	0.00	0.00	1.00	11
Chardonnay	100.00	0.00	0.00	1.00	8
Red	100.00	0.00	0.00	1.00	14
White	100.00	0.00	0.00	1.00	8
All	97,83	0.68	3.85	0,99	14

Table 6.21 presents the result of calculating the optimal reflectance features from August and September measurements (slightly and highly infected leaves). Very good results again achieved (>94.4%). Here there was no FNR reported except for Vermentino.

Table 6.21 Results of using GA-KNN in classifying different groups of reflectance data acquired in the August and September acquisition campaigns for FD detection (Severity of infestation = 1&2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Marselan	98,42	0.00	3.03	0,99	5
Grenache	96,77	0.00	5.97	0,99	31
Vermentino	96,55	2.17	4.88	0,99	35
Chardonnay	98,68	0.00	2.86	0,99	9
Red	97,61	0.00	4.51	0,99	24
White	97,55	0.00	5.13	0,99	24
All	94,44	0.00	10.27	0,98	15

From the last section, we conclude that the FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA region using GA-KNN was overall

94

satisfying. The feature selection method outperformed complete spectral data in almost all cases. The GA-KNN also outperformed common SVIs, except for one case where all varieties are combined when FD had a severity stage of 1. Here, TCARI/OSAVI gave better accuracy. Furthermore, GA-KNN outperformed SPA-KNN in absolutely all cases. The highest detection accuracies for GA-KNN were at the end of the season (severity stage 2).

The number of features selected was smaller than SPA-KNN in general (maximum of 35 features).

6.1.4.2 *APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA*

Table 6.22 presents the result of calculating the optimal reflectance features from July measurements (slightly infected leaves). Very good precision is achieved (98.5%) with no FNR.

Table 6.22 Results of using GA-KNN in classifying different groups of reflectance data acquired in the July acquisition campaign for Esca detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Cabernet Sauvignon	98,48	0.00	2.70	1.00	30

Table 6.23 presents the result of calculating the optimal reflectance features from September measurements (highly infected leaves). Similarly to the previous case, a very good precision is achieved (98.5%) with no FNR.

Table 6.23 Results of using GA-KNN in classifying different groups of reflectance data acquired in the September acquisition campaigns for Esca detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Cabernet Sauvignon	98,48	0.00	2.70	0,99	28

Table 6.24 presents the result of calculating the optimal reflectance features from July and September measurements (slightly infected leaves) and the result is satisfying (97.5%). The FNR reported was null.

Table 6.24 Results of using GA-KNN in classifying different groups of reflectance data acquired in the July and september acquisition campaigns for Esca detection (Severity of infestation = 1&2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Cabernet Sauvignon	97,43	0.00	4.70	0,99	35

We conclude that the Esca detection on red (Cabernet-Sauvignon) grapvines in the Nouvelle-Aquitaine region using GA-KNN was overall satisfying. The features from GA-KNN gave better accuracy than the complete spectra data. They also behaved better than common SVIs for all cases.

With respect to SPA-KNN, GA-KNN achieved better accuracies. The Esca detection on Cabernet-Sauvignon was in general similar to the FD detection accuracy on Marselan and Grenache.

The number of features selected was also lower than SPA-KNN (maximum of 35 features).

6.1.4.3 *APPLICATION TO BURGUNDY 2017 DATA*

Table 6.25 presents the result of calculating the optimal reflectance features from August measurements (highly infected leaves). A perfect classification is achieved (100%) with no FPR and no FNR.

Table 6.25 Results of using GA-KNN in classifying different groups of reflectance data acquired in the August acquisition campaign for BN detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Chardonnay	100.00	0.00	0.00	1.00	7

Table 6.26 presents the result of calculating the optimal reflectance features from September measurements (highly infected leaves). A very good classification is achieved (98%) with no FNR.

Table 6.26 Results of using GA-KNN in classifying different groups of reflectance data acquired in the September acquisition campaign for BN detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Chardonnay	98,33	0.00	3.22	0,99	8

Table 6.27 presents the result of calculating the optimal reflectance features from August and September measurements (slightly & highly infected leaves). The classification precision is nearly perfect (99.5%) and the FNR was null.

Table 6.27 Results of using GA-KNN in classifying different groups of reflectance data acquired in the August and September acquisition campaigns for BN detection (Severity of infestation = 1&2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	Nb of selected features
Chardonnay	99,52	0.00	0.94	1.00	35

We can conclude from the previous section that the BN detection on white (Chardonnay) grapvines in the Burgundy region using GA-KNN was overall satisfying. The features from GA-KNN gave better accuracy than the complete spectra data. They also behaved better than common SVIs for all cases. With respect to SPA-KNN, GA-KNN achieved better accuracies. The BN detection on Chardonnay was in general similar to the FD detection accuracy on Vermentino and Chardonnay.

The number of features selected was also lower than SPA-KNN (maximum of 35 features).

6.1.5   PROPOSED PRACTICAL DEFINITION OF OPTIMAL REFLECTANCE FEATURES FOR MS SENSOR BY SDI BASED ON SPA-KNN

In the following tables, wavelengths in the visible range that form the SDIs are in bold italic, those in the near infrared are in italic and those in the infrared are underligned.

6.1.5.1   *APPLICATION TO PACA 2016 DATA*

Table 6.28 presents the result of calculating the best SDI from August measurements (slightly infected leaves). The classification accuracies were satisfying (>85%). Perfect classification accuracy was achieved for Grenache and Vermentino measurements (100%). The majority of the wavelengths chosen are in the visible range, less wavelenghts are chosen in the short-wave infrared range and even less wavelengths are the near infrared range. In fact, at least two of the four chosen wavelengths are in the visible range.

Table 6.28 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the August acquisition campaign for FD detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by SPA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Marselan	93.55	0.00	11.76	0.97	*402*	2130	*681*	1827	-1
Grenache	100.00	0.00	0.00	1.00	*495*	*1128*	1716	2091	-1
Vermentino	100.00	0.00	0.00	1.00	*402*	*411*	*753*	1623	-1
Chardonnay	96.87	0.00	5.88	0.94	*423*	*402*	1524	*936*	-1
Red	85.71	15.15	13.33	0.86	*423*	*495*	2004	*1128*	-0.5
White	96.87	5.88	0.00	0.96	*402*	*474*	*1464*	*708*	0.5
All	93.70	6.25	6.35	0.93	*687*	2118	*447*	1824	1

Table 6.29 presents the result of calculating the best SDI from September measurements (highly infected leaves). The classification accuracies were satisfying (>85%). Perfect classification accuracy was achieved for Marselan, Vermentino, Chardonnay and white-combined measurements (100%). The majority of the wavelengths chosen is in the visible range, in the second place come wavelengths in the near infrared region and in the third place come wavelengths in the short-wave infrared region. Two or three of the wavlengths are in the visible region but this wasn't the case for combined measurements where no visible wavelength is selected.

Table 6.29 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the September acquisition campaign for FD detection (Severity of infestation = 2).

				AUC	SDI by SPA-KNN	

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)		λ_1	λ_2	λ_3	λ_4	w
Marselan	100.00	0.00	0.00	1.00	*402*	*471*	*558*	*762*	-0.5
Grenache	95.06	7.89	2.32	0.95	*459*	*489*	*1434*	*777*	0.5
Vermentino	100.00	0.00	0.00	1.00	*402*	*441*	*750*	*552*	-1
Chardonnay	100.00	0.00	0.00	1.00	*402*	*414*	*429*	*693*	-1
Red	86.44	14.28	12.90	0.89	<u>*2031*</u>	*1362*	*465*	*1005*	-1
White	100.00	0.00	0.00	1.00	*402*	*471*	*525*	<u>*2286*</u>	-0.5
All	85.87	15.20	13.24	0.89	<u>*1623*</u>	<u>*2292*</u>	<u>*2055*</u>	*1110*	1

Table 6.30 presents the result of calculating the best SDI from August and September measurements (slightly and highly infected leaves). The classification accuracies were satisfying (>92%). The majority of the wavelengths chosen are in the visible range, less wavelenghts are chosen in the short-wave infrared range and even less wavelengths are in the near infrared range. Two or more wavelengths are in the visible range but this wasn't the case for red combined varieties and when all measurements are mixed.

Table 6.30 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the August and the September acquisition campaigns for FD detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by SPA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Marselan	97.64	3.08	1.61	0.99	*495*	*447*	*696*	*1131*	0.5
Grenache	94.35	7.57	3.45	0.94	*444*	*648*	*1344*	<u>*2283*</u>	1
Vermentino	97.70	4.76	0.00	0.96	*402*	*432*	*498*	*465*	-1
Chardonnay	98.68	2.86	0.00	0.98	*402*	*426*	*453*	*651*	0.5
Red	93.23	7.69	5.78	0.92	<u>*2082*</u>	<u>*1698*</u>	<u>*1944*</u>	*1119*	1
White	97.55	3.95	1.15	0.98	*459*	*402*	*534*	<u>*2292*</u>	0.5
All	92.99	8.65	5.34	0.94	<u>*1650*</u>	<u>*2145*</u>	<u>*1971*</u>	*1026*	-1

The FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA region using SDIs from SPA-KNN was overall very satisfying. The SDIs performed better than complete spectral data in general, but there are five cases where that wasn't true (red varieties for the first severity stage, all measurements for the second severity stage; Grenache, red varieties and all measurements for the mixed severity stages). The SDIs from SPA-KNN performed in general better than SVIs in all cases except for three cases where ARI and mCAI performed better

than the chosen SDI (Grenache, red-varieties, all combined varieties in the second severity stage). The SDI based on SPA-KNN was in general better than the generalized SPA-KNN but this wasn't the case for red and all combined varieties where FD was in its second severity stage. The results of SDI based on SPA-KNN were contrasting with respect to the generalized GA-KNN. In fact, in seven cases, the general GA-KNN outperformed the SDI based on SPA-KNN. The highest detection accuracies for SDIs from SPA-KNN were at the end of the season (severity stage 2).

The wavelengths chosen are mainly in the visible range (at least two from four). In addition to visible wavelengths, for the first severity stage and when the infestation levels are combined, mainly wavelengths from the short-wave infrared region are selected, but for the second severity stage wavelengths were essentially from the near infrared region.

6.1.5.2 *APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA*

Table 6.31 presents the result of calculating the best SDI from July measurements (slightly infected leaves). The classification accuracy was satisfying (94% classification accuracy). Selected wavelengths are from the visible range mainly in addition to one other in the short-wave infrared region.

Table 6.31 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the July acquisition campaign for Esca detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by SPA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Cabernet-Sauvignon	93.94	7.89	3.57	0.92	*447*	*498*	*612*	<u>1638</u>	-0.5

Table 6.32 presents the result of calculating the best SDI from September measurements (highly infected leaves). The classification accuracy was good (82%). Selected wavelengths are from the visible range mainly in addition to one other in the near infrared region.

Table 6.32 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the September acquisition campaign for Esca detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by SPA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Cabernet-Sauvignon	82.22	20.41	14.63	0.85	*582*	*432*	*1419*	*498*	-1

Table 6.33 presents the result of calculating the best traditional SDI from July and September measurements (slightly and highly infected leaves). The classification accuracy was good (91%). Selected wavelengths are from the visible range mainly in addition to one other in the short-wave infrared region.

Table 6.33 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the July and the September acquisition campaigns for Esca detection (Severity of infestation = 1 & 2).

Grapevine Variety				AUC	SDI by SPA-KNN

	Accuracy (%)	FNR (%)	FPR (%)		λ_1	λ_2	λ_3	λ_4	w
Cabernet-Sauvignon	91.02	7.59	10.39	0.94	*672*	*456*	*402*	2286	-0.5

The Esca detection on red (Cabernet-Sauvignon) grapvines in the Nouvelle-Aquitaine region using SDIs from SPA-KNN was overall very satisfying (accuracy>82%). The SDIs from SPA-KNN had less accuracy than the complete spectra data except for the first severity of infestation. The SVIs (ARI and PRI) behaved better than the chosen SDIs at the beginning and when infestation stages are combined. The SDI based on SPA-KNN was in general less accurate than the generalized SPA-KNN. The accuracies of SDI based on SPA-KNN were also less than the generalized GA-KNN. The best SDIs worked better for the beginning of the season than the end of the season. The Esca detection on Cabernet-Sauvignon was in general less accurate than the FD detection accuracy on Marselan and Grenache.

The wavelengths chosen are mainly in the visible range (at least three from four). When leaves are slightly infested or when slightly and highly infested leaves are mixed, the fourth chosen wavelength is from the short-wave infrared region. For the case of highly infected samples, the fourth wavelength chosen is from the near infrared region.

6.1.5.3 *APPLICATION TO BURGUNDY 2017 DATA*

Table 6.34 presents the result of calculating the best SDI from August measurements (slightly infected leaves). The classification was satisfying and perfect classification accuracy was achieved (99.5%). Chosen wavelengths are from the visible range mainly in addition to one other in the short-wave infrared region.

Table 6.34 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the August acquisition campaign for BN detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by SPA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Chardonnay	99.52	0.00	0.94	0.99	*402*	*477*	*693*	2100	-1

Table 6.35 presents the result of calculating the best SDI from September measurements (highly infected leaves). The classification accuracy was satisfying and a perfect accuracy was achieved (100%). Selected wavelengths are from the visible range mainly in addition to one in the near infrared region and one short-wave infrared region.

Table 6.35 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the September acquisition campaign for BN detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by SPA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Chardonnay	100.00	0.00	0.00	1.00	*489*	*702*	*1302*	1788	0.5

Table 6.36 presents the result of calculating the best SDI from August and September measurements (slightly and highly infected leaves). The classification accuracy was satisfying and a perfect accuracy was achieved (100%). Chosen wavelengths are from the visible range mainly in addition to one other in the short-wave infrared region.

Table 6.36 Results of using the best SDI by SPA in classifying different groups of reflectance data acquired in the August and the September acquisition campaigns for BN detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by SPA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Chardonnay	100.00	0.00	0.00	1.00	*693*	*477*	2106	*402*	0.5

The BN detection on white (Chardonnay) grapvines in the Burgundy region using SDIs from SPA-KNN was perfect at all moments. The SDIs performed in general better for white-varieties than red ones. The SDIs performed in general better than complete spectra and better than SVIs in all cases. The SDI based on SPA-KNN was more accurate than the generalized SPA-KNN. The accuracies of SDI based on SPA-KNN were also better than the generalized GA-KNN except for the first level of severity. The BN detection accuracy on Chardonnay was similar or higher than the FD detection accuracy on Vermentino and Chardonnay leaves at all moments.

The wavelengths picked are mainly in the visible range (at least two from four). When leaves slightly infested are considered, a wavelength in the short-wave infrared region is also selected. This is the same for the case where highly and slightly leaves are mixed. However, when leaves highly infested are tested, one wavelength is picked in the near infrared region and another one is chosen in the short-wave infrared region.

6.1.6 PROPOSED PRACTICAL DEFINITION OF OPTIMAL REFLECTANCE FEATURES FOR MS SENSOR BY SDI BASED ON GA-KNN

In the following, wavelengths in the visible range that form the SDIs are in bold italic, those in the near infrared are in italic and those in the infrared are underligned.

6.1.6.1 APPLICATION TO PACA 2016 DATA

Table 6.37 presents the result of calculating the best SDI from August measurements (slightly infected leaves). The classification accuracies were satisfying (>90%). Perfect classification accuracy was achieved for Vermentino and for combined white-berried measurements (100%). The majority of the wavelengths chosen is in the short-wave infrared region, in the second place come the wavelengths in the visible region and in the third place come wavelengths in the near infra-red.

Table 6.37 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the August acquisition campaign for FD detection (Severity of infestation = 1).

Grapevine Variety				AUC	SDI by GA-KNN

	Accuracy (%)	FNR (%)	FPR (%)		λ_1	λ_2	λ_3	λ_4	w
Marselan	96.77	5.88	0.00	1.00	*402*	*717*	*915*	*720*	-1
Grenache	90.62	11.76	6.66	1.00	*576*	*795*	*1053*	*765*	1
Vermentino	100.00	0.00	0.00	1.00	*477*	*498*	*846*	*1497*	0.5
Chardonnay	96.87	0.00	5.88	0.95	*585*	*1863*	*1611*	*2274*	-1
Red	98.41	3.03	0.00	1.00	*402*	*1032*	*2097*	*723*	1
White	100.00	0.00	0.00	1.00	*1842*	*2121*	*2253*	*1233*	-0.5
All	97.64	4.48	0.00	0.97	*717*	*2139*	*1512*	*2136*	-0.5

Table 6.38 presents the result of calculating the best SDI from September measurements (highly infected leaves). The classification accuracies were very satisfying (>98%). Perfect classification accuracy was achieved for almost all cases (100%). The majority of the chosen wavelengths are in the visible range; no wavelengths were picked in the near infrared region and only for Chardonnay four wavelengths in the short-wave infrared region are selected.

Table 6.38 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the September acquisition campaign for FD detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Marselan	100.00	0.00	0.00	1.00	*540*	*615*	*711*	*714*	-1
Grenache	100.00	0.00	0.00	1.00	*540*	*543*	*546*	*711*	-1
Vermentino	100.00	0.00	0.00	1.00	*474*	*549*	*570*	*621*	-1
Chardonnay	97.73	5.26	0.00	1.00	*1167*	*1455*	*1134*	*1563*	0.5
Red	100.00	0.00	0.00	1.00	*537*	*642*	*552*	*693*	1
White	100.00	0.00	0.00	1.00	*444*	*522*	*591*	*714*	1
All	98.55	2.34	0.67	1.00	*552*	*699*	*627*	*711*	1

Table 6.39 presents the result of calculating the best SDI from August and September measurements (slightly infected leaves). The classification accuracies were very good (>96%). The majority of the chosen wavelengths are in the visible range, one wavelength was picked in the near infrared region and only for white mixed samples; the rest of the picked wavelengths were in the short-wave infrared region.

Table 6.39 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the August and the September acquisition campaigns for FD detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Marselan	98.42	3.03	0.00.	1.00	*411*	*627*	*705*	*546*	-1
Grenache	99.19	1.56	0.00	1.00	*426*	*519*	*726*	2127	-0.5
Vermentino	98.85	24.39	0.00	1.00	*708*	1917	*402*	2175	-1
Chardonnay	98.68	2.86	0.00	1.00	*462*	*537*	*723*	1395	1
Red	97.61	4.51	0.00	1.00	*435*	*549*	*705*	1494	-0.5
White	98.16	2.67	1.14	1.00	*726*	*717*	1602	*861*	-1
All	96.13	6.57	0.99	1.00	*513*	*678*	2145	*708*	-1

Briefly, the FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA region using SDIs from GA-KNN was overall very satisfying (accuracy>90%). The SDIs performed in general better than complete spectra for all cases and better than SVIs in all cases except for Grenache leaves at severity stage 1 where ARI performed better than the chosen SDI. SDIs from GA-KNN gave superior accuracies than SDIs from SPA-KNN except for Grenache leaves in the first severity stage and Chardonnay in the second severity stage. SDIs from GA-KNN also gave in general better results than the generalized SPA-KNN except for the Chardonnay case when the severity stage of FD is 2. SDIs from GA-KNN outperformed the generalized GA-KNN except for two cases (Grenache in severity stage 1 and Chardonnay in severity stage 2). The highest detection accuracies for SDIs from GA-KNN were at the end of the season (severity stage 2).

The wavelengths chosen were by majority in the visible region, fewer were picked in the short-wave infrared region and less were chosen in the near infrared region.

The wavelengths chosen were mainly in the visible range (at least two from four) especially for the second severity stage (four from four for majority of the cases). In addition to visible wavelengths, for the first severity stage and when the infestation levels are combined, mainly wavelengths from the short-wave infrared and fewer from the near infrared regions were selected.

6.1.6.2 *APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA*

Table 6.40 presents the result of calculating the best SDI from August measurements (slightly infected leaves). A perfect accuracy was achieved (100%). Chosen wavelengths are from the visible range mainly in addition to one in the near infrared region and another in the short-wave infrared region.

Table 6.40 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the July acquisition campaign for Esca detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Cabernet-Sauvignon	100.00	0.00	0.00	1.00	*537*	*618*	855	<u>1236</u>	-0.5

Table 6.41 presents the result of calculating the best SDI from September measurements (slightly infected leaves). The classification was perfect (100%). Picked wavelengths are from the visible range mainly in addition to one in the near infrared region and another in the short-wave infrared region.

Table 6.41 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the September acquisition campaign for Esca detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Cabernet-Sauvignon	100.00	0.00	0.00	1.00	*588*	<u>1590</u>	*705*	924	1

Table 6.42 presents the result of calculating the best traditional SDI from August and September measurements (slightly and highly infected leaves). The classification accuracy was satisfying (90%). Here, only one wavelength was in the visible range and the other three picked wavelengths were in the short wave infra-red region.

Table 6.42 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the July and the September acquisition campaigns for Esca detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Cabernet-Sauvignon	90.38	8.75	10.53	1.00	<u>1860</u>	*711*	<u>2052</u>	<u>1674</u>	1

In summary, the Esca detection on red (Cabernet-Sauvignon) grapevines in the Nouvelle-Aquitaine region using SDIs from GA-KNN was overall very satisfying (accuracy>90%) and perfect at the begining and the end of the season. The SDIs were less accurate than using complete spectral data in only one case, when severity stages were mixed. The SDIs performed better than SVIs in all cases except for the case where severity stages are combined, here, PRI performed better than the chosen SDI. SDIs from GA-KNN gave in general better results than the generalized SPA-KNN and the generalized GA-KNN also except for the combined severity stages. The SDIs from GA-KNN resulted in better classifications than SDIs from SPA-KNN, this wasn't true for only the case where severity stage are mixed.

The Esca detection accuracy on Cabernet-Sauvignon was higher for the first severity stage than the FD detection accuracy on Marselan and Grenache leaves but similar for the second severity stage and lower when stages are combined.

In general, mainly wavelengths from the visible range are chosen (two from four). For slightly and highly infested leaves, two of the wavelengths were from the visible range another is from the near infrared range and the last is from the short-wave infrared region. For the combined severity stage, in addition to the two visible wavelengths, another two in the short-wave infrared region are picked.

6.1.6.3 *APPLICATION TO BURGUNDY 2017 DATA*

Table 6.43 presents the result of calculating the best SDI from August measurements (slightly infected leaves). The classification was satisfying and a perfect accuracy was achieved (100%). Chosen wavelengths are from the visible range mainly in addition to one in the near infrared region and another in the short-wave infrared region.

Table 6.43 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the August acquisition campaign for BN detection (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Chardonnay	100.00	0.00	0.00	1.00	585	594	1617	726	0.5

Table 6.44 presents the result of calculating the best SDI from September measurements (highly infected leaves). The classification was satisfying and a perfect accuracy was achieved (100%). To wavelengths were picked from the visible range and another two were chosen in the short-wave infrared region.

Table 6.44 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the September acquisition campaign for BN detection (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Chardonnay	100.00	0.00	0.00	1.00	564	780	1200	1755	-0.5

Table 6.45 presents the result of calculating the best SVI from August and September measurements (slightly and highly infected leaves). The classification was satisfying and a perfect accuracy was achieved (100%). Selected wavelengths are from the visible range mainly in addition to one in the short-wave infrared region.

Table 6.45 Results of using the best SDI by GA in classifying different groups of reflectance data acquired in the August and the September acquisition campaigns for BN detection (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC	SDI by GA-KNN				
					λ_1	λ_2	λ_3	λ_4	w
Chardonnay	100.00	0.00	0.00	1.00	654	708	2094	615	-1

The BN detection on white (Chardonnay) grapvines in the Burgundy region using SDIs from GA-KNN was perfect at all moments. The SDIs performed in general better than complete spectra and better than SVIs in all cases. SDIs from GA-KNN gave in general better results than the generalized SPA-KNN and the generalized GA-KNN in all cases. The SDIs calculated from GA-KNN had similar perfect behavior than those resulting from SPA-KNN. The BN detection accuracy on Chardonnay was higher than the FD detection accuracy on Chardonnay leaves for all cases but similar to Vermentino for the first and second stage of infestation.

In general, mainly wavelengths from the visible range are chosen (at least two from four). For slightly infested leaves, three of the wavelengths were from the visible range and the last is from the short-wave infrared region. The same can be concluded for the combined severity stage. For the second infestation level, in addition to the two visible wavelengths, another two in the short-wave infrared region are picked.

6.1.7 DISCUSSION

Reflectance data were acquired using a spectro-radiometer from three different regions. Three diseases are considered on different grapevine varieties. The spectral reflectance characteristics change when the cellular leaf structure is altered due to the presence of a pathogen [172] and this is in accordance with our results. In fact, some differences in leaf reflectance between FD, Esca and BN infected leaves on one side and healthy leaves on another side were distinguished. This was also compatible with the results found by [46], where grapevine leaf-roll infection caused metabolic and pigment changes and the leaf spectral properties were not the same in healthy and infected leaves.

The grapevine varieties responded to the same virus in different manners. This is most likely due to other specific nuances in addition to the variety itself. We mentioned previously that many parameters can affect the reflectance data (ANNEX A). Some authors highlighted that factors such as leaf illumination and surface texture [173] that differ among vineyard sites due to plant height and health, vegetative growth stages, trellis system, soil properties (type, moisture and available nutrients) may result in spectral variations that affect sensor data and detection sensitivity. Moreover, the development of symptoms may be affected by scion-rootstock combinations and environmental factors and can therefore vary between vineyards and growing seasons [174]. Our reflectance signatures have the potential to detect diseases. However, they were unable to distinguish among different infections and had also poor potential in grading them.

