



ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR



Doctoral Dissertation
Doctoral Program in Electronics Engineering (32.nd cycle)

Advancements in Multi-Temporal Remote Sensing Data Analysis Techniques for Precision Agriculture

Aleem Khaliq

* * * * *

Supervisor

Prof. Marcello Chiaberge

Politecnico di Torino

This thesis is licensed under a Creative Commons License, Attribution - Noncommercial-NoDerivative Works 4.0 International: see www.creativecommons.org. The text may be reproduced for non-commercial purposes, provided that credit is given to the original author.

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

.....
Aleem Khaliq
Turin,

Summary

Demand in agriculture production will increase more than 50% by 2050 according to the Food and Agriculture Organization (FAO), which will create an imbalance between food demand and supply. In this regard, an essential aspect of FAO's manifesto, defined by its strategic aims, is focused on contributing towards more productive and sustainable agriculture. Precision agriculture has the potential to deal with this demand, which includes a set of tools that incorporate information acquisition, analysis, management, and deployment to help in making site-specific decisions, with the objective of maximizing production while reducing environmental noise generated by overuse of fertilizers and chemicals. Recent developments in Remote Sensing (RS) sensors and geospatial technologies are the leading sources of agriculture-related information - in the form of vegetation indexes. This goes on wards in maintaining the near real-time constraints whilst supplementing cost-effective methods of monitoring agriculture resources. There has been a rising demand for multi-temporal remotely sensed imagery to exploit vegetation dynamics. Among crops, grape production has been considered one of the most profitable asset in the agriculture sector and developing in many countries, as well organized competitive and high-value industry. Proper knowledge of the spatial variability in vineyards is considered as a key factor for vine growers to estimate the outcomes in terms of yield and quality. Monitoring vegetation status during the crop phenological cycle through means of remotely sensed data is beneficial for the winegrowers. This thesis aims at the multi-temporal analysis of remote sensing data and introducing new methodologies to address critical applications in the precision agriculture domain.

The first part is focused on the viticulture related application in which the three main problems are discussed which are 1) Vigor variability assessment of vineyards is performed by using Unmanned ground vehicle (UAV) and satellite-based multi-spectral imagery with a comparison. Advanced vegetation indices are extracted from UAV imagery to understand the influence of inter-row and vine canopies contribution in the moderate resolution satellite imagery. 2) RGB imagery acquired from UAV is used to generate a Path plan for the Unmanned Ground Vehicle (UGV) to move autonomously in a vineyard. 3) Vegetation index derived from freely available moderate spatial resolution satellite imagery refined by using UAV imagery for

the vineyard environment. Results of proposed advanced remote sensing methods proved their potential in improving the agricultural practices and decision making for site-specific vineyards, and these can be integrated with the existing practices to improve the production and quality.

The second part of the dissertation is to exploit the use of freely available sentinel-2 satellite imagery to improve the previous remote sensing methods used to address applications in precision agriculture. Two tasks are discussed in this part. 1) A new Recurrent-Convolution Neural Network (R-CNN) based approach is presented to classify different land covers and crop types by learning temporal features from sentinel-2 time series data. The results show considerable improvement as compared with the other mainstream methods. 2) A multi-temporal data analysis is performed to find a relationship between several vegetation indices derived from sentinel-2 image time series and biophysical parameters such as Above Ground Biomass and height of Maize crop.

Acknowledgements

First and foremost, I would like to thank Allah Almighty for all His blessings bestowed upon me. Without His will, I would not have made it this far.

I would like to express my sincere gratitude to my advisor Prof. Marcello Chiaberge for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides, I would also like to thank all the faculty members and researchers from the *Interdisciplinary Center of the Polytechnic of Turin, PIC4SeR*, who supported me in many ways during the course of my PhD. I would also like to take this opportunity to acknowledge the support of technical staff and my group mates (past and present) at Polito, especially Gianluca Dara, Luca Navelli, Vittorio Mazzia, and Mario Silvagni. The discussions with them have always been extremely rewarding and I learned a lot from them.

My immense thankfulness goes for my family members including my mother, wife, siblings and others for their unconditional support and prayers.

I would also like to acknowledge the support and help i received from the *Dipartimento di Scienze Agrarie, Forestali e Alimentari (DiSAFA)*, especially Lorenzo Comba from whom i learned a lot.

Finally, I am extremely thankful to the higher education commission (HEC) Pakistan for funding my doctorate at the Politecnico di Torino. This is indeed a great initiative by the government of Pakistan to prepare a group of quality researcher with the international exposure who can ultimately contribute to the prosperity of the nation.

*I would like to dedicate
this thesis to my late
Father who wanted me
to achieve this mile
stone*

Contents

List of Tables	XI
List of Figures	XII
1 Introduction	1
1.1 Precision Agriculture and Remote Sensing time series data	1
1.2 Precison Viticulture	4
1.3 Problem Statement	5
1.4 Contributions	7
1.5 Organization of the thesis	7
2 Background on Remote Sensing	9
2.1 Remote Sensing Platforms for Precision Agriculture	9
2.1.1 Satellites	11
2.1.2 Airborne	14
2.1.3 Unmanned Aerial Vehicles (UAVs)	15
2.1.4 Proximal Remote Sensing	16
2.2 Spectral Vegetation Indices	16
3 Machine Learning for Remote Sensing Data Analysis	19
3.1 Machine Learning	20
3.2 Neural Networks	22
3.3 Machine Learning paradigms for Remote Sensing	26
3.3.1 Classification	26
3.3.2 Clustering	30
3.3.3 Regression	30
3.3.4 Dimension Reduction	32
4 Research contributions related to Precision Viticulture using UAV and Satellite Imagery	33
4.1 Variability Assessment of Vineyard using UAVs and Satellite Imagery	34
4.1.1 Background and Related Work	34
4.1.2 Materials and Methods	36

4.1.3	Satellite Time Series Images	36
4.1.4	UAV-Based Imagery	38
4.1.5	In-Field Vigour Assessment	39
4.1.6	Data Processing	41
4.1.7	Results	44
4.1.8	Discussion	48
4.1.9	Conclusions	50
4.2	Automatic Path Planning for Unmanned Ground Vehicle using UAV imagery	52
4.2.1	Background	52
4.2.2	Study Area	53
4.2.3	Fieldwork and data acquisition for base map generation	53
4.2.4	Data Processing and Validation	54
4.2.5	Mask generation for path planning	55
4.2.6	Steps to generate path plan	56
4.2.7	Parcel recognition and rows clustering	56
4.2.8	Steps generation and path planning	58
4.2.9	Conclusion and future work	60
4.3	Refining Satellite based Vegetation Indices using UAV Imagery	61
4.3.1	Background	61
4.3.2	Study Area and Data Set	63
4.3.3	Data Pre-processing	63
4.3.4	Proposed Method	64
4.3.5	RarefyNet Architecture	64
4.3.6	Experimental Discussion and Results	66
4.3.7	Data set Creation	66
4.3.8	Experimental setting	67
4.3.9	Quantitative Results	68
4.3.10	Conclusions	68
5	Research contributions in Multi Temporal approaches for Agricultural Remote Sensing	69
5.1	Improvement in Land Cover and Crop Classification based on Temporal Features Learning	69
5.1.1	Background and Related Work	70
5.1.2	Study Area and Data	73
5.1.3	Convolutional and Recurrent Neural Networks for pixel based crops classification	75
5.1.4	Network Training	80
5.1.5	Data set preparation	81
5.1.6	Dataset Visualization	83
5.1.7	Experimental settings	84

5.1.8	Classification Results	85
5.1.9	Non deep learning classifiers	88
5.1.10	Conclusion and future works	90
5.2	Assessment of Biophysical Parameters of Crops using Sentinel-2 Imagery	91
5.2.1	Introduction and Background	91
5.2.2	Study Area and Data Set	93
5.2.3	Methodology	94
5.2.4	Vegetation Indexes	95
5.2.5	Field Measurements	95
5.2.6	Results Assessment	96
6	Overall Conclusions and Future Work	101
	Bibliography	105

List of Tables

2.1	Operational aspects of Remote Sensing Platforms	10
2.2	Most common satellite missions for remote sensing and their technical details	13
2.3	Types of UAVs and their discription	15
4.1	Information on satellite and UAV based acquired datasets.	38
4.2	Technical details of the considered and adopted platforms and sensors.	39
4.3	Pearson correlation coefficients results of the NDVI maps comparison procedure.	48
4.4	Results of the ANOVA of UAV based $NDVI_{vin}$ map in relation to the three vigour classes from in-field assessment.	50
4.5	Results of the ANOVA of satellite based $NDVI_{sat}$ map in relation to three vigour classes from in-field assessment.	50
4.6	Correlation of train set and test set for both network architectures.	66
5.1	Bands used in this study.	73
5.2	Sentinel-2 data acquisition.	75
5.3	Land cover types contribution in the reference data.	82
5.4	Obtained confusion matrix.	87
5.5	Comparison with recent studies.	87
5.6	Comparison of Pixel R-CNN with non deep learning classifiers.	89
5.7	Spectral bands of Sentinel-2.	94
5.8	Vegetation Indices Formulation	95
5.9	Basic statistics of the plot-wise spectral vegetation indices (VIs)	96
5.10	Basic statistics of field measurements.	97
5.11	Pearson correlation between vegetation indices derived and height and biomass field measurements for four different maize growth stages.	100

List of Figures

1.1	Illustration of Evolving Precision Agriculture [1]	2
2.1	Illustration of remote sensing platform for precision agriculture [1]. .	10
2.2	Spectrum	11
2.3	Illustration of observational cube of remote sensing platforms [218].	12
3.1	General overview of classifier categories [23].	21
3.2	Basic units on neural networks.	22
3.3	Back propagation neural network (BPNN) structure.	23
3.4	convolutional neural network (CNN) structure.	24
3.5	Recurrent neural network (RNN) structure.	25
3.6	Most common machine learning methods for remote sensing data analysis.	27
4.1	Selected test field located in Serralunga d’Alba (Piedmont, northwest of Italy). The boundaries of the three considered parcels, named “Parcel-A”, ”Parcel-B” and “Parcel-C”, are marked with solid green polygons. The cropland region, represented by pixel $s_{8,20}$ of the Sentinel-2 tile, is highlighted by a yellow square. The map is represented in false colours (NIR, Red and Green channels).	37
4.2	Enlargement of UAV-based multispectral imagery, represented in false colours (NIR, Red and Green channels), of: (a) : “Parcel-A”; (b) : ”Parcel-B”; and (c) : “Parcel-C”.	40
4.3	(a) Ordered grid of pixels $s(i, j)$ belonging to satellite tile \mathcal{S} , located at latitude and longitude coordinates $\alpha_s(i, j)$ and $\beta_s(i, j)$; and (b) ordered grid of pixels $d(u, v)$ belonging to satellite imagery \mathcal{D} , located at $\alpha_d(i, j)$ and $\beta_d(i, j)$. Selected UAV pixels belonging to $\mathcal{G}(i, j)$, used for comparison to satellite pixel $s(i, j)$, are highlighted in light green.	41
4.4	(a) Enlargement of subset $\mathcal{G}(8,20)$ of UAV map \mathcal{D}_2 , highlighted by a yellow square in Figure.4.1, is represented in false colours (NIR, Red and Green channels); (b) classification of pixels $d(u, v) \subset \mathcal{G}(8,20)$ into two classes: \mathcal{G}_{vin} , representing vine canopies (green), and \mathcal{G}_{int} , representing inter-row surfaces (brown); (c) computed NDVI values of vine canopies pixels \mathcal{G}_{vin} ; and (d) inter-row surface \mathcal{G}_{int}	43

4.5	Comprehensive (a) $NDVI_{sat}$ map, computed from satellite imagery \mathcal{S}_2 , and (b) $NDVI_{uav}$ derived from UAV imagery \mathcal{D}_2 . (c) Enhanced vineyard $NDVI_{vin}$ map, processing UAV imagery \mathcal{D}_2 by considering only canopy pixels \mathcal{G}_{vin} and (d) $NDVI_{int}$ map considering only inter-row surface \mathcal{G}_{int} . In all represented NDVI maps, only pixels (i, j) completely included within “Parcel A”, “Parcel B” and “Parcel C” boundaries are shown.	45
4.6	Scatter plots of NDVI values from $NDVI_{sat}$ map (x-axis) and: (a) the comprehensive $NDVI_{uav}$ map (y-axis); (b) the enhanced NDVI values of map $NDVI_{vin}$ (y-axis); and (c) the enhanced NDVI values of $NDVI_{int}$ map (y-axis), using imagery pair data $\mathcal{D}_2/\mathcal{S}_2$. The regression model and data pair correlation coefficients are also reported.	46
4.7	Vineyard test site classification into three vigour classes on the basis of the observed in-field vigour assessment. Classes “L”, “M” and “H” refer to low, medium and high vigour, respectively.	47
4.8	Box plots representation of: (a) $NDVI_{sat}$ values derived from satellite imagery; and (b) enhanced $NDVI_{vin}$ values derived from UAV imagery, considering only canopy pixels, divided into three groups on the basis of the observed in-field vigour classes “L”, “M” and “H”.	49
4.9	UAV imagery acquisition points and positions.	54
4.10	Ground Control Points placements for georeferencing.	55
4.11	Binary mask representing vine rows generated by 3D geometric information. points and positions.	55
4.12	Workflow of UGV path generation.	56
4.13	Clustering.	57
4.14	Mean points of rows.	58
4.15	Zoomed in view of rows connection.	59
4.16	Case Study. From left to right: overview, vineyard enhancement, path plan.	60
4.17	Study site located in Serralunga d’Alba (Piedmont, northwest of Italy): map derived from UAV is shown and Sentinel pixels are shown as overlapped region for vineyard fields.	62
4.18	Overview of the designed network, named RarefyNet, used for refining Sentinel pixels.	64
4.19	Detailed view of proposed Inception block used in the first two layer of the overall architecture mentioned in Fig.4.18.	65
4.20	Scatter plot with regression equation and correlation coefficient (a) relation between Sentinel raw pixels and ground truth is presented, (b) coherence trend between predicted Sentinel pixels and ground truth.	66

5.1	The study site is located in Carpi, region Emilia-Romagna is shown with the geo-coordinates (WGS84). RGB image composite derived from sentinel-2 imagery acquired in August-2015 is shown and the yellow marker showing geo-locations of ground truth land cover extracted from Land Use and Coverage Area frame Survey (LUCAS-2015).	74
5.2	An overview of Pixel R-CNN model used for classification. Given a multi-temporal, multi-spectral input pixel $X^{(i)}$, the first layer of LSTM units extracts sequences of temporal patterns. A stack of convolutional layers hierarchically processes the temporal information.	76
5.3	A recurrent layer and its unrolled through time representation. A multi-temporal, multi-spectral pixel $X^{(i)}$ is made by a sequence of time steps, $x_t^{(i)}$, that along the previous output $h^{(i)}$ feed the next iteration of the network.	77
5.4	LSTM with peephole connections. A time step t of a multi-spectral pixel $x_t^{(i)}$ is processed by the memory cell which decides what to add and forgot in the long-term state $c_{(t)}$ and what discard for the present state $y_t^{(i)}$	78
5.5	Overview of the tensor $\mathbf{X} \in \mathbb{R}^{i \times t \times b}$ generation. The first dimension i represents the collected instances $X^{(i)}$, the second t the different time steps and b the number of spectral bands.	82
5.6	Visual representation of the data points projected in the tri-dimensional space using PCA. The three principal components taken into account preserve 64.5% of the original data set variance.	83
5.7	Pareto Chart of the explained variance as a function of the number of components.	84
5.8	(a). Final classified map using Pixel R-CNN, (b). zoomed in region of the classified map, and (c). Raw Sentinel-2 RGB composite of the zoomed region.	85
5.9	The study site of city carpi, region emilia-romagna is shown with the geo-coordinates (WGS84). RGB image composite derived from sentinel-2 imagery acquired in August 2018 is shown and the red marker showing location points where physical observations were taken.	93
5.10	Adopted methodology for comparing spectral vegetation indices (VIs) derived from satellite imagery with biophysical parameters of maize crop.	94
5.11	Maize growth development stages with satellite data acquisition dates along with field crop height and biomass measurements.	96
5.12	False color representation of Vegetation indices (a) GLA ,(b) GNDVI, (c) RVI, (d) NDVI, (e) EVI, (f) WI derived from sentinel-2 imagery for the study area.	97

5.13 Relationship between field-measured canopy height $H_{canopy}(m)$ and VIs (NDVI, GNDVI, RVI, EVI, WI and GLA) derived from satellite imagery acquired on maize growth stage T3.	98
5.14 Relationship between field-measured biomass (AGB kg/m^2) and VIs (NDVI, GNDVI, RVI, EVI, WI and GLA) derived from satellite imagery acquired on maize growth stage T3	99

Chapter 1

Introduction

The global human population is rapidly growing and expected to be 10 billion by 2050, demanding a significant boost in agriculture sector to meet the food requirements [54]. Meeting the increased demand is a big concern with the modest economy growth. Another challenge for the agriculture sector is to meet the food demand of growing population whilst finding more efficient ways of managing the involved key resources such as soil fertility, efficient land use, and water in more efficient way. Another crucial factor at play here is climate change. In this regard, a key aspect of FAO's manifesto, defined by its strategic aims, is to make a contribution towards more productive and sustainable agriculture. In this challenging context for agriculture, monitoring crop growth and status with multi temporal approach in spatially diverse locations and different environmental conditions is strongly needed. Monitoring is required in near real-time to deal with the extreme change in climate conditions which can be helpful to minimize their impact on the food production by optimizing agricultural practices in a sustainable way [232]. Yield prediction is one of the valuable factor that can anticipate food production and ensures food security in case of food shortage. Precision agriculture (PA) is an evolving farm management technique (Figure. 1.1), that has the potential to deal with above requirements. This is discussed further in the following section.

1.1 Precision Agriculture and Remote Sensing time series data

Precision agriculture (PA) is a farming management technique which includes a set of tools that incorporate information acquisition, analysis, management and deployment to help in making site-specific decisions, with the objective of maximizing production while reducing environmental noise generated by overuse of fertilizers and chemicals. Information acquisition is the essential step in PA. Remote sensing is well established and widely used tool in PA that provides solutions to address

the above mentioned requirements since it provides recurrent information with out physical interaction from the small to large scale agricultural areas in a concise way. Use of remote sensing time series data in PA has been investigated many times in numerous studies, specifically in crops growth assessment, yield estimation, irrigation management, and weed detection using one of the remote sensing platforms such as Satellites, Airborne, Unmanned Aerial Vehicles (UAVs), Unmanned Ground Vehicles (UGVs) and Ground based sensors mounted on the agricultural machines. Multi spectral time series data acquired from MODIS and LANDSAT have been widely used in many agricultural applications such as crop yield prediction [95], landcover and crop classification (LC&CC) [192], leaf area index estimation [92], plant height estimation [121], vegetation variability assessment [191] and many more. Two different data sources can also be used together to extract more features that lead to improve results . For example, Landsat-8 and Sentinel-1 used together for LC&CC [112]. The increasing trend of the remote sensing applications in the agriculture sector is evident from significant progress in technological aspects of remote sensing platform that include availability of data from numerous sensors with the unprecedented spatial, spectral and temporal resolution. The emergence of platforms such as nano satellites or UAVs offer more flexibility in terms of data acquisition planning.

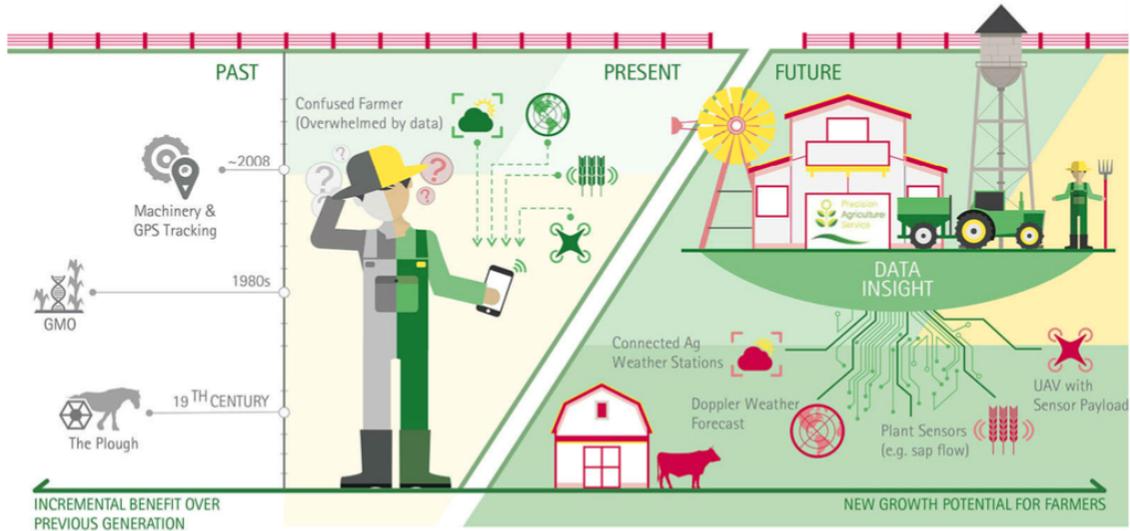


Figure 1.1: Illustration of Evolving Precision Agriculture [1]

Over the years, statistical analysis of data derived from remote sensing measurements has been in practice to measure changes in nature and manage land cover such as water bodies, crops and forests. Special attention has been paid to remote sensing data analysis in order to use in production of environmental and agricultural statistics. Statistical machine learning methods in remote sensing data

analysis can be defined as tool in such scenarios where a statistical relationship between spectral information coming from remote sensing platforms and ground/field measurements is established. Depending on the cases, derived relationship can be parametric, semi parametric or non-parametric. For example, Logistic regression [14], Support vector machines [213], Decision tree [155], K nearest neighbor [140], Random Forest [217], have been used in the analysis of agriculture remote sensing applications.

Image classification of remote sensing images plays an essential role in a broad range of applications and therefore receiving remarkable attention. Through the last decades, considerable efforts have been made to develop various methods for the task of scene classification using satellite or aerial images. As scene classification is usually carried out in feature space, effective feature description plays a vital role in creating high-performance scene classification methods. We can generally divide existing scene classification methods into three main categories according to the features they used: handcrafted feature-based methods [29, 28], unsupervised feature learning-based methods [30, 254], and deep learning based methods [259, 257].

Most of the current state-of-the-art approaches generally rely on supervised learning to obtain good feature representations. Especially in 2006, a breakthrough in deep feature learning was made by Hinton and Salakhutdinov [83]. Since then, researchers have aimed to replace hand-engineered features with trainable multi-layer networks, and several deep learning models have shown impressive feature representation capability for a wide range of applications including remote sensing image scene classification [30, 254, 132, 137]. Traditional supervised methods are based on handcrafted features that require a considerable amount of engineering skill and domain expertise; however, deep learning features are automatically learned from data using a general-purpose learning procedure via deep-architecture neural networks. On the other hand, compared with aforementioned unsupervised feature learning methods that are generally shallow-structured models (e.g., sparse coding), deep learning models that are composed of multiple processing layers can learn more powerful feature representations of data with multiple levels of abstraction [116]. In addition, deep feature learning methods have also turned out to be very good at discovering intricate structures and discriminative information hidden in high-dimensional data, and the features from top layers of the deep neural network show semantic abstracting properties. All of these make deep features more applicable for semantic-level scene classification.