Remote sensing-based methods enable the calculation of various SVIs that relate to biophysical and biochemical crop variables. The SVIs are considered as somewhat traditional since they were extensively applied in the agricultural field. Thirteen SVIs from the scientific literature were computed and were shown to be more advantageous than using the complete spectra. The classification accuracy for late season measurements, in general, was found to be better than that of early season for all kinds of measurements. This seems logical, since symptoms at the end of the season become well established, so diseased spectral reflectance are more influenced by the disease and can be more easily discriminated from healthy ones. The top best SVIs were ARI, ZTM, GM1, mCAI and TCARI/OSAVI.

ARI was found interesting in particular for red-varieties. The ARI index is documented as a relevant feature in many studies. Scientists in [175] concluded that ARI had a persistent response to yellow

rust disease at 4 out of 5 growth stages and mentioned that the ARI index was selected for diagnosis of yellow rust in other studies like the one conducted in [176]. Among the indices investigated in the research [177], only the ARI index could differentiate healthy from rust infected leaves. However, it was not capable of distinguishing stem rust from leaf rust pustules. ARI tracks changes in photosynthetic efficiency and was shown to be significant in the study of Devadas et al. (2009). ARI was identified and considered in [178] as an ideal candidate to diagnose yellow rust disease in winter wheat owing to its sustained sensitivity to the disease presence. The index was also found to be efficient in further studies such as in [176].

ZTM and GM1 were found particularly adapted for white-berried grapevine varieties. Chlorophyll content is a potential indicator of vegetation stress because of its direct role in the photosynthesis process of light harvesting, initiation of electron transport [93]. This was confirmed in our study as the ZTM was chosen as the best SVI for FD detection on Vermentino and Chardonnay data. Loss of chlorophyll in response to infestation by sap feeding insects like aphids [179] and leafhoppers [180] has been reported earlier. Vegetation stress can be assessed by measuring the chlorophyll because it is directly related to the photosynthesis process of light harvesting, initiation of electron transport and it is responsive to a range of stresses [181]. A reduction in chlorophyll levels as a reaction to an infestation induced by sap feeding insects like aphids ([179], [182], [183]), phylloxera [184] and leafhoppers [180] have been reported. GM1 was selected also for white-grape data in this study, in fact, differences in reflectance between healthy and stressed vegetation due to changes in Chlorophyll a+b levels have been detected previously in the green peak and along the red-edge spectral region of 690–750 nm [185].

For the case where red and white varieties are tested together, mCAI and TCARI/OSAVI proved to be robust enough for disease detection. In addition, TCARI/OSAVI was found to be a good SVI in our study especially for BN detection on Chardonnay leaves. The ratio of the TCARI and the OSAVI indices formed a good Chlorophyll estimator; this was done independently of Leaf Area Index (LAI) and illumination state. The ratio demonstrated good results not only in continuous closed crop canopies [96] but also in open tree canopy orchards [186]. In [177], it was found that the TCARI index, was the only index capable of discriminating stem and leaf rust, among all others. The CAI index indicates exposed surfaces containing dried plant material [187]. Absorptions in the 2000 nm to 2200 nm range are sensitive to cellulose. It was stated in [188] that the CAI index is useful to monitor vegetation coverage for biomass estimation. mCAI was found interesting in our study for detecting FD when all measurements are combined and for detecting BN on Chardonnay leaves late in the season.

NDVI and PRI were less important as SVIs for our application. Authors in [189] confirmed that NDVI was correlated to the presence of Russian wheat aphid in four out of six tested fields. In addition to that, study [10] identified the NDVI as the most discriminating feature since it was capable of discriminating between three kinds of wheat stressors (yellow rust, powdery mildew, wheat aphid). In our study, the popular NDVI performed well in only few cases: for FD detection on Vermentino and Chardonnay leaves early or late in the season. This observation was also reported in [177], where NDVI was not capable of discriminating rust infection in wheat leaves. The PRI was used to assess the infestation of the grapevine leaf roll-associated virus [46], and pest damage [190]. PRI was also proven to be sensitive to different infestation scales of rice leaves damaged by the rice leaf holder pest [8], but in the present study it was only an indicator of Esca for the Cabernet-Sauvignon leaves when measurements from different moments of the season are combined.

Other SVIs such as GM2, SIPI, WI and PSSR were not relevant for contamination detection in our specific study. Eventually, we can say that the ARI vegetation index was found interesting for detecting diseases on red varieties, GM1 for white varieties and mCAI for combined red and white varieties. A further detailed calculation can be found in ANNEX D for the three interesting SVIs.

The conclusion about the most convenient SVIs cannot however be generalized since only three diseases are studied in our work and only two colors of grapevines were considered. More tests need to be done in order to verify our claim about SVIs. Since SVIs were general indices, we put in place two feature selection techniques to improve and simplify disease detection in grapevines based on hyperspectral data. GA-KNN based selection takes much longer than SPA-KNN based selection. In fact, GA took 200 times longer than SPA and this ratio tends to rise dramatically when the number of wavelengths increases. The Successive Projection Algorithm (SPA) acts as a faster technique that could select interesting spectral bands for damage detection with a low computational load. General optimal features from GA-KNN and SPA-KNN performed better than the complete reflectance data. GA-KNN features were better than the best SVI but this wasn't the case for SPA-KNN. The number of bands selected here can't be defined in advance.

For a particular application that considers only four bands, SDIs are created after the choice of some features from the GA-KNN or SPA-KNN pool of features. The SDIs were normalized in order to reduce the impact of change in lighting, land, crop variety or sensor specific effects. This helped producing more robust indices. SDIs from GA-KNN and SPA-KNN were both more advantageous than using complete reflectance data and than the best SVI. SDIs from GA-KNN were better than those from SPA-KNN.

The proposed indices proved high accuracy in general with the advantage of reducing data dimensionality by speeding up the disease detection. Both types of SDIs are good but if the computation time isn't a problem, we advise using the GA-KNN based approach. The results obtained in this study, confirmed the effectiveness of the genetic algorithms in the feature selection procedure. In fact, GA was able to reach a global optimum despite local peaks that might be caused by noise or interdependencies in the data set. This conclusion was also confirmed in other studies in different fields. In [191], GA selected the best subset of features for breast cancer diagnosis system. Furthermore, in [192], GA feature selection algorithm was applied for hand writing recognition. The complexity of the feature set was reduced using fewer features and achieved recognition rates similar to those reached when no feature selection is applied.

The SVM classifier behaved very well. This was in accordance with the conclusions made in [108], who found that by comparing different classifiers, SVM use the inherent information of the vegetation indices in an optimal way.

For the rest, we basically consider the SDIs based on GA-KNN since they exhibited the best results.

When we investigate the wavelengths chosen for the red varieties at least one or two that are very close to the wavelengths considered in the ARI index (550 nm and 700 nm). A clear example would be the SDI selected for red combined varieties when the infestation level is mixed; here two of the selected wavelengths were 549 nm and 705 nm. When we look at the wavelengths chosen for the white varieties, for the majority of the cases at least one or two that are very close to the wavelengths considered in the GM1 index (550 nm and 750 nm). A clear example would be the SDI selected for

Chardonnay samples in PACA region when the infestation level is mixed; here two of the selected wavelengths were 537 nm and 723 nm. When considering the wavelengths chosen for the white varieties, for the majority of the cases at least one or two that are very close to the wavelengths considered in the mCAI index (545 nm and 752 nm). A clear example would be the SDI selected for combined samples in PACA region for the second infestation level, here two of the selected wavelengths were 552 nm and 711 nm.

In general, when the spectral band is lower than 1000 nm, we can find affordable filters to cover it, this is however not the case for wavelengths that are higher than 1000 nm. The reason for this price difference is that inexpensive and ubiquitous silicon photodetectors are sensitive to short wave near infrared light, and provide the opportunity to bring consumers invisible information upto 1000 nm. We asked Midopt for more information about the subject and they confirmed that they can make customized filters upon request but a large quantity (thousands should be ordered). Maybe the future will uncover affordable filters that could surpass 1000 nm. Working on the full electromagnetic spectrum seems indeed promising but for a practical cheap solution, we advise to restrain the solution to the [450 nm - 1000 nm] range which is the actual range that could be scanned by affordable narrow-band filters. One solution is to to remove the reflectances on wavelengths that are less than 450 nm or higher than 1000 nm after selecting the bands from GA-KNN and SPA-KNN and design the SDI based on this limited range. Another solution would be to restrain the spectral range to [450 nm - 1000 nm] instead of [400-2400 nm] since the beginning of the reflectance analysis.

For our August data, the majority of the selected bands were found in the near infrared region; however, for September data, visible bands were mostly selected since symptoms became more visually pronounced. In this case, bands from blue (450–520 nm), green (530–570 nm), red (580–700 nm) were selected. This was in accordance with [46], in fact, the two maximum differences in the visible region appeared at the green peak (550 nm) and in the red peak (680 nm) indicating less chlorophyll absorption in the infected leaves. Furthermore, changes in Cab levels were translated as modifications occurring over the spectral red edge region; this explains why many optimal bands were selected in the specific range of 690–750 nm. Reflectance near 700 nm was pointed out in [105] as an essential feature of green vegetation produced by an equilibrium between biochemical and biophysical plant characteristics. Since plant diseases influence the chlorophyll content of crop plants, increased reflectance around 700 nm can be a first but unspecific indicator to detect diseased crops. Many chosen bands were also mixed with the water absorption bands; the research conducted in [193], demonstrated that the sensitivity to water content was greatest in spectral bands centered at 1450, 1940 where water has its major absorption features.

It can be concluded, that the SDIs were dependent on the disease and the grapevine variety considered; the best variables selected were different from one case to another depending on the samples used in the learning task. Our results suggest that changes in reflectance differ by cultivar and therefore the technique can be optimized for each variety in order to obtain ideal results. Since more than one variety can be grown in the same vineyard area, different measurements were combined to simulate the case where grapevine varieties of the same group were grown together and the approach was tested (Marselan and Grenache measurements were combined in the "Red" group / Vermentino and Chardonnay measurements were combined in the "White" group). In the case where different grapevine varieties were grown together (Marselan, Grenache, Vermentino and Chardonnay were

combined in the "All" group), we combined all the measurements. More than 90% accuracy was still achieved using our new SDIs.

6.2 TEXTURE DATA ANALYSIS

The texture features tested are traditional ones: Hist, Grad, LoG, Wav, GLCM, GLRM features and Fractal dimension. Each texture technique is considered solely and their concatenation is also tested. When each technique is considered alone, only the one having the best accuracy was displayed.

A feature selection technique is added for each technique alone and when all techniques are combined. Actually, it is better to incorporate a feature selection step in the analysis. In fact, the results before and after feature selection are very similar but with ReliefF, we are sure that there is no dimensionality problem, so we will mainly emphasize results obtained after ReliefF application.

Only the results obtained from the Lab color space are reported next since it performed better than the HSV color space, which in turn performed a bit better that the RGB color space in all possible configurations.

In the PACA region, the texture data were acquired in 2017 and the disease studied was FD. In the Nouvelle-Aquitaine region, the texture data were measured in 2017 and the infection tested was Esca. In the Burgundy region, the texture data used were recorded in 2017 and the infestation considered was BN.

The classifier used for texture data analysis is SVM with cross-validation. For texture analysis, we test the potential of different methods in disease detection, which means that a binary configuration is used (healthy vs diseased), only healthy and infested groups exist. Each group contained reflectance acquired at the beginning and the end of the season; this means that there was no distinction between the infestation level. Here, one disease is tested at a time but a combination of varieties is investigated.

In addition to this, we assess the potential of texture analysis in identifying the infection (multi-class configuration: disease 1 vs disease 2 vs healthy). Here, many diseases are tested together from the same colored variety.

Furthermore, texture was investigated for grading the severity stage of a disease (multi-class configuration: disease 1 stage 1 vs disease 1 stage 2 vs healthy). Here, one disease is tested at a time but a combination of varieties is investigated.

The accuracy and the AUC are displayed for each case.

6.2.1 DATA CONFIGURATION (DISEASE DETECTION, BINARY CLASSIFICATION)

The same binary data configuration used for reflectance analysis was also adopted for texture analysis for the three regions.

6.2.2 DATA CONFIGURATION (DISEASE SEVERITY STAGE IDENTIFICATION, MULTI-CLASS CLASSIFICATION)

6.2.2.1 *PACA REGION*

The data considered for the texture analysis in the PACA region correspond to those acquired in 2017. The disease tested here was FD. Two dimensions can be used in the analysis of these data. The first dimension is the severity of infestation which implies the number of classes used. Here, a multi-class is chosen (shown in *Figure 6.4*); in fact, it is possible to study the healthy group versus the slightly infested group versus the highly infested group as different classes.

The second dimension is the type of measurements. In the testing campaigns, four varieties were taken into consideration: Marselan, Grenache, Vermentino and Chardonnay. They each could have been studied separately or combined which means that the analysis could be conducted based on the grapevine color; so, two groups were considered: Red (measurements from Marselan + measurements from Grenache) vs White (measurements from Vermentino + measurements from Chardonnay). In the last case, all the grapevine measurements were combined together (measurements from Marselan + measurements from Grenache + measurements from Vermentino + measurements from Chardonnay).

Multi-class classification	Marselan	Grenache	Vermentino	Chardonnay
Severity of infestation 0 (Healthy)				
Severity of FD infestation 1 (Slightly infested)				
Severity of FD infestation 2 (Highly infested)				

Marselan ———   Grenache - - - -   Vermentino — · —   Chardonnay — · · —

White ············   Red - - - - -   All ———
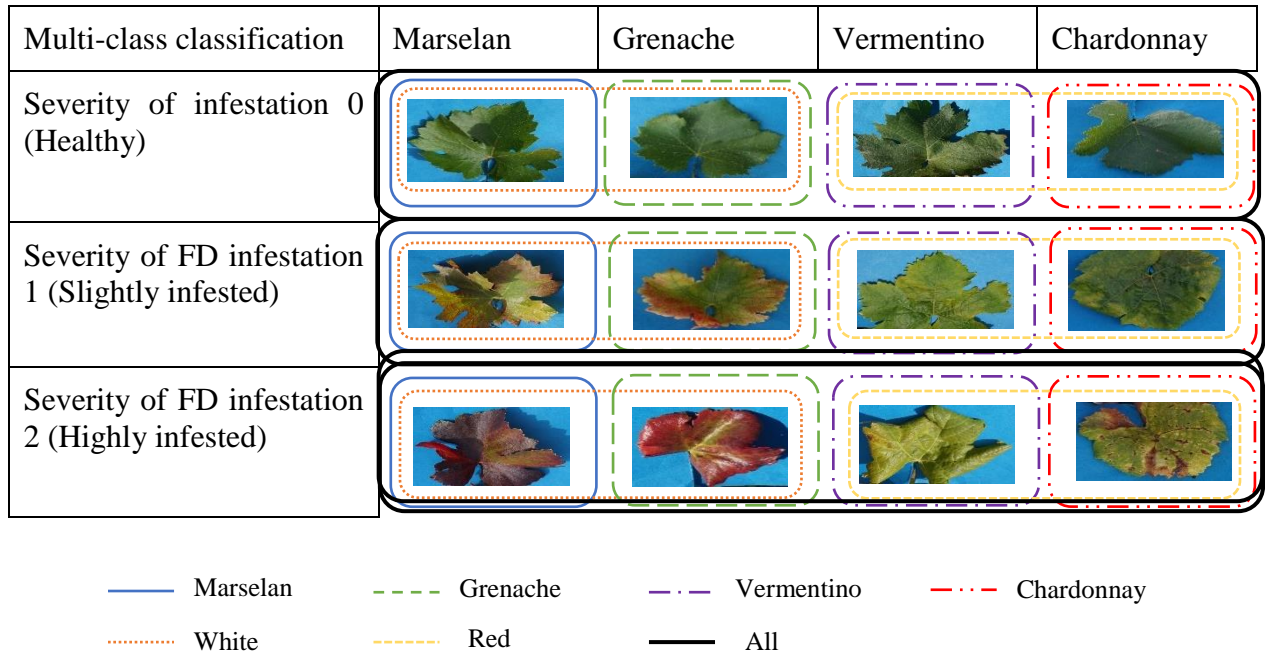
Figure 6.4 PACA data configuration relative to a multi-class classification.

6.2.2.2 *NOUVELLE-ACQUITAINE REGION*

The data considered for the tetxure analysis in the Nouvelle-Acquitaine region correspond to those acquired in 2017. The disease tested here was Esca. One dimension can be used in the analysis of these data; it is the severity of infestation which implies the number of classes used. Here, we used a dimension of three, which implies a multi-class classification (shown *in Figure 6.5*); in fact, we studied the healthy group versus the slightly infested measurements from the July acquisition campaign versus the highly infested measurements from the September acquisition campaign).

Binary classification	Cabernet-Sauvignon
Severity of infestation 0 (Healthy)	
Severity of Esca infestation 1 (Slightly infested)	
Severity of Esca infestation 2 (Highly infested)	

Figure 6.5 Nouvelle-Aquitaine data configuration relative to a multi-class classification.

6.2.2.3 *BURGUNDY REGION*

The data considered for the reflectance analysis in the Burgundy region correspond to those acquired in 2017. The disease tested here was BN. One dimension can be used in the analysis of these data; it is the severity of infestation which implies the number of classes used. Here, we used a dimension of three, which implies a multiclass classification (shown in *Figure 6.6*); in fact, we studied the healthy group versus the slightly infested measurements from the August acquisition campaign versus the highly infested measurements from the September acquisition campaign.

Binary classification	Chardonnay
Severity of infestation 0 (Healthy)	
Severity of BN infestation 1 (Slightly infested)	
Severity of BN infestation 2 (Highly infested)	

Figure 6.6 Burgundy data configuration relative to a multi-class classification.

6.2.3 DATA CONFIGURATION (DISEASE IDENTIFICATION, MULTI-CLASS CLASSIFICATION)

6.2.3.1 *PACA REGION VS NOUVELLE-AQUITAINE REGION*

The data considered for the texture analysis correspond to those acquired from PACA and Nouvelle-Aquitaine regions in 2017. The grapevines considered are Marselan and Cabernet-Sauvignon. The

diseases tested here were FD and Esca. One dimension can be used in the analysis of these data; it is the nature of the infestation which implies the number of classes used. Here, we used a dimension of three, which implies a multiclass classification (shown *in Figure 6.7*); in fact, we studied the healthy group versus the FD infested measurements (from the August acquisition and from the September acquisition campaign) versus the Esca infested measurements (from the July acquisition and from the September acquisition campaign).

Multi-class classification	Severity of infestation 0 (Healthy)	Severity of infestation 1 (Slightly infested)	Severity of infestation 2 (Highly infested)
Marselan - PACA - FD			
Cabernet Sauvignon - Nouvelle Aquitaine - Esca			

Figure 6.7 PACA and Nouvelle-Aquitaine data configuration relative to a multi-class classification.

6.2.3.2 *PACA REGION VS BURGUNDY REGION*

The data considered for the texture analysis correspond to those acquired from PACA and Burgundy regions in 2017. The grapevines considered are Chardonnay. The diseases tested here were FD and BN. One dimension can be used in the analysis of these data; it is the nature of the infestation which implies the number of classes used. Here, we used a dimension of three, which implies a multiclass classification (shown in *Figure 6.8*); in fact, we studied the healthy group versus the FD infested measurements (from the August acquisition and from the September acquisition campaign) versus the BN infested measurements (from the August acquisition and from the September acquisition campaign).
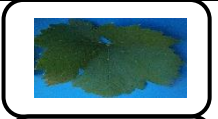
Multi-class classification	Severity of infestation 0 (Healthy)	Severity of infestation 1 (Slightly infested)	Severity of infestation 2 (Highly infested)
Chardonnay - PACA - FD			
Chardonnay - Burgundy-BN			

Figure 6.8 PACA and Burgundy data configuration relative to a multi-class classification.

6.2.4 TEXTURE CALCULATION AND DEFINITION OF OPTIMAL TEXTURE FEATURES WITH RELIEFF

6.2.4.1 *APPLICATION TO PACA 2017 DATA*

Features from leaves acquired in PACA region (Marselan, Grenache, Vermentino, Chardonnay) in 2017 are considered, which means that the FD disease is exclusively studied.

From *Table 6.46*, we can see that the results are generally satisfying; an accuracy higher than 83% is reached. The best texture features in a binary context for the detection of FD when the symptoms start to appearar are: histogram moments for Marselan (100%), fractal dimension for Grenache, for red and white varities (97%, 98% and 90%), GLCM for Vermentino and Chardonnay (90% and 100%), GLRM is the best when all measurements are concatenated (83%). When the top 10 texture features are selected, the results were good, the highest was for Marselan (100%) and the lowest was for white varieties (85%). In general, when choosing only the best technique or a concatenation of them, the results were very similar.

Table 6.46 Results of using texture features in the Lab color space acquired in the August acquisition campaign for FD detection (Severity of infestation = 1). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Grapevine variety	Accuracy (%)	AUC	
Best texture technique	Marselan	100.00	1.00	**Mom**
	Grenache	96.77	0.98	**Frac**
	Vermentino	90.00	0.97	**GLCM, GLRM, Grad**
	Chardonnay	100.00	1.00	**GLCM**
	Red	98.43	1.00	**Frac**

	White	90.32	0.94	**Frac**
	All	83.33	0.87	**GLRM**
Concatenation of all texture techniques	Marselan	100.00	1.00	
	Grenache	96.77	0.98	
	Vermentino	96.67	0.98	
	Chardonnay	93.75	0.99	
	Red	98.44	0.99	
	White	85.48	0.92	
	All	86.51	0.91	

From *Table 6.47*, the best texture features for FD detection when symptoms become clearly visible are GLRM for Marselan and Vermentino (100%), Fractal dimension for Grenache (100%), LoG moments for Chardonnay (100%), histogram moments for red varieties (100%) and wavelet parmaeters for white varieties and for all measurements (100% and 98%). When the top best texture features are only selected the results were also very good (>96%) for all cases. Here when a texture feature is considered solely the results were slightly better than combining all texture features. With respect to the severity stage 1, when the symptoms are clearer, the accuracy is higher.

Table 6.47 Results of using texture features in the Lab color space acquired in the September acquisition campaign for FD detection (Severity of infestation = 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Grapevine variety	Accuracy (%)	AUC	
Best texture technique	Marselan	100.00	1.00	**GLRM, Wav**
	Grenache	100.00	1.00	**Mom, GLCM, Grad, Wav, Frac**
	Vermentino	100.00	1.00	**GLRM, Grad, Wav**
	Chardonnay	100.00	1.00	**LoG**
	Red	100.00	1.00	**Mom**
	White	100.00	1.00	**Wav**
	All	98.44	0.99	**Wav**
Concatenation of all texture techniques	Marselan	97.06	1.00	
	Grenache	100.00	1.00	
	Vermentino	100.00	1.00	

	Chardonnay	96.67	1.00
	Red	98.48	1.00
	White	96.77	0.99
	All	96.09	0.97

From *Table 6.48*, we can see that the results are generally satisfying; an accuracy higher than 86% is reached. The best texture features in a binary context for the detection of FD wherever the degree of severity are: histogram moments for Marselan, Grenache and red-berried samples (100%, 94% and 97%), GLRM moments for Vermentino and Chardonnay (97% and 88%), LoG parameters for white-berried samples (85%) and Fractal dimension for the different samples together (88%). When the texture features are concatenated and the best ten are kept, the results were also very good (>87%). When choosing only the best technique or a concatenation of them, the results were very similar.

Table 6.48 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns for FD detection (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Grapevine variety	**Accuracy (%)**	**AUC**	
Best texture technique	Marselan	100.00	1.00	**Mom**
	Grenache	94.56	0.95	**Mom, Frac**
	Vermentino	96.77	0.98	**GLRM**
	Chardonnay	88.04	0.95	**GLRM**
	Red	97.35	0.98	**Mom**
	White	85.06	0.96	**LoG**
	All	87.75	0.91	**Frac**
Concatenation of all texture techniques	Marselan	100.00	1.00	
	Grenache	94.56	0.97	
	Vermentino	88.71	0.95	
	Chardonnay	94.56	0.97	
	Red	97.35	0.99	
	White	88.31	0.95	
	All	87.46	0.93	

In *Table 6.49*, a distinction between the different levels of FD infestation is made in the PACA region. GLRM was the best for Marselan and red samples (98% and 94%), the histogram moments performed best for Grenache (85%), same goes to the gradient parameters for Vermentino, white samples and all combined samples (87%, 75% and 77%). Wavelet parameters on the other hand were the most

116

advantageous for grading the Chardonnay variety (87%). When many texture features are combined and arranged, the results were higher than 73%. Grading the severity stage of FD was more difficult for white varieties and was overall more difficult than simply detecting the presence of the infection.

Table 6.49 Classification results from texture features in PACA after feature selection using the Lab color space acquired in the August and the September acquisition campaigns for FD severity stage identification. Each texture technique is considered solely and a concatenation of some of them is also presented (multi-class approach)

	Grapevine variety	Accuracy (%)	AUC	
Best texture technique	Marselan	97.92	0.99	**GLRM**
	Grenache	85.43	0.92	**Mom**
	Vermentino	87.05	0.87	**Grad**
	Chardonnay	86.90	0.92	**Wav**
	Red	94.53	0.97	**GLRM**
	White	74.60	0.82	**Grad**
	All	76.84	0.85	**Grad**
Concatenation of all texture techniques	Marselan	97.78	0.99	
	Grenache	88.81	0.93	
	Vermentino	81.64	0.89	
	Chardonnay	73.03	0.85	
	Red	92.35	0.98	
	White	76.79	0.83	
	All	79.87	0.89	

The FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA region was feasible early in the season and nearly perfect at the end of the season. The most frequently used texture features for the FD detection were histogram moments, fractal dimension and GLRM parameters. The results due to severity stage discrimination (multi-class configuration) were somehow less than the simple disease detection with no grading (binary configuration) specially for the white varieties, it seemed difficult to distinguish early from late FD symptoms.