In this dissertation, learning based methods have been employed for classification problem and refinement of satellite imagery. A detailed discussion about the learning based methods is provided in state of the art methods for remote sensing Chapter. 3. The main goal of this thesis is to analyze and use multi-temporal remote sensing data from multiple sources for several agricultural applications to further improve the efficacy of practices used in Precision Agriculture. In first part of work,

high resolution UAV and moderate resolution satellite imagery is processed and analyzed to address three research problems related to viticulture which are further highlighted in the following section. The second part presents (i) methodology and results for the landcover and crop classification, (ii) the performed analysis on assessment of relationship between remotely sensed spectral vegetation indices and crops biophysical parameters.

1.2 Precison Viticulture

Precision Viticulture (PV) is an emerging practicing technique in vineyards based on the same principal as precision agriculture. Precision farming is required to maximize vineyard efficiency, which guarantees to achieve optimal grape yield and quality while mitigating environmental impacts and risks. Grape growers are actively searching the ways to increase their profits by maximizing yield while optimizing fertilization and reducing environmental impact [8, 199]. Precision viticulture relies on rapidly evolving technologies for evaluating vineyard variability such as global positioning systems (GPS), meteorological, satellite and airborne remote sensing, and geographical information systems (GIS). A proper knowledge of the spatial variability between and within crop parcels is considered as a key factor for vine growers to estimate the outcomes in terms of yield and quality [20, 205, 169]. In this context, remote sensing (RS) has already proved its potential and effectiveness in spatio-temporal vegetation monitoring [74, 114, 72, 168], describing several crop biophysical characteristics such as biomass and height of the crop [94, 136]. There are types of crops which are grown in a row formation in particular, vineyards. Agronomists adopt the traditional practices such as keeping certain inter row spacing between the vinerows to achieve optimal yield. However in PV related past works, contribution of inter-row soil in remotely sensed data has not been properly addressed, particularly when the moderate resolution satellite based sensors are used. Very few studies compared the satellite and UAV based data for variability assessment in vineyard environments with the consideration of inter-row soil contribution, a detail discussion can be found in section 4.1. Therefore, it is indeed a research problem which has to be addressed in the context of viticulture. In this dissertation, a detailed analysis and comparison of vineyards multi spectral imagery, provided by a decametric resolution satellite and low altitude UAV platforms, is presented in section 4.1.

Moreover, increase in the use of automated machines in agricultural environment to minimize human work-load to perform related tasks, has garnered significant attention from farmers and industries alike. Automated vehicles need to have an automatic navigation system and they have to be able to autonomously follow the path according to the specific area and obstacles. Several advancements in path planning methods for agricultural productions using different optimization

techniques for single or multiple coordinated vehicles which span the entire field, are discussed in [57]. However, the path planning for autonomous vehicle for a hilly vineyard environment is challenging. Section 4.2 of this thesis is dedicated to the path planning research problem in which UAV imagery over a vineyard is used to generate a geo-coordinated path plan for automated UGV.

Satellite based imagery with moderate spatial resolution may not be used directly to reliably describe the vegetation variability in case of crops where inter-row surface involves considerable portion of cropland. Few studies have used centimetric imagery acquired from UAVs to improve satellite imagery based on remotely sensed spectral indices. In [91], estimates of canopy structure and biochemical parameters derived from moderate spatial resolution satellite imagery were improved using high resolution UAV multispectral imagery. T.W. Cui et al [44] described how high resolution airborne SAR imagery can be used to refine satellite imagery driven estimates of macroalgal coverage in the yellow sea. It is worth taking into account that frequent availability of free satellite imagery can be very useful if vegetation indices derived from it, could be refined. Therefore, section 4.3 of this thesis is dedicated to satellite imagery refinement part in which UAV imagery with centimetric spatial resolution is used to refine the vegetation index derived from decametric sentinel-2 imagery.

Second part of this work is based on sentinel-2 image time series data and its application to support the main objective which is to improve performance of agricultural practices. Many active satellites missions, equipped with the optical sensors such as multi spectral and hyper spectral sensors are providing data with unprecedented temporal and spectral resolutions. However freely available satellite data sources do not offer high spatial resolution data product, for example, MODIS can provide from 250m to 1000m per pixel Ground Sampling Distance (GSD), whereas LANDSAT-8 optical land imager (OLI) offers 15m to 30m GSD. Recently launched sentinel-2 satellite, equipped with optical multi spectral imager that can provide up to 10m per pixel spatial resolution with the revisit time of 5 days and with almost the same spectral resolution as provided by LANDSAT-8 OLI, which offers great opportunity to be exploited in remote sensing domain. Taking advantage of these features, dense time series of sentinel-2 can be very useful in near-real time crop growth monitoring, yield prediction, land cover and crop classification, biophysical parameters estimation and other agricultural applications to achieve the goal of sustainable agriculture.

1.3 Problem Statement

Time series remote sensing data is one of the most valuable resource to address the wide range of agricultural applications to achieve the goal of sustainable agriculture set by the FOA and UN. Moreover, there is a rapid increase in the volume and

availability of multi temporal images. This is due to the growing number of space missions, the improvement in temporal resolution and the free data access policy introduced for satellite missions such as Sentinel and Landsat. In this context, there is room for the development of both novel methodologies and applications for image time series employment.

This thesis aims at introducing and implementing new methodologies to address key applications related to precision agriculture by performing multi temporal analysis of remote sensing data. Precision farming is essential to maximize vineyard efficiency, which guarantees to achieve optimal grape yield and quality while mitigating environmental impacts and risks. A proper knowledge of the spatial variability between and within crop parcels is considered as a key factor for vine growers to estimate the outcomes in terms of yield and quality. Any contribution in this field of study will have valuable impact on the sustainable agriculture goal as well as in economic returns. Therefore, a detailed analysis is needed to compare variability assessment of vineyards using time series of Satellite and UAV-Based imagery. Indeed, many research works have been carried out in the past related to growth monitoring, yield prediction, biophysical parameters estimation of various crops like Maize, Soya, Wheat using freely available multitemporal satellite imagery. However, the same imagery is effected by the inter row surfaces in case of vineyards. Therefore, refinement is needed to use freely available satellite imagery. Another research problem related to viticulture, discussed in this thesis, is the path planning of unmanned ground vehicles in the challenging vineyard environment.

Second part of the thesis work is only associated with the Sentinel-2 multi temporal analysis and its applications in precision agriculture. Understanding the use of current land cover, along with monitoring change over time, is vital for agronomists and agricultural agencies responsible for land management. Land cover and crop classification is widely addressed topic in field of agricultural remote sensing. However there is still room to address this problem for the large scale classification. Moreover, advancements in machine learning techniques has shown great potential to deal with the huge data volume, for example multi temporal satellite data for large scale area. Therefore, machine learning techniques can be employed to improve the performance for large scale land cover and crop classification. Furthermore, with the availability of reasonably wide spectral range of on-board sensor of sentinel-2, many spectral vegetation indices can be derived and then used in various agricultural applications discussed in section 1.1. Various spectral indices derived from sentinel-2 will be exploited in this thesis to assess the relationship with the crop's biophysical features such as Above Ground Biomass (AGB) and the Leaf Area Index (LAI).

1.4 Contributions

In this dissertation, the focus of the research is to exploit multi-temporal remote sensing data from different sources for agricultural monitoring applications in order to improve the efficacy of practices used in precision agriculture. This work was mainly carried out for two research themes. First, to address the applications of UAV and satellite based remotely sensed imagery in the field of viticulture. The second is to use sentinel-2 multispectral image time series for land cover and crop classification, and to estimate the crop biophysical parameters. Major contributions in this research are listed as follows.

- A detailed analysis and comparison between UAV and satellite based multi-spectral imagery is performed to assess the variability in vineyard. An effective approach is adapted to investigate the effect of spectral reflectances from inter-row soil on the satellite driven vegetation index. For robust analysis, three different NDVI maps are derived by considering: (i) the whole cropland surface; (ii) only the vine canopies; and (iii) only the inter-row terrain.
- A frame work is devised to generate an automatic coverage path plan for agricultural unmanned ground vehicles (UGV) by using the RGB imagery acquired from unmanned aerial vehicle (UAV) of hilly vineyard environment.
- CNN based approach is proposed to refine the moderate resolution satellite driven vegetation maps by using spatial information from its neighbors and high resolution UAV imagery.
- An improved land cover and crop classification method is developed based on temporal feature learning from multi-temporal satellite imagery. A supervised deep learning based frame work is adopted that comprises of Recurrent Neural Network (RNN) that extracts temporal correlations from time series of sentinel-2 data in combination with Convolutional Neural Network (CNN) to analyzes and encapsulate the crops pattern through its filters.
- Several spectral vegetation indices are derived from sentinel-2 multi-temporal images for maize crop and their relationship are assessed with the in-field measurements of biophysical parameters such as Above Ground Biomass (AGB) and plant height.

1.5 Organization of the thesis

This thesis presents a collection of work done in the field of Precision Agriculture by analyzing multi temporal remote sensing data and developing methodologies for agricultural applications to further improve the efficacy of practices used in farming systems. This thesis is organized as under:

- **Chapter 2** provides a brief introduction to Remote sensing and its role in the precision agriculture applications. It also provides technical background about various remote sensing platforms like satellites, Airborne, UAVs and their contribution in agricultural applications. Furthermore, it gives introduction to several spectral vegetation indices derived from multi spectral imagery including their applications in the precision agricultural domain.
- **Chapter 3** discusses state of the art methods and their contributions to exploit the potential of huge volume of remote sensing data. It provides discussion on several machine learning methods, which includes regression, classification, and clustering. Brief introduction on neural networks and their role in supervised classification is provided.
- **Chapter 4** presents research contributions during doctoral studies in which several methods are proposed to analyze remote sensing data and its applications in the field of Precision Viticulture. This chapter is divided into three sections, and each section is dedicated for one addressed application. Section 4.1 will provide a detailed analysis and comparison of UAV and satellite based imagery to assess the variability in vineyard. Section 4.2 discusses an application of UAV imagery to generate a path plan for ground vehicle for a challenging vineyard environment. Finally in section 4.3, a methodology to refine satellite based vegetation indices using high resolution UAV imagery is presented.
- **Chapter 5** presents multi temporal approaches to exploit the potential of free available moderate spatial resolution satellite imagery. First, a novel methodology for land cover and crop classification based on Recurrent-Convolutions Neural Network is presented in section 5.1. Moreover in section 5.2, relationship between satellite driven spectral vegetation indices and biophysical parameters of crop is analyzed and presented.
- **Chapter 6** concludes all the work done during the doctoral studies with some recommendations for future work.

Chapter 2

Background on Remote Sensing

Remote sensing is a widely used technique which acquires information of an object or observes phenomenon from a distance using sensors without making physical contact. A few decades back, it was initially known for acquiring data from spaceborne satellite and airborne platforms equipped with the optical and radar sensors [22]. In recent years, image and spatial data acquisitions from Unmanned Aerial Vehicles, Unmanned Ground Vehicles and Ground sensors are also considered as remote sensing methods. Because of exceptional technological developments, the remote sensing has significantly used for various applications, and also, the performance has drastically improved [218]. It is now hard to delimit the boundaries of remote sensing applications as the observation space is not limited to only the land cover classifications and traditional topographic mapping but also being used in other disciplines such as agronomy, hydrology, meteorology, geology, oceanography, etc. Remote sensing is gaining immense importance with the advancements in technological aspects, such as imaging sensor developments and exponential rise in information infrastructure, including processing data, data storage and communication among the technological devices. Moreover, with the advancements in aerial platforms in particular, UAS (Unmanned Aircraft System) also known as UAV (Unmanned Aerial Vehicle), have increased the performance of remote sensing methods and broaden up the application areas [156].

2.1 Remote Sensing Platforms for Precision Agriculture

Remote sensing technique aims to observe the physical parameters of the desired object space with time. There are four key characteristics of any remote sensing platform that define its applicability and performance: (1) the sensor type: sensors are mainly categorized as active for example Synthetic Aperture Radar (SAR) and passive such as optical imager and thermal imager, (2) spatial resolution also

known as GSD (Ground Sampling Distance), (3) temporal resolution: revisit time of the platform which determine the data acquisition frequency, (4) object range, the distance between the sensing platform and observed object. There are several operational aspects associated with the existing remote sensing platforms are listed in Table.2.1 that are considered before selection.

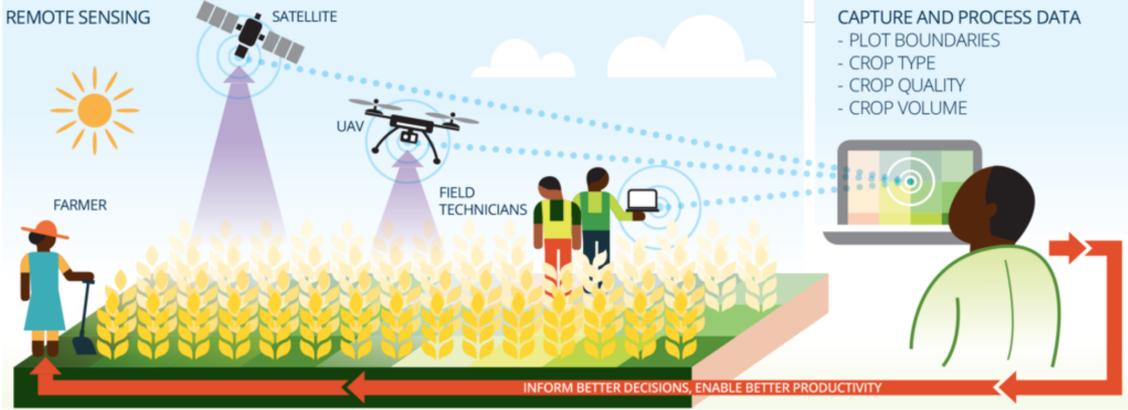


Figure 2.1: Illustration of remote sensing platform for precision agriculture [1].

Table 2.1: Operational aspects of Remote Sensing Platforms

Operational Aspects	Remote Sensing Platforms			
	Satellite	Airborne	UAS	Mobile/static (ground)
Observation space	Worldwide	Regional	Local	Local
Sensor diversity	MS/HSI/SAR	MS/HSI/LiDAR/SAR	MS/LiDAR/HIS	MS/LiDAR (HSI)
Maneuverability	No/limited	Moderate	High	Limited
Ground coverage	Large (10 km)	Medium (1 km)	Small (100 m)	Small (50 m)
FOV	Narrow	Wide	Wide/super wide	Wide/super wide
Revisit rate	Day	Hours	Minutes	Minutes
Spatial resolution	0.30–300 m	5–25 cm	1–5 cm	1–5 cm
Spatial accuracy	1–3 m	5–10 cm	1–25 cm	3–50 cm
Deployability	Difficult	Complex	Easy	Moderate
Operational risk	Moderate	High	Low	Moderate
Cost	High	Medium	Low	Low

Remote sensing platforms are typically categorized according to the sensor system type, including platforms based on satellite, aerial, and ground sensors (Figure. 2.3). These platforms with their accompanying imaging systems can be distinguished based on the altitude of the platform, the spatial resolution of the equipped sensor, and the minimum revisit time of the platform. Spatial resolution determines the area that defines the smallest pixel of the imaging systems. Low spatial resolution images imply large pixels with increased soil or plant heterogeneity. For the evaluation of temporal patterns in soil or plant features, platform revisit time

is significant. Images acquired by satellite and aerial platforms are often severely affected by cloud cover [147], while the ground-based remote sensing is not relevant.

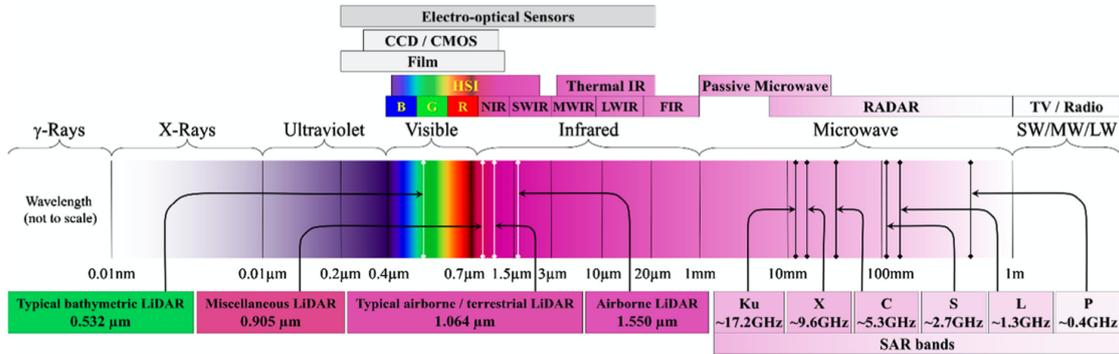


Figure 2.2: Spectrum .

2.1.1 Satellites

Satellites equipped with sensors have been the source of information mainly for earth observation programs since the past four decades. NASA launched its first Landsat-1 mission in 1972. Later, with the technological developments in satellite systems, a new era of commercial satellites was begun after the launch of IKONOS in 1999. Role of satellite-based remote sensing can be realized by the fact that more than 50 countries are presently operating and nearly 600 satellites carry sensors that are used in Earth observation missions. Mounted sensors on satellites capture data that span the electromagnetic spectrum, detect visible light, thermal infrared, and SAR (Synthetic Aperature Radar), spectrum is shown in Figure. 2.2. The type of sensor and the type of orbit that the platform travels is based on the monitoring mode and data collection. Satellite-based remote sensing platforms are considered to be the most stable and reliable platforms as compared with others such as airborne, UAS and ground-based.

From the beginning era of satellite missions, satellites with the on-board sensor continue to progress in imaging systems by their performance, and incorporating better spatial and spectral resolution [166]. Multiple sensors were also considered in a cooperative sense and gained significant attention of the remote sensing research community. The type of satellite orbit generally determines the monitoring capabilities of a satellite. The orbit that satellite follows is a circular or elliptical designated by the satellite system. Different orbital path (polar orbit, sun-synchronous orbit, geostationary orbit) are followed to achieve continuous monitoring (meteorology) and mapping (land cover mapping). Satellites fly in cycles for specific sensors [182],

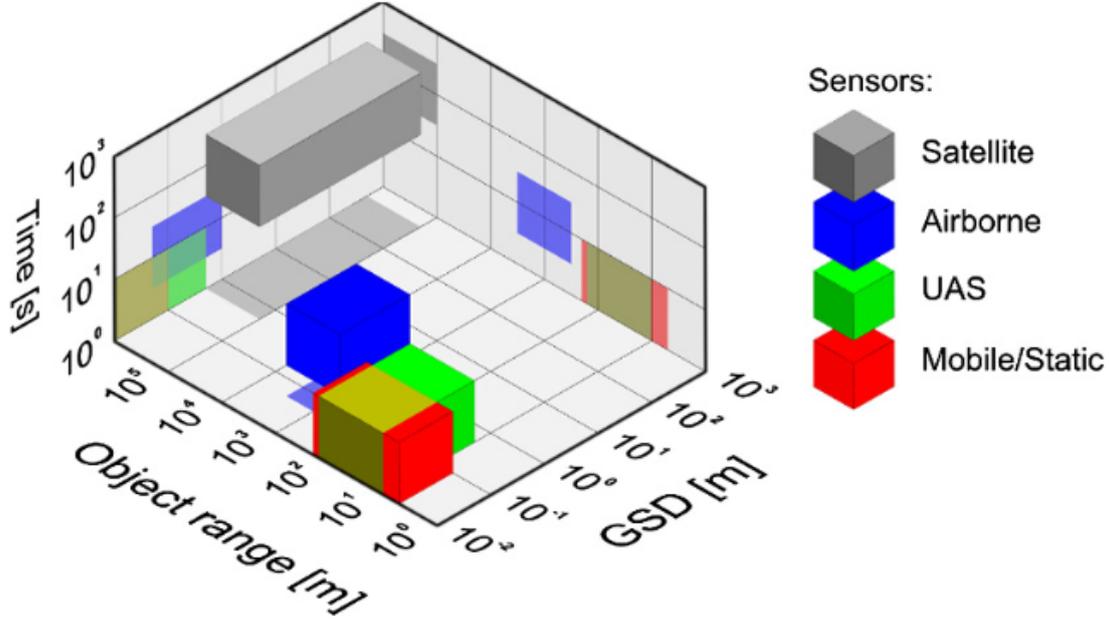


Figure 2.3: Illustration of observational cube of remote sensing platforms [218].

commonly in constellations, results in shorter revisit times, and the Earth globally may be observed several times a day.

Recent satellite missions, such as Landsat, Sentinel-2, offer high temporal resolution as compare to other freely available multispectral data sources, therefore these are preferred. Due to this fact, extensive time-series data can be acquired for numerous applications such as vegetation monitoring, ocean monitoring, post-earthquake analysis, urban land change etc. The images captured by satellite-based sensors cover a large area that is used to observe the desired phenomena at a global scale. There are external sources that act as noise while acquiring images by the remote sensing platforms such as atmospheric interferences, geometric noise. However, state of the art hardware and software (atmospheric corrections, geometric correction, georeferencing, sensor calibration etc.) resources are used in satellite-based remote sensing to cope with the interferences which make this platform as the most stable and reliable among all others. Drawbacks of satellite-based platforms are their high cost and their strictly fixed revisit schedule, so data cannot be acquired at desired days and timings as needed in some applications. The revisit time of the platforms varies from twice in one day to 16 days, depending on the orbital path of the satellite. The other limitation is with those satellite platforms which are equipped with optical sensors are sensitive to weather conditions, especially when the weather is cloudy. Many studies have addressed this issue and proposed various cloud masking methods and achieved some excellent results but not suffice the limitation [236].

Table 2.2: Most common satellite missions for remote sensing and their technical details

Name	Launch	Country	Sensor	GSD (m)	Range	Swath Width (km)	Revisit Time (day)
Aqua (MODIS)	2002	USA	36 HSI	1000 x 1000 500 x 500		2330	1–2
Ikonos-2	1999	USA	PAN 4MS	0.8 x 0.8, 3.2 x 250 x 250		11.3	3
RapidEye	2008	Germany	5 MS	6.5 x 6.5		77	1–5.5
QuickBird-2	2001	USA	PAN 4 MS	0.7 x 0.7 2.6 x 2.6		16.8–18	1–3.5
Pleiades 1	2011 2012	France	PAN 4 MS	0.5 x 0.5 2 x 2		20	1
Sentinel-1	2014 2016	EU	C-band SAR	5 x 5 5 x 20 25 x 40		80 250 400	12 6 (dual)
WorldView-3	2014	USA	PAN, 8 MS, 8 MS (SWIR), 12 MS	0.3 x 0.3 1.2 x 1.2 , 3.7 x 3.7		13.1	1–4.5
Landsat-8	2013	USA	PAN, 11 MS	15 x 15 , 30 x 30		185	16
Sentinel-2	2015 2016	EU	13 MS	10 x 10 , 20 x 20 60 x 60		290	10 5 (dual)
EnMap	2017	Germany	232 HIS	30 x 30		30	4
ICESat	2003	USA	2 HIS	70 (footprint)		N/A	8
TanDEM-X	2007	Germany	X-band SAR	1 x 1 3 x 3 16 x 16		5 x 10 1500 x 30 1500 x 100	11
SkySat	2013 2014 2015	USA	PAN Video PAN 4 MS	1.1 x 1.1 0.9 x 0.9 2 x 2		2 x 1 8	0.5 (2015) 0.12 (2017)
ICESat-2	2018	USA	1 HIS (9-beam)	10 (footprint)		N/A	N/A
Sentinel-3	2015 2017	EU	21 MS 11 MS (IR)	300 x 300 500 x 500 1000 x 1000		1270 1420 750 (nadir)	0.25
RADARSAT-2	2007	Canada	C-band SAR	3 x 3 100 x 100		20 500	24 (OR)
SPOT-6 SPOT-7	2012 2014	France	PAN 4 MS	1.5 x 1.5 6 x 6		60	1–5
TerraSAR-X	2007	Germany	X-band SAR	1 x 1 16 x 16		5 x 10 1500 x 100	11

List of satellites which are more common in the remote sensing research community is presented in Table.2.2 with their specifications [187]. Among all these, some satellite provides data freely, which is available online, while others offer a commercial solution. The commercial satellites such as Pleiades-1 launched in 2011 equipped with panchromatic (PAN) sensor (GSD 0.5x0.5m) and multi spectral (MS) sensor (GSD 2x2m), and revisit time of one day that makes this platform promising where finer details with the high temporal resolution is needed. QuickBird-2 is another example of a commercial solution that comes with PAN and four MS imaging sensors with a GSD of $0.7 \times 0.7\text{m}$ and $2.6 \times 2.6 \text{ m}$ respectively with a temporal frequency of 1–3.5 days. Indeed, these commercial solutions offer data products with better spatial and temporal resolution but at the expense of high cost, which may not be suitable to promote the research culture and its implementation for remote sensing applications. In contrast to commercially available satellites, there are some satellite missions come with open access data policy, which for example, data products of Landsat-8 and Sentinel are freely available with moderate spatial and temporal resolutions. Landsat-8 was launched in 2013 by the United States Geological Survey (USGS) that provides PAN and 11-MS data with 16-day revisit time.

Senitnel-2(A/B) mission was launched by European Space Agency in 2015 and 2017 respectively which comprises a constellation of two polar-orbiting satellites. Key features of sentinel-2 satellite such as wide swath with (290km), High revisit time (5 days), fine spatial resolution (10m) and spectral resolution (13 bands) make it promising among other freely available satellite platforms. It is widely used in various applications such as vegetation monitoring, forest management, maritime monitoring.

2.1.2 Airborne

The airborne platform had been the main source of geospatial data in the past. Since the launch of commercial satellites for land remote sensing systems with 30cm imagery became commercially available, the difference between spaceborne and airborne imagery is becoming negligible. However, with the introduction of a new state of the art sensing technology, LiDAR sensors can be effectively used with the airborne platform at a lower speed to offer better point distribution. Airborne platforms are flexible compared to satellite platforms as the flight schedule, types of sensors, speed of the aircraft are in human control. Nevertheless, the coverage of this platform can be at a regional level which is smaller than satellite-based ones, but rather greater than the UAV platforms. Some examples of airborne platforms used for remote sensing [187] are mentioned in Table.2.3.

2.1.3 Unmanned Aerial Vehicles (UAVs)

Unmanned Aerial Vehicles also frequently termed as drones, are controlled remotely by an operator, currently gaining immense attention among the remote sensing research community due to recent technological advancements. Over the past few years, research and development in UAV systems have broadened up the application areas. UAVs in variant sizes with different weights are available, the remote sensing community, however, is mostly dealing with the lightweight systems. Among all the UAV applications, remote sensing shares a little percentage of the market at the moment; however, it will increase with the reduction of cost of the sensors and with relaxed aviation/airspace policies.

Table 2.3: Types of UAVs and their discription

Weight (kg)	UAV Type	Manufacturer /Model	Flying Time (min)	Flying Speed (m/s)	Sensors
0.7	Fixed-wing	senseFly/eBee RTK	40	11–25	Camera
1.3	Quadro copter	DJI/Phantom 2	25	15	Camera
2.5	Fixed-wing	Trimble/UX5	50	22	Camera
2.7	Fixed-wing	Topcon/SIRIUS PRO	50	18	Camera
6.1	Fixed-wing	AeroVironment/Puma AE	210	23	Camera
6	Quadro copter	Microdrones/MD4-1000	90	12	Camera/LiDAR
4.6–6.6	Hexacopter	Aibotix/Aibot X6	30	14	Camera
5	Fixed-wing	Trigger Composites/Pteryx	120	12.5–15	Camera
5.1–5.8	Fixed-wing	Hawkeye UAV/AeroHawk	90	16.5–19.5	Camera
6.9–9.5	Hexacopter	TRGS/Li-AIR	15	8	LiDAR
9.5	Octocopter	Altus UAS/Delta X8	10–14	12	Camera/LiDAR
25	Octocopter	Riegl/Ricopter	30	22	LiDAR/camera
38	Fixed-wing	American Aerospace/RS-16	720–960	33	Camera

UAV platforms are a dynamic, customized and budget-effective alternative to satellite and airborne platforms. Communication and navigation system are considered as main functioning blocks along with the sensors. There are mainly two types of UAVs are used in remote sensing applications, a) Fixed-wing, b) Multi-rotor. Aspects that need to be considered before selection of the UAV type for a particular application are flying time, pay load carrying capability, the total weight of the UAV including all mounted sensors, the maximum attainable altitude of the UAV, flying speed and control.

Generally, fixed-winged solutions deliver better in terms of flight time and are used primarily in agricultural monitoring. On the other hand, multirotor type like quadcopters, hexacopters and octocopters, are more manoeuvrable and can be used virtually anywhere and therefore are preferred for research purposes. Typically, GPS and IMUs are used for navigation and flight control and multi-rotor solutions support a number of flying operations. In some cases, GPS accuracy is compromised as the high-performance GPS device with lightweight are extremely expensive with respect to the total cost of the UAV. There are three main aspects 1) on-board power available on UAV, 2) regulations to fly UAV in the

area, 3) communication range between the UAV and operator, that determines the flight height and distance of a UAV for remote sensing. In most of the countries, commonly used rules (without a license) allow up to 150 m of flying height and line of site operations. With these limitations, flying campaign for the only small area can be carried out, and thus communication system range becomes less relevant. On the other hand, the most appropriate regulations require pilot license and the local aviation authorities to apply and approve the advance flight plan. In this case, there are no such hard limitations on flying height and range. Specific country wise regulatory information can be found at <http://www2.isprs.org/com-missions/comm1/icwg15b/resources.html>. At the moment, there are many UAV based solutions available in the market that are applicable in the agricultural remote sensing. Few examples are listed in Table.2.3 with attributes associated with the UAV based remote sensing.

2.1.4 Proximal Remote Sensing

There are limitations of satellite-based remote sensing for precision agriculture summarized in [147]. Images in the visible and near-infrared bands, acquired by the satellite platforms are often get affected by the clouds. However, SAR based imagery acquired by satellite or airborne is not affected by the cloud cover. There are also other challenges such as calibration of the raw data to true surface reflectance, correction of atmospheric interferences, and georeferencing the imagery. Given the above limitations, proximal remote sensing technique was introduced for crop growth monitoring since the past few decades that involves agricultural vehicles mounted with the sensors and handheld sensors. It can be used for real-time site specific management. Scheprs and francis [221] performed an analysis to measure chlorophyll contents in maize crop during the silking stage under the variable Nitrogen fertilizer rate using soil plant analysis development (SPAD) meter. Usage of SPAD meter and its role in precision agriculture are discussed in [177, 144, 219]. Other types of sensors, such as RGB optical sensors, multi-spectral sensors and thermal sensors, mounted on the ground vehicle or platform, have also been used in numerous studies. The main purpose of proximal remote sensing is a real-time assessment of crops to aid in decision-making practices.

2.2 Spectral Vegetation Indices

Remote sensing of vegetation is frequently performed by capturing the electromagnetic wave reflectance information from plant's canopies using passive sensors. It is a well-established phenomenon that the reflectance of light spectra from plants changes with plant type, water content within tissues, and other inherent factors.

The reflectance from vegetation or emission characteristics of vegetation is determined by chemical and morphological properties of the canopies or leaves [250]. The primary uses for remote sensing of vegetation are based on the following light spectra: (i) the ultraviolet region (UV), ranges from 10 to 380 nm; (ii) the visible spectra, which are comprised of the blue (450–495 nm), green (495–570 nm), and red (620–750 nm) wavelength ranges; and (iii) the near and mid infrared bands (850–1700 nm) [43]. Advancements in high resolution spectral instrumentation have brought an expansion in the number of bands captured by the remote sensing platforms, which are employed to determine various indices to characterise vegetation status. One of the most used indexes derived from multi-spectral information is the Normalized Difference Vegetation Index (NDVI) which is formed by the normalized ratio between the red and near infrared bands [102]. Direct use of NDVI is to describe canopy growth or vigour; hence, numerous works have made a comparison of it with the Leaf Area Index (LAI) [206], where LAI can be expressed as the measured area of one-sided leaves per area of soil [61]. Vegetation information from remote sensing images is described by differences and changes of the green leaves from plants and canopy spectral characteristics. The most popular validation process is through direct or indirect correlations between VIs obtained and the vegetation characteristics of interest by infield measurements, such as vegetation cover, canopy size, LAI, biomass, growth, and vigour assessment. Besides, more established approaches are used to assess VIs using direct and georeferenced methods by monitoring plants to be matched with VIs obtained from the same plants for calibration purposes. Jordan [99] introduced in 1969 one of the first VIs named Ratio Vegetation Index (RVI), which is based on the principle that plant's canopies or leaves absorb relatively more red than infrared wavelength; RVI can be represented mathematically as

$$RVI = \frac{R}{NIR} \quad (2.1)$$

Where NIR is the near infrared band and R is the red band. Because of the spectral characteristics of vegetation, bushy plants have low reflectances for the red band and shows a high correlation with LAI, and chlorophyll content of leaves [170]. The RVI has been used extensively for vegetation monitoring and green biomass estimations, particularly, at high-density vegetation coverage, since this index is susceptible to vegetation and shows good correlation with plant biomass. Richardson et al. [179] proposed the Difference Vegetation Index (DVI) that can be expressed as

$$DVI = NIR - R \quad (2.2)$$

The DVI can be applied to monitoring the vegetation ecological environment due to its sensitivity for change in the background soil. As previously mentioned, the Normalized Difference Vegetation Index (NDVI) is the most extensively used for

agricultural monitoring applications; proposed by Rouse Jr. et al. [186], that can be expressed as

$$NDVI = \frac{(NIR - R)}{(NIR + R)} \quad (2.3)$$

NDVI values ranges between 0 and 1, because the index is calculated by a normalization procedure, it is sensitive to green vegetation even for low vegetation covered areas. NDVI is frequently used in research related to regional and global vegetation monitoring and has found to be associated with not only to the plant's biophysical parameters such as canopy structure and LAI but also to canopy photosynthesis [60, 69]. Nevertheless, it is sensitive to the reflectances from background soil, soil colour, atmosphere, cloud and cloud shadow, and leaf canopy shadow, which further requires remote sensing calibration. To overcome some deficiencies of NDVI and RVI in describing the spectral behaviour of vegetation and soil background, Huete [88] proposed the Soil-Adjusted Vegetation Index (SAVI), which can be expressed as follows:

$$SAVI = \frac{(NIR - R)(1 + L)}{(NIR + R + L)} \quad (2.4)$$

The model as mentioned above of a soil vegetation index was proposed to improve the sensitivity of NDVI to soil backgrounds, where L is the soil conditioning index, which improves the sensitivity of NDVI to soil background. With the rising interest in remote sensing applications in agriculture, many vegetation indices have been introduced which can be found in [241].

Since the past few decades, several techniques have been developed in the field of remote sensing image analysis. Besides well known statistical approaches, many recent methods, based on techniques taken from the field of machine learning, are successfully developed and deployed. Remote sensing based vegetation indices are used as features in many most of the classification and regression tasks. A significant aim of machine learning algorithms in remote sensing is supervised classification, which is perhaps the most extensively used technique for classification of remotely sensed imagery. In the following chapter, a brief introduction to machine learning and the different paradigms in remote sensing is given.

Chapter 3

Machine Learning for Remote Sensing Data Analysis

With the recent development of several Earth observation platforms with increased spatial and spectral resolution along with higher revisit times, remote sensing gives more accurate information on land cover and the environmental status than ever before. Furthermore, different Earth-observation systems, equipped multi spectral and SAR systems, operate in different wavelengths, ranging from visible to microwave. The data acquired from the platforms as mentioned earlier, consequently provide different but complementary information. The classification of such huge data might be considered complex; however, on the other, concerning recent and upcoming satellite missions, remote sensing applications become even more attractive.

In some cases, comprehensive knowledge of the system is required to model the systems, which are often known as physics-based modelling methods. These modelling methods do not use data directly; however, data can be used to calibrate the parameters of the model. These methods are instrumental in modelling large areas provided that the model is accurately calibrated and parameterized [84].

For over a decade, biophysical and geophysical models using remote sensing data were developed and employed in many applications. For instance, the addition of remote sensing data sets, acquired from NOAA and NASA satellite platforms, into PHYSGROW plant growth model presented an opportunity to generate forage production maps even for large areas. Input parameters of PHYSGROW such as daily temperature, precipitation rate, vegetation indices in particular normalized vegetation index (NDVI) derived from remotely sensed data were girded to analyze the spatial autocorrelation [6]. Furthermore, maps produced from remote sensing data and physics-based models are used in drought detection and can be augmented with the geographic information system (GIS), which further could be linked with the agro-economic models, natural resources monitoring and management. In [68], the land surface temperature derived from satellite observations is used along with

the surface energy balance model to evaluate groundwater consumption by vegetative land cover.

Over the years, remote sensing data analysis has been in practice to measure changes in natural and manage landcover such as water bodies, crops and forests. Special attention has been paid to fully exploit the potential of remote sensing data in order to use in the production of environmental and agricultural statistics. Statistical machine learning methods in remote sensing data analysis can be defined as a tool in such scenarios where a statistical relationship between spectral information coming from remote sensing platforms and ground/field measurements is established.

Many application areas have been addressed with the use of remote sensing images during the past few decades. Apart from well-known statistical approaches, numerous recent methods are established based on machine learning techniques. Significant use of machine learning algorithms in remote sensing is supervised classification, which is conceivably the most extensively used image classification method. A brief description of machine learning and the different paradigms in remote sensing is given in the following sections.

3.1 Machine Learning

Machine Learning is a sub-field of artificial intelligence and usually refers to the evolution of methods that optimize their performance in an iterative way by learning from the data. These methods can be predictive (e.g., a regression model) and describe a particular phenomenon (e.g., a classification model) and identify, for example, between different classes of patterns. In remote sensing domain, machine learning methods often focus on land cover classifications and consequently contribute to valuable input information for various environmental monitoring systems, for example, in the flood forecast system and urban degradation. As an example, to differentiate various land cover classes, the machine learning algorithm learns to distinguish different types of the class pattern (i.e. land cover classes). Let us consider the detection of a particular land cover type, e.g., artificials such as buildings, to be a generic machine learning problem. The corresponding machine learning formulation will be:

$$f : \mathbf{x} \rightarrow y = f(\mathbf{x}) \tag{3.1}$$

where \mathbf{x} is a feature vector derived from the image and y represents a scalar which shows the occurrence ($y = 1$) or non occurrence ($y = 0$) of a building \mathbf{x} in that image. Note that the values chosen for y are arbitrary. Generally, machine learning aims at predicting the function f from some prior data. To this end, numerous methods have been introduced to achieve this goal. These methods can be further categorized in several ways, e.g., mainly depending on the type of the training data,

the form of function $f(x)$ and inherent assumptions on the underlying density function (Figure 3.1). There are two main categories, such as unsupervised and supervised, based on whether these approaches incorporate a priori knowledge while the decision making process by using labelled training data or not. In contrast, the semi-supervised methods belong to a particular type that relies on information coming from both labelled and unlabelled data or samples.

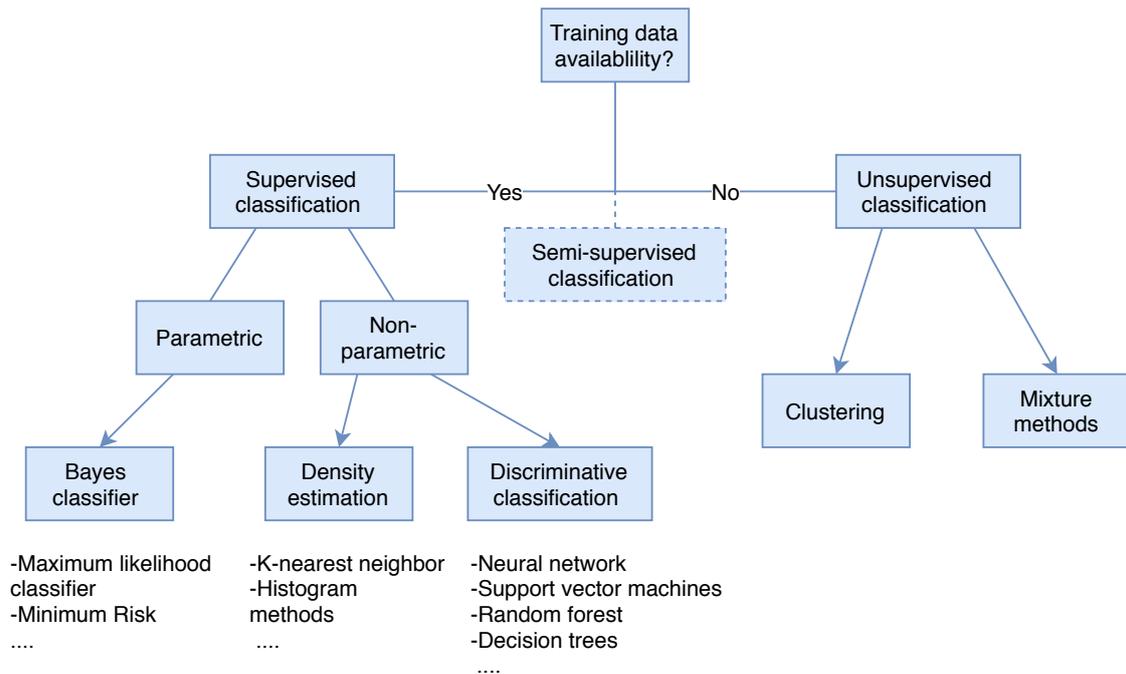


Figure 3.1: General overview of classifier categories [23].

- **Supervised algorithms:** These algorithms can be used in a situation when a set of labelled data or samples (\mathbf{x}_i, y_i) where $i = 1$ to l , and training data is available. The goal is to predict the value y corresponding to a new sample \mathbf{x} , i.e. determining the class membership of \mathbf{x} .
- **Unsupervised algorithms:** These correspond to the situation where only the data, i.e., (\mathbf{x}_j) $j= 1$ to l , is known and the main goal is to describe how the data is spread in the form of several clusters.
- **Semi-supervised algorithms:** In this case, the two previous approaches are combined. The learning is based on both the available set of labelled samples (\mathbf{x}_i, y_i) and also on some additional data (\mathbf{x}_j) where $j=l+1$ to $l+u$ for which no prior knowledge y_j is available.

3.2 Neural Networks

A neural network (NN) sometimes referred as an artificial neural network (ANN) can be defined as “a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs” [24]. NNs are generally arranged in layers. Layers are created from several interconnected ‘nodes’ which include an ‘activation function’. Input sequences are fed to the network through the ‘input layer’, which communicates to single or multiple ‘hidden layers’ where the actual processing is performed via a scheme of weighted ‘connections’. The hidden layers are then connected to an ‘output layer’ where the network’s output can be taken as shown in the Figure. 3.2. Even though the NN has good universality, a single network framework does not

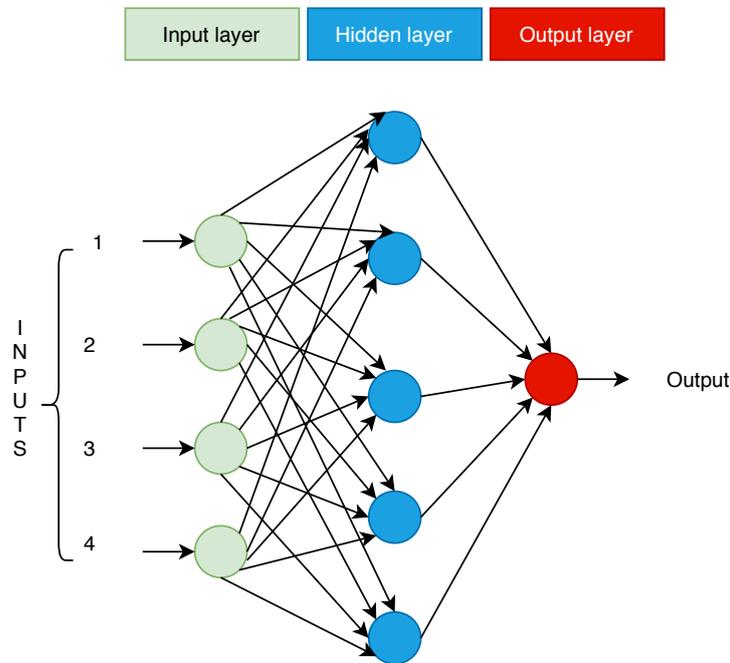


Figure 3.2: Basic units on neural networks.

have the potential to address all problems. Many different types of NN frameworks have been developed to address various types of research problems, whereby proving the significance of network frameworks. The back-propagation NN (BPNN) and generalized regression NN (GRNN) are two common examples of the classical neural network framework. Generally, the four established NN architectures are the autoencoder (AE), convolutional neural network (CNN), deep belief network (DBN), and recurrent NN (RNN).

One of the basic NNs architecture is back-propagation NN. In BPNN (Figure. 3.3), at the minimum one hidden layer is present which connects one input and one

output layer. Layers are formed with several neurons (also known as nodes). The BPNN mainly involves forward and backward propagation: the nodes in the input layer are carried through each hidden layer consecutively to reach the final output layer. In case the required results are not achieved in the output layer; then the errors are back-propagated to update the weights of the nodes of each hidden layer iteratively to minimize the error.

The NNs applications in remotely sensed images differ from those in natural

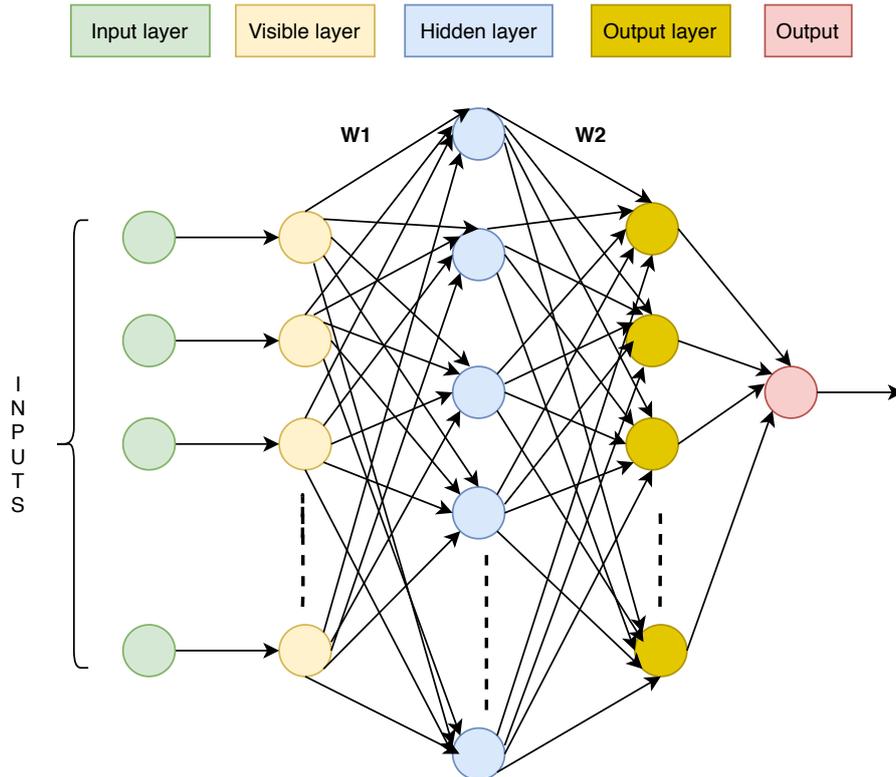


Figure 3.3: Back propagation neural network (BPNN) structure.

images, the remotely sensed images usually have more complicated and diverse patterns, as well as richer spatio-temporal and spectral information that can be used, thus higher requirements are required on the processing methods of remotely sensed images. Thanks to the introduction of deep learning (DL) with strong ability of in feature representation, DL has been introduced into environmental remote sensing and applied in many aspects, including land cover mapping, environmental parameter retrieval, data fusion and downscaling, and information construction and prediction. The following sections will further discuss CNN and RNN architectures.