6.2.4.2 *APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA*

Features from leaves acquired in Nouvelle-Aquitaine region (Cabernet-Sauvignon) in 2017 are considered, which means that the Esca disease is exclusively studied. Texture features are studied in the Lab color space.

Table 6.50 shows that the fractal dimensions performs perfectly for detecting Esca on Cabernet-Sauvignon leaves when symptoms start to show (100%). When many top texture features are used the accuracy was also perfect (100%).

Table 6.50 Results of using texture features in the Lab color space acquired in the July acquisition campaign for Esca detection (Severity of infestation = 1). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	100.00	1.00	**GLCM, Wav, Frac**
Concatenation of all texture techniques	100.00	1.00	

Fom *Table 6.51*, we can say that the Fractal dimension was perfect for detecting Esca in the Nouvelle-Aquitaine region (100%) when symptoms are already established. When many texture features are employed and the best 10 are used, also a perfect accuracy was reached.

Table 6.51 Results of using texture features in the Lab color space acquired in the September acquisition campaign for Esca detection (Severity of infestation = 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	100.00	1.00	**GLCM, Frac, GLRM, LoG, Mom, Wav**
Concatenation of all texture techniques	100.00	1.00	

From *Table 6.52* we can say that the results are very good (accuracy of 100%). The best texture features for Esca detection no matter what the symptoms were on a red grapevine-variety are Fractal dimension. When choosing only the best technique or a concatenation of some of them, the results were similar.

Table 6.52 Results of using texture features in the Lab color space acquired in the July and the September acquisition campaigns for esca detection (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	100.00	1.00	**Mom, Wav, Frac**
Concatenation of all texture techniques	100.00	1.00	

From *Table 6.53* we can say that the results are very good and a little less than the binary case. The best texture features for Esca detection in a multi-class context on a red grapevine-variety (Cabernet-

Sauvignon) are the GLRM parameters. When choosing only the best technique or a concatenation of some of them, the results were similar (94%).

Table 6.53 Classification results from texture features in Nouvelle-Aquitaine after feature selection using the Lab color space for Esca severity stage identification. Each texture technique is considered solely and a concatenation of some of them is also presented (multi-class approach)

	Accuracy (%)	AUC	
Best texture technique	93.72	0.97	**GLRM**
Concatenation of all texture techniques	93.72	0.97	

The Esca detection on a red grapevine variety (Cabernet-Sauvignon) in the Nouvelle-Aquitaine region was perfect when tested early or late in the season and the best texture feature was the Fractal dimension. When discriminating between the different intensity of the disease, GLRM parameters were the best features and the results for a multi-class case were overall a bit smaller than the simple detection of the disease (binary case).

The detection of Esca on Cabernet-Sauvignon leaves was quite similar to FD detection on Marselan leaves and a bit better than FD detection on Grenache leaves in most cases using one or all texture features. The severity stage discrimination of Esca on Cabernet-Sauvignon was less than FD on Marselan but higher than FD on Grenache.

6.2.4.3 *APPLICATION TO BURGUNDY 2017 DATA*

Features from leaves acquired in Burgundy region (Chardonnay) in 2017 are considered, which means that the BN disease is exclusively studied. Texture features are studied in the Lab color space.

From *Table 6.54*, we can say that the Fractal dimension was the best choice for BN detection on Chardonnay leaves when symptoms are still appearing. When ReliefF was used and the top features are defined, the accuracy was also very good (97%).

Table 6.54 Results of using texture features in the Lab color space acquired in the August acquisition campaign for BN detection (Severity of infestation = 1). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	100.00	1.00	**Frac**
Concatenation of all texture techniques	96.66	1.00	

From *Table 6.55*, we can say that when BN symptoms become very clear in the Burgundy region in September, Fractal dimenison was also the best texture feature and it gave perfect classification

accuracy (100%). The same can be concluded for the case where top 10 texture features are only considered.

Table 6.55 Results of using texture features in the Lab color space acquired in the September acquisition campaign for BN detection (Severity of infestation = 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	100.00	1.00	**Frac**
Concatenation of all texture techniques	100.00	1.00	

From *Table 6.56* we can say that the results are promising. The best texture features for BN detection in a binary context on a white grapevine-variety (Chardonnay) are the GLRM parameters (94%). A concatenation of some of the texture features resulted in a slightly lesser accuracy than using the best features (90%).

Table 6.56 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns for BN detection (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	94.28	0.97	**GLRM**
Concatenation of all texture techniques	90.00	0.98	

From *Table 6.57* we can say that the results are promising. The best texture feature for BN detection in a multi-class context on a white grapevine-variety (Chardonnay) are again the GLRM features (88%). When choosing only the best technique or a concatenation of some of them, the results were similar (89%)

Table 6.57 Classification results from texture features in Burgundy after feature selection using the Lab color space for BN severity stage identification. Each texture technique is considered solely and a concatenation of some of them is also presented (multi-class approach)

	Accuracy (%)	AUC	
Best texture technique	88.18	0.87	**GLRM**
Concatenation of all texture techniques	89.18	0.91	

The BN detection on a white grapevine variety (Chardonnay) in the Burgundy region was perfect when tested early or late in the season and the best texture feature was the Fractal dimension. The

detection was also very good when samples of different infection intensity are tested together, here the GLRM features performed best. When discriminating between the different intensity of the disease, GLRM parameters were also the best features and the results for a multi-class case were overall smaller than the simple detection of the disease (binary case).

The detection of BN on Chardonnay leaves was quite similar to FD detection on Veremntino and on Chardonnay in most cases using one or combined texture features. The severity stage discrimination of BN on Chardonnay was also similar to FD on Vermentino and FD on Chardonnay (~87%). The symptoms of FD and BN are indistinguishable to the naked eye which explains why the classification models had the same behavior.

6.2.4.4 *APPLICATION TO PACA VS NOUVELLE-AQUITAINE 2017 DATA*

In *Table 6.58*, when samples of different infestation levels and affected by FD are compared to those affected by Esca on red-berried grapevines (Marselan and Cabernet-Sauvignon), texture parameters were found capable of discriminating the FD samples from the Esa samples with an accuracy of 92% when fractal dimension is used. The accuracy was even higher (95%) when all texture features are combined and the top 10 are selected.

Table 6.58 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns for FD and Esca disease identification. Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	92.34	0.96	**Frac**
Concatenation of all texture techniques	94.73	0.97	

6.2.4.5 *APPLICATION TO PACA VS BURGUNDY 2017 DATA*

In *Table 6.59*, we test the possibility of discriminating symptoms of BN and FD on white grapevine-varieties, from PACA (Chardonnay) and from Burgundy (Chardonnay). The results are lower than all other cases when one texture technique is used (Fractal dimension: 76%) and when many are combined (72%).

Table 6.59 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns for FD and BN disease identification. Each texture technique is considered solely and a concatenation of some of them is also presented.

	Accuracy (%)	AUC	
Best texture technique	75.78	0.95	**Frac**
Concatenation of all texture techniques	72.16	0.85	

We can say that the fractal dimenison is the best texture feature when it comes to identifying a disease. Here samples from the same colored variety are considered and FD is evaluated with respect to Esca and with respect to BN. The accuracies reached could be not only due to the disease differences but also to varieties differences and regions differences. The discrimination between FD and BN was the worst which is expected since as mentioned before the visible symptoms of FD and BN are very similar.

6.2.5   DISCUSSION

RGB images acquired directly in the field from three different regions and containing samples from different varieties affected by various diseases are used to calculate a set of traditional texture features.

We have found in our analysis (not shown here) that the Lab color space performed better than the other ones (HSV and RGB). This was in accordance with the results found in [194]. In fact, the disease was accurately detected by Lab in color model and wasn't affected by background, type of leaf or distance to the leaf. We found a foliar disease classification accuracy for Lab color space that was higher than HSV, this result wasn't supported in [195], authors proved that HSV color space performed better than Lab for their application. However, their conclusion was valid only for segmenting noisy colored images and they used the PSNR and MSE as tools for comparison. We noticed that when a feature selection step is used, the most chosen features are from the band A (for Lab) and H (for HSV), this detail wasn't shown here. Bands A and H give information about color and not about intensity. This conclusion was further supported in [196], authors after the transformation process, considered only the H component for their analysis. S and I are dropped since they claimed that they don't give extra information. Moreover, in [131], the gray co-occcurence matrices are generated also for H image only.

As mentioned in [125], K-means is one of the simplest unsupervised learning algorithms that solve the clustering problem and showed good leaf segmentation for our case. Segmented images in the Lab color space are used to calculate the following texture parameters: histogram moments, wavelet, LoG, Gradient, GLCM, GLRM features and fractal dimension.

We mentioned before (ANNEX A) that many parameters could affect the accuracy of our image recordings altering the accuracy of our texture calculations. Many studies confirmed that acquisition conditions can alter texture analysis. The authors in [197] studied the dynamic range influence on texture classification results of brain white matter. They concluded that in order to obtain more reliable classification results, the dynamic range must be optimized to avoid too small or sparse matrices. In fact, larger dynamic ranges do not necessarily give better texture results; they might increase the computation costs and limit the method's performance. Another study [198] assessed how sensitive texture features on MR images are to changes in some parameters related to image acquisition and preprocessing: noise, resolution, the choice of quantization method and the number of gray levels in the quantized image. Authors found that noise, resolution, choice of quantization method and the number of gray levels in the quantized images had a significant influence on most Haralick texture features, and that the effect size varied between different features. In summary, to meaningfully compare texture feature values, scientists recommended to use images with similar resolution and noise levels, to use one quantization method, to find the range of intensities that are

common for all data sets and to use that as a common limit for all images and finally to use one GLCM size.

In general, our traditional texture features yielded good accuracies for detection, identification and grading of grapevine diseases.

Haralick parameters gave promising results in our analysis. This conclusion was also demonstrated in [130] who, found that the leaf sample discriminant analysis using Haralick textural features achieved classification accuracies of over 95%. In addition, study [199] reported that texture features calculated from the the gray co-occcurence were able to classify Black Rot with an accuracy of 95%.

LoG parameters proved good accuracy in classifying grapevine diseases through color images. Another study [200] concluded that the analysis of liver texture by LoG on portal phase CT images is promising and may has potential in predicting the survival for patients with colorectal cancer than CT with better accuracy than perfusion imaging. Parameters extracted from filtered images in [201] showed that there are differences in fine to medium texture ratios in apparently disease-free areas of the liver in patients with hepatic metastases as compared to patients with no tumour and patients with extra-hepatic disease. The LoG filtration technique presented is flexible and allows extracting image features at particular scales that have relevance. The filtration step also enables heterogeneity to be quantified using first-order statistical parameters such as entropy and uniformity, which benefit from being directionally independent and less computationally expensive [202].

Wavelet features were also interesting for distinguishing healthy from diseased leaves. This is not the first time such a result occurs. In fact, in [148] MRI images were decomposed using discrete wavelet transformation (DWT). In the classification stage, two classifiers have been developed. The first classifier based on feed forward backpropagation artificial neural network and the second classifier is based on k-nearest neighbor, a classification with a success of 97% and 98% has been achieved, respectively.

GLRM features reported good accuracy in our case. They even behaved better than the GLCM features. This question was raised in [203], authors classified and compared different rice grain types using GLRLM features and GLCM features. Results show that GLRM features outperforms GLCM features for bulk rice grain classification and this conclusion was made for three various types of classifiers.

In general, Fractal dimension gave satisfying results for our problem. When comparing texture features, it seems that the fractal dimension was the best and worked for the majority of the cases for detecting diseases, grading them and it was also shown that the fractal dimenison is the best texture feature when it comes to identifying a disease. This finding was reported in another study from the medical field [204]. Authors claimed that on its own, fractal dimension analysis gave highly significant results for the separation between colon cancer and normal tissue. Since, the Fractal dimension was the best technique among the other traditional texture features; we detailed in ANNEX E its calculation.

Eventhough the Fractal dimension behaved very well, it wasn't the best for different cases. In practice, we couldn't find one type of texture parameters that could be the top best for all grapevine varieties, for all diseases and for all stages. Our textural processing was, similarly to our previous reflectance

analysis, case-dependent. GLCM, LoG, Wav, Fractal, GLRM, Grad and Hist were only successful in certain configurations. Instead of checking what parameter could work for what case and building a classification model for each type of texture, we suggested a method that concatenates texture parameters, selects the best among them and builds one classification model. The results were similar to the case where each texture technique is considered solely.

To avoid a dimensionality problem, we inserted a variable selection stage with the ReliefF algorithm and kept only the best 10 texture features from the three bands (L, a, b). In general, the simplification of the features largely reduces the calculation burden, and it also helps improve the performance of classification models. A study on citrus diseases [130] showed that a reduced hue and saturation feature set, emerged as the best data model for the task of classification achieving accuracies of 96%. When texture features corresponding to H, S, I, HIS and HS are compared in [205], results concluded that Model HS achieves the highest overall classification accuracy. One more advantage of using HS is the decrease in computation time for training and classification because of the elimination of intensity features. A similar result was demonstrated in [206], where 13 texture features were selected for the color combinations of HSI. The model using 13 selected HIS texture features achieved the best classification accuracy (96.7%) with respect to the one using all 39 HSI texture features. Since Lab was found to be the best color space, the results suggest that it would be best to use a reduced L, a, b texture feature set to detect grapevine leaf conditions.

The texture feature concatenation and reduction procedure worked perfectly for FD, Esca and BN detection. Results were good for grading disease severity stages. The method also gave promising results when trying to identify diseases (FD vs Esca; FD vs BN).

The FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA region was feasible using texture features early in the season and nearly perfect at the end of the season. The Esca detection using texture data on a red grapevine variety (Cabernet-Sauvignon) in the Nouvelle-Aquitaine region was perfect when tested early or late in the season. Similarly, the BN detection with texture on a white grapevine variety (Chardonnay) in the Burgundy region was perfect when tested early or late in the season. In general, binary classifications gave better results than multi-class classifications, the classifiers could easily detect the presence of a disease but found it more difficult to determine the stage of an infection or to differentiate many infections. The FD and Esca were successfully separated on red-varieties (Cabernet-Sauvignon from Nouvelle-Aquitaine and Marselan from PACA). The attempt to separate FD from BN was more complicated and this is somewhat expected since the symptoms of FD and BN are indistinguishable to the naked eye.

6.3 GLOBAL DISCUSSION

This Chapter demonstrated that in addition to reflectance, texture analysis could very interesting for grapevine disease classification. Our finding was in accordance with many scientific research studies.

First, in Oklahoma and Texas, two winter wheat fields, three field experiments and one green house were studied in [27]. Authors were looking to quantify damage by greenbugs by investigating the use of spectral reflectance together with digital image analysis. The reflectance data was used to calculate a set of spectral vegetation indices deduced from the literature. The image data were used to estimate the mean percentage of damage caused by greenbugs. Correlation analysis showed strong

relationships between the vegetation indices and the severity of damage by greenbugs and the correlation coefficients reached 0.98.

Second, a hyperspectral imaging system collected images of healthy and diseased tomato sample leaves. The idea of [207] was to check whether this technique could determine the presence of tomato yellow leaf curl disease. Sensitive wavelengths and band ratios were defined by reflectance data analysis. Then, 24 texture features were extracted using grey level co-occurrence matrix and evaluated by a receiver operator characteristic curve analysis. The result for a validation set was excellent (100%).

Finally, a spectro-radiometer collected the reflectance spectrum in the 450-1000 nm range for healthy and rust diseased wheat. Work [208] took also RGB digital images for each sample; ratio of the area affected by the infection with respect to the total area was calculated. The reflectance spectra were strongly affected by the presence of the diseases; this was clearly shown in the first-derivative of the spectrum. The research discovered a direct correlation between the first derivative of the spectral data and the severity of the infection.

Other applications from the agricultural field but not related to disease detection also had similar results when it comes to joining reflectance information and image analysis. Visible and near infra-red corn seed hyperspectral images were classified by their varieties based on combining many features in [99]. In fact, authors developed classification models based on a combination of reflectance, morphological and texture features. They tested Support vector machines and partial least squares–discriminant analysis. Produced results demonstrated that this new procedure enhanced the correct recognition rate. In fact, the SVM model was more efficient than PLS-DA and scored higher than 96.3%.

The idea of testing reflectance and textural features was also found in other fields than agriculture. For example, in the food industry, heme and non-heme iron contents in meat products were estimated [209]. Authors calculated texture features from MS images (405-970 nm). Results showed that the combination of reflectance and textural features was more effective than when the spectra or textures were used alone for evaluating heme and non-heme iron contents. This finding suggests that MSI technology might be a good alternative to the time-consuming and destructive chemistry method and might become a valuable direction in the food industry.

In the scientific literature, texture features weren't only associated to reflectance data in the visible and infra-red regions but also to fluorescence information. Cases of biotic and abiotic stress were evaluated in [210]. At the beginning, barley was subjected to different nitrogen supply and fluorescence ratios were identified to characterize the stress and draw conclusions about leaf pigments status. Later, shape analysis described the appearance of three structural different kinds of fungal symptoms. Shape descriptors demonstrated great ability in successfully discriminating fungal infections. An advanced multi-color imaging system, based on chlorophyll fluorescence and blue-green fluorescence, in addition to some shape descriptors, is capable of detecting and characterizing local stress symptoms in leaves.

6.4 CONCLUSIONS

FD infected samples are studied for four different kinds of varieties (Marselan, Grenache, Vermentino and Chardonnay) in PACA region. Esca and BN affected leaves from Cabernet Sauvignon in Nouvelle-Aquitaine and Chardonnay in Burgundy, respectively, were also checked.

Complete spectral data and a set of traditional SVIs were assessed for disease detection. Two generalized feature selection methods were proposed: SPA-KNN and GA-KNN. Both methods were better than the use of complete spectral data. GA-KNN was also better than traditional SVIs but this wasn't the case for SPA-KNN. GA-KNN also proved better accuracy than SPA-KNN but had the disadvantage of being slow and computationally expensive.

For a practical case of MS imaging, we proposed the design of specific spectra disease indices based on the generalized feature selection methods already introduced. SDIs based on SPA-KNN outperformed complete spectral data, SVIs and the generalized SPA-KNN and GA-KNN for most cases. SDIs based on GA-KNN on the other hand didn't only outperform complete spectral data, SVIs and the generalized SPA-KNN and GA-KNN but also achieved better accuracies than the SDIs based on SPA-KNN. Hence, SDIs that are GA-KNN based proved to be the best approach. Here, four effective wavelengths were selected and the accuracy of classifying reflectance data in a binary context was higher than 90%.

RGB color images are taken for each sample leaf, leaf pixels are segmented and texture features are extracted from the Lab color model but only the optimal top 10 features are kept, like previously explained. The accuracy reached by the combination of these 10 texture features for all cases was higher than 85% in detecting disease. The accuracy was also higher than 73% in grading the infestation and higher than 72% in identifying the infection.

The results of this study indicated that remotely sensed data recorded by a hyperspectral spectrometer and a digital camera have the potential to aid in monitoring damage in grapevines growing under field conditions. Furthermore, by combining different features, we were able to distinguish between Esca and FD on one hand and BN and FD on another hand, with very promising degrees of accuracy.

At the end, we can say that grapevine disease detection seems to be possible using a MS sensor. In addition to the reflectance data, texture features from colored images were found very interesting; these could also be extracted from the MS images and used to further refine the classification.

CHAPTER 7

7 CONCLUSIONS AND PERSPECTIVES

7.1 CONCLUSIONS AND PARTICULAR PERSPECTIVES

In CHAPTER 2, we mentioned that visual inspection and biological techniques are frequently used to monitor disease occurrence in crop fields. However, they have some limitations. For example, visual inspection requires experts to check each plant looking for symptoms which is time-consuming and expensive. Biological techniques, on the other hand, like PCR and ELISA, require specific procedures or the results will be unreliable. Hence, these techniques can be complicated.

Highly accurate and reliable systems are required to replace traditional ones and to respond to the worldwide spread of crop diseases. In fact, non-contact systems are particularly needed to help farmers and producers in automatically identifying initial symptoms of plant diseases. Among non-contact systems, RS tools are of great interest, as they can acquire and analyze energy reflected by a phenomenon and hence monitor the status of vegetation in a non-destructive way. RS devices can go even further and identify infected zones in a field by defining the spatial locations of the damage spots in an early manner. Thus, RS may provide a proper tool to analyze large fields quickly and to allow the eradication of primary infections early, avoiding secondary spread. Moreover, nowadays, tremendous progress is achieved in designing different types of platforms, highly sensitive sensors and implementing multiple data analysis techniques.

Several studies were found in the literature discussing the use of RS in viticulture and some of them investigated in-field disease detection. The symptoms which appear on an infected plant are the result of the physiological alteration induced by the pathogen. The amount of water, the concentration of pigments and the functionality of the tissue, are all factors that may vary depending on the interaction between the pathogen and the host and may also modify the reflectance signatures of the plants. Hence, besides color imaging, visible/infra-red spectroscopy and multispectral/hyperspectral imaging were all tested. However, it seems that multispectral or hyperspectral imaging systems are probably the most convenient means for grapevine disease detection under field conditions.

In general, the imaging technology has emerged as a powerful tool in agriculture but it is still not much exploited for vines. The effort that has been undergone for the moment to accomplish early detection of grapevine diseases automatically is minimal, with respect to the major impact of those diseases on grapes yield and wine quality. The diseases need to be detected not only in symptomatic hosts in their natural production environment but also in asymptomatic hosts when the symptoms are not so obvious. A vine disease must be studied in many cultivars including those with white and red-grapes and also at different stages of severity. Many studies targeted the detection of powdery and downy mildew and grapevine leaf-roll in grapes and so it seems essential to explore other diseases and stresses.

Many grapevine diseases exist, a special emphasis was given to grapevine yellows ('Flavescence Dorée' and 'Bois Noir') and to Esca in this thesis. 'Flavescence Dorée' (FD) is the most dangerous among them since it is contagious and incurable.

Many projects aimed to diagnose grapevine diseases; among them a French research project called 'DAMAV' was initiated. It gathered academic and industrial entities in order to provide an innovative tool that will enable the diagnosis of the symptoms of FD (and other diseases that have an impact on the foliage) without mobilizing significant human resources and limiting phytosanitary treatments. A dedicated MS high-resolution sensor will be designed and embedded on a UAV as a specific solution for the automatic grapevine disease detection. This PhD, which is a part of the DAMAV project, aimed to participate in the design and implementation of the acquisition system and to develop algorithms based on the most relevant spectral and texture features related to specific grapevine diseases.

A phase of data acquisition is first performed and explained in CHAPTER 3. Hardwares, softwares and overall protocols for correctly acquiring hyperspectral reflectance and RGB digital images were explained. Three French regions (Provence-Alpes-Côtes d'Azur, Nouvelle-Aquitaine, Burgundy) were considered during two consecutive years (2016-217). Red (Marselan, Grenache, Cabernet-Sauvignon) and white grapevine (Vermentino, Chardonnay) varieties were both tested. Many factors such as the operator, the instrument, the vegetation itself and the atmospheric conditions can compromise the accuracy of the acquired spectral measurements; most of these factors seem to be uncontrollable but they should be noted anyways during the acquisition phase. Parameters affecting the accuracy of the acquired image data were also mentioned (ANNEX A).

In ANNEX B, a visual and statistical anlaysis of reflectance data are presented. From a visual point of view, when FD-infected samples from red and white grape-varieties were considered, we found that the variation of infected spectra with respect to healthy ones was variety-dependent. This is somewhat predictable since symptoms of FD don't look the same for both varieties. We also concluded that the patterns of FD and Esca on a red variety are similar; same goes to FD and BN on a white variety. The potential of our reflectance data was assessed from a statistical point of view using a PCA. Globally, we found that the variation of spectral data due to FD, BN and Esca with respect to healthy cases was significant. However, the variation between FD and BN on one hand and FD and Esca on another hand wasn't significant. Furthermore, the variation between two infestation levels of FD wasn't significant. This implies that spectral analysis could indicate the presence of a disease but isn't reliable when it comes to identifying the infection or grading it. Our finding was in accordance with biologists who claimed that the three studied infections cause common non-specific responses in leaves. Since the changes in contents are very similar for the three infections, spectral analysis can't be expected to successfully discriminate among them; so additional parameters should be considered for this purpose. An enhancement to be considered is replacing the statistical PCA analysis with an ANOVA-Simultaneous PCA analysis (ASCA). It would have given in addition statistical indications of the significance of the factors studied and their interactions.

Traditional spectral data analysis was first tested. Using the complete spectral data (ANNEX C) raised data dimensionality problems. Thus, a dimension reduction step is needed. In order to select the most interesting spectral features, that will be implemented later as filters on the Multi-Spectral (MS) sensor, we needed a variable selection technique. We tested traditional SVIs first in CHAPTER 4. Based on the results in CHAPTER 6, we recommend using the ARI vegetation index for detecting diseases on red varieties, GM1 for white varieties and mCAI for combined red and white varieties. The best SVIs

are shown in ANNEX D. These indices are however general and we wanted to conceive better ones for our disease detection task.

Two new dimension reduction techniques were proposed next in CHAPTER 4. For the practical MS Imaging case, we calculated new SDIs that challenge SVIs. The first method uses SPA. It achieved better results in general with respect to common SVIs. Advantages of the SPA technique include robustness and speed of computation. The second method relied on using GA for the variable selection task. This technique gave promising accuracies with respect to common SVIs. The GA's limitation included high computational load, processing time and complexity. Possible improvements that might be considered are the following. SPA selects variables based on their variance, regardless of their class (sick or healthy), which means that the inherent nature of SPA is unsupervised. This might explain why the GA performed better than SPA. In the future, we think that it is advisable to test another algorithm such as CovSel which is a modified SPA. CovSel (Covariance selection) is a method for variable selection [211], it can be considered a hybrid method as SPA, from which it takes inspiration. CovSel is well adapted to multi-response calibration of spectrometers and can be applied to the problem of discrimination considering indicator variables as responses. For GA, an improvement could consist on reducing the search space or maybe choosing better non-random initial values closer to the solution so that the convergence can be achieved faster. Moreover, many parameters for GA algorithms were chosen empirically; another improvement could be also to look for automatizing this choice.

The SDI we used was designed as a combination of a weighted ratio of two reflectances and a normalized difference of two reflectances. Additional possible improvement would be trying another SDI formula; one obvious example is replacing the weighted ratio of two reflectances with a normalized difference of two reflectances. Such new SDI formula might be assessed and compared to the one proposed in this thesis in terms of effciency.

The final selected spectral features were indeed a function of the grapevine variety and of the disease itself. Theoretically, they can be anywhere from 400 nm till 2400 nm. Affordable filters can go upto 1000 nm. To ensure a practical use, we advise to only study the range from 460 nm and 955 nm either from the beginning or just before applying the exhaustive search for the SDI calculation.

Since the analysis was done on a leaf-level scale and that using a UAV might introduce noise on spectral data, other features related to the appearance of the disease need to be taken into account in order to enhance the classification results. Color images were recorded for each sample leaf. Among RGB, Lab and HSV, the Lab color space was the most convenient one for texture analysis in the context of our application. The band 'a' that holds information related to color was the most informative one. For each Lab-colored image, many texture features were calculated and then significant eliminations of redundant texture features were accomplished through the use of the ReliefF technique (CHAPTER 5). In fact, only the best ten features were kept each time, this resolved the curse of dimensionality and offered satisfying results. The texture features adopted were: histogram moments, Haralick moments, gradient parameters, GLRM, fractal dimension, Wavelet parameters and LoG parameters. These were also effective in detecting many diseases in the medical field.

The potential of our texture data was assessed from a statistical point of view using a PCA. We found that texture data extracted from colored images were in general capable of not only detecting

infections but also grading them and differentiating them. Disease classification and severity grading of FD, BN and Esca was possible. The discrimination between FD and Esca was satisfying but the discrimination between BN & FD was difficult and this was expected since both induce same visible symptoms. This statistical analysis should be tested on texture data acquired from MS images.

If only one texture parameter needs to be used, we recommend the fractal dimension; it performed the best in the majority of the cases in CHAPTER 6. The detailed results of the fractal dimension calculation are exposed in ANNEX E. To improve the results, one can imagine testing other color spaces than RGB, Lab and HSV that might be more convenient for grapevine disease detection on colored images. Furthermore, other texture parameters exist in the scientific literature and we weren't able to test them all. One could also consider nesting some texture features, for example, calculating second order parameters (Haralick) on the resulting image of the gradient, the Laplacian of gaussian or the Wavelet. Another possible texture analysis track could consist on a classification of the colored images at a pixel-level scale. For example, we could verify the performance of the first or second order moments in a classical or variable neighborhood. As a result, instead of having one vector of features per image, we would have one vector of features per pixel.

Paper [212] proved that color balancing models allow for a significant improvement in the accuracy of texture recognition for many CNN architectures. The feasibility of the approach is demonstrated by the experimental results obtained on the RawFooT dataset. Texture images were acquired under several different lighting conditions and the study concluded that color balancing do have an effect on texture calculation and on classification results. During our color image acquisitions, the white balance was put into automatic mode. This means that the color intensities were changing depending on the lighting conditions. As a result, the texture calculation and the classifications might be affected. We advise to manually tune this parameter so that no modification can occur under natural illumination conditions. In the research [213], the luminance was shown as a function of daytime during the month of June. If we limited the duration from 10 o'clock in the morning till 4 o'clock in the afternoon, the period in which we acquired our data in the field; we can find a luminance variation of +/-20%. In the research [214], several features have been proposed ranging from traditional statistical descriptors to features extracted with neural networks. An extensive comparison of old and new texture features with a particular focus on how these features are affected by small and large variations in the lighting conditions was given. The evaluation is performed on a texture database, which includes 68 samples of raw food acquired under 46 conditions that present variations of light color, direction, and intensity. The database allows to investigate the robustness of texture descriptors across large variations of imaging conditions. Shots have been taken while illuminating with neutral light (D65) at different levels of intensity (100%, 75%, 50%, and 25% of the maximum achievable level). Authors found that when the intensity of illumination varies of +/-20%, the accuracy of traditional texture parameters (histogram moments, LBP and Gabor features) in some classification tasks varied about +/-10%. A +/-10% variation in the classification accuracy isn't critic for our case since the texture accuracy was quite high in our study. Anyway, in the future, more tests need to be done for assessing the effect of illumination variation on the performance of our exact texture parameters.

Other type of features different than texture such as the spatial propagation of the disease or maybe the shape of the disease form interesting factors that could be considered in the future for the aim of grapevine disease identification.

The SVM classifier was used due to its popularity in the field of spectral and image data processing and gave good results.

Strong correlations between spectral/textural data and grapevine damage suggest that multispectral imaging sensors can be used as quick, nondestructive, repeatable, and cost-effective techniques to detect many types of damage in vines. This study was thus the first step to use reflectance measurements along with digital image analysis to estimate damage by Yellowing and Esca on many grapevine varieties. The proposed methods can be transferred to other crops varieties and can used for different kinds of diseases or biotic and abiotic stresses in crops.

7.2 GENERAL PERSPECTIVES

The correct discrimination between different vine diseases is a difficult task. Indeed, different diseases can cause very similar symptoms, and sometimes, the same virus can cause different symptoms depending on the grape variety; symptoms can also be due to a combination of diseases occurring at the same time. Some factors such as bad weather conditions, variance in soil minerals and nutrients, pollution, and pesticides can induce virus-like symptoms; the time of infection and the overall environment can also affect the appearance of symptoms, and finally some viruses might not cause any visual symptoms of infection. As a result, a wide range of phenomena can give identical leaf symptoms (water stress, nutrient deficiencies, broken branches, mechanical operations, different diseases such as leaf-roll, red blotch, and others). Expected accuracies are true in the ideal case where the only symptoms to be detected are those of the considered diseases. Hence, suggested analysis need to be tested in different environments, at different field conditions; other types of diseases or stresses and grapevine varieties must also be taken into consideration.

Our study focused on disease detection at a leaf-scale level. In fact, measurements were taken through direct contact with the sample. Based on these findings, it might be possible to conceive a multispectral camera and mount the sensor on any movable platform to localize infection foci in a field. A preliminary version of the MS sensor was made in our laboratory (ANNEX F), this version isn't optimal and therefore not ready to be embarked on the drone. Future research, part of them being included in the DAMAV project, will mainly focus on incorporating the MS sensor on a UAV platform to expand the study to the whole field or landscape level. The full system (UAV + sensor) is in the finalization phase and new acquisition campaigns are expected in the summer of 2019. Additional work will address the processing of the associated multispectral images. If we want to examine a complete branch, or maybe the whole grapevine, additional corrections need to be taken into account. Geometric and radiometric improvements capable of solving shadowing problems, branch structures and interfering reflectance from other surrounding objects are necessary.

The utility of the developed texture analysis will be assessed on a set of multispectral images obtained from the new sensor instead of RGB colored ones. It should be mentioned that the texture processing employed in this study includes traditional feature extraction algorithms along with a traditional machine learning task. New approaches such as deep learning could be tested. The advantage of this technique is that we, no longer, have to calculate, select the features and inject them to the classifier. Instead, with deep learning, the neural network chooses the important features and trains itself automatically. The main drawback of this technique is that the learning time is extra-long, many samples are needed for training (millions sometimes) but the classification task itself is very fast and precise.

Furthermore, when adopting a multispectral sensor, we will not only have spectral information but also spatial data. This available spatial data will enable adding advanced image processing algorithms

other than texture analysis, to yield a more robust detection of FD. Actually, two other FD symptoms cannot be detected spectrally; so, in order to better conclude the presence of the condition, additional pattern recognition algorithms can be integrated directly in the sensor to detect the absence of lignification and the berry mortality that are additional indicators of FD presence.

Finally, to our knowledge, this PhD is the first to investigate the potential of the combination of textural and reflectance information for identifying foliar symptoms of Yellowing and Esca vine diseases directly under field-conditions. This study is done at a leaf-level scale but will be extended to the canopy level in the future.

Intensive research will identify new sensor-based methods for the detection, identification and quantification of grapevine diseases. Modern progress in RS techniques has the aptitude for boosting vineyard management and ensures that the vineyards management of the future become scientifically and automatically handled. Therefore, it is crucial to focus current research on developing real-time, universal algorithms that are rapid and precise, aiding in the localization task of established diseased foci in canopies.

Developing and testing new and emerging RS-based technologies are the current main focus areas of automating crops in general and vineyards in particular. Areas of application may not only include disease detection but also fruit harvesting, fungicide application scheduling, agricultural practices evaluation and others.

ANNEX A

A. PARAMETERS AFFECTING DATA

PARAMETERS AFFECTING THE REFLECTANCE DATA

THE OPERATOR

To prevent errors caused by the operator, he must carefully respect the data acquisition protocol mentioned in CHAPTER 3.

Instrument settings parameters like the dark current integration, the white reference integration and the number of averages per sample should be the same for each acquisition. The spectrometer should be warmed-up before use. It is important to use the same warm-up time (at least 15 minutes) for each acquisition. Operations like optimization and white reference measurements should be performed each time the illumination conditions change. The probe must be well-maintained during acquistions. It is advised to use a leaf clip, if it is not available then use a black support behind the leaf and hold the probe in a stable perpendicular position to the target during measurement.

THE MEASUREMENT DEVICE

The spectro-radiometer contains an inherent error which is related to the instrument's precision. The ambient temperature is a factor that can affect the instrument's accuracy since DC noise is sensitive to temperature. It is recommended not to take measurements when the temperature is above 40 degrees.

THE OBJECT UNDER STUDY

Samples from different plants or even from the same one can produce contrasting reflectance measurements. The dissimilarity between the data obtained is due to spatial and temporal variability of the sample characteristics on one hand and to the soil characteristics on the other hand.

The reflectance data measured on a canopy scale (using a UAV equipped with a multispectral or hyperspectral camera for example) depends on the branching angles and the texture of the target (diffuse, specular) because a surface with rough texture will tend to have a higher proportion of illumination from the diffuse and scattered surrounding sources relative to the direct solar illumination when compared with smooth surfaces. Homogeneity or heterogeneity of the canopy (monoculture or mixed community), described by percent cover of each component as well as differences in phenology of the target species, presence of a single layer or multiple layer, a uniform cover or distinct clumps across the plot, are all factors that can influence the measured spectra. The height of ground cover and phenology (vegetation health and growth stage) are also to be taken into consideration. Finally, the type of canopy and its closure, systematic or unsystematic orientation of rows are additional factors that can affect the precision of the measurements.

The reflectance data measured on a leaf scale (using a spectro-radiometer for example) depends on the leaf area index which defines the area that interacts with solar radiation, on the leaf size (age), the leaf density at the branch level and the leaf angle distribution as this may change during the day while leaves orient themselves toward or away from the incident radiation. Leaf optical properties may further vary during the day and during the seasons; indeed, variations in light intensity, air temperature

133

and relative air humidity leads to diurnal variations of photosynthetic rate and leaf relative water content.

Soil characteristics such as its texture, color and moisture content can affect received reflectance data on both a canopy and leaf level.

ENVIRONMENTAL CONDITIONS

The illumination can influence the accuracy of reflectance measurements. If it is natural, then it is uncontrollable and two measurements are required: the reflectance response of a reference sample and the one of the target sample. The reflectance is computed by dividing the spectral response of the target by that of the reference sample. An inherent assumption is that the characteristics of the illumination are the same for the reference and target materials.

The direct solar illumination is the dominant source of illumination of a target, but it is possible to have several sources of light scattered from surrounding objects each with its own unique reflectance distribution. The illumination of the target surface could be affected by different factors: the operator, the measurement instrument and the surrounding objects. These could form a source of scattered light, diffusing light on the target or could also cover the target preventing direct illumination.

The sun angle and position are important factors to consider. Between the highest position of the sun and that of the sun lying low in the horizon, irradiance varies. Solar zenith angle can become a critical parameter because the column density of water vapor in a given atmosphere increases rapidly as zenith angle increases. This last one decreases from its minimum at vertical either with time of day or season because, as water vapor absorption increases, solar irradiance decreases and this means greater difficulty in detecting spectral features throughout the short wave infra-red but especially near water band locations.

Absorbing molecules in the atmosphere strongly modifies the incoming solar irradiance. By far, water vapor has the strongest impact on the incoming solar spectrum. Other major atmospheric components that influence the atmospheric transmission are carbon dioxide, ozone, oxygen, carbon monoxide, methane, nitrous oxide …

Field measurements are therefore commonly restricted to a period around solar noon when the solar geometry is changing least and when the errors due to the angular response of the reflectance panel are at a minimum

Clouds, smoke and haze also attenuate solar irradiance by absorption which results in scattering that contributes to the secondary source of illumination and variable irradiance as a result of changing conditions between target and reference measurements. The presence of a cloud cover could boost the diffusion of solar illumination since it creates shadows decreasing the contrast between surfaces with different surface textures.

Spectra should be measured under illumination conditions comparable to those at the time the image was captured.

PARAMETERS AFFECTING THE IMAGE DATA

THE OPERATOR

Many techniques available in digital cameras can alter the appearance of an image. Image processing operations may be used to enhance original images, to convert master images to derivative images of

reduced quality, or to prepare master or derivative images for display or printing. Image processing operations include but are not limited to:

- Resampling-changing the image's pixel dimensions
- Adjustment of brightness and contrast
- Color correction
- Rotation
- Compression

Some camera features like color balancing for example can be manually tuned by the user and affect the contrast and hence the overall appearance of objects in the images. The distance from which the image is taken, which is defined by the user, affects the scale and then how objects look like in the images.

The parameters manually tuned and the distance to the object of interest, should be maintained the same during the image acquisition protocole in order to correctly interpret the images.

THE IMAGING DEVICE

There are several parameters that characterise the quality of digital images. Resolution, noise, and artefacts are the main parameters of image quality.

Resolution is of three main categories, spatial resolution (space), contrast resolution (intensity) and temporal resolution (time). Noise is produced by the statistical fluctuation of value from pixel to pixel. The noise sources are of three types, namely quantum noise, electronic or detector noise and computational or quantisation noise. Digital image artefacts can be caused by hardware or image receptor artefacts, and software artefacts

Those parameters are defined by the designers of the sensor. Like any other sensor, digital cameras can also have an internal error that may affect the fidelity of the image to the real life.

THE OBJECT UNDER STUDY

Other digital image artefacts are caused by object artefacts. Specular, glossy or reflective surfaces that could be found in real scenes aren't perceived the same way. The object imaged plays also a role in how the image looks like. The surfaces studied in our case are vegetative living tissues which can change over time with age meaning that their appearance in the image could change too.

ENVIRONMENTAL CONDITIONS

The environment can have a surprisingly large effect on perceived colors in fact, the ambient illumination can affect the color appearance in images. A color checker must be used to calibrate the images before processing them.

ANNEX B

B. ASSESSMENT OF THE POTENTIAL OF REFLECTANCE DATA FROM A VISUAL AND STATISTICAL POINT OF VIEW

VISUAL EXAMINATION OF REFLECTANCE DATA PATTERNS

In the general case explained in CHAPTER 2 the spectum has a strong absorption in the visible region due to photochemical pigments (chlorophylls and carotenoids). However, in the near infra-red region, the reflectance is higher; this is the result of multiple internal light scattering that is a function of the complexity of the leaf's structure. In the next paragraph, the mean reflectances of data acquired from different grapevine varieties and affected by multiple diseases are displayed. The sample leaves considered are ones starting to show visible symptoms. How the grapevine specie and the type of disease can alter the spectral reflectance, is discussed next. We used reflectance average calculation only for a visualization purpose. Despite this, we made sure in all cases that the distributions under study are all unimodal. A statistical test called the DIP test [215] was employed in order to check for the unimodality of the data, we also made sure that the histograms showing the data distribution were unimodal.

REFLECTANCE OF FD ON RED (MARSELAN) AND WHITE GRAPEVINE VARIETIES (CHARDONNAY) IN PACA REGION

When compared to healthy reflectance signatures, it can be clearly seen from *Figure B.1* that the reflectance for both diseased red and white berried leaves were somewhat different. Thus, this clearly shows that the reflectance response was affected by the infestation.

For the Marselan variety (red-grape), the peak in the green range appearing in the visible part of the infected spectra moved towards longer wavelengths. Moreover, the healthy spectrum was higher than the infested one in the visible region (mainly between 500-700 nm), the opposite occurred in the near infra-red region (800-1300 nm) but then in the infra-red region (>1300 nm) the healthy spectra raised a little above the infested one. It would seem that when the infestation level increases, the reflectance signature decreases in the visible region and increases in the near infra-red region. The same trend was also observed for Grenache white grape leaves (spectra not shown).

On the other hand, for the Chardonnay variety (white-grape), the peak in the green line in the visible part of the infected leaf remained in position for infected leaves. Furthermore, the healthy spectrum was lower than the infested one in the visible region (mainly between 500-700 nm), the opposite occurred in the near infra-red region (800-1300 nm) and in the infra-red region (>1300 nm). It would therefore seem that when the infestation level increases, the reflectance signature increases in the visible region and decreases in the near infra-red region. The same trend was also been observed for Vermentino red grape leaves (spectra not shown).

(a)   **(b)**
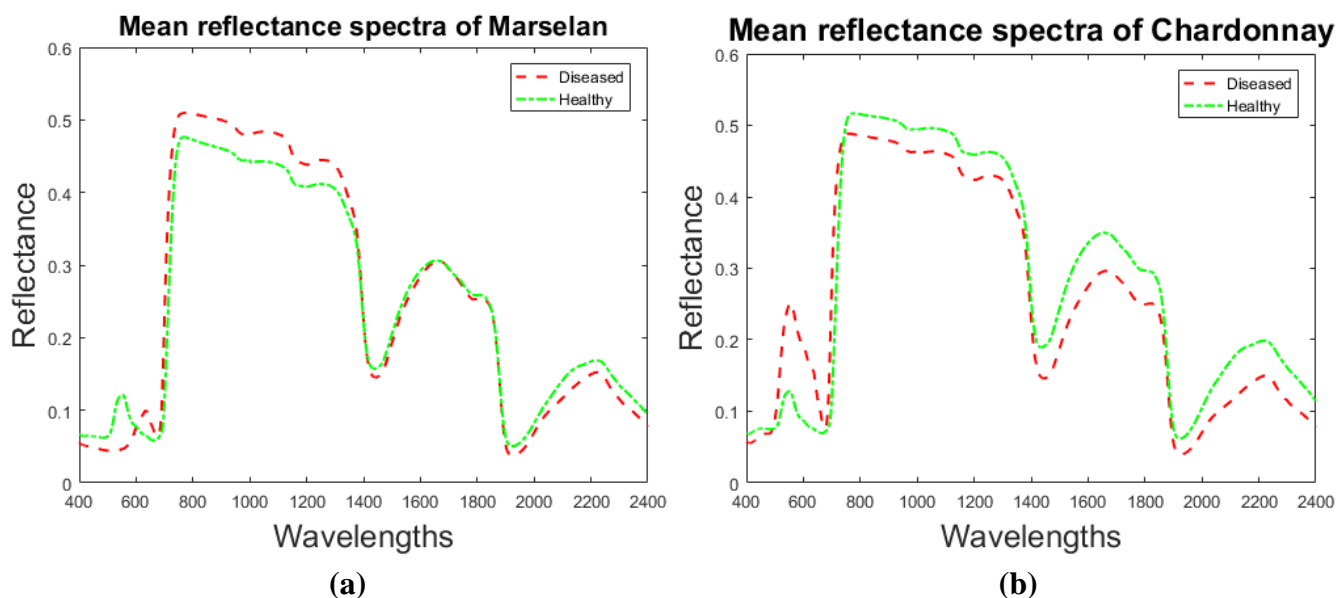
Figure B.1 Mean reflectance of FD infected and healthy for **(a)** Marselan grapevine leaves and **(b)** Chardonnay grapevine leaves. Diseased and healthy spectra are presented in red and green respectively.

We can conclude that above 1400 nm, the difference between the healthy and diseased spectra was almost the same in the red and white varieties. This overall tendency observed due to FD in red and white varieties was verified and found to be the same in the acquisition campaigns during 2016 and 2017 in PACA region.

Note that between 800-1400 nm, the reflectance changes depending on the structure of the leaf. For red varieties, the absorbance increased, but for white varieties, the opposite occurred. This is basically due to modified leaf tissue and the stunting of plants due to the 'Flavescence dorée' infection. The results obtained for the Chardonnay white-berried case, were also found for Russian wheat aphid damage in winter wheat [189]. For the Marselan red-berried case, damaged leaves reflected more energy than the control leaves in the near infra-red wavelengths, this was also found in Russian aphid-damaged leaves compared to healthy ones by [216].

REFLECTANCE OF ESCA ON RED GRAPEVINE VARIETY (CABERNET SAUVIGNON) IN NOUVELLE-AQUITAINE REGION

For the Cabernet Sauvignon samples infested by Esca from the Nouvelle-Aquitaine region, the signatures can be depicted in *Figure B.2*.

Figure B.2 Mean reflectance of Esca infected and healthy Cabernet Sauvignon grapevine leaves. Diseased and healthy spectra are presented in red and green respectively.

Here again, the spectral reflectance of healthy leaves wasn't similar to that of leaves affected by Esca, particularly in the visible and near infrared regions.

The peak in the green range appearing in the visible part of the infected spectra moved towards longer wavelengths and was reduced in intensity. In addition, the infected spectra were higher than the infested one in the near infrared region (mainly between 800-1300 nm). However, in the infrared regions (>1300 nm) both infested and normal spectra were very similar, indicating that the infrared regions might not be interesting for detecting Esca in red grapevine varieties.

REFLECTANCE OF BN ON WHITE GRAPEVINE VARIETY (CHARDONNAY) IN BURGUNDY REGION

For the Chardonnay variety infected with BN from the Burgundy region, the signatures can be depicted in *Figure B.3.*

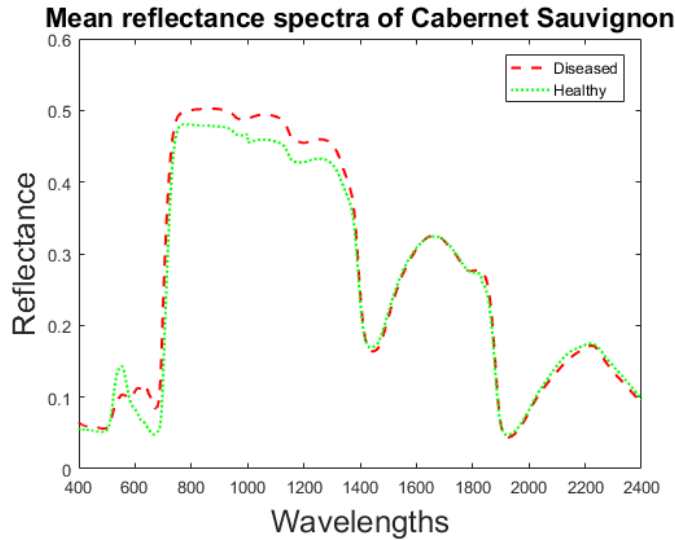*Figure B.3* Mean reflectance of BN infected and healthy Chardonnay grapevine leaves. Diseased and healthy spectra are presented in red and green respectively.

Here also, the spectral reflectance of healthy leaves presented a pattern slightly different than those affected by BN, in all parts of the spectrum.

The peak in the green area of the visible part of the infected leaf remained in position for infected leaves. Furthermore, the healthy spectra was lower than the infested one in the visible region (mainly between 500-700 nm), but the opposite occurred in the near infra-red region (800-1300 nm) and in the infra-red region (>1300 nm).

REFLECTANCE OF FD VS ESCA ON RED GRAPEVINE VARIETIES (MARSELAN AND CABERNET SAUVIGNON) IN PACA AND NOUVELLE-AQUITAINE REGIONS

When comparing the Esca disease reflectance signature on Cabernet Sauvignon leaves from Nouvelle-Aquitaine with respect to the FD reflectance signature on Marselan leaves from PACA region, we can say that both display the same trend along the spectrum.

There are very small differences in the visible part of the spectrum (*Figure B.4*).