- **Convolutional Neural Networks (CNNs):** Convolutional Neural Networks (CNNs), dating back decades, is one of the classical forms of modern

DL [115]. CNNs were inspired from classical concepts of computer vision theory [212], [46] and a study of the brain's visual cortex [87]. CNN involves a series of basic units stacked between the input and output layers Figure. 3.4. Each basic unit may incorporate the subsequent operational layers: convolutional, pooling, and activation layers. Firstly, in the convolution layer, input data is fed through numerous local filters which perform convolution operation [246]. The output data with the same dimension as the input data are usually obtained. The pooling layer extracts the low-dimensional data from the input data during multiple operations, such as max-pooling and average-pooling operations. The nonlinear operations are adapted in the activation layer to improve the nonlinear fitting ability of CNN. The dimension of the output and label data should be as close as possible to each other throughout the set of shrink operations.

Although these have been applied practically and successfully in image classi-

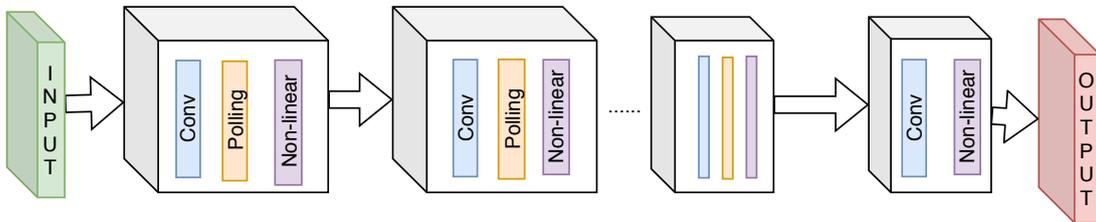


Figure 3.4: convolutional neural network (CNN) structure.

fication [115] ever since the 1990s, CNNs were not scaled to large applications because of technical constraints such as large volumes of data, theoretical limitations, and lack of hardware performance. Despite that, at the ImageNet ILSVRC competition [110] Geoffrey Hinton and his team showed the possibility to train large architectures which can learn several layers of features, each with increasingly more abstract internal representations. This was a breakthrough moment in DL history and following this, CNNs became the fundamental iconic symbol for the Deep Learning revolution [116].

Nowadays, CNNs have been successfully used in wide range of fields, such as object detection [50, 64, 123], semantic segmentation [185, 27], signals processing [171], face recognition [158] and feature extraction [234, 184]. So, over the last decade, with the significant advancements of modern powerful general purpose units (GP-GPUs), slightly improvements in the techniques (like ReLu as activation functions) and the easy access to large quantity of data [48, 124] encouraged researchers of all around the world in investigating and discover increasingly more efficient and powerful architecture [81, 211].

In recent years, DL was widely used in data mining and remote sensing applications. In particular, image classification studies exploited several DL architectures due to their flexibility in feature representation, and automation

capability for end-to-end learning. In DL models, features can be automatically extracted for classification tasks without feature crafting algorithms by integrating autoencoders [228, 148]. 2D CNNs have been broadly used in remote sensing studies to extract spatial features from high resolution images for object detection and image segmentation [101, 135, 10]. In crop classification, 2D convolution in spatial domain performed better than 1D convolution in spectral domain [111]. These studies formed multiple convolutional layers to extract spatial and spectral features from remotely sensed imagery.

- **Recurrent Neural Networks (RNNs):** A prerequisite of the model definition for an ordinary NN is the independence of the input/input, output/output, and input/output samples from each other. This condition makes it impossible to learn the relationship between samples. However, many samples are correlated in real datasets, such as sequential data, and the traditional network model cannot effectively cope with this correlation. RNN, LSTM and

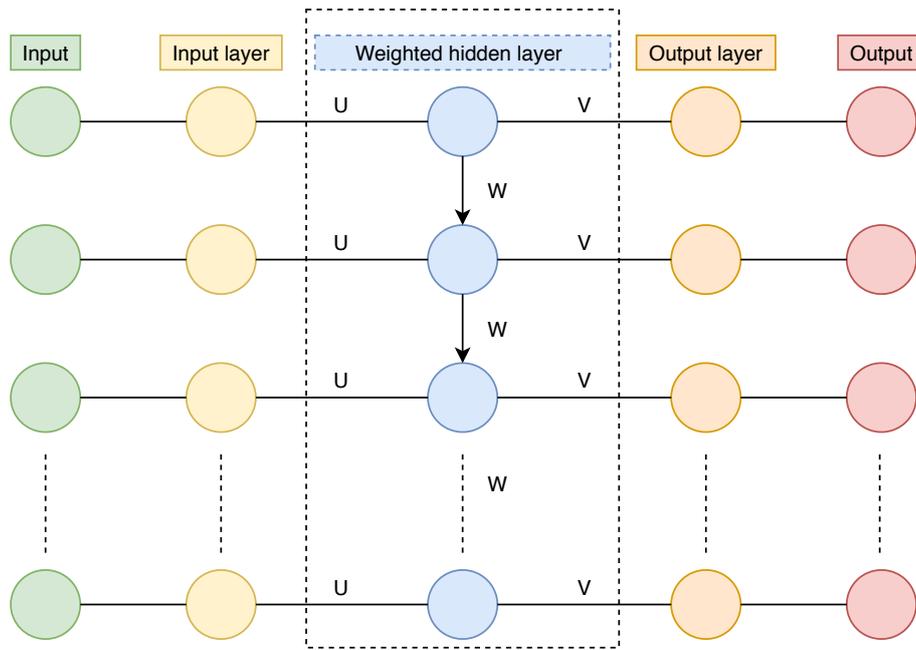


Figure 3.5: Recurrent neural network (RNN) structure.

GRU are models that can deal with sequences, utilize the correlation between sequence data, and even generate sequences. As depicted in Figure. 3.5, RNN has three main parts with several hidden layers. Each unit of the input sequence is successively inputted into RNN to obtain the corresponding output sequence unit of this phase and the information to transmit to the next phase. Such a task is performed to utilize the correlation within sequences. As an improved version of RNN, LSTM contains forget, input, and output

gates, which are used to control the filtering of the previous status. This structure aims to obtain previous statuses that are influential to the present instead of the most recent ones.

For many domains, ranging from natural language processing, handwriting recognition, robot automation, and image captioning, sequence data analysis plays an important role. In recent years, Recurrent Neural Networks has been the primary tool for sequence learning [209]. RNNs allow to compile and represent information from context windows of hundreds of elements. Moreover, over the course of recent years, the research community has come up to different techniques over many time steps to overcome training difficulty. This can be seen by contrasting RNN with, for example, an architecture based on Long Short-Term Memory (LSTM) [63] and Gated Recurrent Unit (GRU) [33] the latter having achieved ground-breaking achievements [32, 11]. RNNs are usually used when Sequential data analysis is needed. An example of such a case is in Lyu et al [134] where RNN was exploited to make use of spectral correlation, intra bands variability of multispectral data, and other such sequential properties. In addition, LSTM model was also employed to extract combined spectral-temporal feature representation from a pair of images ,acquired at different times and dates – to detect changes in remotely sensed images [133].

3.3 Machine Learning paradigms for Remote Sensing

Machine learning in remote sensing involves several paradigms such as classification, regression, clustering, feature extraction, and dimensionality reduction shown in Figure. 3.6. These aspects are often interdependent, e.g., before performing a classification one might extract some additional texture features and also reduce the dimensionality of the data set with feature selection techniques. Perhaps the most commonly undertaken applications in remote sensing are feature reduction, clustering and classification.

3.3.1 Classification

Classification is a process in which things or objects are recognized, differentiated and divided into discrete set of groups are known as classes. Assignment of classes is based on set of input variables also referred as attributes, factors, explanatory variables. The relationship between the data and the classes into which they are classified need to be well understood in order to split a set of data into different groups or categories. In order to achieve this, computers are trained with the known data set. Training is considered to be the most critical phase during

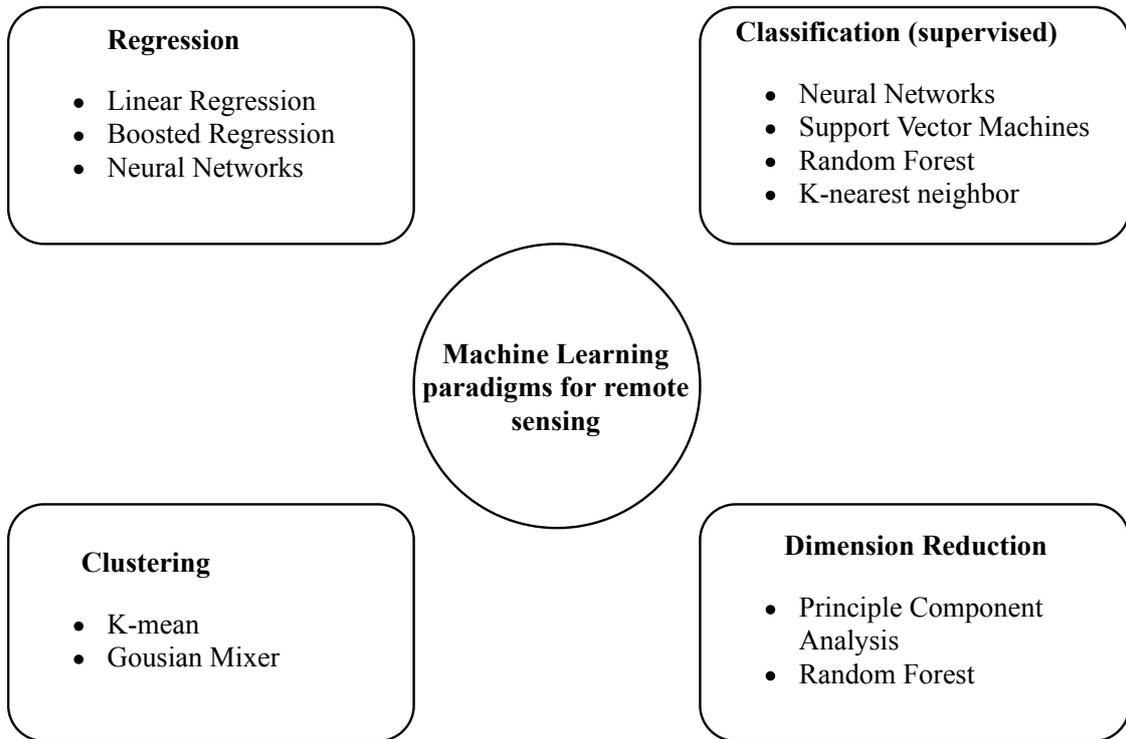


Figure 3.6: Most common machine learning methods for remote sensing data analysis.

the classification process. Automated classification of remotely sensed images is carried out by computer programs by learning the relationship between the data and information associated with the classes of interest. Numerous learning techniques have been used in the literature to improve the performance of programs used for classification. Performance of the computer programs known as classifier is generally assessed by the accuracy they achieve. Classification methods used in remote sensing data analysis can be categorized as follows.

The major part of machine learning algorithms for remote sensing image analysis is aiming perhaps for a supervised classification of the data. The machine learning problem is to learn the function f , from a set of labelled examples, i.e. our training samples whose class membership is known:

$$f : \mathbb{R}^n \rightarrow \mathbb{N}, \mathbf{x} \rightarrow \hat{c} \quad (3.2)$$

which assigns each sample to a particular class \hat{c} . In general such learning problems are ill-posed and it is necessary to restrict the space of possible functions f . In the following paragraphs, widely used *supervised classification algorithms* and their applications in RS domain are discussed.

- **Neural Networks (NNs):** To able to deal with large data sets, recent

developments in machine learning and particularly on deep learning applications have emerged in resulting huge improvements on taking account of remote sensing based large-scale data. To this end, CNNs are devised to process data that come in the form of multiple sequences, for instance, images acquired from multi-spectral sensors comprised of multiple 2D arrays comprising pixel intensities in the multi spectral channels. To start with the remarkable success achieved by AlexNet [5], numerous representation of CNN models such as VGGNet [229], GoogLeNet [12], SPPNet [125], and ResNet [230] have been introduced in the literature.

Besides, most of the CNN models used feed-forward approach in remote sensing and computer vision applications as demonstrated by Ishii et al. and alike tasks for satellite image classification approaches by using state-of-the-art approaches [90, 162, 85, 62]. Nevertheless, numerous studies [151, 146] showed a significant impact of low-quality or noisy data on the accuracy of CNN models. Consequently, inspired by biological perception systems, some recent studies introduced the use of recurrent and feedback features in CNNs, that bring significant improvement to the existing feed-forward CNNs [126, 154, 248]. The combination of recurrence and/or feedback to CNNs often tested with deep networks on natural scenes with comparatively good resolution color images.

Most of the applications of CNNs in remote sensing domain is Multi / hyper-spectral image classification, in which CNNs were applied to obtain spatial-spectral features, for either 1D convolution for the spectral domain [86], 2D across the spatial dimensions [247, 253], or 3D across the spectral and the spatial dimensions simultaneously [119]. Kussul et al. [111] discovered that results in terms of classification accuracy could be improved with 2D convolution in the spatial domain in contrast to 1D convolution in the spectral domain. Further, 1D convolution was applied by Guidici and Clark [71] to multi temporal hyperspectral images in the spectral domain for land cover classification. In these studies, convolutional layers in CNNs used as feature extractors in the spatial or the spectral dimension, though rarely for the temporal domain of remotely sensed image time series.

Recurrent neural networks (RNNs) are an advanced category of ANNs in which the extended form of the conventional networks with loops in connections are adopted [41, 249]. RNNs are devised explicitly for sequential data interpretation or analysis and have recently proved to be a successful form of NNs in numerous remote sensing applications. Lyu et al. [133] used a RNN to take full advantage of the sequential features of multispectral data, like spectral correlations and band-to-band variability across the spectral dimension. In [149], authors used an encoder to generate multi-level convolutional feature maps from shallow to deep; conversely, an RNN based decoder was used to recursively obtain multi-scale features and accumulated those features

sequentially to generate a high-resolution semantic segmentation image. The key feature of RNNs is the capability of representing data in continuous dimensions with sequential dependency, and the frequent use of RNNs is in the remote sensing to extract features from multi-temporal observations.

- **Support Vector Machines (SVMs):** Support vector machines is a supervised machine learning method and can be defined by a separating hyperplane in an N-dimensional space which specifically classifies the data points. Hyperplanes are referred as decision boundaries that is useful in classifying the data points. Data points that fall on either side of the plane is recognized to various classes. It is suitable for both regression and classification. Nonetheless, it is widely used in classification tasks, such as Szuster et. al [213] performed land use and land cover classification in tropical coastal zones. A hierarchical approach of SVMs has been compared with other classifiers to map nine summer crops from ASTER satellite imagery and claimed overall accuracy of 88% [161]. Kernel selection for SVMs has been an important issue which is addressed in detail in [235, 80].
- **K nearest neighbor (K-nn):** K-nn is well known non-parametric machine learning technique used in classification and regression tasks. In the simplest form of k-nn, image classification is performed by taking into account of the majority vote of its neighbor pixels. The neighbors are usually assigned weights so that the nearest neighbor contribute more with respect to the one that are more distant. For instance, common scheme in which a weight $1/d$ is assigned to each neighbor, where d is the distance to the neighbor. The neighbors are chosen in such as way from the objects for which the class or value is known. Although this method is easy to implement and fast, but these are not always considered to be the best for remote sensing data. In [140], k-nn method was adopted for forest inventory mapping and estimation based on Landsat satellite imagery. Synergy of k-nn and SVM methods have been exploited to perform classification on satellite images, it was found that the accuracy and predictive ability is considerably improved [222].
- **Random Forest (RF):** It is an ensemble machine learning method that is widely used for classification and regression. Random forest method is a type of classification tree methods under the well-known decision trees which are discussed earlier in this section. It consists of set of trees constructed from random sample of the training data. Keichi et. al [217] adopted RF method for crop classification of upland fields using time series data acquired from Landsat 7 ETM data. They claimed 81% overall classification accuracy with the Kappa 0.70. It is indeed an important step for any classifier to select the input features to be used to perform classification task, in this regard, Pengyu et. al [76] proposed a method of feature selection of MODIS temporal

data using random forest for early crop classification in Kanas, USA. In some studies [189, 174], comparison has been made between the RF classifier and other traditional classifier for classification of remotely sensed imagery. A detailed discussion on the applications of RF in the field of remote sensing can be found in [16].

3.3.2 Clustering

The concept of clustering involves the grouping of objects with similar characteristics into set of clusters. Objects within the same cluster are similar to one another, while different to the objects in other clusters. An example of clustering is the identification of the zones with similar land use in an earth observation database. Wide range of approaches for clustering are being widely used in many different fields. Major approaches include partitioning algorithms, density based, grid based, model based and hierarchy algorithms. For example, pixels of an image are allocated into groups based on the similarity measures extracted by the clustering algorithm. Performance of the clustering method is generally assessed by the intra-class and inter-class similarity. High intra-class similarity will produce better clustering results while inter-class similarity should be low to achieve good results.

- **K-means:** It is most commonly used method for clustering. It is based on iterative process that break down the data set into K predefined non-overlapping subgroups also known as sub clusters where each data point is the member of single group. It keeps trying to make inter cluster data points as homogeneous as it could be, at the same time it also tries to maintain considerable distance among the clusters so that these can be distinguished from each other. One of the examples of k-mean in the literature is to classify high resolution satellite data into several classes, that is further used to map the different land cover [220].
- **Gaussian Mixture Models:** A Gaussian mixture model belongs to a category of probabilistic model that implies all the data points are generated from a mixture with unknown parameters of a finite number of Gaussian distributions. For example, Gaussian mixture model was used to map urban areas using very high resolution remotely sensed imagery [216].

3.3.3 Regression

In statistical modelling, regression approach is used when the relationship is estimated between dependent variable (often called as outcome variable) and independent variable (often known as predictor). Similar to classification approach, the regression model is trained by means of set of input attributes or variables for

which the response is known. An example of regression from remote sensing domain is the estimation of crop yield based on variables (vegetation indices) derived from remote sensed imagery. There are many regression methods ranging from simple linear regression to currently famous methods inspired by neural networks as mentioned in Figure. 3.6.

- **Linear Regression:** In this type of regression, linear approach is followed to model the relationship between a dependent variable and explanatory (often known as predictor) variable. The relationship is modelled based on linear predictor function whose model parameters are estimated from the data. In case of dealing with only one explanatory variable, the process is called simple linear regression. Linear regression was the first form of regression analysis to be extensively investigated, and commonly used in practical uses [yan2009linear]. Anup et. al [167] produced a model to estimate crop yield by employing linear regression on the remote sensing data. In another example, linear regression was used to predict reflectance of synthetic spectral imagery using archived satellite images [189].
- **Boosted Regression Trees:** Boosted Regression Tree (BRT) models incorporate two techniques: decision tree and methods for boosting. BRTs repeatedly attempts to fit many decision trees to boost the accuracy of the model. This modeling method has the ability to fit complex nonlinear relationships and it does not require the pre data transformation [52]. Zhang et. al used BRTs model to assess the impacts of eleven metrological factors of hand, foot and mouth disease in china [251].
- **Neural Networks:** Neural networks are commonly famous for classification tasks, however, can be used in regression problems. Neural network is composed of number of layers; an input layer which represents the input or predictor variable, one or more number of hidden layers that are represented by an activation function for each hidden node acted on the weighted input coming from previous layer, and a single output layer in case of regression and may have multiple output layers in case of classification. Few examples of neural networks used for regression tasks includes Shao et. al [195] tested and compared neural network with traditional methods (SVM and classification trees) to classify land cover with the limited training data, xiao et. al [237] used a general regression neural network to estimate the leaf area index (LAI) from multi temporal remote sensing data. In past few years, neural networks were progressively being used for the estimation and prediction of LAI from remote sensed data [120], [243]. Recently, Reddy et. al [173] used LSTM to predict NDVI across different regions.

3.3.4 Dimension Reduction

Usually data acquired from remote sensing sources involves wide range of variables, dimensionality reduction is used to create small set of new variables that represents most of the information contained in the large (original) set of variables. These new variables are also referred as features and can be used as input of any other process for analysis. Dimensionality reduction is used prior to deploy machine learning model on the original data. A machine learning model which is trained on a large number of features or variables, becomes increasingly dependent on the data and model gets over fitted, leading to poor performance of the model. One of the major motivations to use dimensionality reduction is to avoid over-fitting by reducing the size of input features. There are several dimensionality reduction methods being used by the researchers. The most common and well-known methods are mentioned in Figure. 3.6.

In this chapter, several machine learning paradigms for remote sensing applications have been discussed. Introduction to NNs and their potential application in the remote sensing domain have also been highlighted. In the next chapters, these ML methods are used to address different research problems considered in this dissertation.

Chapter 4

Research contributions related to Precision Viticulture using UAV and Satellite Imagery

Since last few decades, precision agriculture has received significant attention in the agricultural community [157, 37]. Precision viticulture (PV) can be defined as sub-domain and emerging field, that address the problems during the grape production cycle and propose appropriate techniques to improve quality and productivity of vineyards while reducing cost and potential damage to the environment due to fertilizers, pesticides, machinery, and fuel [8, 199]. Inter- and intra-vineyard spatial variability is considered as a critical factor for vine-growers to describe and predict the outcomes in terms of yield and quality [20, 205, 169]. Nowadays, the evolution of new technological instruments provides us with possibilities to consider and adopt new methods to solve viticulture related problems [74, 114, 72, 168]. To this end, remote sensing (RS) has already demonstrated its potential and effectiveness in spatio-temporal vegetation monitoring by describing biophysical characteristics of plants such as vigour that can be used to assess the quality and yield of crop field [94, 136].

In this chapter, research contributions related to precision viticulture are presented in the following sections which includes (i) variability assessment of vineyard using UAVs and satellite imagery (4.1), (ii) automatic coverage path plan for agricultural unmanned ground vehicles (UGV) by using the UAV imagery (4.2) and (iii) CNN based approach to refine the moderate resolution satellite driven vegetation maps by using high resolution UAV imagery (4.3).

4.1 Variability Assessment of Vineyard using UAVs and Satellite Imagery

In agriculture, remotely sensed data plays a crucial role in providing valuable information regarding the status of crops, soil and other relevant objects involved in effective crop management. In particular, various spectral indexes have been proved to be useful tools in describing crops variability, both in term of spatial and temporal dimensions. In this work, detailed analysis and comparison of multispectral imagery of vineyards, acquired from decametric-resolution satellite and low altitude Unmanned Aerial Vehicle (UAV) platforms, is presented. The effectiveness of Sentinel-2 imagery and high-resolution UAV aerial images was evaluated considering the well-known relation between Normalized Difference Vegetation Index (NDVI) and vineyard vigour. To be compared with satellite imagery (10 m of resolution), high-resolution data from UAV were pre-processed. In addition, three different NDVI indexes were defined considering the (i) whole cropland surface, the (ii) sole vines canopy and (iii) sole inter-row terrain, respectively, to properly analyses the unbundled spectral contribution of the different elements in the vineyard environment. Results show that raw satellite imagery, with moderate spatial resolution, cannot be directly used to describe the variability of vines rows in vineyards reliably. Indeed, the contribution of inter-row surfaces to the remotely sensed dataset may mainly/deeply affect the NDVI computation, leading to biased crop descriptors. On the contrary, vigour maps computed from UAV imagery, considering only pixels representing the crop canopies, results to be more related to the in-field assessment concerning the considered satellite ones. The proposed approach can be extended to other crops typologies which are grown by rows or without intensive layout, where the crop canopies do not extend on the whole surface or where the presence of weeds is relevant.