Mean reflectance spectra of Cabernet Sauvignon and Marselan
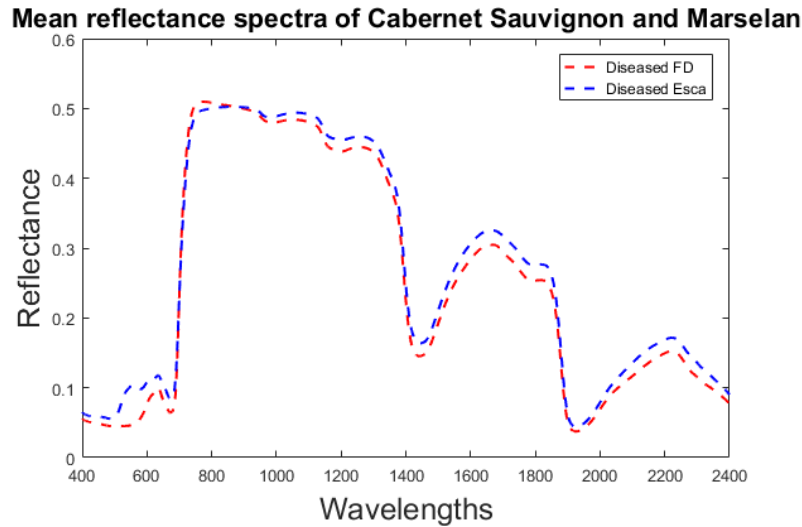
Figure B.4 Mean reflectance of Esca and FD infected and healthy Cabernet Sauvignon and Marselan grapevine leaves. FD and Esca diseased spectra are presented in red and blue respectively.

The green peak was shifted and reduced for both diseased signatures. This indicates that the visible region is poorly useful for discriminating both infections, this is somewhat predictable since both infections cause a reddish discoloration on concerned leaves.

Other tiny deviations can be found between the two cases in the near infra-red and infra-red regions. The difference between both cases is hence not significant; so even considering bands from the near infra-red or infra-red regions might not be sufficient for distinguishing FD from Esca on red-berried leaves.

REFLECTANCE OF FD VS BN ON WHITE GRAPEVINE VARIETY (CHARDONNAY) IN PACA AND BURGUNDY REGIONS

When comparing the BN disease reflectance signature on Chardonnay leaves from the Burgundy region with respect to the FD reflectance signature on Chardonnay leaves from the PACA region, we can say that both display the same trend along the spectrum. The remarkable resemblance between the two average spectra, particularly in the visible part of the spectrum (*Figure B.5*), explains why FD and BN exhibit very similar symptoms that can't be distinguished by the naked eye. This finding further demonstrates that the visible spectrum isn't suitable for FD and BN discrimination.
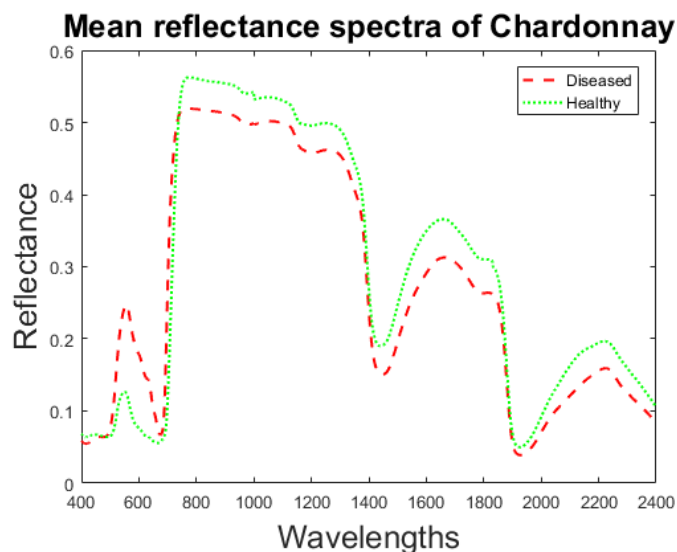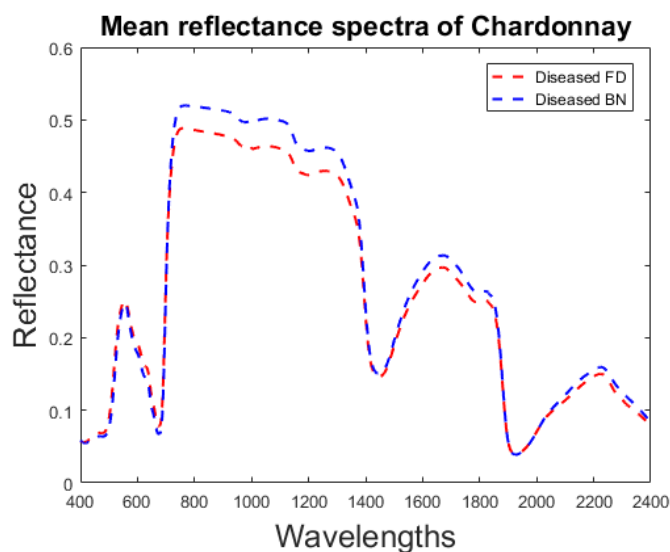
Figure B.5 Mean reflectance of BN and FD infected and healthy Chardonnay grapevine leaves. FD and BN diseased spectra are presented in red and blue respectively.

A very small difference between the two cases can be depicted in the near infra-red and infra-red regions. The difference between both cases is hence not significant, so even considering bands from the near infra-red or infra-red regions might not be sufficient for discriminating BN from FD on white-berried leaves.

STATISTICAL EXAMINATION OF REFLECTANCE DATA PATTERNS

Intra-class distance was calculated as the mean disctance between each point in the cluster and the centroid of the cluster. Inter-class disctance was calculated as the mean distance between all points in a cluster and the centroid of the other cluster.

DISEASE DETECTION

REFLECTANCE OF FD ON RED (GRENACHE) AND WHITE GRAPEVINE VARIETIES (VERMENTINO) IN PACA REGION

In *Figure B.6 (a)* is displayed the PCA analysis using only two principal components on healthy and FD diseased Grenache leaves from PACA region.

Four PCs explain more than 95% of the variance. Intra-class distances are calculated. For the FD diseased cluster it was found to be 0.39 and for the healthy one it was 0.84. Inter-class distance was also calculated, we found that the distance between the FD diseased and healthy cluster is 1.16.

In *Figure B.6 (b)* is also displayed the PCA analysis using only two principal components on healthy and FD diseased Vermentino leaves from PACA region.

Four PCs explain more than 95% of the variance. Intra-class distances were computed. For the diseased FD cluster it was 0.29 and for the healthy cluster it was 0.59. The Inter-class distance was also computed and we found that the distance between the FD Diseased cluster and the healthy cluster is 1.99. Same tests are applied on the the two other grapevine varieties (Marselan and Chardonnay) and we found similar results.

141

The Inter-class distance was higher than the Intra-class distances by a factor of two for both cases. This implicates that the FD disease detection using reflectance data is possible from a statistical point of view.
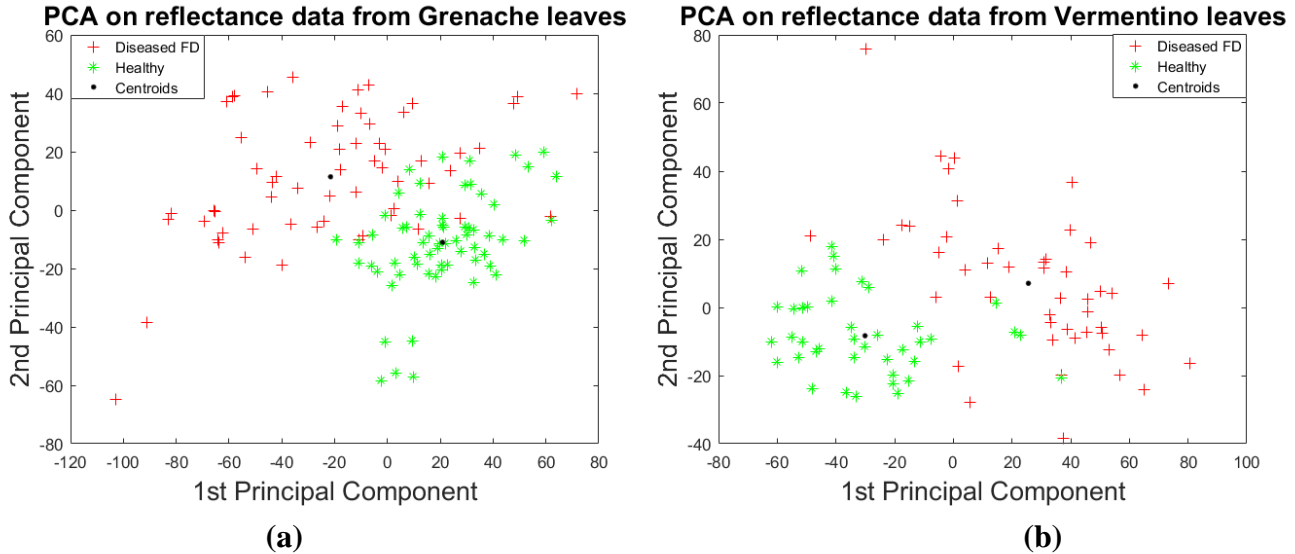
Figure B.6 PCA analysis on **(a)** Grenache reflectance data and **(b)** Vermentino reflectance data in PACA region. FD diseased and healthy data points are presented in red and green respectively. Centroids of the two healthy and diseased clusters are marked in black.

REFLECTANCE OF ESCA ON RED GRAPEVINE VARIETY (CABERNET SAUVIGNON) IN NOUVELLE-AQUITAINE REGION

In *Figure B.7* is displayed the PCA analysis using only two principal components on healthy and Esca diseased Cabernet Sauvignon leaves from Nouvelle-Aquitaine region.

Five PCs explain more than 95% of the variance. Intra-class distances were found to be 0.44 for the diseased Esca cluster and 0.15 for the healthy cluster. However, for the Inter-class distance, we found 0.81 from the Esca diseased to the healthy clusters.

The Inter-class distance was higher than the Intra-class distances for this case. This suggests that the Esca disease detection using reflectance data is possible from a statistical point of view.
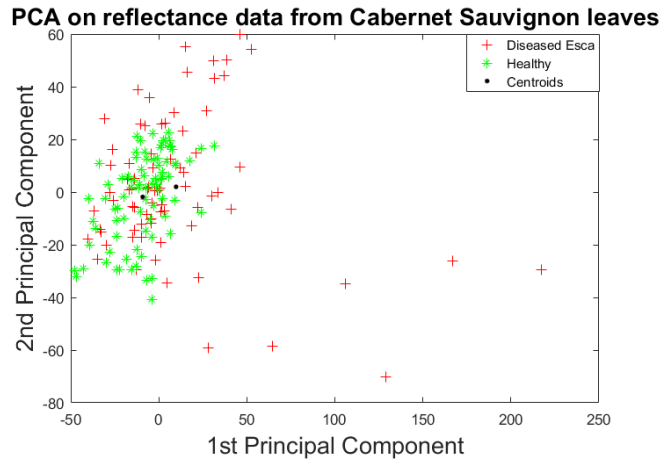
Figure B.7 PCA analysis on Cabernet Sauvignon reflectance data in Bordeaux region. Esca diseased and healthy data points are presented in red and green respectively. Centroids of the two healthy and diseased clusters are marked in black.

REFLECTANCE OF BN ON WHITE GRAPEVINE VARIETY (CHARDONNAY) IN BURGUNDY REGION

In *Figure B.8* is shown the PCA analysis using only two principal components on healthy and BN affected Chardonnay leaves in Burgundy region.

Four PCs explain more than 95% of the variance. The Intra-class distances were found to be 0.07 for the diseased BN cluster and 0.24 for the healthy cluster. For the Inter-class distance, we found a distance of 0.47 from the BN diseased to the healthy clusters.
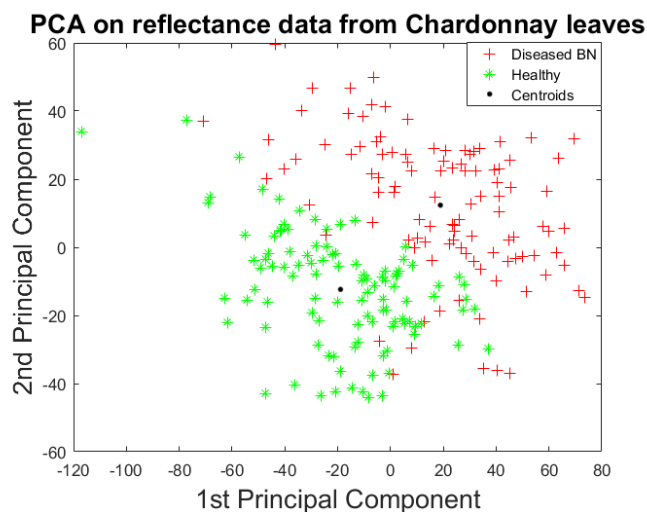
Figure B.8 PCA analysis on Chardonnay reflectance data in Burgundy region. BN diseased and healthy data points are presented in red and green respectively. Centroids of the two healthy and diseased clusters are marked in black.

143

The Inter-class distance was higher than the Intra-class distances by a factor of two for this case. This means that the BN disease detection using reflectance data is possible from a statistical point of view.

DISEASE SEVERITY STAGE DISCRIMIATION

REFLECTANCE OF FD ON RED (GRENACHE) AND WHITE GRAPEVINE VARIETIES (CHARDONNAY) IN PACA REGION

We can see in *Figure B.9 (a)*.the PCA analysis using only two principal components on FD diseased Grenache leaves with different infestation levels from PACA region.

Five PCs are used to obtain 95% of the variance. Intra-class distances were calculated. The distance inside the FD diseased Level 1 cluster is 2.48; for the FD diseased Level 2 cluster it was 0.74. Inter-class distance was also calculated and we foud a distance between the FD diseased Level 1 cluster to the FD diseased Level 2 cluster equal to 0.66.

Similarly, in *Figure B.9 (b)* , we can find the PCA analysis using only two principal components of FD diseased Chardonnay leaves having two infestation stages from PACA region.

Four PCs are used to obtain 95% of the variance. Here, the Intra-class distances were 0.97 and 0.92 for FD diseased Level 1 and FD diseased Level 2 clusters respectively. For the Inter-class distance, we found a distance of 1.04 between the FD diseased Level 1 and FD diseased Level 2 clusters.

Same tests are applied on the the two other grapevine varieties (Marselan and Vermentino) and we found similar results.

The Inter-class distance was just a bit higher than the Intra-class distances for both cases. This suggests that the FD disease severity stage identification using reflectance data doesn't seem quite feasible from a statistical point of view.

| (a) | (b) |

Figure B.9 PCA analysis on **(a)** Grenache reflectance data and **(b)** Chardonnay reflectance data in PACA region. FD diseased infestation level 1 and infestation level 2 data points are presented in blue and red respectively. Centroids of the two FD diseased severity 1 and FD diseased severity 2 clusters are marked in black.

REFLECTANCE OF ESCA ON RED GRAPEVINE VARIETY (CABERNET SAUVIGNON) IN NOUVELLE-AQUITAINE REGION

In *Figure B.10,* we can see the PCA analysis using only two principal components of Cabernet Sauvignon leaves affected by two levels of Esca in Nouvelle-Aquitaine region.

Five PCs are used to obtain 95% of the variance. For Intra-class distances, we obtained a distance inside the Esca diseased Level 1 cluster of 0.71; for the Esca diseased Level 2 cluster, it was 0.78. The Inter-class distance obtained between the Esca diseased Level 1 and the Esca diseased Level 2 was 1.01.

The Inter-class distance was just a little bit higher than the Intra-class distances. This means that the Esca disease severity stage identification using reflectance data doesn't seem completely possible from a statistical point of view.
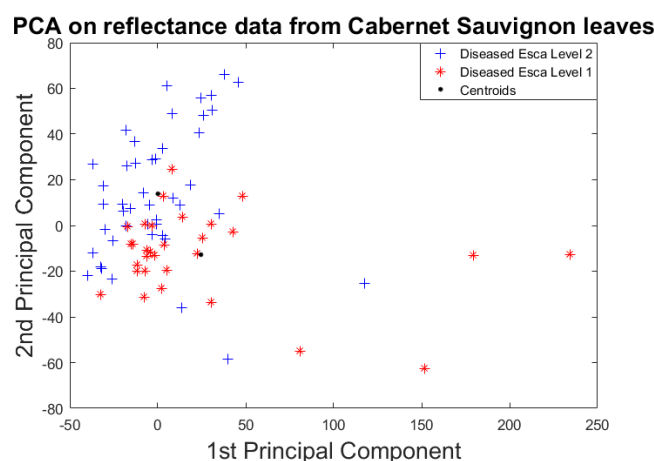
Figure B.10 PCA analysis on Cabernet Sauvignon reflectance data in Bordeaux region. Esca diseased infestation level 1 and infestation level 2 data points are presented in blue and red respectively. Centroids of the two Esca diseased severity 1 and Esca diseased severity 2 clusters are marked in black.

REFLECTANCE OF BN ON WHITE GRAPEVINE VARIETY (CHARDONNAY) IN BURGUNDY REGION

Figure B.11 displays the PCA analysis using only two principal components on Chardonnay leaves affected by two stages of BN in Burgundy region.

Four PCs are needed to explain 95% of the variance. The Intra-class distances obtained for each cluster were 0.26 and 0.32 for BN diseased Level 1 and BN diseased Level 2, respectively. Inter-class distance on the other hand was obtained from the BN diseased BN Level 1 and the BN Diseased Level two clusters and was equal to 0.42.

The Inter-class distance was just slightly higher than the Intra-class distances. This implies that the BN disease severity stage identification using reflectance data sdoesn't seem entirely feasible from a statistical point of view.

145

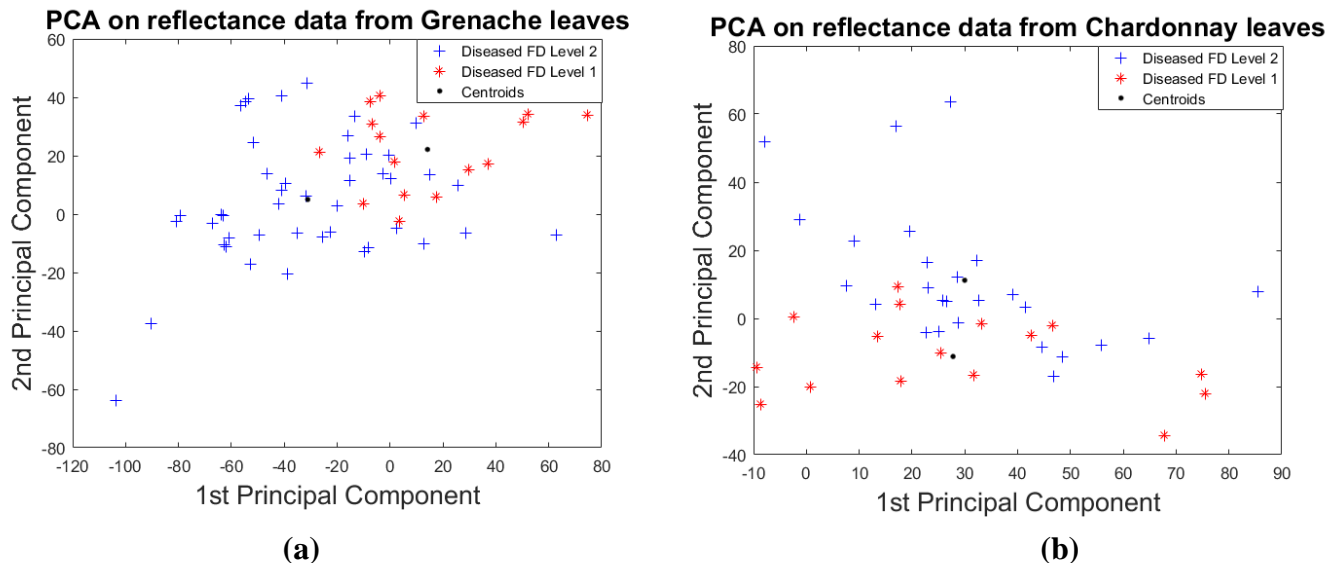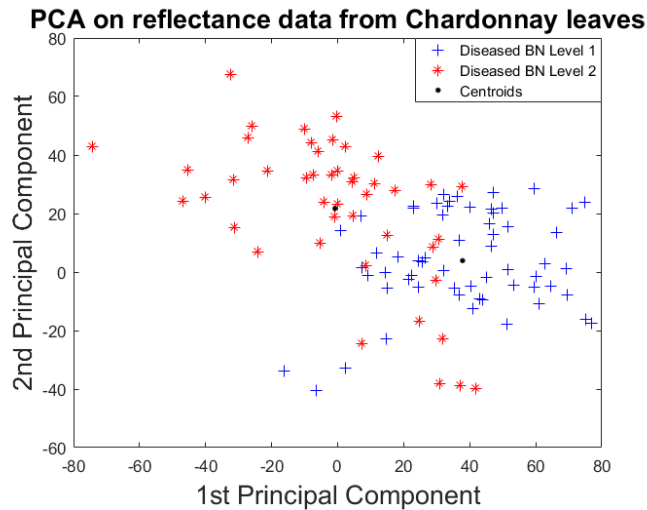*Figure B.11* PCA analysis on Chardonnay reflectance data in Burgundy region. BN diseased infestation level 1 and infestation level 2 data points are presented in blue and red respectively. Centroids of the two BN diseased severity 1 and BN diseased severity 2 clusters are marked in black.

DISEASE DISCRIMINATION

REFLECTANCE OF FD VS ESCA ON RED GRAPEVINE VARIETIES (MARSELAN AND CABERNET SAUVIGNON) IN PACA AND NOUVELLE-AQUITAINE REGIONS

In *Figure B.12* we can find the PCA analysis using only two principal components of Marselan (from PACA region) vs Cabernet Sauvignon (from Nouvelle-Aquitaine region) leaves affected by FD and Esca respectively.

Five PCs are used to explain 95% of the variance. The cluster of FD diseased samples had and intra-class distance of 0.45, the cluster of Esca diseased cluster had on the other hand an intra-class distance of 0.24. The Inter-class between the FD diseased and Esca diseased clusters is 0.44.
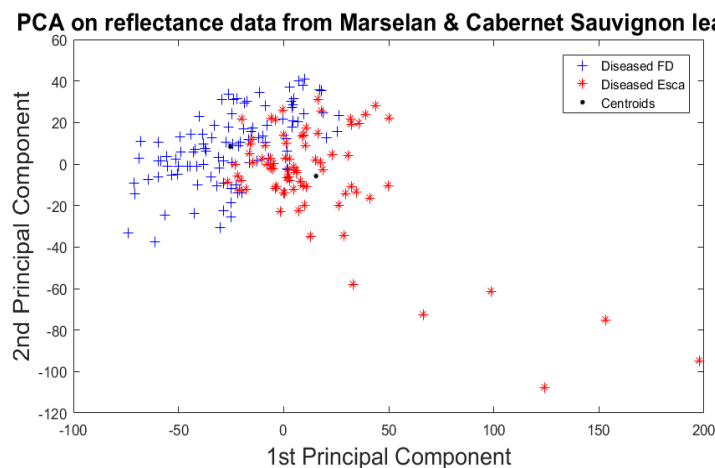
Figure B.12 PCA analysis on Marselan and Cabernet Sauvignon reflectance data in PACA and Bordeaux regions. FD diseased and Esca diseased data points are presented in blue and red respectively. Centroids of the two FD diseased and Esca diseased clusters are marked in black.

146

The Inter-class distance was very similar to the Intra-class distances. This indicates that the discrimination between FD and Esca diseases using reflectance data only seems difficult from a statistical point of view.

REFLECTANCE OF FD VS BN ON WHITE GRAPEVINE VARIETY (CHARDONNAY) IN PACA AND BURGUNDY REGIONS

Figure B.13 demonstrated the PCA analysis using only two principal components of Chardonnay leaves affected by FD and BN from PACA and Burgundy regions respectively.

Three PCs are necessary to explain 95% of the variance. The Intra-class distance of the FD diseased cluster was found to be 0.19; the same distance for the BN diseased cluster was found to be 0.04. The Inter-class distance between the FD diseased and BN diseased clusters was calculated and was equal to 0.19.

The Inter-class distance was very similar to the Intra-class distances. This indicates that the discrimination between FD and BN diseases using reflectance data only seems complicated from a statistical point of view.
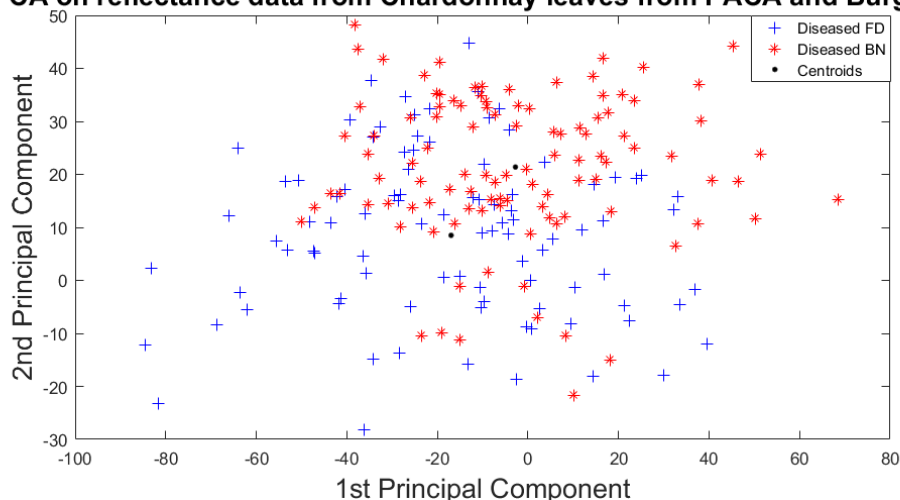
Figure B.13 PCA analysis on Chardonnay reflectance data in PACA and Burgundy regions. FD diseased and BN diseased data points are presented in blue and red respectively. Centroids of the two FD diseased and BN diseased clusters are marked in black.

147

ANNEX C

C. ANALYSIS OF COMPLETE REFLECTANCE DATA

Here, complete reflectance data are considered (400–2400 nm) in the analysis; this means that no dimension reduction method is applied.

APPLICATION TO PACA 2016 DATA

Table C.1 Results of using complete spectra in classifying different groups of spectral data acquired in the August acquisition campaign (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Marselan	70.97	27.27	30.00	0.82
Grenache	90.63	7.14	11.11	0.73
Vermentino	93.75	6.67	5.88	0.93
Chardonnay	90.63	12.50	6.25	0.95
Red	87.30	7.69	16.22	0.96
White	92.19	9.38	6.25	0.93
All	88.19	13.85	9.68	0.95

Table C.2 Results of using complete spectra in classifying different groups of spectral data acquired in the September acquisition campaign (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Marselan	94.79	4.35	6.00	0.99
Grenache	95.06	2.38	7.69	0.97
Vermentino	98.18	3.23	0.00	0.96
Chardonnay	97.73	3.85	0.00	0.96
Red	96.61	2.22	4.60	0.99
White	98.99	1.75	0.00	0.93
All	96.01	2.08	6.06	0.99

Table C.3 Results of using complete spectra in classifying different groups of spectral data acquired in the August and September acquisition campaigns (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Marselan	92.91	9.23	4.84	0.98

Grenache	96.77	1.72	4.55	0.89
Vermentino	96.55	2.22	4.76	0.97
Chardonnay	97.37	4.65	0.00	0.97
Red	96.41	3.28	3.88	0.99
White	95.09	5.56	4.11	0.93
All	95.65	1.98	6.60	0.99

APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA

Table C.4 Results of using complete spectra in classifying different groups of reflectance data acquired in the July acquisition campaign (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Cabernet-Sauvignon	89.39	7.69	12.5	0.98

Table C.5 Results of using complete spectra in classifying different groups of reflectance data acquired in the September acquisition campaign (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Cabernet-Sauvignon	94.44	8.51	2.33	0.98

Table C.6 Results of using complete spectra in classifying different groups of reflectance data acquired in the July and September acquisition campaigns (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Cabernet-Sauvignon	96.79	5.19	1.27	0.99

APPLICATION TO BURGUNDY 2017 DATA

Table C.7 Results of using complete spectra in classifying different groups of reflectance data acquired in the August acquisition campaign (Severity of infestation = 1).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Chardonnay	97.77	4.35	0.00	0.99

Table C.8 Results of using complete spectra in classifying different groups of reflectance data acquired in the September acquisition campaign (Severity of infestation = 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Chardonnay	99.17	1.67	0.00	0.99

Table C.9 Results of using complete spectra in classifying different groups of reflectance data acquired in the August and September acquisition campaigns (Severity of infestation = 1 & 2).

Grapevine Variety	Accuracy (%)	FNR (%)	FPR (%)	AUC
Chardonnay	98.57	1.90	0.95	0.99

From the previous, we can say that the FD detection on red (Marselan, Grenache) and white-berried (Vermentino, Grenache) grapvines in the PACA, the Esca detection on red Cabernet-Sauvignon grapevines in Nouvelle-Aquitaine and the BN detection on white Chardonnay grapevines in Burgundy with the use of complete spectra gave good results (accuracy>71%).

The results were in general better for a higher severity stage of the disease (stage 2). When levels of infestation are combined, the accuracy is somewhere in the middle between accuracies obtained from the first and second stages. We also noted that the classification models based on using complete reflectance data and obtained from white grapevines measurements exhibited relatively higher accuracy than red grapevines measurements. The last observation is that the classification models for Esca and BN detection had similar accuracies than FD detection on same colored grapevines.

Eventhough the results are good, however this technique is not acceptable because we have a dimensionality issue.

ANNEX D

D. BEST SVIS DETAILED CALCULATION

APPLICATION TO PACA DATA FROM 2016

Table D.1 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the August acquisition campaign (Severity of infestation = 1).

Grapevine Variety	SVIs	Accuracy (%)	FNR (%)	FPR (%)	AUC
Marselan	**ARI**	**90.32**	**11.11**	**7.69**	**0.94**
Grenache	**ARI**	**96.88**	**0.00**	**6.25**	**1.00**
Vermentino	GM1	90.63	11.11	7.14	0.96
Chardonnay	GM1	87.50	11.76	13.33	0.90
Red	**ARI**	**95.24**	**5.88**	**3.45**	**0.99**
White	GM1	90.63	8.82	6.67	0.95
All	mCAI	88.98	14.71	11.86	0.88

Table D.2 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the September acquisition campaign (Severity of infestation = 2).

Grapevine Variety	SVIs	Accuracy (%)	FNR (%)	FPR (%)	AUC
Marselan	**ARI**	**96.88**	**4.00**	**2.17**	**0.98**
Grenache	**ARI**	**97.53**	**2.70**	**2.27**	**0.99**
Vermentino	**GM1**	94.55	7.69	3.45	0.95
Chardonnay	**GM1**	97.73	0.00	3.85	0.94
Red	**ARI**	**98.31**	**2.33**	**1.10**	**0.99**
White	**GM1**	98.99	0.00	1.75	0.97
All	**mCAI**	**94.20**	**7.52**	**2.80**	**0.95**

Table D.3 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the August and the September acquisition campaigns (Severity of infestation = 1 & 2).

Grapevine Variety	SVIs	Accuracy (%)	FNR (%)	FPR (%)	AUC
Marselan	**ARI**	**93.70**	**8.70**	**3.45**	**0.98**
Grenache	**ARI**	**91.13**	**13.70**	**1.96**	**0.97**

Vermentino	**GM1**	**95.40**	**6.98**	**2.27**	**0.97**
Chardonnay	GM1	92.11	10.81	5.13	0.95
Red	**ARI**	**93.23**	**11.19**	**0.93**	**0.98**
White	GM1	93.87	9.88	2.44	0.96
All	**mCAI**	**88.41**	**16.96**	**5.98**	**0.91**

APPLICATION TO NOUVELLE-AQUITAINE DATA FROM 2017

Table D.4 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the July acquisition campaign (Severity of infestation = 1)

Grapevine variety	SVIs	Accuracy (%)	AUC
Cabernet Sauvignon	**ARI**	**98.48**	**0.99**

Table D.5 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the September acquisition campaign (Severity of infestation = 2)

Grapevine variety	SVIs	Accuracy (%)	AUC
Cabernet Sauvignon	**ARI**	**98.89**	**0.98**

Table D.6 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the July and the September acquisition campaigns (Severity of infestation = 1 & 2)

Grapevine variety	SVIs	Accuracy (%)	AUC
Cabernet Sauvignon	ARI	93.46	0.67

APPLICATION TO BURGUNDY DATA FROM 2017

Table D.7 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the August acquisition campaigns (Severity of infestation = 1).

Grapevine variety	SVIs	Accuracy (%)	AUC
Chardonnay	**GM1**	**97.78**	**0.99**

Table D.8 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the September acquisition campaigns (Severity of infestation = 2).

Grapevine variety	SVIs	Accuracy (%)	AUC
	GM1	95.83	0.96

Table D.9 Detailed classification results of SVIs in classifying different groups of reflectance data acquired in the August and the September acquisition campaigns (Severity of infestation = 1 & 2).

Grapevine variety	SVIs	Accuracy (%)	AUC
Chardonnay	GM1	93.76	0.95

ANNEX E

E. BEST TEXTURE PARAMETER DETAILED CALCULATION

APPLICATION TO PACA 2017 DATA

Table E.1 Results of using texture features in the Lab color space acquired in the August acquisition campaign (Severity of infestation = 1). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Grapevine variety	Accuracy (%)	AUC
Frac	Marselan	96.97	1.00
	Grenache	**96.77**	0.98
	Vermentino	86.67	0.91
	Chardonnay	87.50	0.95
	Red	**98.43**	0.99
	White	**90.32**	0.93
	All	82.54	0.89

Table E.2 Results of using texture features in the Lab color space acquired in the September acquisition campaign (Severity of infestation = 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Grapevine variety	Accuracy (%)	AUC
Frac	Marselan	97.06	1.00
	Grenache	**100**	1.00
	Vermentino	81.256	0.91
	Chardonnay	96.67	0.99
	Red	98.48	1.00
	White	93.55	0.99
	All	96.09	0.99

Table E.3 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Grapevine variety	Accuracy (%)	AUC
Frac	Marselan	86.59	0.89
	Grenache	**94.56**	0.99
	Vermentino	77.42	0.86
	Chardonnay	73.91	0.79
	Red	76.19	0.86
	White	77.92	0.82
	All	**87.75**	0.91

Table E.4 Classification results from texture features in PACA after feature selection using the Lab color space. Each texture technique is considered solely and a concatenation of some of them is also presented (multi-class approach)

Texture technique	Grapevine variety	Accuracy (%)	AUC
Frac	Marselan	95.69	0.972
	Grenache	79.79	0.88
	Vermentino	78.35	0.76
	Chardonnay	71.91	0.84
	Red	83.56	0.92
	White	71.90	0.73
	All	72.02	0.73

APPLICATION TO NOUVELLE-AQUITAINE 2017 DATA

Table E.5 Results of using texture features in the Lab color space acquired in the July acquisition campaign (Severity of infestation = 1). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	**100.00**	1.00

Table E.6 Results of using texture features in the Lab color space acquired in the September acquisition campaign (Severity of infestation = 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	**100.00**	1.00

Table E.7 Results of using texture features in the Lab color space acquired in the July and the September acquisition campaigns (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	**100.00**	1.00

Table E.8 Classification results from texture features in Nouvelle-Aquitaine after feature selection using the Lab color space. Each texture technique is considered solely and a concatenation of some of them is also presented (multi-class approach)

Texture technique	Accuracy (%)	AUC
Frac	94.75	0.98

APPLICATION TO BURGUNDY 2017 DATA

Table E.9 Results of using texture features in the Lab color space acquired in the August acquisition campaign (Severity of infestation = 1). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	**100.00**	1.00

Table E.10 Results of using texture features in the Lab color space acquired in the September acquisition campaign (Severity of infestation = 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	**100.00**	1.00

Table E.11 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	95.71	0.97

Table E.12 Classification results from texture features in Burgundy after feature selection using the Lab color space. Each texture technique is considered solely and a concatenation of some of them is also presented (multi-class approach)

Texture technique	Accuracy (%)	AUC
Frac	73.82	0.88

APPLICATION TO PACA VS NOUVELLE-AQUITAINE 2017 DATA

Table E.13 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	**92.34**	0.96

APPLICATION TO PACA VS BURGUNDY 2017 DATA

Table E.14 Results of using texture features in the Lab color space acquired in the August and the September acquisition campaigns (Severity of infestation = 1 & 2). Each texture technique is considered solely and a concatenation of some of them is also presented.

Texture technique	Accuracy (%)	AUC
Frac	**75.78**	0.94

ANNEX F

F. MS SENSOR PROTOTYPE

PRESENTATION

The aim of the following part is to show the potential product of the work. Afterall, the DAMAV project is a FUI project based on applied research. This kind of projects develop new products and services that could be commericialized in the short or medium term.

A preliminary uniptimized MS test sensor (*Figure F.1*) was prepared. We have chosen four IDS identical cameras UI-3370CP-M-GL and assembled them to create a unique and constrained block. The cameras were CMOS monochrome, with a resolution of 4.19 Mpx and a pixel size of 5.5 µm. The resolution chosen is very close to that defined by the Global sensing technologies Company, responsible of the MS sensor design at the beginning of the DAMAV project. The four cameras were equipped with Fujinon CF16HA-1/F1.4/16mm lens.

Figure F.1 Designed MS test sensor consisting of four identical IDS cameras equipped with 16 mm lens and four different filters {530, 63, 740, 810 nm} chosen for FD disease detection.

The geometric deformation of this type of focal length is relatively important, imposing mandatory corrections. An example of four filters are mounted on the four cameras, they are midopt narrow band filters (with narrow bandwidth): NB-532 {525-550 nm}, NB-630 {625-645 nm}, NB-740 {730-755 nm}, NB-810 {798-820 nm}. These are chosen only for clarification purpose and to show how things could be done in practical.

The quantum efficiency is the ratio of the number of photoelectrons released by a photoelectric process to the number of incident photons absorbed. For the set of four filters used in this example, the quantum efficiency is given in *Figure F.2*.

We have used a multi-processing technique in order to collect the flux relative to each of the four cameras. A program was implemented to visualize the four cameras independently one from another but at a common instant t, MS images can hence be acquired from the four sensors and stored in memory.
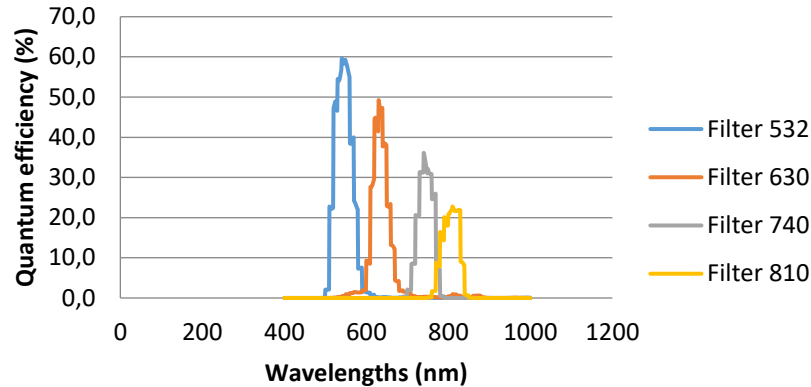

Figure F.2 Quantum efficiency of the four selected filters in the case of FD disease.

The spectral bands considered don't present any or little overlapping. We should also consider while choosing the filters if the desired wavelengths is in the center of the spectral band covered by the sensor. If this is not the case, this means that some of the information will be lost.

GEOMETRIC CORRECTIONS

Two kinds of sensor corrections have to be considered. The first one involves the effect of vignetting and the second one covers the lens distortion, which is generally the most relevant.

The vignetting effect is a radial falloff in illumination strength which, appears as a shadowing increasing towards the image periphery. It is due to an occlusion of the detector plane [217]. This wasn't present in our case so no correction was needed.

The lens distortion ([218], [219]) is produced by a non-uniform magnification across the lens surface; measurements are radially shifted in consequence. Furthermore, lens distortion can also be generated by a misalignment between the lens and the detector plane; measurements in that case undergo a planar shift. The Brown-Conrady lens distortion model can compensate the effects of both uneven magnification and lens/detector plane misalignment. The model, however, requires the calculation of sensor specific intrinsic and extrinsic coefficients. A common approach for the calculation of these coefficients is through the use of calibration panels; these are typically planar grids of known geometric properties. For each camera, the lens distortion model is constructed using the extracted coefficients. Once the four models are found, corresponding MS images are corrected.

RADIOMETRIC CORRECTIONS

Remotely sensed data are affected by environmental conditions, which include atmospheric conditions, surface conditions and illumination changes [220]. In most cases, only temporal change in light, matters and radiometric calibration is performed to provide consistency in images taken under variable environmental conditions. The empirical line method is the most used approach to radiometric correction. A relationship is assumed to exist between the collected digital numbers measurements from the MS sensor and the reflectance of the surface of the imaged object [221]. For this purpose, a set of targets are used and known as pseudo invariant features. The method requires taking at least measurements at two conditions one light and another one dark. A color chart

containing 24 regions is used as pseudo invariant features in our case; it is a spectrally homogeneous surface, invariant with environmental conditions. The at-surface reflectance measurements from each region of the pseudo invariant features were registered using the FieldSpec3 Portable Spectrometer and the contact probe. Then, MS images and radiance of the incident light (acquired using the Photosynthetically Active Radiation sensor) are taken under different illumination conditions. For each illumination condition, the relationship between the digital numbers measurements recorded within the imagery and the at-surface reflectance acquired from measurements, are calculated and a linear relationship is derived for each band. The slope and the offset of each line obtained are clearly incident light-dependent. Two linear relationships were used to calculate both the slope and the offset, as function of the intensity of the incident light. Once the equations are concatenated, a final linear relationship is found between the digital numbers and the at-surface reflectance. An example of what we obtained is displayed in *Figure F.3*. The final relationships between real reflectance and predicted reflectance by the four sensors, indicating the reliability of each one of them are detailed; R-squared corresponding to each camera are also presented, it is a statistical measure of how close the data are to the fitted regression line. Once these equations are established, MS images can be converted to at-surface reflectance.

Filter 810 nm gave the worst result, since we were limited in terms of the lens opening and many images were dark. The dispersion that we obtained for each filter can be reduced if the encoding of the data could be done more finely. In fact, the digital numbers in the MS images were encoded on only 8 bits, giving intensities between 0 and 255. However, this will be taken into consideration in the enhanced MS sensor version.
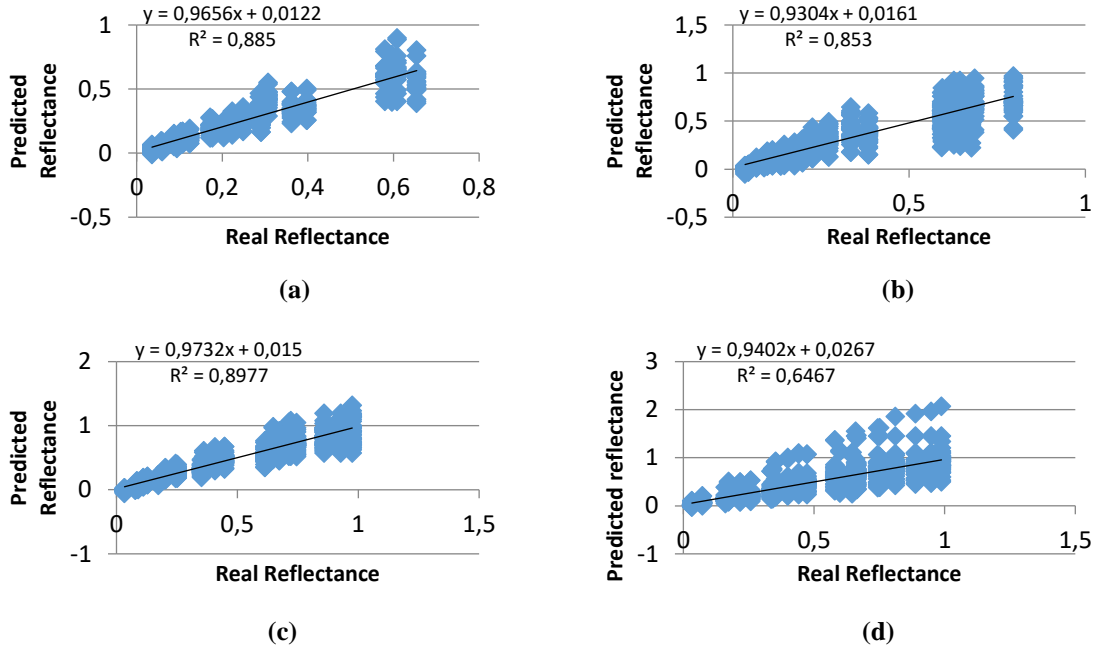
Figure F.3 Graphs presenting the regression line between the real and the predicted reflectance of the four sensors. **(a), (b), (c), (d)**, correspond respectively to cameras equipped with 530,630, 740, 810 filters, respectively.

IMAGE REGISTRATION

Image registration is the process of aligning multiple images of the same scene, taken at slightly different times or from different viewpoints or by different sensors into a single integrated image. In our study, the same scene is taken with a little difference in viewpoints. The registration geometrically overlays two images, the sensed image undergoes geometric transformations or local displacements to fit the reference or fixed image. The majority of the registration methods consist of the following four steps [222]: feature detection where characteristic objects are selected; feature matching, here, features found in the sensed image are matched to those from the reference image; transform model estimation which consists on estimating the parameters capable of overlaying the sensed and the reference images; image transformation, which means that the sensed image is transformed via the estimated mapping functions. In our work, we calculated the image registration based on a similarity transformation model, it is a non-reflective similarity transformation, consisting of translation, rotation, and scale. After geometrically and radiometrically calibrating the images, images are registered and the superimposed set can be used for further processing. *Figure F.4* presents the superimposition of an example set of four MS images corresponding to the same scene and acquired by the four sensors.
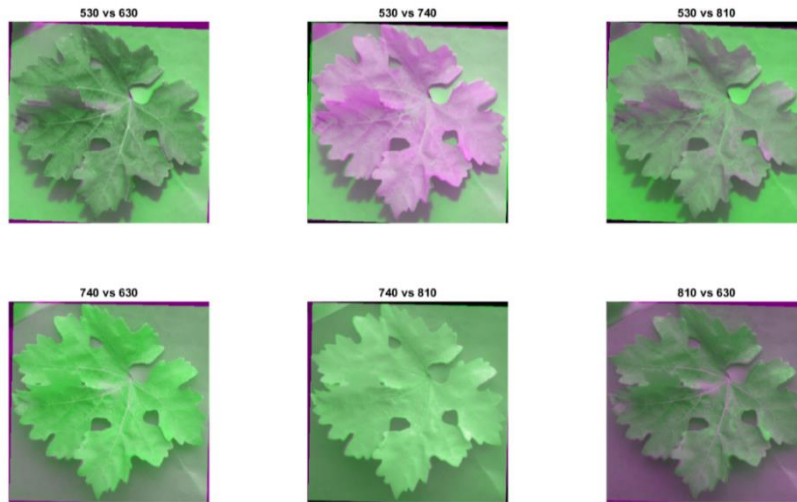
Figure F.4 Image registration on corrected images presented as superimposed pairs.

Similar geometric and radiometric corrections approaches need to be conducted on the optimized MS sensor that will be conceived by GST. Other corrections will definitely be needed when moving from the leaf scale to the grapevine or the whole field scale. Moreover, the simple registration algorithm worked for our current images that include 2D structures (leaves) taken at different points of view by the four cameras but another more sophisticated algorithm will definitely be needed when dealing with images containing 3D structures such as grapevines in the field.

PUBLICATIONS

INTERNATIONAL

- Hania AL SADDIK, Jean-Claude SIMON and Frederic COINTAULT, "Development of Spectral Indices for 'Flavescence Dorée' grapevine disease identification", *Sensors*, 17(12), 2017.
- Hania AL SADDIK, Anthony LAYBROS, Bastien BILLIOT and Frederic COINTAULT, "Using Image Texture and Spectral Reflectance Analysis to Detect Yellowness and Esca in Grapevines at leaf-level", *Remote Sensing*, 10(4), 2018.
- Hania AL SADDIK, Jean-Claude SIMON, and Frederic COINTAULT, "A Comparative Analysis on the Assessment of Optimal Spectral Bands for Vineyard Disease Detection Sensor Design: Case of Flavescence Dorée", *Precision Agriculture*, 2018.

NATIONAL

- Hania AL SADDIK, Simeng HAN, Jean-Claude SIMON, Frederic COINTAULT, Solution de détection des maladies de la vigne par imagerie de drone, *Revue des Œnologues et des techniques vitivinicoles et oenologiques*, 44(162), 2017.

BOOK

- Hania AL SADDIK, Anthony LAYBROS, Jean-Clause SIMON, Frederic COINTAULT, "Protocol for the definition of a Multi-Spectral sensor for specific foliar disease detection: case of Flavescence Dorée", *Methods in Molecular Biology, Phytoplasma: Methods and Protocols, Springer*, 2018.

CONFERENCES

INTERNATIONAL

- International Conference on Image Processing (ICIP 2018). Between the 7th and the 10th of Octobre 2018, in Athens, Greece.
- International Conference on Agricultural Engineering (CIGR 2016). Between the 26th and the 29th of June 2016, at Aarhus University, Denmark.
  - ➢ Oral presentation
  - ➢ Article : *http://conferences.au.dk/cigr-2016/full-papers/*

EUROPEAN

- European Conference on Precision Agriculture (ECPA 2017). Between the 16th and the 20th of July 2017, at John McIntyre Center, Edinburgh, UK.
  - ➢ Oral presentation

- ➤ Article: <u>Hania AL SADDIK</u>, Jean-Claude SIMON, Olivier BROUSSE and Frederic COINTAULT, Multispectral band selection for imaging sensor design for vineyard disease detection: case of 'Flavescence Dorée', *Advances in Animal Biosciences*, 8(2), 2017.

TECHNICAL SEMINARS

- GREV (Group of Wine Experiments in Burgundy) reunion. Tuesday 14th of February 2017, at the BIVB (Interprofessional Office of Burgundy Wines), Chablis.
  - ➤ Oral presentation
- Preciison Viticulture « Les capteurs à la loupe ». Monday 2nd of February 2016 at Bordeaux Science Agro, Bordeaux, France.
  - ➤ Oral presentation
  - ➤ Article

RESEARCH SEMINARS

- Scientific imaging forum, when cells meet space. Monday 13th of February 2017, at « la Maison des Sciences de l'homme de Dijon ».
  - ➤ Oral discussion
- Vine and Wine Research Days in Burgundy. Friday 16th of December 2016, « Palais des Congrés », Beaune.
  - ➤ Oral presentation

BIBLIOGRAPHY

[1]   J. Pitte, "Implantation de vignes: phénomène mondial," 2008. [Online]. Available: https://www.larvf.com/,vins-implantation-de-vignes-monde-international-chine,4517106.asp. [Accessed 17 09 2018].