Part of the work described in this section. 4.1 has been previously published in "Comparison of Satellite and UAV-Based Multispectral Imagery for Vineyard Variability Assessment" [107].

4.1.1 Background and Related Work

Over the last couple of decades, PA has earned significant attention in the agricultural community [157, 37]. In viticulture, addressing difficulties during the production cycles by defining appropriate crop management, the PA approach has the final aim to improve vineyard yield and grape quality while reducing waste, costs and environmental impact [8, 199]. A proper knowledge of the spatial variability between and within crop parcels is considered as a key factor for vine growers to estimate the outcomes in terms of yield and quality [20, 205, 169]. In this context, remote sensing (RS) has already proved its potential and effectiveness in spatiotemporal vegetation monitoring [74, 114, 72, 168]. Indeed, data provided by optical

sensors of multispectral and hyperspectral imagery systems are profitably exploited to compute a wide set of indices (such as the wide dynamic range vegetation index, the normalised difference red-edge index Index, etc.) by properly describing several crop biophysical characteristics [94, 136].

Furthermore, multispectral and hyperspectral imagery has been the source the red and near infrared bands used for vegetation monitoring and these bands are captured by the optical sensors [13, 252]. The combination of red and near infrared already showed the potential to approximate plant vegetation vigour, that describes biophysical parameters [145, 72]. Among the wide set of defined spectral indices, the normalized difference vegetation index (NDVI) is one of the most extensively used, since it is strictly related to crop vigour and, thus, to the estimated quality and quantity of field production [55, 49, 208, 96, 98].

Satellite multispectral imagery, due to sensors features and platforms altitude, covers extensive areas. In addition, many satellite programmes (such as Landsat, MODIS, ASTER, SPOT, Sentinel-1 and Sentinel-2) are nowadays providing free datasets, thus promoting satellite imagery exploitation for many agricultural applications [164, 82, 183, 244, 200], even with the multi-sensor data fusion approach [190]. Examples of valuable research contributions are the low-resolution Modis and high-resolution IKONOS satellite imagery exploitation for mapping vineyard leaf area [208, 97].

However, dealing at vineyard level is quite challenging for the RS applications because of physically discontinuous nature of vine canopies, medium size cover with respect to interrow soil, interrow reflected signal and shadow [49]. These factors must be considered and processed to separate vegetation from interrow soil. Therefore, processing steps are needed such as masking (to separate vine row and interrow), shadow minimization in order to calculate the reliable spectral index for vegetated area. After these processing steps, one can exploit spatial variability of index/vigour and interpret to generate prescription maps. The ultimate goal is to use these maps in precision viticulture practices for maximizing grape quality and yield while minimizing the environmental impact and net production cost [8, 199]. At the moment, UAVs and Airborne are considered to be the most effective remote sensing platforms for agricultural monitoring due to providing high spatial resolution imagery and flexible flight scheduling, that leads to having accurate and timely monitoring and decision making of crops. In the PV domain, with high resolution imagery, it is possible to differentiate between vegetated/canopies and non-vegetated/interrow pixels, that certainly improves the separation and interpretation of canopies and interrow soil. However, during the UAVs image acquisition campaign of a vineyard, due to the limitation on field of view of the optical sensors, many images are required to cover the whole area to generate an ortho-mosaicked image.

However, when considering crops with discontinuous layouts, such as vineyards and orchards, remote sensing becomes more challenging [18]. Indeed, the presence of

inter-row paths and weed vegetation within the cropland may deeply affect the overall spectral indices computation, leading to a biased crop status assessment. Indeed, novel approaches and algorithms using Unmanned Aerial Vehicle (UAV) or satellite based multispectral imaging have been developed for vegetation pixels classification [193, 160, 163]. Low altitude platforms, such as UAV and airborne sensors, by providing imagery with a high spatial resolution (even a few centimetres) and a flexible flight scheduling [91], allow differentiating between pure canopy pixel and other objects in the scene [128, 39, 130, 93, 73] or even to classify different details within canopies [178, 4, 51].

In this work, a detailed analysis and comparison of vineyards MSI, provided by a decametric resolution satellite and low altitude UAV platforms, is presented. The effectiveness of the MSI from Sentinel-2 and from the UAV airborne sensors, with very high resolution, was evaluated by considering the well-known relation between the NDVI and crop vigour. In particular, this work is structured as follows: Section.4.1.2 to Section.4.1.6 reports information on the considered study area, on data acquisition from the satellite and the UAV platforms, and on the performed data processing to allow a comparison of the NDVI computed from different imagery sources. The results obtained by the data processing and comparison are discussed in Section.4.1.7 and Section.4.1.8. Section.4.1.9 reports the conclusions and future developments.

4.1.2 Materials and Methods

A vineyard located in Serralunga d’Alba (Piedmont, northwest of Italy), covering a surface of about 2.5 ha, was selected as a field test. The cropland, whose latitude and longitude positions range between $[44.62334^{\circ}44.62539^{\circ}]$ and $[7.99855^{\circ}8.00250^{\circ}]$ (World Geodetic System 1984-WGS84), includes three vineyard parcels (named “Parcel A”, “Parcel B” and “Parcel C”) cultivated with the Nebbiolo grapevine, with an area of around 0.36, 0.69 and 0.19 ha, respectively shown in Figure.4.1. The vineyard is on a sloped land conformation, with an elevation ranging from 330 to 420 m above sea level and a predominantly southwest orientation. Due to the irregularity of the terrain morphology in terms of altitude and soil properties, the selected vineyard is characterised by a great variation in vine vigour within and between parcels. To extend the study to several vine phenological phases, the acquisition campaigns were performed from April to September 2017.

4.1.3 Satellite Time Series Images

The Sentinel-2 satellite is equipped with a multi-spectral imaging sensor that measures Earth’s Top of Atmosphere (TOA) reflected radiance in 13 spectral bands ranging from 443 nm to 2190 nm. The technical details of Sentinel-2 along with its

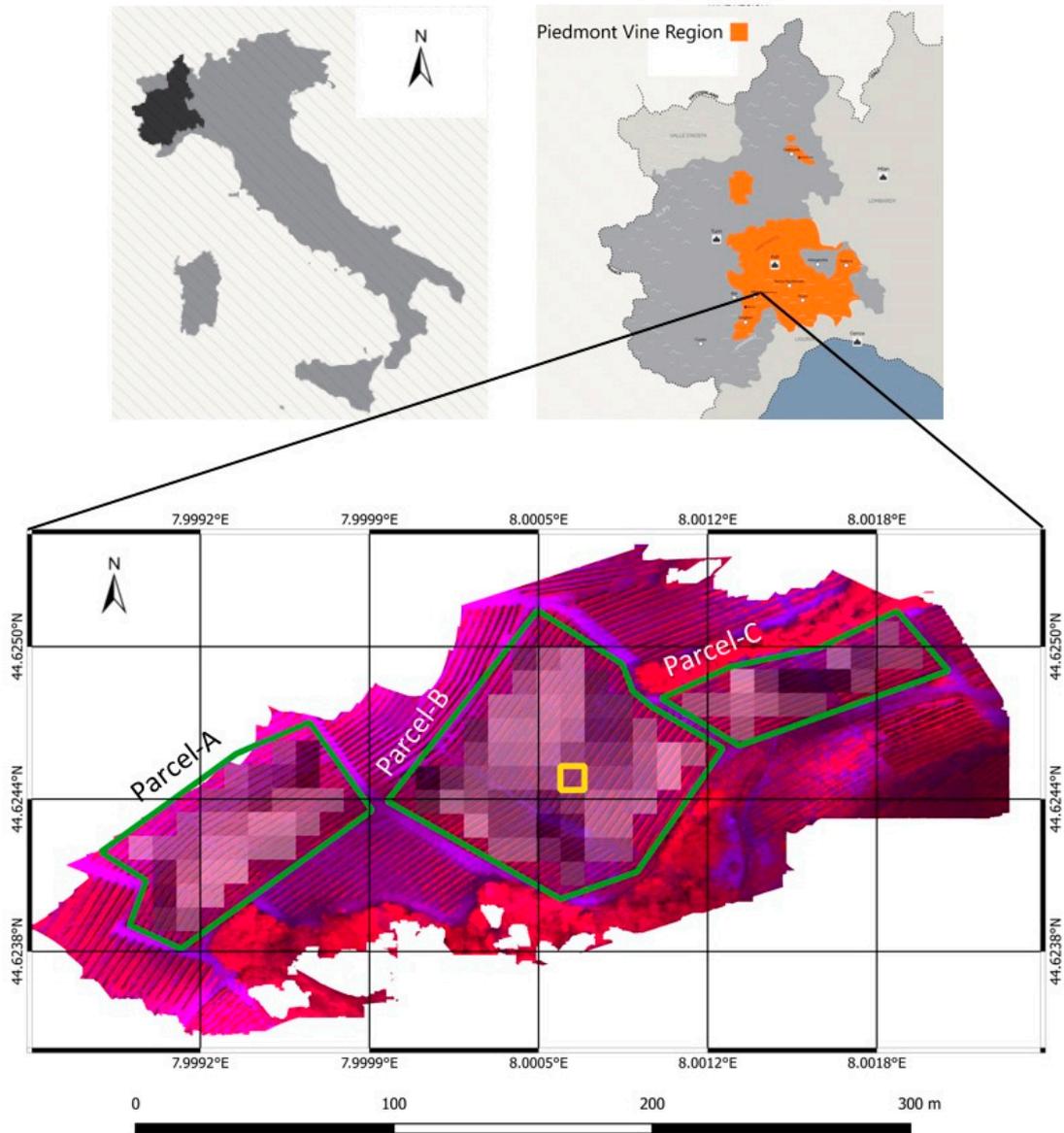


Figure 4.1: Selected test field located in Serralunga d'Alba (Piedmont, northwest of Italy). The boundaries of the three considered parcels, named “Parcel-A”, “Parcel-B” and “Parcel-C”, are marked with solid green polygons. The cropland region, represented by pixel $s_{8,20}$ of the Sentinel-2 tile, is highlighted by a yellow square. The map is represented in false colours (NIR, Red and Green channels).

spatial resolution and spectral ranges are summarised in Table.4.1. The Sentinel-2 imagery database, processed at different levels, can be downloaded from <https://scihub.copernicus.eu/dhus/#/home>. In this study, cloud-free level-2A Sentinel-2 Bottom of Atmosphere (BOA) reflectance images were used. The Level-2A imagery was derived from Level 1 by applying scene classification, atmospheric and BDRF correction algorithms, using SNAP toolbox (6.0) and sen2core processor (2.5.5) provided by the ESA [181, 103]. Additional details about the Sentinel-2 MSI products can be found in [194].

Table 4.1: Information on satellite and UAV based acquired datasets.

Dataset Name	Acquisition Date	Data Source	Time Difference (days)
\mathcal{D}_1	5 May 2017	UAV	+5
\mathcal{S}_1	30 April 2017	Satellite	-5
\mathcal{D}_2	29 June 2017	UAV	-7
\mathcal{S}_2	6 July 2017	Satellite	+7
\mathcal{D}_3	1 August 2017	UAV	-4
\mathcal{S}_3	5 August 2017	Satellite	+4
\mathcal{D}_4	13 September 2017	UAV	-4
\mathcal{S}_4	17 September 2017	Satellite	+4

The selected satellite tiles were acquired on four dates during the 2017 growing season (see Table.4.2) to consider different vegetative vine status. Only red and near infrared bands (bands 4 and 8, respectively) were used in this study. The pixels completely included within the boundaries of the three considered “Parcel A”, “Parcel B” and “Parcel C” were selected, as shown in Figure.4.1. All relevant information regarding the satellite imagery processed in this study is organised and summarised in Table.4.1 and Table.4.2.

4.1.4 UAV-Based Imagery

The UAV-based MSI were generated with the Agisoft PhotoScan[®] software (Agisoft[®], 2018) <https://www.agisoft.com/> processing imagery blocks of more than 1000 aerial images acquired with an airborne Parrot Sequoia[®] multispectral camera (Parrot[®] SA, 2017 [2]). The UAV path was planned to maintain the flight height close to 35 m with respect to the terrain by properly defining waypoint sets for each mission block on the drone guidance platform on the base of the GIS cropland map. With this specification, the aerial images GSD resulted to be 5 cm can be seen in Figure.4.2.

A camera geometric calibration procedure was performed before the image alignment task; moreover, a radiometric calibration was applied to the image blocks by using the reference images of a Micasense calibrated reflectance panel [21] acquired before and after each UAV flight. A set of 12 ground control points, whose positions

were determined with a differential GNSS system (with an accuracy of 0.1 m), was placed on selected vine trellis poles within the vineyard to georeference the MSI in a geodetic coordinates frame.

The UAV flights were performed on four different dates over the 2017 crop season (15 May, 29 June, 1 August and 23 September), according to the satellite visiting dates mentioned in Table.4.2.

4.1.5 In-Field Vigour Assessment

The vigour of the vines within the three considered parcels in the study site was evaluated based on the results of a specific in-field survey performed by trained operators and on the past experience of the farmer. In the considered study site, the vigour variability is mainly related to the pedological soil conformation and to the water availability, since irrigation is not allowed by Piedmont regulation. The vigour classification was performed by defining three classes: low “L”, medium “M” and high “H”. A specific data processing was performed to make the in-field vigour assessment comparable to decametric resolution imagery. In particular, a 10 m × 10 m map was obtained by rastering and clustering a vector GIS map provided by expert agronomists made by a set of three vigour class layers, according to Sentinel-2 pixel location.

Table 4.2: Technical details of the considered and adopted platforms and sensors.

	Satellite	UAV
Platform	Sentinel-2	8-rotors custom UAV
Sensors	 <p>Multispectral Imager</p>	 <p>Parrot sequoia Multispectral camera</p>
Number of channels	13	4
Spectral band details	Band name	Range
	B4-Red	650–680 nm
	B8-NIR	785–900 nm
GSD per band	B4, B8 = 10 m	5 cm
Flight altitude	786 km	35 m
Field of view	290 km	70.6° HFOV
Image Ground Dimension	100 km × 100 km	64 m × 48 m
Number of images to cover vineyards test site	1	>1000

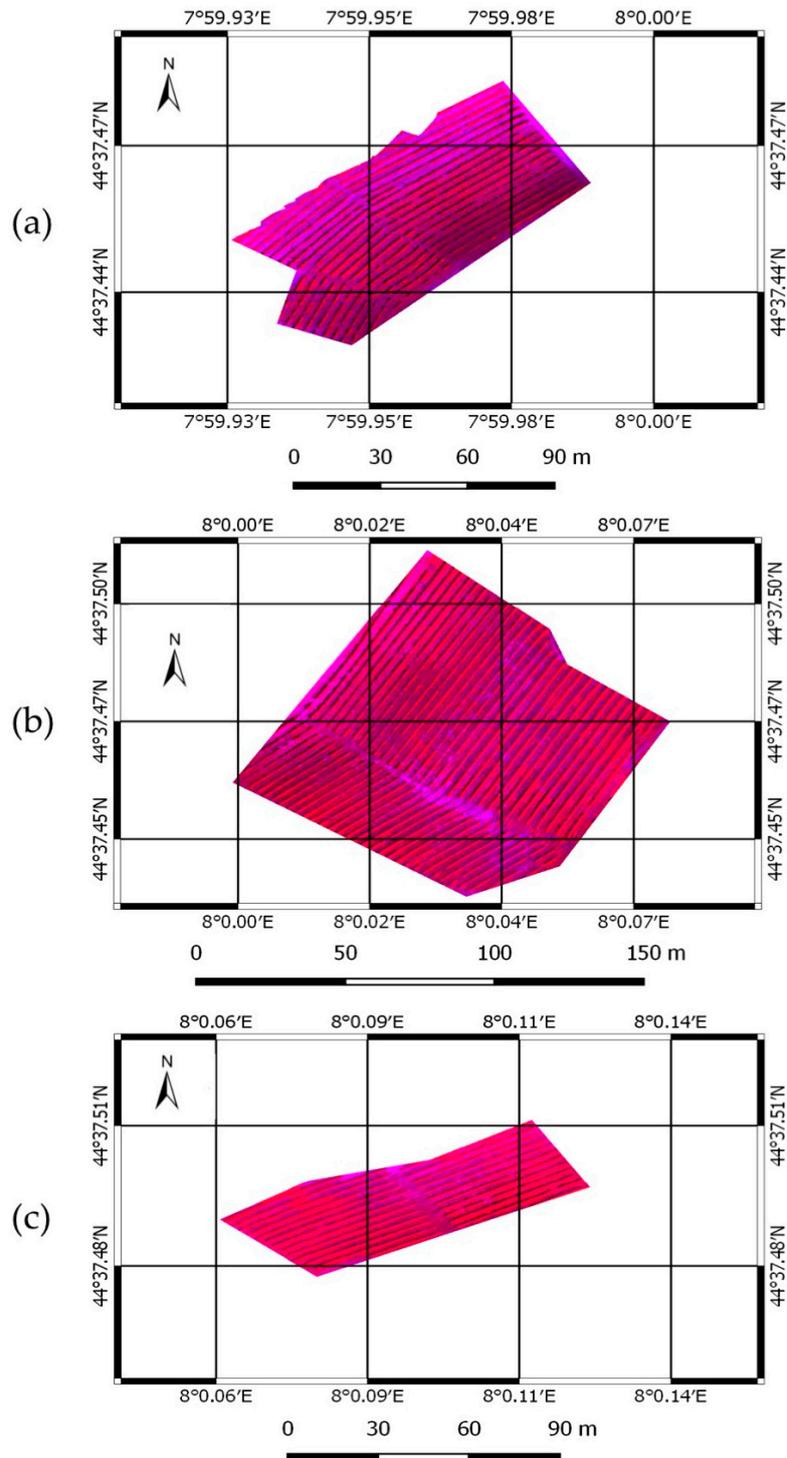


Figure 4.2: Enlargement of UAV-based multispectral imagery, represented in false colours (NIR, Red and Green channels), of: (a): “Parcel-A”; (b): ”Parcel-B”; and (c): “Parcel-C”.

4.1.6 Data Processing

In this section, specific methods for data processing, developed to compare and investigate the imagery derived from the two platforms with different spatial resolutions, are presented and discussed in detail.

A tile \mathcal{S} , derived from the satellite platform, can be considered as an ordered grid of pixels $s(i, j)$, with indices i and j representing the pixel row and column locations in the raster matrix, respectively. Each pixel $s(i, j)$ was here defined as $s(i, j) = [\alpha_s(i, j), \beta_s(i, j), n_R(i, j), n_N(i, j)]^T \subset \mathcal{S}$, where $\alpha_s(i, j)$ and $\beta_s(i, j)$ are the latitude and longitude coordinates (expressed in WGS84) of the upper left corner of pixel $s(i, j)$, respectively, and $n_R(i, j)$ and $n_N(i, j)$ are the pixel digital numbers in the red and near infrared bands (12 bit representation), respectively.

Data \mathcal{D} derived from the UAV flights were defined as an ordered grid of pixels $d(u, v) = [\alpha_d(u, v), \beta_d(u, v), m_R(u, v), m_N(u, v)]^T \subset \mathcal{D}$, where pixel $d(i, j)$ coordinates $\alpha_d(i, j)$ and $\beta_d(i, j)$ (latitude and longitude in WGS84) are related to the pixels centre and $m_R(i, j)$ and $m_N(i, j)$ are the pixel digital numbers in the red and near infrared bands (16 bit representation), respectively.

A graphical representation of the defined parameters for the satellite and UAV-based datasets is shown in Figure.4.3.

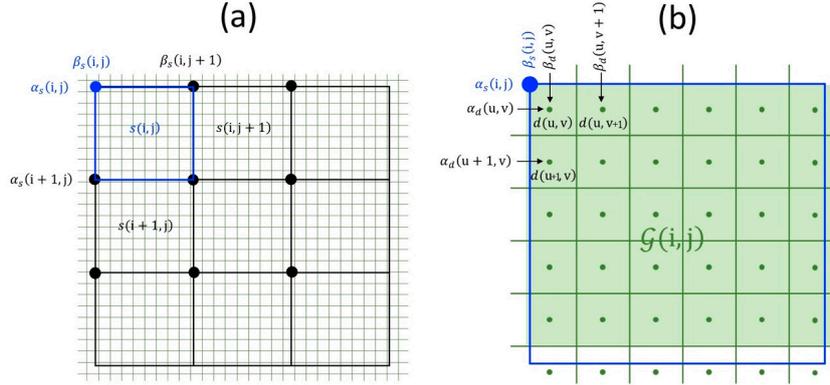


Figure 4.3: (a) Ordered grid of pixels $s(i, j)$ belonging to satellite tile \mathcal{S} , located at latitude and longitude coordinates $\alpha_s(i, j)$ and $\beta_s(i, j)$; and (b) ordered grid of pixels $d(u, v)$ belonging to satellite imagery \mathcal{D} , located at $\alpha_d(i, j)$ and $\beta_d(i, j)$. Selected UAV pixels belonging to $\mathcal{G}(i, j)$, used for comparison to satellite pixel $s(i, j)$, are highlighted in light green.

The evaluation of effectiveness in describing vineyard variability by the satellite and UAV multispectral imagery was focused on the plants vigour assessment by using the NDVI. The NDVI value for satellite pixel $s(i, j)$ can be easily computed as

$$NDVI_{\text{sat}}(i, j) = \frac{n_N(i, j) - n_R(i, j)}{n_N(i, j) + n_R(i, j)} \quad (4.1)$$

by using the spectral information provided by the digital numbers $n_N(i, j)$ and $n_R(i, j)$ of the red and near infrared bands. Figure.4.5a shows the $NDVI_{\text{sat}}$ map obtained by applying Equation (1) to the entire set of selected pixels representing “Parcel A”, “Parcel B” and “Parcel C” of the Sentinel-2 tile of 7 July.

To allow the comparison of the UAV-based MSI and of the Satellite imagery, a preliminary downsampling procedure of the high-resolution UAV imagery was performed. A portion of UAV dataset \mathcal{D} , made by pixel cluster $\mathcal{G}(i, j)$, which is related to satellite pixel $s(i, j)$, was defined as

$$\mathcal{G}(i, j) = \left\{ d(u, v) \in \mathcal{D} \mid \begin{array}{l} \alpha_s(i, j+1) \leq \alpha_d(u, v) < \alpha_s(i, j), \\ \beta_s(i, j) \leq \beta_d(u, v) < \beta_s(i+1, j) \end{array}, \forall u, v \right\} \quad (4.2)$$

With this approach, satellite pixel $s(i, j)$ and the portion of UAV map $\mathcal{G}(i, j)$ represent the same section of vineyard cropland, with latitude and longitude coordinates ranging between $[\alpha_s(i, j+1) \alpha_s(i, j)]$ and $[\beta_s(i, j) \beta_s(i+1, j)]$. As an example, an enlargement of UAV map subset $\mathcal{G}(8, 20)$, related to satellite pixel $s(8, 20)$ and highlighted in Figure.4.1 by a yellow square on the field test map, is displayed in Figure.4.4.

Three specific NDVI indices were defined to perform a detailed analysis of the radiometric information provided by the UAV-based MSI, and then to compare it with the satellite one. In detail, they were computed from the UAV high-resolution data by considering: (i) the whole cropland surface represented by $\mathcal{G}(i, j)$; (ii) only the crop canopy pixels and, for completeness; and (iii) only the pixels representing the inter-row terrain. Using all pixels in subset $\mathcal{G}(i, j)$, the comprehensive $NDVI_{\text{uav}}(i, j)$ for the UAV imagery was defined as

$$NDVI_{\text{uav}}(i, j) = \frac{\sum_u \sum_v \frac{m_N(u, v) - m_R(u, v)}{m_N(u, v) + m_R(u, v)} \forall d(u, v) \in \mathcal{G}(i, j)}{\text{card}\mathcal{G}(i, j)} \quad (4.3)$$

By applying Equation (2) to raw UAV map \mathcal{D} , an $NDVI_{\text{uav}}$ map congruent (properly aligned and with the same spatial resolution) to the ones derived from the satellite imagery ($NDVI_{\text{sat}}$) can be obtained, as shown in Figure.4.5b for the UAV imagery acquired on 29 June 2017.

Since within a vineyard UAV orthophoto, with a GSD of 5 cm, the pixels representing the vine canopies can be detected, a more accurate crop NDVI computation with respect to $NDVI_{\text{uav}}$ can be performed. For this task, a pixel classification procedure is thus required for each subset $\mathcal{G}(i, j)$ to define two different groups of pixels $\mathcal{G}_{\text{vin}}(i, j)$ and $\mathcal{G}_{\text{int}}(i, j)$, with $\mathcal{G}_{\text{vin}}(i, j) \cup \mathcal{G}_{\text{int}}(i, j) = \mathcal{G}(i, j)$ and $\mathcal{G}_{\text{vin}}(i, j) \cap \mathcal{G}_{\text{int}}(i, j) = \emptyset$, representing crop canopies and inter-row surfaces, respectively. The automatic classification procedure described in Comba et al. (2015) [38] was adopted. Figure.4.4b reports the obtained pixel classification belonging to subset $\mathcal{G}(8, 20)$ into the two groups $\mathcal{G}_{\text{vin}}(8, 20)$ and $\mathcal{G}_{\text{int}}(8, 20)$. By exploiting the spatial information concerning

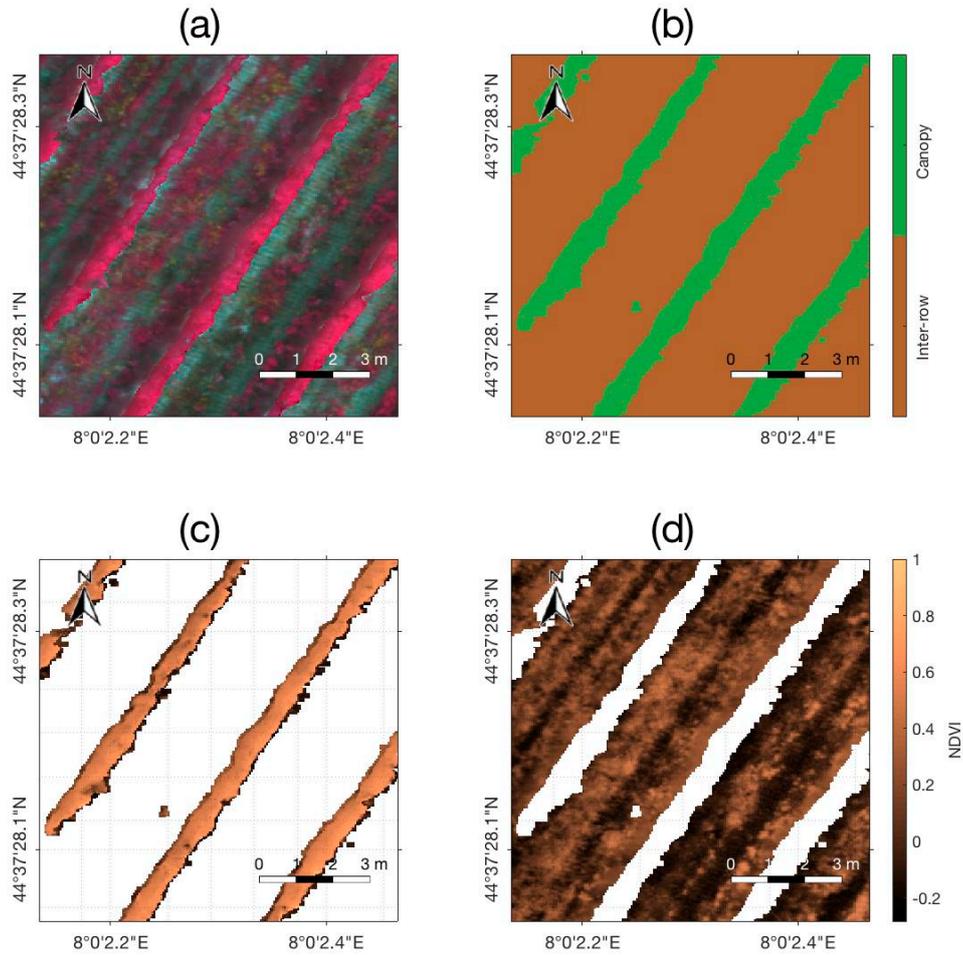


Figure 4.4: (a) Enlargement of subset $\mathcal{G}(8,20)$ of UAV map \mathcal{D}_2 , highlighted by a yellow square in Figure.4.1, is represented in false colours (NIR, Red and Green channels); (b) classification of pixels $d(u, v) \subset \mathcal{G}(8,20)$ into two classes: \mathcal{G}_{vin} , representing vine canopies (green), and \mathcal{G}_{int} , representing inter-row surfaces (brown); (c) computed NDVI values of vine canopies pixels \mathcal{G}_{vin} ; and (d) inter-row surface \mathcal{G}_{int} .

the location and extension of the vine canopies, an enhanced NDVI computation can be defined as

$$NDVI_{\text{vin}}(i, j) = \frac{\sum_u \sum_v \frac{m_N(u,v) - m_R(u,v)}{m_N(u,v) + m_R(u,v)}}{\text{card}\mathcal{G}(i, j)} \forall d(u, v) \in \mathcal{G}_{\text{vin}}(i, j) \quad (4.4)$$

An example of the enhanced NDVI definition, by considering only the NDVI of the pixels representing the vine canopies, is reported in Figure.4.4c, while the complete $NDVI_{\text{vin}}$ map for the June dataset is shown in Figure.4.5c. For completeness, the NDVI index was computed also for the vegetation in the inter-row, such as weed or grass, as

$$NDVI_{\text{int}}(i, j) = \frac{\sum_u \sum_v \frac{m_N(u,v) - m_R(u,v)}{m_N(u,v) + m_R(u,v)}}{\text{card}\mathcal{G}(i, j)} \quad \forall d(u, v) \in \mathcal{G}_{\text{int}}(i, j) \quad (4.5)$$

to further evaluate the contribution of no-canopy reflectance to the comprehensive NDVI computed from satellite imagery. The obtained NDVI map for the inter-row areas, obtained by processing the UAV imagery of 29 June, is shown in Figure.4.5d.

4.1.7 Results

Three vineyard parcels (named “Parcel-A”, “Parcel-B” and “Parcel-C”), selected for their peculiar spatial distributions and different micro-climate conditions, were considered in this study Figure.4.1.

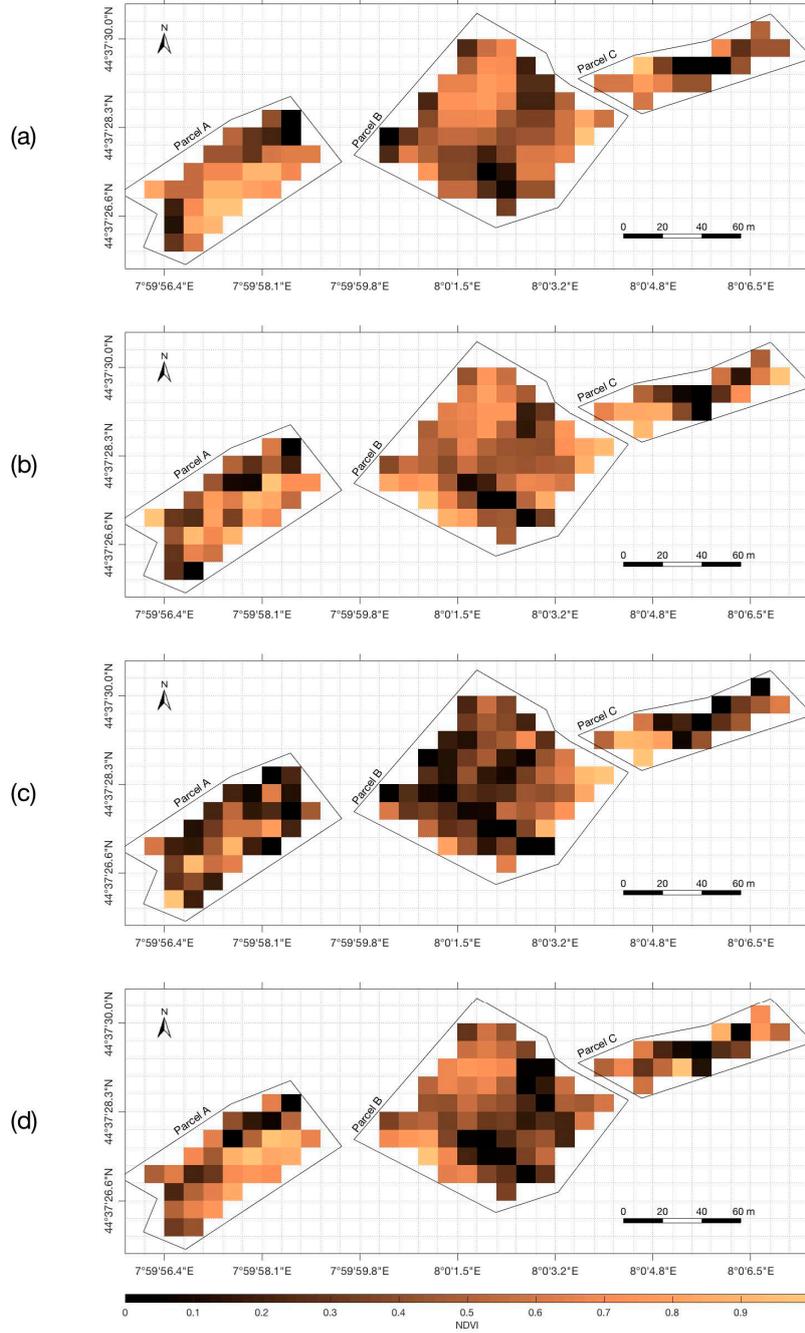


Figure 4.5: Comprehensive (a) $NDVI_{sat}$ map, computed from satellite imagery \mathcal{S}_2 , and (b) $NDVI_{uav}$ derived from UAV imagery \mathcal{D}_2 . (c) Enhanced vineyard $NDVI_{vin}$ map, processing UAV imagery \mathcal{D}_2 by considering only canopy pixels \mathcal{G}_{vin} and (d) $NDVI_{int}$ map considering only inter-row surface \mathcal{G}_{int} . In all represented NDVI maps, only pixels (i, j) completely included within “Parcel A”, “Parcel B” and “Parcel C” boundaries are shown.

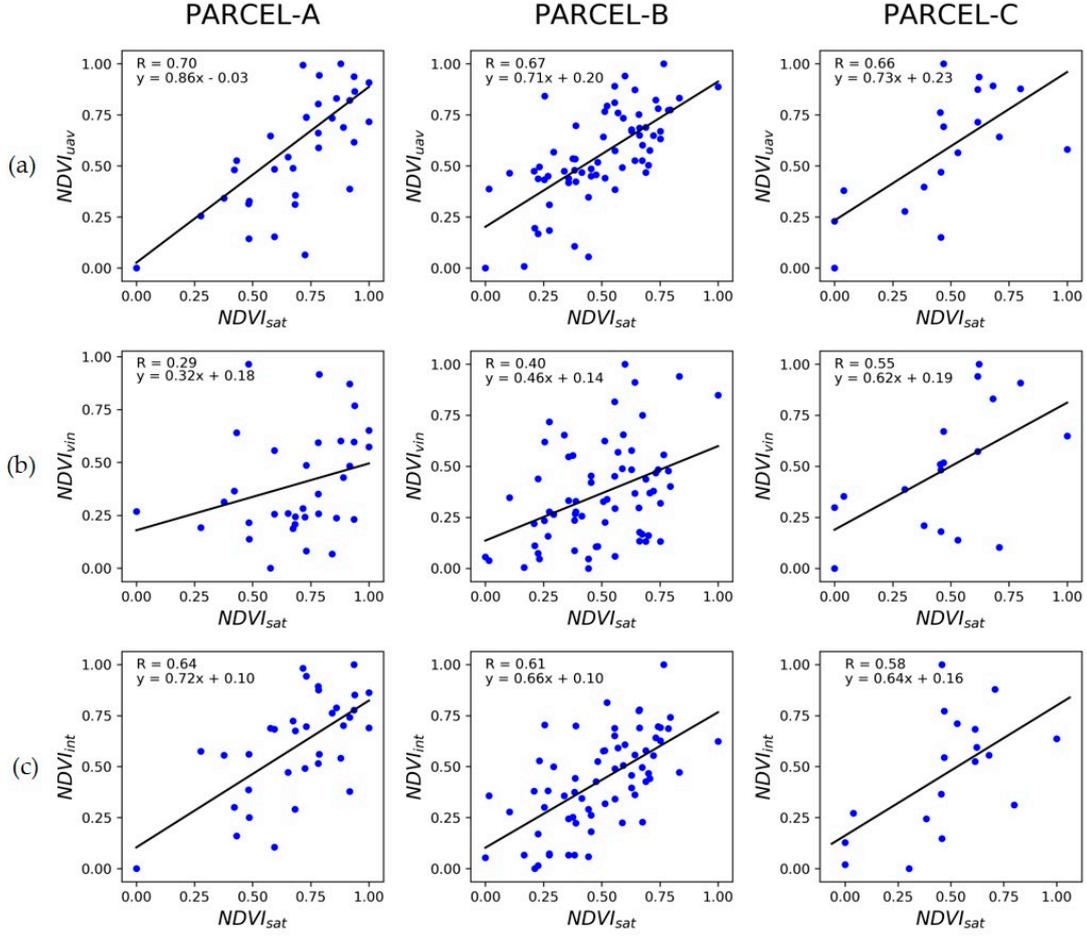


Figure 4.6: Scatter plots of NDVI values from $NDVI_{sat}$ map (x-axis) and: (a) the comprehensive $NDVI_{uav}$ map (y-axis); (b) the enhanced NDVI values of map $NDVI_{vin}$ (y-axis); and (c) the enhanced NDVI values of $NDVI_{int}$ map (y-axis), using imagery pair data $\mathcal{D}_2/\mathcal{S}_2$. The regression model and data pair correlation coefficients are also reported.

To extend the performed analysis to different vineyard phenological phases, four imagery acquisitions were performed (UAV airborne campaigns) and considered (satellite Sentinel-2 platform) during the 2017 crop growing season reported in Table.4.2. Regarding the evaluation and the comparison of the effectiveness of the satellite and UAV-based imagery in describing and assessing the variability within and between vineyard parcels, four different NDVI maps were computed. More in detail, the $NDVI_{sat}$ map (Equation (1)) was derived from satellite imagery while imagery acquired with the UAV platform was processed to obtain three different NDVI maps: (i) a comprehensive $NDVI_{uav}$ map (Equation (3)) by considering the

spectral information provided by both pixels representing vine canopies and inter-row surfaces; (ii) an $NDVI_{\text{vin}}$ map for vineyard canopies (Equation (4)); and (iii) an $NDVI_{\text{int}}$ map for inter-row paths (Equation (5)), by considering only one group of pixels at a time. To be able to compare UAV-based imagery with the satellite imagery, with a GSD of 10 m, high-resolution imagery from the UAV airborne platform was downsampled using Equation (2). The coherence between each spatiotemporal map pair was investigated using the Pearson correlation coefficients, adopted as a map similarity measure [214], after performing a normalisation procedure to focus on the relative differences of each map pair.

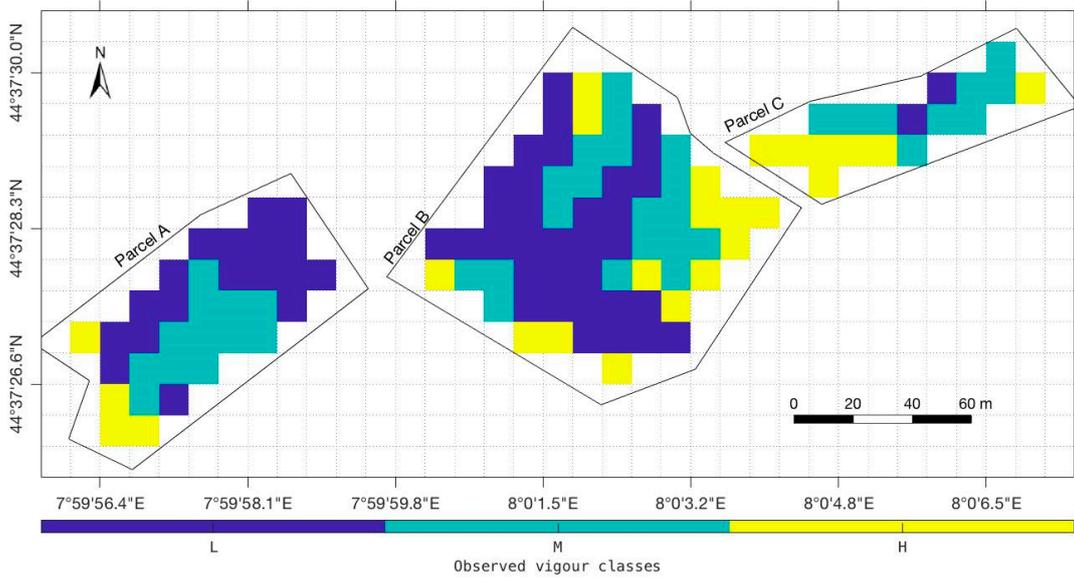


Figure 4.7: Vineyard test site classification into three vigour classes on the basis of the observed in-field vigour assessment. Classes “L”, “M” and “H” refer to low, medium and high vigour, respectively.

A preliminary analysis investigated the coherence of the adopted dataset by comparing the $NDVI_{\text{sat}}$ map and the properly downsampled comprehensive $NDVI_{\text{uav}}$ map for all three considered parcels and for the four acquisition campaigns. The coherence between the information provided by the two platforms was confirmed by the values of the obtained Pearson correlation coefficients, named $R_{\text{Sat/UAV}}$, which were higher than 0.6 for more than 75% of the performed comparison, and never lower than 0.55. All the obtained $R_{\text{Sat/UAV}}$ values, which showed a considerable similarity between the maps derived from the satellite and UAV raw imagery, are organised in Table.4.3. The correlation plots for imagery pair $\mathcal{D}_2/\mathcal{S}_2$, detailed for “Parcel A”, “Parcel B” and “Parcel C”, are shown in Figure. 4.6(a–c), respectively.

Table 4.3: Pearson correlation coefficients results of the NDVI maps comparison procedure.

Map pair	$R_{Sat/UAV}$				$R_{Sat/vin}$				$R_{Sat/int}$			
	$\mathcal{D}_1/\mathcal{S}_1$	$\mathcal{D}_2/\mathcal{S}_2$	$\mathcal{D}_3/\mathcal{S}_3$	$\mathcal{D}_4/\mathcal{S}_4$	$\mathcal{D}_1/\mathcal{S}_1$	$\mathcal{D}_2/\mathcal{S}_2$	$\mathcal{D}_3/\mathcal{S}_3$	$\mathcal{D}_4/\mathcal{S}_4$	$\mathcal{D}_1/\mathcal{S}_1$	$\mathcal{D}_2/\mathcal{S}_2$	$\mathcal{D}_3/\mathcal{S}_3$	$\mathcal{D}_4/\mathcal{S}_4$
Parcel A	0.63	0.71	0.58	0.55	0.31	0.33	0.45	0.40	0.52	0.65	0.56	0.49
Parcel B	0.60	0.68	0.62	0.65	0.39	0.40	0.37	0.38	0.56	0.61	0.60	0.62
Parcel C	0.64	0.67	0.60	0.72	0.41	0.61	0.28	0.51	0.59	0.67	0.54	0.66

Once the coherence of the adopted dataset was verified, the quality of the information provided by the two typologies of NDVI maps was investigated, focusing on the well-known relationship between the considered index and the vegetative condition of the vineyard. This task was performed by comparing $NDVI_{sat}$ and $NDVI_{uav}$ maps to the in-field vigour assessment provided by expert operators, which classified the different regions of the considered vineyard into three different vigour classes shown in Figure. 4.7. The results of the performed analysis of variance (ANOVA), which provided p -values from 0.04 to 0.26, did not prove significant difference between vigour groups means for all the considered parcels (Figure.4.8a).

4.1.8 Discussion

The decametric resolution satellite imagery revealed some limitations in describing substantial information regarding the status of vineyards where the crop radiometric information could be altered by other sources (e.g., inter-row paths) that, in the case of crops grown by rows, could be influential and could negatively affect the overall assessment. This effect was confirmed by the strong relation found between the satellite $NDVI_{sat}$ map and the $NDVI_{int}$ map, derived by the UAV imagery considering only the inter-row pixels, with Pearson correlation coefficients $R_{Sat/int}$ of the 12 performed comparisons ranging from 0.49 to 0.67 (Table.4.3) and with more than 65% of spatiotemporal map pairs achieving $R_{Sat/int} > 0.60$. On the contrary, the relation between the satellite $NDVI_{sat}$ map and the enhanced $NDVI_{vin}$ map, produced by considering only pixels representing vine canopies within the UAV imagery, appeared to be weak. Further analysis revealed that about 75% of the Pearson correlation coefficients $R_{Sat/vin}$, obtained by comparing $NDVI_{sat}$ and $NDVI_{vin}$ maps of all three considered parcels for four acquisition campaigns, were found to be lower than 0.41. All the obtained values of $R_{Sat/vin}$ are organised in Table.4.3. This analysis shows that in those crop types where the inter-row surfaces and paths involve a relevant portion of the cropland, such as vineyards, the radiometric information obtained by moderate spatial resolution satellite platforms are not satisfactory to assess crop status and variability properly. Indeed, depending on the specific adopted crop management strategy, the inter-row surface can be covered by grass, by other crops for pest and disease integrated control or

it could be bare soil. In all these conditions, the vineyard vigour could often be in discord with the inter-row areas, leading to biased vineyard vigour assessments from decametric spatial resolution imagery.