[2]   B. Stenne, "Dossier de presse: les exportations françaises de vin et spiritueux bilan 2011 perspectives 2012," 2012. [Online]. Available: http://www.fevs.com/files/actu/28_210_1.pdf. [Accessed 17 09 2018].

[3]   K. Alemu, "Detection of diseases, identification and diversity of viruses: a review," *Journal of Biology, Agriculture and Healthcare,* vol. 5, no. 1, pp. 204-213, 2015.

[4]   M. Lopez, E. Bertolini, A. Olmos, P. Caruso, M. Gorris, P. Llop, R. Penyalver and M. Cambra, "Innovative tools for detection of plant pathogenic viruses and bacteria," *International Microbiology,* vol. 6, p. 233–243, 2003.

[5]   E. Knipling, "Physical and Physiological Basis for the Reflectance of Visible and Near-Infrared Radiation from vegetation," *Remote Sensing of Environment,* vol. 1, pp. 155-159, 1970.

[6]   M. Elowitz, "What is Imaging Spectroscopy (Hyperspectral Imaging)?," [Online]. Available: http://www.markelowitz.com/Hyperspectral.html. [Accessed 3 09 2018].

[7]   M. Prabhakar, Y. Prasad, M. Thirupathi, G. Sreedevi, B. Dharajothi and B. Venkateswarlu, "Use of ground based hyperspectral remote sensing for detection of stress in cotton caused by leafhopper (Hemiptera: Cicadellidae)," *Computers and Electronics in Agriculture,* vol. 79, p. 189–198, 2011.

[8]   J. Huang, H. Liao, Y. Zhu, J. Sun, Q. Sun and X. Liu, "Hyperspectral detection of rice damaged by rice leaf folder (Cnaphalocrocis medinalis)," *Computers and Electronics in Agriculture,* vol. 82, p. 100–107, 2012.

[9]   A. Stilwell, G. Hein, A. Zygielbaum and D. Rundquist, "Proximal sensing to detect symptoms associated with wheat curl mite-vectored viruses," *International Journal of Remote Sensing,* vol. 34, no. 14, p. 4951–4966, 2013.

[10]   L. Yuan, Y. Huang, R. Loraamm, C. Nie, J. Wang and J. Zhang, "Spectral analysis of winter wheat leaves for detection and differentiation of diseases and insects," *Field Crops Research,* vol. 156, p. 199–207, 2014.

[11]   G. Verhoeven, Multispectral and hyperspectral imaging, John Wiley and Sons, 2018.

[12]   A. Mahlein, "Plant disease detection by imaging sensors-Parallels and specific demands for precision agriculture and plant phenotyping," *Plant disease,* vol. 100, no. 2, pp. 241-251, 2016.

[13] M. Neumann, L. Hallau, B. Klatt, K. Kersting and C. Bauckhage, "Erosion Band Features for Cell Phone Image Based Plant Disease Classification," in *Proceedings of the 22nd International Conference on Pattern Recognition*, Stockholm, Sweden, 2014.

[14] P. Rothe and R. Kshirsagar, "Cotton leaf disease identification using Pattern Recognition techniques," *International Conference on Pervasive Computing (ICPC),* 2015.

[15] S. Mohanty, D. Hughes and M. Salathe, "Using deep learning for image-based plant disease detection," 2016. [Online]. Available: http://arxiv.org/abs/1604.03169. [Accessed 20 06 2018].

[16] G. Elmasry, M. Kamruzzaman , D. Sun and P. Allen, "Principles and applications of Hyperspectral Imaging in quality evaluation of Agro-Food products: A review," *Critical Reviews in Food Science and Nutrition,* vol. 52, p. 999–1023, 2012.

[17] G. Reynolds, "Remote Sensing for Assessing Rhizoctonia Crown and Root Rot Severity in Sugar Beet," *Plant disease,* pp. 497-505, 2012.

[18] M. Mirik, R. Ansley, J. Price, F. Workneh and C. Rush, "Remote monitoring of Wheat Streak Mosaic progression using sub-pixel classification of Landsat 5 TM Imagery for site specific disease management in Winter Wheat," *Advances in Remote Sensing,* vol. 2, pp. 16-28, 2013.

[19] C. Hillnhütter, A. Mahlein, R. Sikora and E. Oerke, "Remote sensing to detect plant stress induced byHeterodera schachtii and Rhizoctonia solaniin sugar beet fields," *Field Crops Research,* vol. 122, p. 70–77, 2011.

[20] T. Mewes, J. Franke and G. Menz, "Spectral requirements on airborne Hyperspectral remote sensing data for wheat disease detection," *Precision Agriculture,* vol. 12, p. 795–812, 2011.

[21] A. Kumar, W. Lee, R. Ehsani, L. Albrigo, C. Yang and R. Mangan, "Citrus greening disease detection using aerial hyperspectral and multispectral imaging techniques," *Journal of Applied Remote Sensing,* vol. 6, 2012.

[22] M. Dubbini, A. Pezzuolo, M. De Giglio, M. Gattelli, L. Curzio, D. Covi, T. Yezekyan and F. Marinello, "Last generation instrument for agriculture multispetcral data collection," in *International Conference on Agricultural Engineering (CIGR)*, 2017.

[23] J. Suomalainen, N. Anders, S. Iqbal, G. Roerink, J. Franke, W. P., D. Hunninger, H. Bartholomeus, R. Becker and L. Kooistra, "A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles," *Remote Sensing,* vol. 6(1), pp. 11013-11030, 2014.

[24] A. Lowe, N. Harrison and A. French, "Hyperspectral image analysis techniques for the detetcion and classification of the early onset of plant disease and stress," *Plant Methods,* vol. 13, pp. 1-12, 2017.

[25] L. Dale, A. Thewis, C. Boudry, I. Rotar, P. Dardenne, V. Baeten and J. Pierna, "Hyperspectral Imaging applications in Agriculture and Agro-Food Product quality and safety control: A review," *Applied Spectroscopy Reviews,* vol. 48, p. 142–159, 2013.

[26]   L. Chaerle, D. Hagenbeek, E. De Bruyne, R. Valcke and D. Van Der Straeten, "Thermal and Chlorophyll-Fluorescence Imaging distinguish plant-pathogen interactions at an early stage," *Plant Cell Physiology,* vol. 45, no. 7, p. 887–896, 2004.

[27]   M. Mirik, G. Michels, S. Kassymzhanova-Mirik, N. Elliott, V. Catana, D. Jones and R. Bowling, "Using digital image analysis and spectral reflectance data to quantify damage by greenbug (Hemitera:Aphididae) in winter wheat," *Computers and Electronics in Agriculture,* vol. 51, pp. 86-98, 2006.

[28]   "Wine industry," [Online]. Available: https://www.businessfrance.fr/Media/Default/PROCOM/Kits/Agroalimentaire/Business_France-Wine_Industry.pdf. [Accessed 17 06 2018].

[29]   A. Carisse, R. Bacon, J. Lasnier and W. McFadden-Smith, "Identification Guide to the major diseases of grapes," *Agriculture and Agri-Food Canada,* 2006.

[30]   "Grapevine yellows: Flavescence Dorée and Bois Noir identification guide," [Online]. Available: http://www.inspection.gc.ca/plants/plant-pests-invasive-species/diseases/grapevine-yellows/identification-guide/eng/1326143947606/1326144092300. [Accessed 17 06 2018].

[31]   J. Chuche and D. Thiery, "Biology and ecology of the Flavescence Dorée vector Scaphoideus Titanus: a review," *Agronometry for Sustainable Development,* vol. 34, pp. 381-403, 2014.

[32]   T. Cverkivic, "Recognize vector of phytoplasma Flavescence Dorée on grapevine," 2004. [Online]. Available: http://www.chem.bg.ac.rs/~mario/scaphoideus/English/side_8_vector_pub.htm. [Accessed 2017 10 2017].

[33]   "Grapevine Flavescence Dorée," [Online]. Available: http://www.inra.fr/en/Scientists-Students/Plant-biology/All-reports/Grapevine-flavescence-doree.

[34]   V. Hofstetter, B. Buyck, D. Croll, O. Viret, A. Couloux and K. Gindro, "What if Esca disease of grapevine were not a fungal disease?," *Fungal diversity,* 2012.

[35]   S. Li, F. Bonneu, J. Chadoeuf, D. Picart, A. Gégout-Petit and L. Dubrana, "Spatial and temporal pattern analysis of esca grapevine disease in vineyards in France," *Ecology and Epidemology,* vol. 107(1), pp. 59-69, 2017.

[36]   P. Lecompte, G. Darrieutort, J. Liminana and G. Comont, "New insights into esca of grapevine: the development of foliar symptoms and their associaion with xylem discoloration," *Plant Disease,* vol. 96(7), pp. 924-934, 2012.

[37]   "Grapevine trunk diseases: A review," 04 2016. [Online]. Available: http://www.oiv.int/public/medias/4650/trunk-diseases-oiv-2016.pdf. [Accessed 16 06 2018].

[38]   G. Diaz and B. Latorre, "Infection caused by Phaeomoniella chlamydospora associated with Esca-like symptoms in grapevine in Chile," *Plant Disease,* vol. 98(3), pp. 351-360, 2014.

[39]   S. Vasquez and M. Fidelibus. [Online]. Available: http://articles.extension.org/pages/64365/grapevine-measles. [Accessed 22 06 2018].

[40]   "Grape, Esca (Black Measles)," [Online]. Available: http://ipm.ucanr.edu/PMG/r302100511.html. [Accessed 22 06 2018].

[41]   L. Mugnai, A. Graniti and G. Surico, "Esca (Black Measles) and Brown Wood-Streaking: two old and elusive diseases of grapevines," *Plant Disease,* vol. 83(5), pp. 404-418, 1999.

[42]   A. Hall, D. Lamb, B. Holzapfel and J. Louis, "Optical remote sensing applications in viticulture: a review," *Australian Journal of Grape and Wine Research,* vol. 8, p. 36–47, 2002.

[43]   C. Santos, M. Lopo, R. Pascoa and J. Lopes, "A review on the applications of Portable Near-Infrared Spectrometers in the Agro-Food Industry," *Applied Spectroscopy OA,* vol. 67, no. 11, 2013.

[44]   J. Whalley and S. Shanmuganathan, "Applications of image processing in viticulture: A review," Adelaide, Australia, 2013.

[45]   A. Calcante, A. Mena and F. Mazzetto, "Evaluation of "ground sensing" optical sensors for diagnosis of Plasmopara viticolaon vines," *Spanish Journal of Agricultural Research,* vol. 10, no. 3, pp. 619-630, 2012.

[46]   R. Naidu, E. Perry, F. Pierce and T. Mekuria, "The potential of spectral reflectance technique for the detection of Grapevine leafroll-associated virus-3 in two red-berried wine grape cultivars," *Computers and Electronics in Agriculture,* vol. 66, p. 38–45, 2009.

[47]   J. Hou, L. Li and J. He, "Detection of grapevine leafroll disease based on 11-index imagery and ant colony clustering algorithm," *Precision Agriculture,* vol. 17, no. 4, pp. 488-505, 2016.

[48]   A. Meunkaewjinda, P. Kumsawat, K. Attakitmongcol and A. Srikaew, "Grape leaf disease detection from color imagery using hybrid inteligent system," *Proceedings of ECTI-CON,* 2008.

[49]   S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience,* vol. 2016, 2016.

[50]   R. Oberti, M. Marchi, P. Tirelli, A. Calcante, M. Iriti and A. Borghese, "Automatic detection of powdery mildew on grapevine leaves by image analysis: Optimal view-angle range to increase the sensitivity," *Computers and Electronics in Agriculture,* vol. 104, p. 1–8, 2014.

[51]   S. MacDonald, M. Staid, M. Staid and M. Cooper, "Remote hyperspectral imaging of grapevine leafroll-associated virus 3 in cabernet sauvignon vineyards," *Computers and Electronics in Agriculture,* vol. 130, p. 109–117, 2016.

[52]   J. Albetis, S. Duthoit, F. Guttler, A. Jacquin, M. Goulard, H. Poilvé, J. Feret and G. Dedieu, "Detection of Flavescence Dorée grapevine disease using unmanned aerial vechicle (UAV) multispectral imagery," *Remote Sensing,* vol. 9(308), pp. 1-20, 2017.

[53] S. Di Gennaro, E. Battiston, S. Di Marco, O. Facini, A. Matese, M. Nocentini, A. Palliotti and L. Mugnai, "Unmanned Aerial Vehicle (UAV)-based remote sensing to monitor grapevine leaf stripe disease within a vineyard affected by esca complex," *Phytopathologia Mediterranea,* vol. 55(2), pp. 262-275, 2016.

[54] F. Vanegas, D. Bratanov, K. Powell, J. Weiss and F. Gonzalez, "A novel methodology for improving plant pest surveillance in vineyards and crops using UAV-based hyperspectral and spatial data," *Sensors,* vol. 18(260), pp. 1-21, 2018.

[55] F. Martinelli, R. Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, M. Boschetti, L. Goulart, C. Davis and A. Dandekar, "Advanced methods of plant disease detection: A review," *Agronomy for Sustainable Development,* vol. 35, pp. 1-25, 2015.

[56] J. Richards, Remote sensing digital image analysis: an introduction, Canberra: Springer, 1999.

[57] R. Olsen, Remote sensing from air and space, vol. 162, Bellingham: Society of Photo Optical, 2007.

[58] A. Goetz, "Three decades of hyperspectral remote sensing of the earth: a personal view," *Remote sensing of environment,* vol. 113, pp. S5-S16, 2009.

[59] R. Musetti and L. Pagliari, Phytoplasmas: methods and protocols, vol. 1875, In Press: Springer - Humana Press, 2019.

[60] D. Clair, J. Larrue, G. Aubert, J. Gillet, G. Cloquemin and E. Boudon-Padieu, "A multiplex nested-PCR assay for sensitive and simultaneous detection and direct identification of phytoplasma in the yellows group and Stolbur group and its use in survey of grapevine yellows in France," *Vitis,* vol. 42(3), pp. 151-157, 2003.

[61] K. Pfitzner, R. Bartolo, G. Carr, A. Esparon and A. Bollhofer, "Standard for reflectance spectral measurements of temporal vegetation plots," 2011. [Online]. Available: https://www.environment.gov.au/system/files/resources/bf8002d0-2582-48a1-820f8e79d056faed/files/ssr195.pdf. [Accessed 11 09 2016].

[62] ASDI, "FieldSpec3 user manual," 2010. [Online]. Available: http://support.asdi.com/Document/Viewer.aspx?id=108. [Accessed 10 09 2016].

[63] M. Bertamini, M. Grando and N. Nedunchezhian, "Effects of phytoplasma infection on pigments, chlorophyll protein complex and photosynthetic activities in field grown apple leaves," *Biologia Plantarum,* vol. 47(2), pp. 237-242, 2003.

[64] M. Bertamini and N. Nedunchezhian, "Efects of phytoplasma [stolbur-subgroup (Bois Noir)] on photosynthetic pigments, saccharides, ribulose, 1.5-biosphate carboxylase, nitrate and nitrite reductase and photosynthetic activity in field-grown grapevine (Vitis vinifera L. cv. Chardonnay) leaves," *Photosynthetica,* vol. 39(1), pp. 119-122, 2001.

[65] M. Himeno, Y. Kitazawa, T. Yohida, K. Maejima, Y. Yamaji, K. Oshima and S. Namba, "Purple top symptoms are associated with reduction of leaf cell death in phytoplasma-infected plants," *Scientific Reports,* vol. 4, pp. 1-7, 2014.

[66]   S. Endeshaw, S. Murolo, G. Romanazzi and D. Neri, "Effects of Bois Noir on carbon assimilation, transpiration, stomatal conducatance of leaves and yields of grapevine (Vitis vinifera) cv. Charodnnay," *Physiologia Plantarum,* vol. 145, pp. 286-295, 2012.

[67]   M. Vitali, W. Chitarra, L. Galetto, D. Bosco, C. Marzach, M. Gullino, F. Spanna and C. Lovisolo, "Flavescence Dorée phytoplasma deregulates stomatal control of photosynthesis in Vitis vinifera," *Annal of applied Biology,* vol. 162, pp. 335-346, 2013.

[68]   D. Rusjan, H. Halbwirth, K. Stich, M. Petkovsek and R. Veberic, "Biochemical response of grapevine variety Chardonnay (Vitis vinifera L.) to infection with grapevine yellows (Bois Noir)," *European Journal of Plant Pathology,* vol. 134, pp. 231-237, 2012.

[69]   Y. Tan, H. Wei, J. Wang, X. Zong, D. Zhu and Q. Liu, "Phytoplasmas change the source-sink relationship of field-grown sweet cherry by disturbing leaf function," *Physiological and Molecular Plant Pathology,* vol. 92, pp. 22-27, 2015.

[70]   R. Musetti, S. Buxa, F. De Marco, A. Loschi, R. Polizzotto, K. Kogel and A. Van Bel, "Phytoplasma-triggered Ca2+ influx is involved in Sieve tube blockage," *Molecular Plant Microbe Interaction,* vol. 26(4), pp. 379-386, 2013.

[71]   A. Petit, N. Vaillant, M. Boulay, C. Clement and F. Fontaine, "Alteration of photosynthesis in grapevines affected by esca," *Phytopathology,* vol. 96(10), pp. 1060-1066, 2006.

[72]   B. Lorrain, I. Ky, G. Pasquier, M. Jourdes, L. Dubrana, L. Geny, P. Rey, B. Doneche and P. Tesseidre, "Effect of Esca disease on the phenolic and sensory attributes of Cabernet Sauvignon grapes, musts and wines," *Australian Journal of Grape and Wine Research,* vol. 18, pp. 64-72, 2012.

[73]   F. Calzarano, L. Seghetti, M. Del Carlo and A. Cichelli, "Effect of Esca on the quality of berries, musts and wines," *Phytopathologia Mediterranea,* vol. 43, pp. 125-135, 2004.

[74]   A. Mahlein, U. Steiner, H. Dehne and E. Oerke, "Spectral signatures of sugar beet leaves for the detection and differentiation of diseases," *Precision Agriculture,* vol. 11, pp. 413-431, 2010.

[75]   H. Santoso, T. Gunawan, R. Jatmiko, W. Darmosarkoro and B. Minasny, "Mapping and identifying basal stem rot disease in oil palms in North Sumatra with QuickBird imagery," *Precision Agriculture,* vol. 12, pp. 233-248, 2011.

[76]   C. Tucker, "Red and photographic infra-red linear combination for monitoring vegetation," *Remote Sensing of Environment,* vol. 8, pp. 127-150, 1979.

[77]   C. Tucker, B. Holben, J. Elgin, H. James, I. McMurtrey and E. James, "Remote sensing of total dry-matter accumulation in winter wheat," *Remote Sensing of Environment,* vol. 11, pp. 171-189, 1981.

[78]   J. Penuelas, I. Filella and J. Gamon, "Assessment of photosynthetic radiation-use efficiency with spectral reflectance," *New Phytologist,* vol. 131, no. 3, pp. 291-296, 1995b.

[79] J. Penuelas, J. Gamon, A. Fredeen, J. Merino and C. Field, "Reflectance indices associated with physiological changes in nitrogen and water limited sunflower leaves," *Remote Sensing of Environment,* vol. 48, pp. 135-146, 1994.

[80] G. Trotter, D. Whitehead and E. Pinkey, "The photochemical reflectance index as a measure of photosynthetic light use efficiency for plants of varying foliar nitrogen contents," *International Journal of Remote Sensing,* vol. 23, pp. 1207-1212, 2002.

[81] A. Gitelson, N. Merzylak and O. Chivkunova, "Optical properties and non-destructive estimation of anthocyanin content in plant leaves," *Photochemistry and Photobiology,* vol. 74, pp. 38-45, 2001.

[82] G. Blackburn, "Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves," *International Journal of Remote Sensing,* vol. 19, no. 4, pp. 657-675, 1998.

[83] J. Penuelas, F. Baret and I. Filella, "Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance," *Photosynthetica,* vol. 31, no. 2, pp. 221-230, 1995a.

[84] R. Laudien, "Development of a field and GIS-based management information system for the sugar beet industry," in *EFITA WWCA Congress*, Vila Real, Portugal, 2005.

[85] A. Richardson, M. Aikens, G. Berlyn and P. Marshall, "Drought stress and paper birch (Betula papyrifera) seedlings: Effectes of an organic biostimulant on plant health and stress tolerance and detection of stress effects with instrument-based, non-invasive methods," *Journal of Arboriculture,* vol. 30, pp. 52-61, 2004.

[86] A. Sims and A. Gamon, "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages," *Remote Sensing of Environment,* vol. 81, pp. 337-354, 2002.

[87] A. Apan, A. Held, S. Phinn and J. Markley, "Formulation and assessment of narrow-band vegetation indices from EO-1 hyperion imagery for discriminating sugarcane disease," in *Spatial Sciences Institute Biennial Conference on Spatial Knowledge without Boundaries*, Canberra, Australia, 2003.

[88] L. Kooistra, R. Leuven, R. Wehrens, P. Nienhuis and L. Buydens, "A comparison of methods to relate grass reflectance to soil contamination," *International Journal of Remote Sensing,* vol. 24, pp. 4995-5010, 2003.

[89] A. Gitelson, J. Kaufman and N. Merzylak, "Use of green channel in remote sensinf of global vegetation EOS-MODIS," *Remote Sensing of Environment,* vol. 58, no. 3, pp. 289-298, 1996a.

[90] A. Gitelson and N. Merzylak, "Remote estimation of cholorophyll content in higher plant leaves," *International Journal of Remote Sensing,* vol. 18, pp. 2691-2697, 1997.

[91] R. Main, A. Cho, R. Mathieu, M. O'Kennedy, A. Ramoelo and S. Koch, "An investigation into robust spectral indices for leaf chlorophyll estimation," *ISRPS Journal of Photogrammetry and Remote Sensing,* vol. 66, pp. 751-761, 2011.

[92]   H. Seelig, A. S. L. Hoehn, D. Klaus, W. Adams and W. Emery, "Plant water parameters and the remote sensing R1300/R1450 leaf water index: Controlled condition dynamics during the development of water dificit stress," *Irrigation Science,* vol. 27, no. 5, pp. 357-365, 2009.

[93]   P. Zarco-Tejada, J. Miller, T. Noland, G. Mohammed and P. Sampson, "Scaling-up and model inversion methods with narrowband optical indices for chloropgyll content estimation in closed forest canopies with hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, pp. 1491-1507, 2001.

[94]   E. Underwood, S. Ustin and D. DiPietro, "Mapping non-native plants using hyperspectral imagery," *Remote Sensing of Environment,* vol. 86, pp. 150-161, 2003.

[95]   G. Rondeaux, M. Steven and F. Baret, "Optimization of soil-adjusted vegetation indices," *Remote Sensing of Environment,* vol. 55, pp. 95-107, 1996.

[96]   D. Haboudane, J. Miller, N. Tremblay, P. Zarco-Tejada and L. Dextraze, "Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture," *Remote Sensing of Environment,* vol. 81, pp. 416-426, 2002.

[97]   M. Araujo, T. Kawakami, R. Galvao, T. Yoneyama, H. Chame and V. Visani, "The successive projection algorithm for variable selection in spectroscopic multicomponent analysis," *Chemometrics and Intelligent Laboratory Systems,* vol. 57, no. 2, pp. 65-73, 2001.

[98]   Y. Zhang, L. Tan, H. Shi and Y. He, "Successive projection algorithm for variable selection on the rapid and non-destrctive classification of coolant," *International Journal of Digital Content Technology and its Applciations,* vol. 7, pp. 386-394, 2013.

[99]   X. Yang, H. Hong, Z. You and F. Cheng, "Spectral and image integrated analysis of hyperspectral data for waxy corn seed variety classification," *Sensors,* vol. 15, pp. 15578-15594, 2015.

[100]   O. Babatunde and L. Armstrong, "A genetic algorithm-based feature selection," *International Journal of Electronic Communication and Computer Engineering,* vol. 5, no. 4, pp. 899-902, 2014.

[101]   O. Babatunde, L. Armstrong, J. Leng and D. Diepeveen, "Zernike moments and genetic algorithm: Tutorial and application," *British Journal of Mathematics and Computer Science,* vol. 4, pp. 2217-2236, 2014.

[102]   S. Sivanandam and S. Deepa, "Introduction to genetic algorithms," Springer, Berlin/Heidelberg, Germany, 2008.

[103]   O. Marek, "Introduction to genetic algorithms," Czech Technical University, 1998. [Online]. Available: http://www.obitko.com/tutorials/genetic-algorithms . [Accessed 05 09 2017].

[104]   Matlab, "GA toolbox," [Online]. Available: https://fr.mathworks.com/help/gads/genetic-algorithm.html. [Accessed 01 09 2016].

[105] A. Mahlein, T. Rumpf, P. Welke, H. Dehne, L. Plumer, U. Steiner and E. Oerke, "Development of spectral indices for detecting and identifying plant diseases," *Remote Sensing of Environment,* vol. 128, pp. 21-30, 2013.

[106] "Support vector machines for machine learning," [Online]. Available: https://machinelearningmastery.com/ support-vector-machines-for-machine-learning/. [Accessed 20 04 2016].

[107] A. Ben-Hur and J. Weston, User's guide to support vector machines, vol. 609, Data Mining Techniques for the Life Sciences, Methods in Molecular Biology, 609, Springer, 2010, pp. 223-239.

[108] T. Rumpf, A. Mahlein, U. Steiner, E. Oerke and H. P. L. Dehne, "Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance," *Computers and Electronics in Agriculture,* vol. 74, pp. 91-99, 2010.