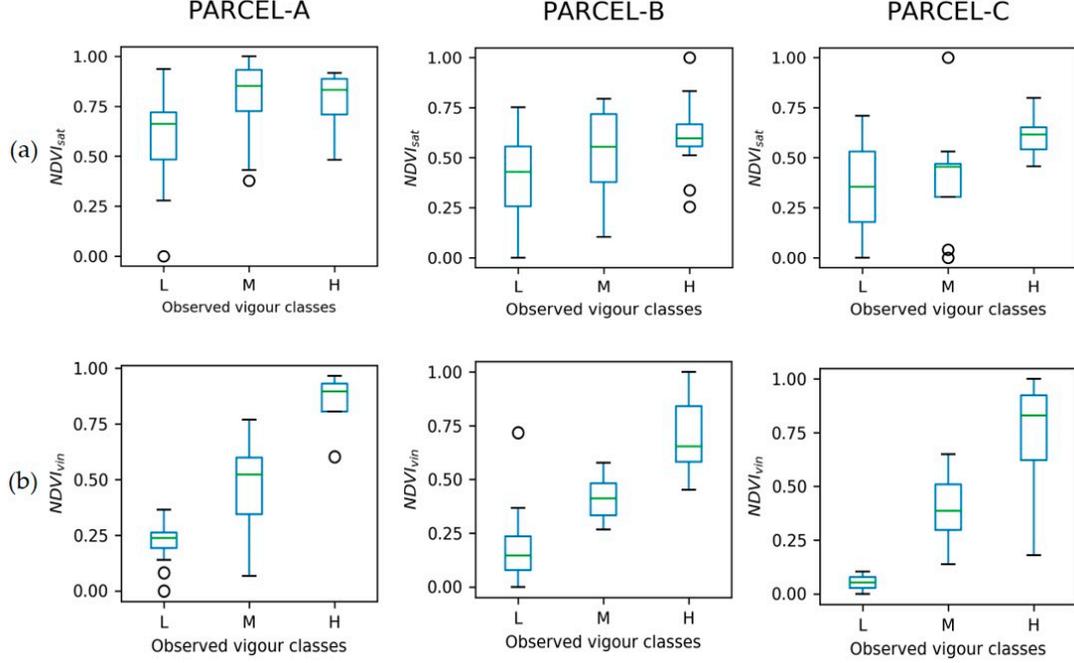


Figure 4.8: Box plots representation of: (a) $NDVI_{sat}$ values derived from satellite imagery; and (b) enhanced $NDVI_{vin}$ values derived from UAV imagery, considering only canopy pixels, divided into three groups on the basis of the observed in-field vigour classes “L”, “M” and “H”.

In addition, the effectiveness of the $NDVI_{vin}$ and $NDVI_{sat}$ maps in discriminating the vigour of the considered parcels in accordance with the experts in-field assessment was investigated with the ANOVA method. For what concern the $NDVI_{vin}$ map, the ANOVA, obtaining p -values ranging from 2.47×10^{-3} to 6.88×10^{-8} (4.4), confirmed that the observed variability of the vineyards within the test site was well described by the $NDVI_{vin}$ map. Boxplots of the $NDVI_{vin}$ values, divided in the three vigour classes, are shown in Figure.4.8b. On the contrary, the ANOVA results considering $NDVI_{sat}$ map showed that the variability of the vineyards within the test site by satellite platform was not significantly in accordance with the experts in-field assessment (Table.4.5). This additional verification confirmed the main result of the presented analysis, proving that, in the case of crops where the inter-row surfaces involve a relevant portion of the cropland, such as vineyards, the radiometric information acquired by satellite platforms can have difficulties properly evaluating crop status and variability. In these situations, imagery with a high spatial resolution is required to properly assess variability within

and between vineyards.

Table 4.4: Results of the ANOVA of UAV based $NDVI_{vin}$ map in relation to the three vigour classes from in-field assessment.

	Source	DF	SS	MS	F-Value	P-Value
Parcel A	classes	2	1.360807	0.680403	30.092543	5.46×10^{-8}
	Error	31	0.700921	0.022610		
	Total	33	2.061721			
Parcel B	classes	2	2.713501	1.356750	71.166427	6.86×10^{-7}
	Error	63	1.201062	0.019064		
	Total	65	3.914563			
Parcel C	classes	2	0.867121	0.433560	9.199357	0.00247
	Error	15	0.706941	0.047129		
	Total	17	1.57406			

DF: degree of freedom, SS: sum of squares, MS: mean square

Table 4.5: Results of the ANOVA of satellite based $NDVI_{sat}$ map in relation to three vigour classes from in-field assessment.

	Source	DF	SS	MS	F-Value	P-Value
Parcel A	classes	2	0.308368	0.154184	3.458293	0.044081
	Error	31	1.382101	0.044584		
	Total	33	1.690464			
Parcel B	classes	2	0.393805	0.196903	4.892817	0.010587
	Error	63	2.535323	0.040243		
	Total	65	2.929128			
Parcel C	classes	2	0.198502	0.099251	1.455564	0.264401
	Error	15	1.022811	0.068187		
	Total	17	1.221313			

4.1.9 Conclusions

In this work, a detailed analysis and comparison of multispectral imagery of vineyards, provided by decametric resolution satellite and low altitude UAV platforms, is presented. The effectiveness of the Sentinel-2 imagery and the high-resolution UAV aerial images was evaluated by considering the well-known relation between the NDVI and vineyard vigour. A cropland located in Piedmont (Serralunga d’Alba, Italy) was selected as the experiment site to perform four image acquisition campaigns, which were properly scheduled according to the main vine phenological stages.

The results show how, in the case of crops where the inter-row surfaces involve a relevant portion of the cropland, such as vineyards, the radiometric information acquired by decametric resolution satellite platforms has difficulties in properly evaluating crop status and variability. In these situations, the vigour of the vineyard could often be in discord with the inter-row areas (e.g., grass, plants for pest and disease integrated control, or soil), leading to biased vineyard vigour assessments from decametric resolution imagery, such as the Sentinel-2 imagery. This was proved by a detailed analysis of the radiometric unbundled contribution of different elements within the cropland, performed by defining three different NDVI indices from the high-resolution UAV imagery, considering: (i) the whole cropland surface; (ii) only the crop canopy pixels; and (iii) only the inter-row terrain pixels.

In this study, the NDVI maps derived from the satellite imagery were found not to be in accordance with the in-field crop vigour assessment. In addition, the satellite-based NDVI maps were found to be more related to the NDVI maps computed by the high-resolution UAV imagery, considering only the pixels representing inter-row surfaces. As a validation, a new type of NDVI map from the UAV imagery, generated by considering only the pixels representing the vine canopies, was defined. The effectiveness of this last type of map in describing the observed vineyard vigour was found to be relevant.

The proposed approach can be extended to other crop typologies that are grown by rows or without intensive layouts, where the crop canopies do not extend on the whole surface or where the presence of weeds is relevant.

4.2 Automatic Path Planning for Unmanned Ground Vehicle using UAV imagery

Over recent years, as farmers and industries have moved towards minimizing human workload on fields, the use of automated machines on fields has grown significantly. These machines help in restricting human labour for tasks such as harvesting, fertilizing, land preparation, plant health monitoring, and seeding. With this development, the role of field machines in management of agricultural environments has become an essential facet of precision agriculture. However, to maximize automated work on the field, agricultural machines need path planning besides being equipped with an autonomous navigation system. Path planning is especially vital and a challenge for regions such as mountain vineyards, due to the terrain morphology.

This work proposes a workflow to generate an automatic coverage path plan for unmanned ground vehicles (UGVs) using georeferenced imagery taken by an unmanned aerial vehicle (UAV). First, image acquisition is performed over a vineyard to generate an orthomosaic and a digital surface model, which are then used to identify the vine rows and inter-row terrain. This information is then used by the algorithm to generate a path plan for UGV.

Part of the work described in this section. 4.2 has been previously published in "Automatic Path Planning for Unmanned Ground Vehicle Using UAV Imagery" [258]. My contributions in this publication as a coauthor are as follows: Conceptualization, Formal analysis, Investigation, Writing—review editing.

4.2.1 Background

Automated machines are being favoured in agriculture by both farmers and industries alike in a bid to minimize human labour in performing tasks on the field. These vehicles have to be equipped with an automatic navigation system and have to be able to select their paths, depending on the specific area as well as to avoid obstacles. Path planning is done – usually by an on-field expert for autonomous agricultural machines – to cover the crop field plots, thus, increasing agricultural productivity. There have been notable advancements in path planning methods, with different optimization methods, for both single and multiple coordinated vehicles. [57] discusses these. Vehicles also make use of greedy algorithms to find efficient routes in order to fulfil base requirements of covering the whole field [153]. Path planning for known environments is done by first portioning the area into smaller polygonal areas called parcels. The best route for each parcel is calculated, and finally, the results are compiled and aggregated. Several optimization techniques have been implemented to find an approximate solution that solves the coverage path planning problem, such as neural networks [245] and genetic algorithms [75]. As such, with a wide variety of available optimization techniques at

disposal, selecting suitable algorithms is an essential step for both proper path planning and environment modelling.

High Spatial Resolution Imagery (HSRI) can be used to characterize areas and distinguish between the different plantations, inter-row terrains, and boundaries [45]. HSRI and emergence of newer image processing methods have also furthered the analysis and classification of vegetative and non-vegetative pixels from textured images. Fast Fourier Transform (FFT) and Garbor Filtering, discussed in [45, 39, 38], determines inter-row width and row orientation however it presents a problem in boundary precision and particularly when missing plants are present.

This work proposes a workflow that automatically generates coverage path plan for UGV using high resolution imagery acquired from UAV. Study area and data acquisition related details are discussed in section 4.2.2 and 4.2.3 respectively. Data processing, mask generation and algorithms used in the study are discussed from section 4 to 8 followed by the conclusion and future work.

4.2.2 Study Area

The survey site was a rural area - near the town of Baldichieri d’Asti (AT) in Piedmont – called “Basso Monferrato Astigiano”, and it was a hilly area dominated by forests and vineyards [202]. The area of interest here specifically, the property of “Azienda Agricola Ciabot”, is around 2 ha and it is present with obstacles both inside and outside. There are also a few areas, and woods overall round the area. Although the path planning was done solely for the small area with only one obstacle – to test, as a first step – the image acquisition and 3d modelling were made for the entire area. (Figure 4.16).

4.2.3 Fieldwork and data acquisition for base map generation

Due to shadows in the area, caused by specific exposures in the area, two acquisition campaigns were carried out in the same day (at 10:30 am and at 5:00 pm) in October 2018. Also, to allow georeferencing, 16 Ground Control Points (GCPs) targets had to be placed because of the lack of natural points in the area. Coordinates for these points were measured using a Network Real-Time Kinematic (NRTK) Global Navigation Satellite System (GNSS) techniques [34]. The NRTK survey was performed by a SP80 Trimble GNSS receiver, using the real time correction of the permanent GNSS station of Canelli. The UAV was selected based on size of the study area, flight times, and projected survey outputs. The used platform was a DJI Phantom 4 Pro, equipped with a RGB camera with focal length 8.8 mm, CMOS sensors 13.2 x 8.8 mm, pixel size of 2.4 μ m. The flights were planned using the open source software Mission Planner that connects the platform to the ground station. This tool was also used to set all the parameters of the flight plan,

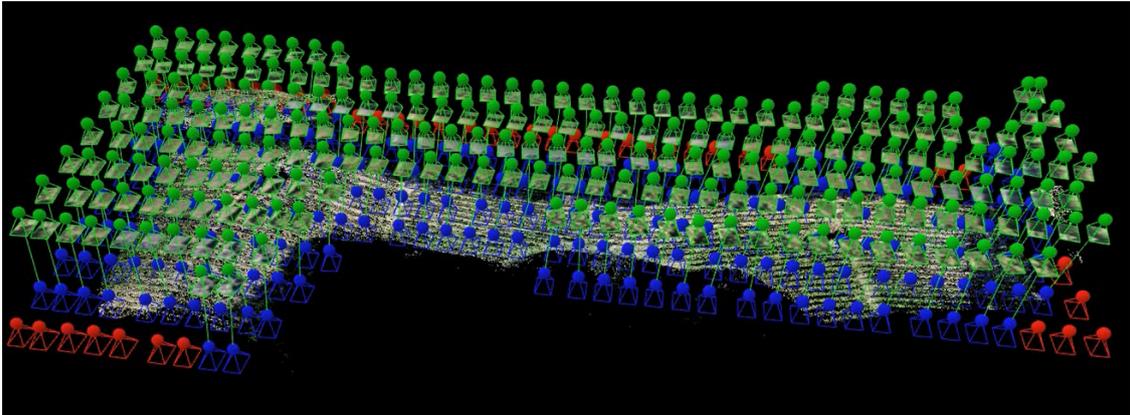


Figure 4.9: UAV imagery acquisition points and positions.

configuration of the flight (nadir images in North-South direction), height of the flight (33m to have a Ground Sample Distance, GSD, of 8mm) and the overlap between different images (80% longitudinal and 60% transverse overlap) Figure. 4.9. The duration of each flight was about 15 minutes and a total of 261 images for the first flight and 209 images for the second flight were acquired.

4.2.4 Data Processing and Validation

In the 3D model generation, data is processed to extract the orthophoto, the Dense Digital Terrain Model (DDTM), and the Dense Digital Surface Model (DDSM) [223]. Commercial software Agisoft PhotoScan 1.2.6 was used for processing the data, with the images acquired from each flight being processed in a single block. These steps were followed routinely: alignment and orientation of the cameras, georeferencing the images by GCPs Figure. 4.15 (horizontal and elevation accuracy are less than 2 cm), point cloud densification (about 80 million points) [3]. The software was used to generate the mesh and the DDSM, and additionally, for this case, the DDTM was also generated. By using the “*Classify ground point*” tool on the software, unsupervised classification of the point cloud could be performed to detect solely the points on the surface the terrain [70]. The classification was performed based on three main parameters:

- max angle: maximum slope of the ground within the scene (20°);
- max distance between the point above the ground and terrain model (0.05 m);
- cell size of the largest area that does not contain any ground points (1 m^2).

The parameters chosen were based off of the features of the area, described in section 4.2.2. The procedure yielded outputs of DDTM, DDSM, and an orthophoto in raster format with cell size of 0.02 m, in WGS84 UTM 32N coordinate system.

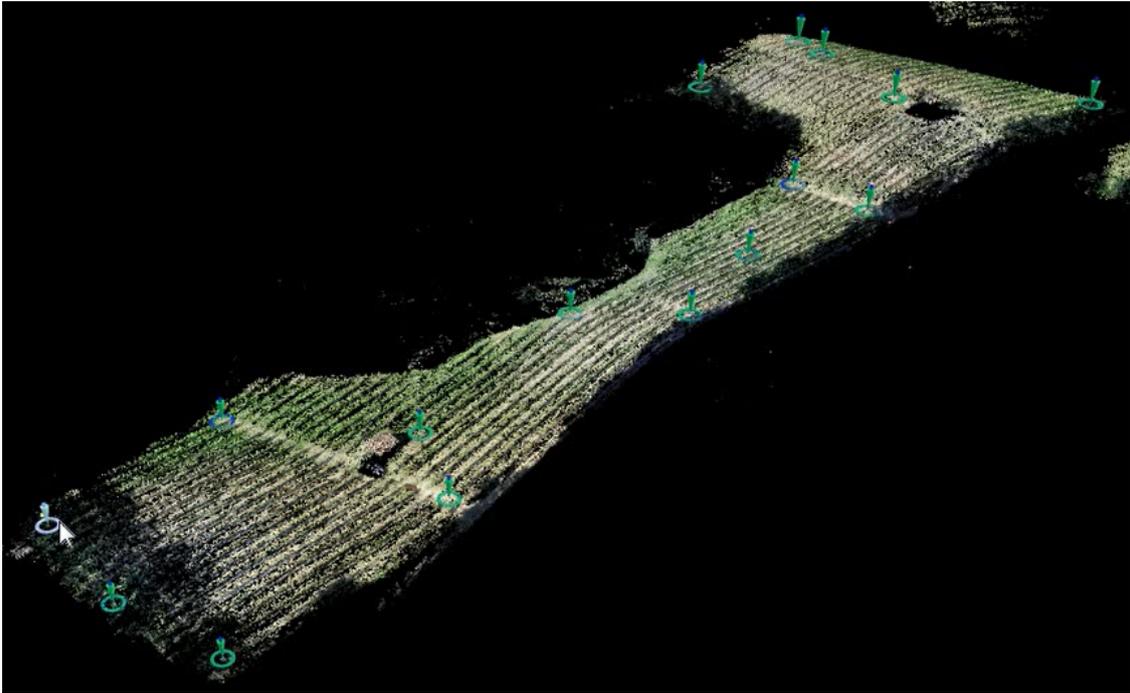


Figure 4.10: Ground Control Points placements for georeferencing.

4.2.5 Mask generation for path planning

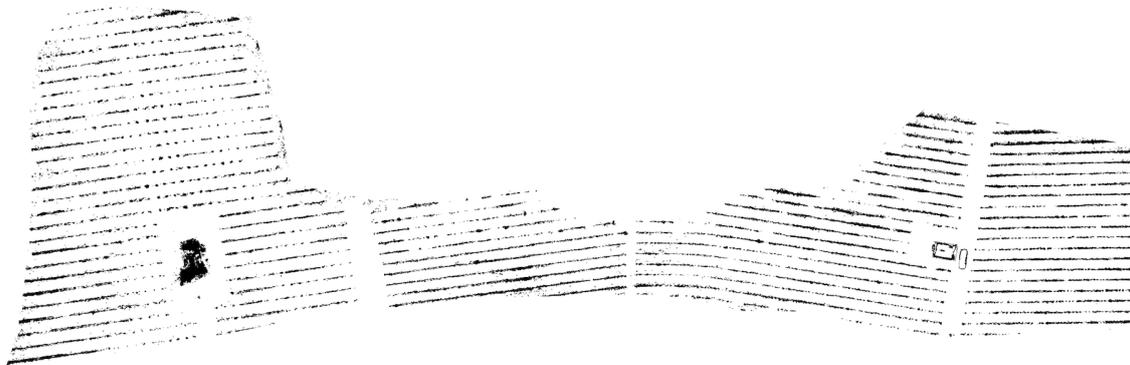


Figure 4.11: Binary mask representing vine rows generated by 3D geometric information. points and positions.

This step defines a typical binary image (see Figure. 4.11) where pixels with non-zero values represent vineyard lines, buildings, and all other potential obstacles for the UGV path. 3D geometric information was used to develop a methodology and by considering the differences between DDSM and DDTM, potential obstacles were mapped. The final map came as a result of splitting the differences into two classes:

(greater than 0.50 cm, and smaller than 0.50 cm). Using the “*Raster Reclass*” tool, non-zero value was assigned to the differences greater than 0.50 cm which represent the presence of obstacles and zero value to the differences smaller than 0.50 cm which represent the ground. This process was carried out for both flights’ data and the masks were merged together to attain a more complete overview.

4.2.6 Steps to generate path plan

From here on the three vineyard fields (Figure. 4.16) focused on in the study area, are referred to as parcels. The algorithm proposed will automatically generate a path plan for the UGV to follow in order to cover a parcel. The algorithm was developed with a generic approach to cater to crops that have considerable inter-row spacing for the UGV movement. This approach is summarized in the flowchart in Figure 4.12.

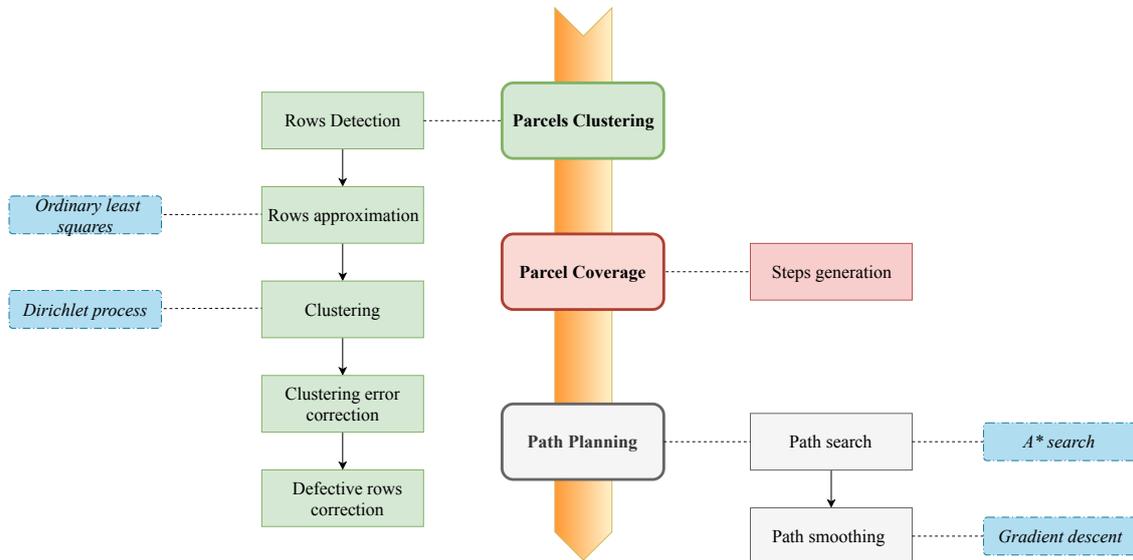


Figure 4.12: Workflow of UGV path generation.

4.2.7 Parcel recognition and rows clustering

The algorithm first detects the rows that represents the vegetation; with each group of the interconnected black pixel representing one canopy row cluster. The steps are based on the concept of visiting masks (section 4.2.5) using sliding windows: proceeding diagonally downwards from the first pixel in the upper left to the pixel in the bottom right. The type of vegetation density is of importance here since it determines the width called ρ , which is used to specify sliding windows. However, since distances differ for different plantation types, a threshold value for

ρ has to be specified. For example, vine rows are more distant as compared to maize rows which are nearer to the vineyards. $\rho = 7$ was used for the environment being investigated in this study. Moreover, due to the acquisition being made in the post-harvest season of vineyards, some missing parts of vine rows were discovered during the mask generation process.

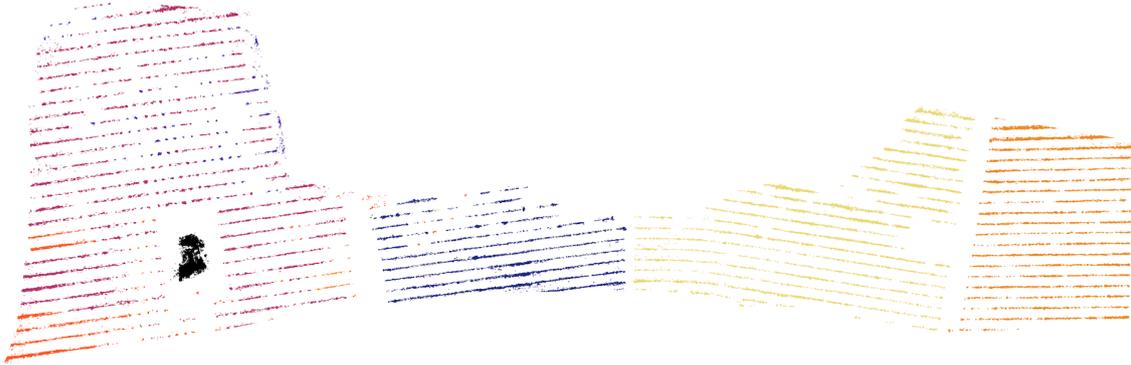


Figure 4.13: Clustering.

In particular, cases where more than one parcel per map is being dealt with, there may be a narrow path between the two parcels through which the farmer or the vehicle can pass, and the problem is then considered twofold. It is then necessary to effectively recombine the segments belonging to the same row. The segments are recombined by finding the line of best fit for each cluster, and by performing a linear regression for each row. The parameterization for this was done by Ordinary Least Squares (*OLS*) method. It uses the given set of variables to determine parameters of a linear function. It minimizes the sum of squares of the difference between dependent variable observed/provided in the dataset and those predicted by the linear function involved.