[109] A. C. J. C. F. N. R. L. K. Marques, "Near-Infrared spectroscopy and variable selection techniques to discriminate Pseudomonas aeruginosa strains in clinical samples," *Microchemical journal,* vol. 124, pp. 306-310, 2016.

[110] A. Lehrer, P. Moore and E. Komor, "Impact of sugarcane yellow leaf virus (ScYLV) on the carbohydrate status of sugarcane: Comparison of virus-free plants with symptomatic and asymptomatic virus-infected plants," *Physiological and Molecular Plant Pathology,* vol. 70, pp. 180-188, 2007.

[111] R. Matthews and R. Hull, Matthews's plant virology, San Diego, CA, USA: Academic Press, 2002.

[112] A. Camargo and J. Smith, "Image pattern classification for the identification of disease agents in plants," *Computers and Electronics in Agriculture,* vol. 66, no. 2, pp. 121-125, 2009.

[113] Q. Yao, Z. Guan, Y. Zhou, J. Tang, Y. Hu and B. Yang, "Application of support vector machine for detecting rice diseases using shape and color texture features," in *International Conference on Engineering Computation*, 2009.

[114] J. Pujari, R. Yakkundimath and A. Byadgi, "Image processing based detection of fungal diseases in plants," *Procedia Computer Science,* vol. 46, pp. 1802-1808, 2015.

[115] A. Rathod, B. Tanawal and V. Shah, "Image processing techniques for detection of leaf disease," *International Journal of Advanced Research in Computer Science and Software Engineering ,* vol. 3, no. 11, pp. 397-399, 2013.

[116] S. Khirade and A. Patil, "Plant disease detection using image processing," in *International Conference on Computer Communication Control and Automation*, 2015.

[117] C. Xie and Y. He, "Spectrum and image texture features analysis for early blight disease detection on Eggplant leaves," *Sensors,* vol. 16(5), pp. 1-15, 2016.

[118] C. Xie, Y. Shao, X. Li and Y. He, "Detetcion of early blight and late blight diseases on tomato leaves using hyperspetcral imaging," *Scientific Reports,* pp. 1-11, 2015.

[119]   S. Dhaygude and N. Kumbhar, "Agricultural plant leaf disease detection using image processing," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,* vol. 2, no. 1, pp. 599-602, 2013.

[120]   H. Bindushree and G. Sivasankari, "Detection of plant disease using image processing techniques," *International Journal of Technology Enhancements and Emerging Engineering Research,* vol. 3, no. 4, pp. 125-128, 2015.

[121]   P. Chaudhary, A. C. A. Chaudhari and S. Godara, "Color transform based approach for disease spot detection on plant leaf," *International Journal of Computer Science and Telecommunications,* vol. 3, no. 6, pp. 65-70, 2012.

[122]   N. Kurniawati, S. A. S. Sheikh Abdullah and S. Abdullah, "Investigation on image processing techniques for diagnosing paddy diseases," in *International Conference of Soft Computing and Pattern Recognition*, 2009.

[123]   L. Luo and G. Zohu, "Extraction of the rice leaf disease image based on BP neural network," in *International Conference on Computational Intelligence and Software Engineering*, 2009.

[124]   A. Kulkarni and A. Patil, "Applying image processing techniques to detect plant diseases," *International Journal of Modern Engineering Research,* vol. 2, no. 5, pp. 3661-3664, 2012.

[125]   K. Renugambal and B. Senthilraja, "Application of image processing techniques in plant disease recognition," *International Journal of Engineering Research and Technology,* vol. 4, no. 3, pp. 919-923, 2015.

[126]   E. Ferdeghini, B. Pinamonti, E. Picano, F. Lattanzi, R. Bussani, G. Slavich, A. Benassi, F. Camerini, L. Landini and A. L'Abbate, "Quantitative texture analysis in echocardiography: Application to the diagnosis of myocarditis," *Journal of Clinical Ultrasound,* vol. 19, no. 5, pp. 263-270, 1991.

[127]   J. Flores, "Statistical moments from the gray-level histograms," 2001. [Online]. Available: http://www.academicos.ccadet.unam.mx/jorge.marquez/cursos/Imagenes/Histogram_features.pdf. [Accessed 10 10 2017].

[128]   D. Vince, K. Dixon, R. Cothren and J. Cornhill, "Comparison of texture analysis methods for the characterization of coronary plaques in intravascular ultrasound images," *Computerized Medical Imaging and Graphics,* vol. 24, pp. 221-229, 2000.

[129]   R. Haralick, K. Shanmugam and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics,* vol. 3, pp. 610-621, 1973.

[130]   R. Pydipati, T. Burks and W. Lee, "Identification of citrus disease using color texture features and discriminant analysis," *Computers and Electronics in Agriculture,* vol. 52, pp. 49-59, 2006.

[131]   S. Arivazhagan, R. A. S. Shebiah and S. Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," *Agricultural Engineering International CIGR Journal,* vol. 15, pp. 211-217, 2013.

[132] S. Sannaki, V. Rajpurohit, V. Nargund and P. Kulkarni, "Diagnosis and classification of grape leaf diseases using neural networks," in *4th International Conference on Computing, Communications and Networking Technologies*, Tiruchengode, India, 2013.

[133] T. Torheim, E. Malinen, K. L. H. Kvaal, U. Indahl, E. Andersen and C. Futsaether, "Classification of dynamic contrast enhanced MR images of cervical cancers using texture analysis and support vector machines," *IEEE Transactions on Medical Imaging,* vol. 33, pp. 1648-1656, 2014.

[134] A. Chaddad, C. Tanougast, A. Dandache and A. Bouridane, "Extracted Haralick's texture features and morphological parameters from segmented multispectral texture bio-images for classification of colon cancer cells," *WSEAS Transactions on Biology and Biomedecine,* vol. 8, pp. 39-50, 2011.

[135] M. Galloway, "Texture analysis using gray level run lengths," *Computer Graphics Image Processing,* vol. 4, pp. 172-179, 1975.

[136] T. Xiaou, "Texture information in run-length matrices," *IEEE Transactions on Image Processing,* vol. 7, no. 11, pp. 1602-1609, 1998.

[137] A. Chu, C. Sehgal and J. Greeleaf, "Use of gray value distribution of run lengths for texture analysis," *Pattern Recognition Letters,* vol. 11, pp. 415-420, 1990.

[138] B. Dasarathy and E. Holder, "Image characterizations based on joint gray-level run-length distributions," *Pattern Recognition Letters,* vol. 12, pp. 497-502, 1991.

[139] F. Chabat, G. Yang and D. Hansell, "Obstructive lung disease: Texture classification for differentiation at CT," *Radiology,* vol. 228, pp. 871-877, 2003.

[140] R. Rangayyanand and T. Nguyen, "Fractal analysis of contours of breast masses in mammograms," *Journal of Digital Imaging,* vol. 20, no. 3, pp. 223-237, 2007.

[141] A. Balan, A. Traina, C. Traina Jr. and P. Marques, "Fractal analysis of image textures for indexing and retrieval by content," in *18th IEEE Symposium on Computer-Based Medical Systems (CBMS 05)*, 2005.

[142] M. Schroeder, Fractals, chaos, power laws: Minutes from an infinite paradise, New York, NY, USA: W.H. Freeman, 1992.

[143] A. Costa, G. Humpire-Mamani and A. Traina, "An efficient algorithm for fractal analysis of textures," in *25th SIBGRAPI Conference on Graphics, Patterns and Images*, 2012.

[144] B. Ganeshan, K. Y. R. Miles and C. Chatwin, "In search of biologic correlates for liver texture on portal-phase CT," *Academic Radiology,* vol. 1, no. 9, pp. 1058-1068, 2007.

[145] S. Rao, D. Lambregts, R. Schnerr, W. Van Omme, T. Van Nijnatten, M. Martens, L. Heijnen, W. Backes, C. Verhoef, M. Zeng, G. Beets and R. Beets-Tan, "Whole-liver CT texture analysis in colorectal cancer: Does the presence of liver metastases affect the texture of the remaining liver ?," *United European Gastroenterology Journal,* vol. 2, no. 6, pp. 530-538, 2014.

[146] F. Ng, B. Ganeshan, R. Kozarski, K. Miles and V. Goh, "Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture," *Radiology,* vol. 266, no. 1, pp. 177-184, 2013.

[147] P. Hiremath, S. Shivashankar and J. Pujari, "Wavelet-based features for color texture classification with application to CBIR," *International Journal on Computer Science and Network Security,* vol. 6, no. 3, pp. 124-133, 2006.

[148] A. El-Dahshan, T. Hosny and A. Salem, "Hybrid intelligent techniques for MRI brain images classification," *Digital Signal Processing,* vol. 20, pp. 433-441, 2010.

[149] M. Kocionek, A. Materka, M. Strzelecki and P. Szcypin ski, "Discrete wavelet transform derived features for digital image texture analysis," in *International Conference on Signals and Electronic Systems*, 2001.

[150] S. Arivazhagan and L. Ganesan, "Texture classification using wavelet transform," *Pattern Recognition Letters,* vol. 24, pp. 1513-1521, 2003.

[151] D. Chen, R. Chang, W. Kuo, M. Chen and Y. Huang, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound Medicine and Biology,* vol. 28, no. 10, pp. 1301-1310, 2002.

[152] S. He, S. Chen, H. Zhai and W. Liu, "Rice paper feature analysis based on texture parameter statistics of multispectral imaging," in *4th IEEE International Conference on Information Science and Technology*, 2014.

[153] F. Rosario and K. Thangadurai, "RELIEF: Feature selection approach," *International Journal of Innovative Research and Development,* vol. 4, no. 11, pp. 218-224, 2015.

[154] R. Durgabai, "Feature selection using ReliefF algorithm," *International Journal of Advanced Research in Computer and Communication Engineering,* vol. 3, no. 10, pp. 8215-8218, 2014.

[155] A. A. H. P. M. Sarrafzadeh and J. Shanbehzadeh, "ReliefF based feature selection in content-based image retrieval," in *Proceeding of the International Multiconference of Engineers and Computer Scientists (IMESCS)*, Hong-Kong, 2012.

[156] Z. D. Z. C. D. L. L. Pang and Y. Shao, "A computer-aided diagnosis system for dynamic contrast-enhanced MR images based on level set segmentation and ReliefF feature selection," *Computational and mathematical methods in medicine,* 2015.

[157] Z. D. B. Shahweli, "Neural network with new Relief feature selection for predicting breast cancer based on TP53 mutation," *International Research Journal of Computer Science,* vol. 4, no. 12, 2017.

[158] K. Gavhale and U. Gawande, "An overview of the research on plant leaves disease detection using image processing techniques," *Journal of Computer Engineering,* vol. 16, no. 1, pp. 10-16, 2014.

[159] K. Khairnar and R. Dagade, "Disease detection and diagnosis on plant using image processing, a review," *International Journal of Computer Applications,* vol. 108, no. 13, pp. 36-38, 2014.

[160] W. Miller and G. Drouillard, "Multiple feature analysis for machine vision grading of Florida citrus," *Applied Engineering in Agriculture,* vol. 17, no. 5, pp. 627-633, 2001.

[161] D. Al Bashish, M. Braik and S. Bani-Ahmad, "A framework for detection and classification of plant leaf and stem diseases," in *International Conference on Signal and Image Processing*, 2012.

[162] S. Kutty, N. Abdullah, H. Hashim, A. Ab Rahim, A. Kusim, T. Yaakub, P. Yunus and M. Abd Rahman, "Classification of watermelon leaf diseases using neural network analysis," in *IEEE Business ENgineering and Industrial Applications Colloquium (BEIAC)*, 2013.

[163] S. Wang and R. Summers, "Machine learning and radiology," *Medical Image Analysis,* vol. 16, no. 5, pp. 933-951, 2012.

[164] J. Juntu, J. Sijbers, S. De Backer, J. Rajan and D. Van Dyck, "Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images," *Journal of Magnetic Resonance,* vol. 31, no. 3, pp. 680-689, 2010.

[165] A. Larroza, D. Moratal, A. Paredes-Sanchez, E. Soria-Olivas, M. Chust, L. Arribas and E. Arana, "Support vector machine classification of brain metastasis and radiation necrosis based on texture analysis in MRI," *Journal of Magnetic Resonance Imaging,* vol. 42, no. 5, pp. 1362-1368, 2015.

[166] S. Tantisatirapong, D. Nigel, D. Rodriguez, L. Abernethy, D. Auer and C. Clark, "Magnetic resonance texture analysis: Optimal feature selection in classifying child brain tumors," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing*, Sevilla, Spain, 2014.

[167] X. Zhang, X. Gao, B. Liu, K. Ma, W. Yan, L. Liling, H. Yuhong and H. Fujita, "Effective staging of liver fibrosis by the selected texture features of liver: Which one is better, CT or MR imaging," *Computerized Medical Imaging and Graphics,* vol. 46, pp. 227-236, 2015.

[168] F. Khalvati, A. Wong and M. Haider, "Automated prostate cancer detection via comprehensive multi-parametric resonance imaging texture feature models," *BMC Medical Imaging,* pp. 15-27, 2015.

[169] K. Urish, K. M.G., J. Durkin, D. Miller, C. Chu and T. Mosher, "T2 Texture index of cartilage can predict early symptomatic OA progression: Data from the osteoarthritis initiative," *Osteoarthritis and Cartilage,* vol. 21, no. 10, pp. 1550-1557, 2013.

[170] F. Wagner, A. Gryanik, R. Schulz-Wedatland, P. Fasching and T. Wittenberg, "3D characterization of texture: Ealuation for the potential application in mammographic mass diagnosis," *Biomedical Engineering,* vol. 57, pp. 490-493, 2012.

[171] D. Mohanty and S. Mishra, "MRI classification of Parkinson's disease using SVM and texture features," in *Proceedings of the Second International Conference on Computer and Communication Technologies*, 2016.

[172] S. Jacquemond and S. Ustin, "Leaf optical properties: A state of the art," in *8th International Symposium Physical Measurements and Signatures in Remote Sensing*, Aussois, France, 2001.

[173]   A. Gitelson, R. Stark, U. Grits, D. Rundquist, Y. Kaufman and D. Derry, "Vegetation and soil lines in visible spectral space: A concept and technique for remote estimation of vegetation fraction," *International Journal of Remote Sensing,* vol. 23, no. 13, pp. 2537-2562, 2002.

[174]   R. Naidu, A. Rowhani, M. Fuchs, D. Golino and G. Martelli, "Grapevine leafroll: A complex viral disease affecting a high-value fruit crop," *Plant Disease,* vol. 98, pp. 1171-1185, 2014.

[175]   Z. Jingcheng, P. Ruiliang, H. Wenjiang, Y. Lin, L. Juhu and W. Jihu, "Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses," *Field Crops Research,* vol. 134, pp. 165-174, 2012.

[176]   W. Huang, D. Lamb, Z. Niu, Y. Zhang, L. Liu and J. Wang, "Identification of yellow rust in wheat using in situ spectral reflectance measurements and airborne hyperspectral imaging," *Precision Agriculture,* vol. 8, pp. 187-197, 2007.

[177]   R. Devadas, D. Lamb, S. Simpfendorfer and D. Backhouse, "Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves," *Precision Agriculture,* vol. 10, pp. 459-470, 2009.

[178]   J. Zhang, R. Pu, W. Huang, L. Yuan, J. Luo and J. Wang, "Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses," *Field Crops Research,* vol. 134, pp. 165-174, 2012.

[179]   E. Khawas and E. Khawas, "Interactions between Aphis gossypii (Glov.) and the common predators in eggplant and squash fields , with evaluating the physiological and biochemical aspects of biotic stress induces by two different aphis species, infesting squash and cabbage plants," *Australian Journal of Basic Applied Sciences,* vol. 2, pp. 183-193, 2008.

[180]   N. Murugesan and A. Kavitha, "Host plant resistance in cotton accessions to the leaf hopper Amrasca devastans (Distant)," *Journal of Biopesticides,* vol. 3, pp. 526-533, 2010.

[181]   P. Zarco-Tejada, J. Miller, G. Mohammed and T. Noland, "Chlorophyll fluorescence effects on vegetation apparent reflectance: Leaf-level measurements and model simulation," *Remote Sensing of Environment,* vol. 74, pp. 582-595, 2000.

[182]   D. Lisa, R. Andrea, M. Tiffany, G. Leon, S. Gautam and D. John, "Physiological and biochemical responses of reisstant and susceptible aheat to injury by Russian wheat aphid," *Journal of Economic Entomology,* vol. 100, pp. 1692-1703, 2007.

[183]   D. John, C. John, T. Williams and R. Leslie, "Cholorophyll loss caused by soybean aphid (Hemiptera: Aphididae) feeding on soybean," *Journal of Economic Entomology,* vol. 100, pp. 1657-1662, 2007.

[184]   A. Blanchfield, S. Robinson, L. Renzullo and K. Powell, "Can leaf pigment composition help us identify grapevines infested with phylloxera ?," *Functional Plant Biology,* vol. 33, pp. 507-517, 2006.

[185]   A. Gitelson and N. Merzylak, "Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll," *Journal of Plant Physiology,* vol. 148, pp. 494-500, 1996b.

[186] P. Zarco-Tejada, J. Miller, A. Morales, A. Berjon and J. Aguera, "Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops," *Remote Sensing of Environment,* vol. 90, pp. 463-476, 2004.

[187] H. Geospatial, "Dry or Senescent Carbon," [Online]. Available: https://www.harrisgeospatial.com/docs/drysenescentcarbon.html#cellulose_absorption_index. [Accessed 28 07 2017].

[188] H. Ren, G. Zhou, F. Zhang and X. Zhang, "Evaluating cellulose absorption (CAI) for non-photosynthetic biomass estmation in the desert steppe of inner Mongolia," *Chinese Science Bulletin,* vol. 57, pp. 1716-1722, 2012.

[189] M. Mirik, G. Michels, S. Kassymzhanova-Mirik and N. Elliott, "Reflectance characteristics of Russian wheat aphid (Hemiptera: Aphididae) stress and abundance in winter wheat," *Computers and Electronics in Agriculture,* vol. 57, no. 2, pp. 123-134, 2007.

[190] E. Luedling, A. Hale, M. Zhang, W. Bentley and L. Dharmasri, "Remote Sensing of spider mite damage in California peach orchards," *International Journal of Applied Earth Observation and Geoinformation ,* vol. 11, no. 4, pp. 244-255, 2009.

[191] A. Shokoufeh, S. Hadi, R. Alireza and E. Saeid, "Feature selection using genetic algorithm for breast cancer diagnosis: Experiment on three different datasets," *Iranian Journal of Basic Medical Sciences,* vol. 19, pp. 476-482, 2016.

[192] L. Oliveira, R. Sabourin, F. Bortolozzi and C. Suen, "Feature selection using multi-objective genetic algorithms for handwritten digit recognition," in *16th International Conference on Pattern Recognition*, Quebec, QC, Canada, 2002.

[193] P. Zarco-Tejada, S. Ustin and M. Whiting, "Temporal and spatial relationships between within field yield variability in cotton and high-spatial hyperspectral remote sensing imagery," *Agronomy journal,* vol. 97, pp. 641-653, 2003.

[194] A. Shire, U. Jawarkar and M. Manmode, "A review paper on agricultural plant leaf disease detection using image processing," *International Journal of Innovative Science, Engineering and Technology,* vol. 2, no. 1, pp. 282-285, 2015.

[195] D. Bora, A. Gupta and F. Khan, "Comparing the performance of LAB and HSV color spaces with respect to color image segmentation," *International Journal of Emerging Technology and Advanced Engineering,* vol. 5, no. 2, pp. 192-203, 2015.

[196] H. Kanjalkar and S. Lokhande, "Feature extraction of leaf diseases," *International Journal of Advanced Research in Computer Engineering and Technology,* vol. 3, no. 1, pp. 153-155, 2014.

[197] D. Ghoneim, M. Alkaabi, J. De Certaines and F. Goettsche, "The impact of image dynamic range on tetxure classification of brain white matter," *BMC Medical Imaging,* vol. 8, no. 18, 2008.

[198]   P. Brynolfsson, D. Nilsson, T. Torheim, T. Asklund, C. Karlsson, J. Trygg and T. G. A. Nyholm, "Haralick texture features from apparent diffusion coefficient (ADC) MRI images depend on imaging and pre-processing parameters," *Scientific Reports,* vol. 7, no. 4041, 2017.

[199]   P. Narvekar and S. Patil, "Novel algorithm for grape leaf disease detection," *International Journal of Engineering Research and General Science,* vol. 3, no. 1, pp. 1240-1244, 2015.

[200]   K. Miles, B. Ganeshan, M. Griffiths, R. Young and C. Chatwin, "Colorectal cancer: Texture analysis of portal phase hepatic CT images as a potential marker of survival," *Radiology,* vol. 250, no. 2, pp. 444-452, 2009.

[201]   B. Ganeshan, K. Miles, R. Young and C. Chatwin, "Texture analysis in non-contrast enhanced CT: Impact of malignancy on texture in apparently disease-free areas of the liver," *European Journal of Radiology,* vol. 70, pp. 101-110, 2009.

[202]   B. Ganeshan, K. Skogen, I. Pressney, D. Coutroubis and K. Miles, "Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: Preliminary evidence of an association with tumour metabolism," *Clinical Radiology,* vol. 67, pp. 157-164, 2012.

[203]   R. Singh, "A comparison of gray-level run length matrix and gray-level co-occurence matrix towards cereal grain classification," *International Journal of Computer Engineering and Technology,* vol. 7, no. 6, pp. 9-17, 2016.

[204]   A. Esgiar, R. Naguib, B. Sharif, M. Bennett and A. Murray, "Fractal analysis in the detection of colonic cancer images," *IEEE Transactions on Information Technology in Biomedecine,* vol. 6, no. 1, pp. 54-58, 2002.

[205]   D. Majumdar, D. Kole, A. Chakraborty and D. Majumder, "Review: Detection and diagnosis of plant leaf disease using integrated image processing approach," *International Journal of Computer Engineering and Applications ,* vol. 6, no. 3, pp. 1-16, 2014.

[206]   D. Kim, T. Burks, J. Qin and D. Bulanon, "Classification of grapefruit peel diseases using color texture feature analysis," *International Journal of Agriculture and Biology Engineering,* vol. 2, no. 3, pp. 41-50, 2009.

[207]   J. Lu, M. Zhou, Y. Gao and H. Jiang, "Using hyperspectral imaging to discriminate yellow leaf curl disease in tomato leaves," *Precision Agriculture,* vol. 19(3), pp. 379-394, 2017.

[208]   D. Ashourloo, M. Mobasheri and A. Huete, "Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements," *Remote Sensing,* vol. 6, pp. 5107-5123, 2014b.

[209]   F. Ma, H. Qin, K. Shi, C. Zhou, C. Chen, X. Hu and L. Zheng, "Feasibility of combining spectra with texture data of multispectral imaging to predict heme and non-heme iron contents in pork sausages," *Food Chemistry,* vol. 190, pp. 142-149, 2016.

[210]   S. Konanz, L. Kocsanyi and C. Buschmann, "Advanced multi-color fluorescence imaging system for detection of biotic and abiotic stresses in leaves," *Agriculture,* vol. 4, pp. 79-95, 2014.

[211]   J. Roger, B. Palagos, D. Bertrand and E. Fernandez Ahumada, "CovSel: Variable selection for highly multivariate and multi-response calibration: Application to IR spectroscopy," *Chemometrics and Intelligent Laboratory Systems ,* vol. 106, pp. 216-223, 2011.

[212]   S. Bianco, C. Cusan, P. Napoletano and R. Schettini, "Improving CNN-based texture classification by color balancing," *Journal of Imaging,* vol. 3, no. 33, 2017.

[213]   J. Granzier and M. Valsecchi, "Variations in daylight as a contextual cue for estimating season time of day and weather conditions," *Journal of Vision,* vol. 14, no. 1, pp. 1-23, 2014.

[214]   C. Cusano, P. Napoletano and R. Schettini, "Evaluating color texture descriptors under large variations of controlled lighting conditions," *Journal of Optical Soceity of America,* vol. 33, no. 1, 2016.

[215]   J. Hartigan and H. P.M., "The DIP test of unimodality," *The Annals of Statistics,* vol. 13, no. 1, pp. 70-84, 1985.

[216]   W. Riedell and T. Blackmer, "Leaf reflectance of cereal aphid-damaged wheat," *Crop Science,* vol. 39, no. 6, pp. 1835-1840, 1999.

[217]   D. Goldman, "Vignette and exposure calibration and compensation," *IEEE transactions on pattern analysis and machine intelligence,* vol. 32(12), pp. 2276-2288, 2010.

[218]   W. Hugemann, "Correcting lens distortions in digital photographs," 2010. [Online]. Available: https://www.imagemagick.org/Usage/lens/correcting_lens_distortions.pdf. [Accessed 12 11 2017].

[219]   A. Wang, T. Qiu and L. Shao, "A simple method of radial distortion correction with center of distortion estimation," *Journal of Mathematical Imaging and Vision,* vol. 35(3), pp. 165-172, 2009.

[220]   D. Hadjimitsis, C. Clayton and V. Hope, "An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs," *International Journal of Remote Sensing,* vol. 25(18), pp. 3651-3674, 2004.

[221]   M. Smith, J. Edward and G. Milton, "The use of empirical line method to calibrate remotely sensed data to reflectance," *International Journal of Remote Sensing,* vol. 20(13), pp. 2653-2662, 1999.

[222]   B. Zitova and F. J., "Image registration methods: a survey," *Image and Vision Computation,* vol. 21, pp. 977-1000, 2003.