The linear function mostly took form as a first-order polynomial, since the canopy rows are, usually, disposed of as straight lines. However, there are limitations to the model which were observed and dealt with accordingly. An instance of this is when the points are distributed along the vertical line and the variance occurs, primarily, in the horizontal direction. In such a case, the *OLS* will not give a suitable parametrization of the model because it is designed for points with variances in the vertical direction. To solve this, before projecting the data points into the *OLS* space, the clusters were rotated through 90° which gave such *verticality*. Finally, when the process has recurred for all the rows, they need to be grouped into separated parcels.

To this end, it is imperative that a proper criterion is set to correctly correlate only the rows that are in the same parcel so that all possible errors are avoided.

The literature analysis [7, 100] reveals the most significant issue to be: not knowing the number of parcels within a field beforehand – unless the final user specifically

feeds this information to the algorithm. Due to this, simple or other hardware-demanding algorithms were discarded [225, 78] and the clustering was implemented using the *Dirichlet Process* model. The discarded models include the *K-means* or *DBSCAN*.

Parcels separated by a free space of a higher dimension with respect to the inter-row distance of the same parcel can be identified uniquely by using their respective *slope*. This will result in a collection of tightly arranged slopes, provided that they belong to the same parcel. Since different slopes may be present, the statistical distribution of angular coefficients, in general, will be multimodal. Behind this is the assumption that all data is generated from a finite number of distributions – with unknown parameters – and this allows us to exploit the *Dirichlet Process*. Although it may be, in some cases, that the slope feature is not enough on its own to restrictively bind together the correct rows, this problem can be solved with the introduction of a new feature correlating rows based on their mutual positioning. Mean point coordinates are picked for each row and distant mean points are ruled to – most likely – not be part of the same cluster.

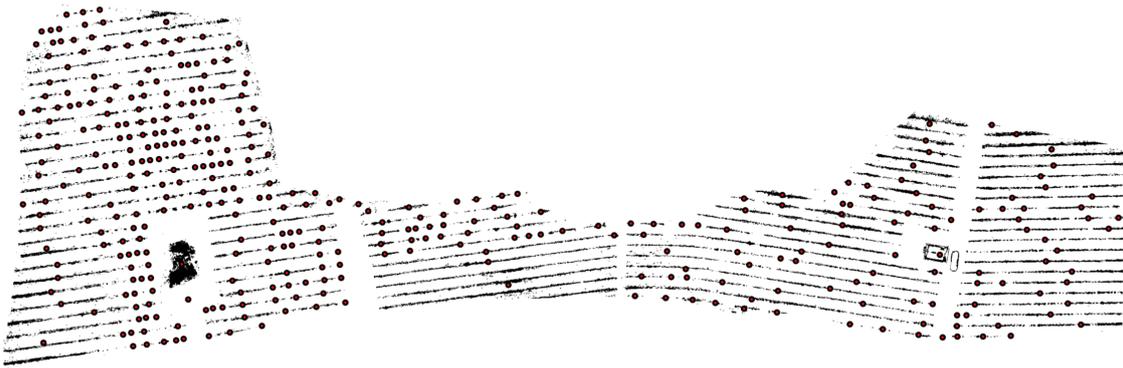


Figure 4.14: Mean points of rows.

There is an apparent drawback to this method though, which arises due to labelling errors occurring in smaller pieces of land with non-contiguous rows. Correction for this is done by considering the majority of row's neighbors, and this process is repeated recursively until no more label changes are needed. Two contiguous rows are connected by the extension of the approximating line from the end of the particular row's extremes until it reaches the neighbor's (Figure 4.16).

4.2.8 Steps generation and path planning

To guarantee the coverage, the UGV needs to receive some nodes along the path between the starting position and the goal. One potential solution to this is to exploit the approximating line between the two close rows. This information can then be used to create a grid that overlays the parcels, and whose dimensions

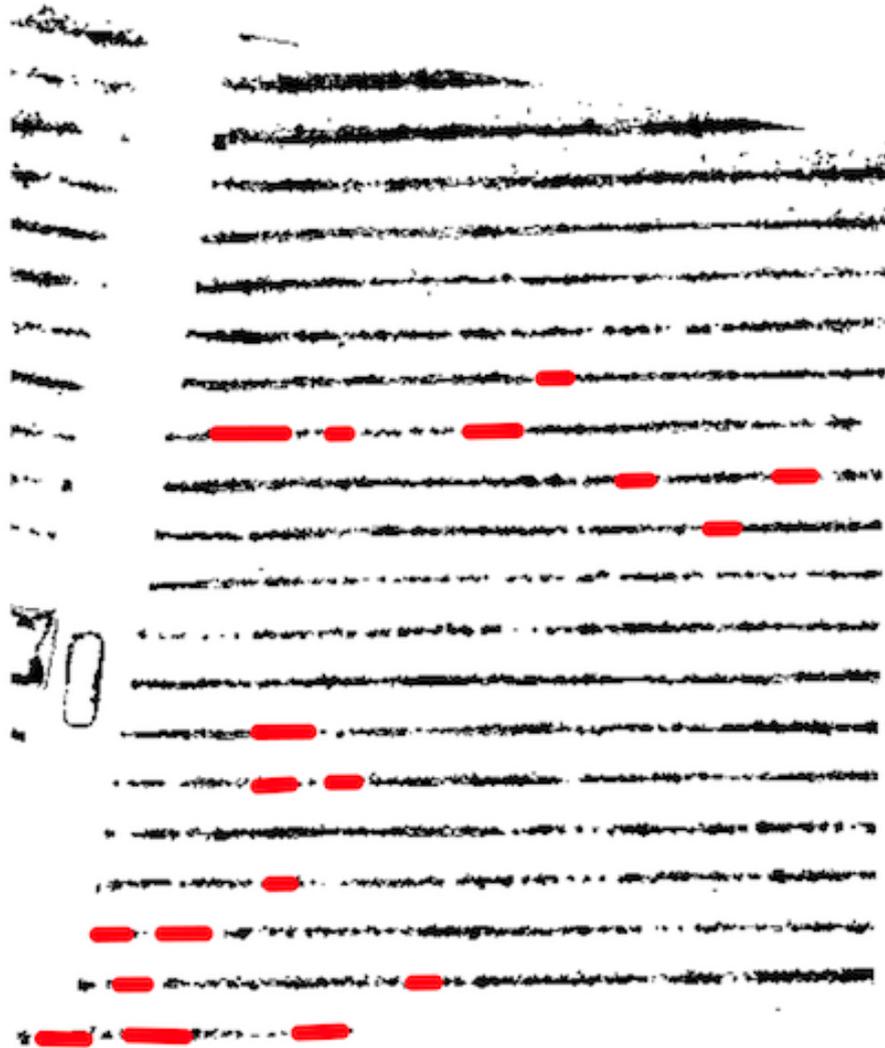


Figure 4.15: Zoomed in view of rows connection.

must be such that is navigable by a robot freely without any cause or incident. This requires the addition of a safety distance. The rows are thus extended in both directions in such a way that the robot can perform the curve. The average point in the curve is computed in such a way so that, in the end, the robot lies precisely in the middle of the two rows.

Still, the robot must be provided with an assigned order by which it must cover the rows. For this, a sorting function has been developed such that it minimizes the *euclidean* distance between consecutive positions. It can also be configured to either cover all the parcels in the map or just a single one.

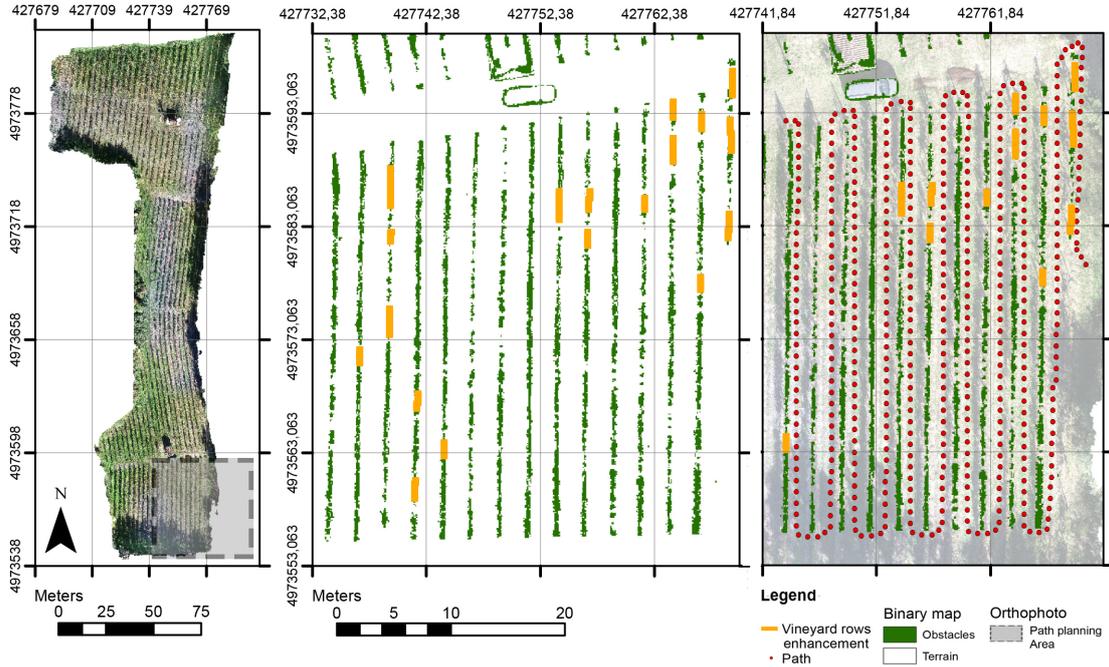


Figure 4.16: Case Study. From left to right: overview, vineyard enhancement, path plan.

Following this, the planner is developed by implementing the A^* algorithm [77], which determines the optimal path between the points by using the previously generated points as the starting and goal nodes, respectively. This also allows the avoidance of obstacles that may be involved. The A^* trajectory is smoothed by running the first-order iterative optimization *Gradient Descent* algorithm.

4.2.9 Conclusion and future work

In this work, an automatic coverage path planning for UGV for hilly vineyard environment is proposed. UAV imagery is used to obtain a DDTM and a DDSM, which are used to build the mask for the path planning. The experimental results show that the work as a whole presents some significant contribution in coverage path planning for UGV in the challenging environment like hilly vineyards that can be useful for the farmers to manage agricultural tasks.

There were some issues encountered when dealing with environments that deviated significantly from one parcel to another while remaining in the same study area. However, we were able to deal with this by considering one parcel at a time in our work. The same approach can be tested and expanded to different vineyard fields in the future, and through considering acquisition campaigns in accordance with the peak phenological stage of the vineyard.

4.3 Refining Satellite based Vegetation Indices using UAV Imagery

Vegetation monitoring of agriculture crops through their phenological cycle assists farmers/agronomist to perform agricultural practices in an effective way to achieve maximum yield while reducing the environmental noise caused by excessive application of chemicals. Numerous remote sensing platforms equipped with optical multispectral sensors such as satellite, airborne and unmanned air vehicles (UAVs) have been in practice for vegetation monitoring. Satellites imagery captured by onboard multispectral sensors is widely used due to their broad coverage and high revisit time. Whereas, UAV imagery is preferred where more detailed information regarding vegetation is needed while it is expensive and time-consuming if more frequent campaigns are to be conducted. In this work, the vineyard site is considered to evaluate the reliability of employing satellite images for vegetation monitoring. Indeed, satellite imagery with moderate spatial resolution cannot describe the vegetation status at vine rows level due to the mixed nature of pixel, expressing the cumulative effect of inter-row terrain and vine rows. Consequently, a pixel refinement is needed to minimize this effect. In this work, a convolutional neural network (CNN) based approach is introduced to gain benefits from high-resolution UAV imagery to refine the vegetation maps derived from moderate resolution satellite images over a vineyard.

Part of the work described in this section. 4.3 has been previously published in "Refining satellite imagery by using UAV imagery for vineyard environment: A CNN Based approach" [104].

4.3.1 Background

Satellite remote sensing data has been widely used since the past two decades for various applications such as agricultural monitoring, forestry, environment monitoring and management, land use monitoring and many more. In recent times, the availability of easily accessible and free multispectral satellite data with reasonable spatial- temporal-resolution has brought new opportunities to researchers and scientist to exploit the possibilities for near real-time agricultural monitoring. Spectral vegetation indices derived from optical imagery through various satellite platforms have been used to assess the vegetative conditions of crops, yield forecast, biophysical parameter estimation of major crop types such as Maize, Wheat, Lucern, Barley etc [15, 175, 25]. Nevertheless, the use of decametric satellite imagery is still challenging for vegetation monitoring for the crops with the row formation having considerable interrow soil separation like vineyards.

Precision agriculture and, in particular, precision viticulture proved to be an imperative approach to improve agricultural products and quality by optimising crop practices [168]. Based on the information of the inherent variability within parcels,

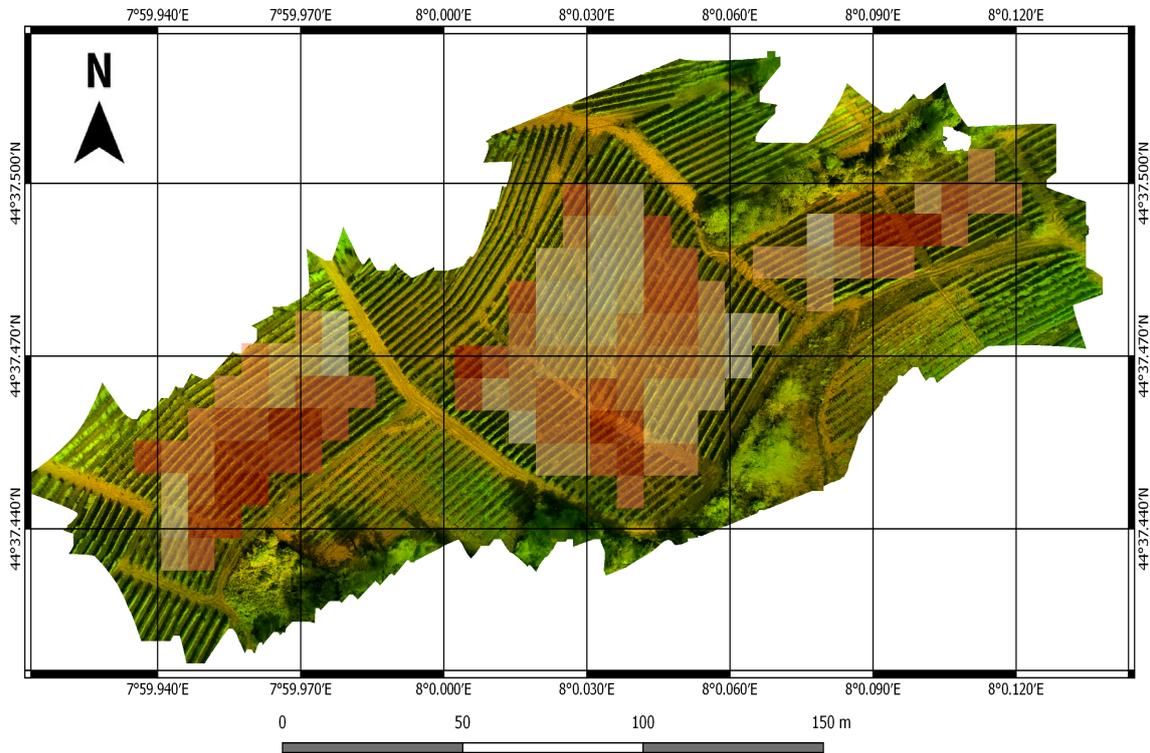


Figure 4.17: Study site located in Serralunga d’Alba (Piedmont, northwest of Italy): map derived from UAV is shown and Sentinel pixels are shown as overlapped region for vineyard fields.

this approach demands new and more reliable methods for crop monitoring goals [107, 169]. In our previous work [107], a time series of sentinel-2 multispectral decametric images and UAV centimetric images acquired between May 2017 and September 2017 were used to compare vigor variability in a vineyard located in Serralunga d’Alba Piedmont, northwest of Italy. Normalized Difference Vegetation Index (NDVI) was calculated for both imageries and comparison had been performed between them, including with the vigour map generated from the field expert. It was concluded that decametric satellite imagery could not be used directly to describe the vigour status of vineyard due to the influence of interrow terrain.

Some past works have used high resolution images acquired from UAVs to refine satellite driven vegetation indices from multispectral imagery. In [91], estimates of canopy structure and biochemical parameters obtained from moderate spatial resolution satellite imagery were improved using high resolution UAV multispectral imagery. In another study, T.W. Cui et al. [44] described how high resolution airborne SAR based images can be used to refine satellite imagery driven estimates of macro-algal coverage in the yellow sea. In the past few years, a branch

of machine learning, known as deep learning, has achieved significant successes in the field of Computer Vision (CV) [81], Natural Language Processing (NLP) [36], speech recognition [31], and automated reasoning [142]. It is made possible due to its intrinsic capability to learn not alone the mapping function from pre-processed input to output, but also the data representations itself. Certainly, representations learned by deep learning architectures are much more disengaged and representative to effectively solve assigned tasks. That is accomplished by hierarchical neural network-based architectures which gradually built more complex and abstract concepts on top of simpler ones. Deep learning provides the ability to learn more robust mapping functions as well as a substantial generalization power than conventional machine learning methods.

In this study, a novel deep learning framework is used to refine satellite driven vegetation maps using high resolution images taken by UAV over vineyard. The remainder of this work is organized as follows. Some details about study site and data set are discussed in Section 4.3.2. Section 4.3.4 and 4.3.5 provides information of the proposed framework and the proposed architecture. Finally, the experimental discussion and results are discussed in Sections 4.3.6, followed by conclusions.

4.3.2 Study Area and Data Set

A vineyard located in Serralunga d’Alba (Piedmont, northwest of Italy), covering a surface of about 2.5 ha, was selected as a field test (see Figure. 4.17). The cropland, whose latitude and longitude positions range between $[44.62334^\circ, 44.62539^\circ]$ and $[7.99855^\circ, 8.00250^\circ]$ (World Geodetic System 1984- WGS84). The Sentinel-2 satellite is equipped with a multi-spectral imaging sensor that measures Earth’s Top of Atmosphere (TOA) reflected radiance in 13 spectral bands ranging from 443 nm to 2190 nm. In this study, cloud-free level-2A Sentinel-2 Bottom of Atmosphere (BOA) reflectance images were used. The UAV-based multispectral imagery were generated with the Agisoft PhotoScan® software (Agisoft©, 2018) processing imagery blocks of more than 1000 aerial images acquired with an airborne Parrot Sequoia® multispectral camera [107].

4.3.3 Data Pre-processing

In our previous study [107], preprocessing steps were performed on UAV imagery such as orthomosaicing, georeferencing, binary mask generation for vine rows. NDVI and two additional up scaled maps were derived representing sole canopies and inter rows. NDVI map was also generated from satellite imagery depicted in Fig. 1. The main objective of this work is to refine the satellite imagery using up scaled map generated from UAV imagery that represents the sole canopies contribution is considered as ground truth in this study.

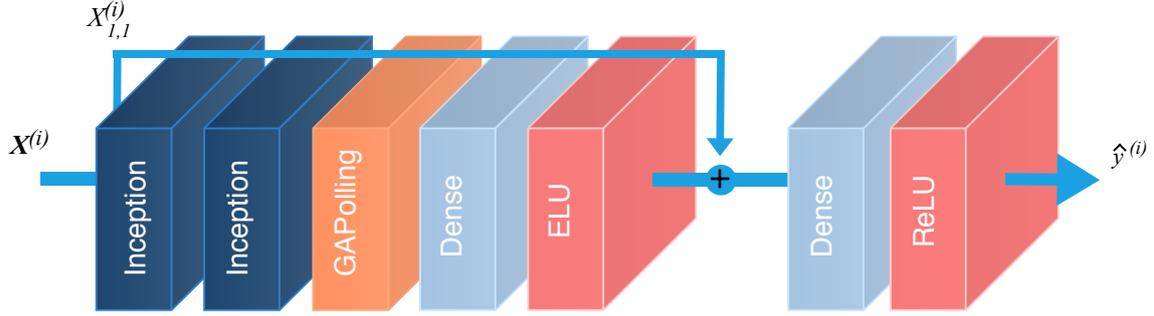


Figure 4.18: Overview of the designed network, named RarefyNet, used for refining Sentinel pixels.

4.3.4 Proposed Method

A convolutional neural network (CNN) with residual connection has been used to generate a non-linear mapping function $F(X^{(i)}, \Theta)$ between $y^{(i)}$ and $X^{(i)}$, where $y^{(i)}$ is the ground truth instance obtained from downscaled UAV images and $X^{(i)}$ is a sample instance acquired from Sentinel-2 images that has to be refined. More formally, $\mathbb{X} = \{\mathbf{X}_i\}$ is a group of all pixels obtained from sentinel NDVI maps depicted in Fig. 5.1 used for training, testing and validation examples, and $\mathbb{Y} = \{\mathbf{y}_i\}$ is a set of downscaled NDVI pixels derived from the UAV images. So, given \mathbb{X} and \mathbb{Y} , optimization method was adopted to train a deep learning-based model, hereafter called RarefyNet, and estimate its parameters Θ using L2 norm $\|x\|_2 = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}}$ as loss function. To provide the architecture with much more exhaustive spatial information and consequently improve the overall model performance, we used for each input data $X^{(i)}$ not only the target Sentinel-2 pixel to enhance, $X_{1,1}^{(i)}$, but also its K neighbourhoods. So, the created training data set is a tri-dimensional tensor $\mathbf{X} \in \mathbb{R}^{i \times j \times k}$ where i are the sample points and 2-D slices $\mathbf{X}_{i,:}$ are the training patches with $K = j \times k$.

4.3.5 RarefyNet Architecture

An overview of the RarefyNet is depicted in Figure. 4.18; each $X^{(i)}$ training sample feeds a stack of two Inception blocks, Figure. 4.19, that gradually extracts spatial correlation between the K neighborhoods pixels and the target central pixel. $X_{1,1}^{(i)}$. Later, a Global Average Pooling (GAP) operation [210] is employed reduce the rank of the processed tensor and construct a 1-D array that feeds a fully connected section with exponential linear units (ELU) as the activation function. GAP operation degrades the spatial dimensions of the secondary Inception block three-dimensional output tensors, explicitly making feature maps to be confidence maps of concepts. Finally, a residual connection blends the processed information with the original Sentinel-2 pixel, and terminate fully connected layer with rectified linear

units (ReLU) function produces the final prediction $\hat{y}^{(i)}$. The residual connection inspired by [215, 122, 143], they implemented super-resolution networks, which plays a primary role inside the overall network; it considerably simplifies the role of the model, moving its objective towards a mere refining operation of pixels derived from satellite imagery. Indeed, the network does not have to recreate the value of the input pixel but progressively learns how to use the starting satellite input value and its spatial information to improve its content of information with its internal learnt representations. In Fig. 5.2, the structure of the two initial Inception blocks is shown; patches $\mathbf{X}_{i,:}$ are concurrently elaborated by a cascade of convolutions with different filters size and dilatation rates. Indeed, distinct kernel sizes exploit the different combination of correlations between data and, on the other hand, atrous convolutions take advantage of non-local spatial correlations. Furthermore, batch normalization [210], as regularization technique, is applied to every branch before ELU activation function. Consequently, in the first block, an input patch $X^{(i)}$ with shape $(j, k, 1)$ is processed in an output tensor of shape (j, k, n_1) where n_1 is the number of derived features maps. The second Inception block builds on top of this tensor, creating more high-level representation and consequently generating a multi-dimensional array with n_2 features.

The complementary use of deep learning-based methods are different regularization techniques to restrain the parameter space and 1×1 convolutions to reduce

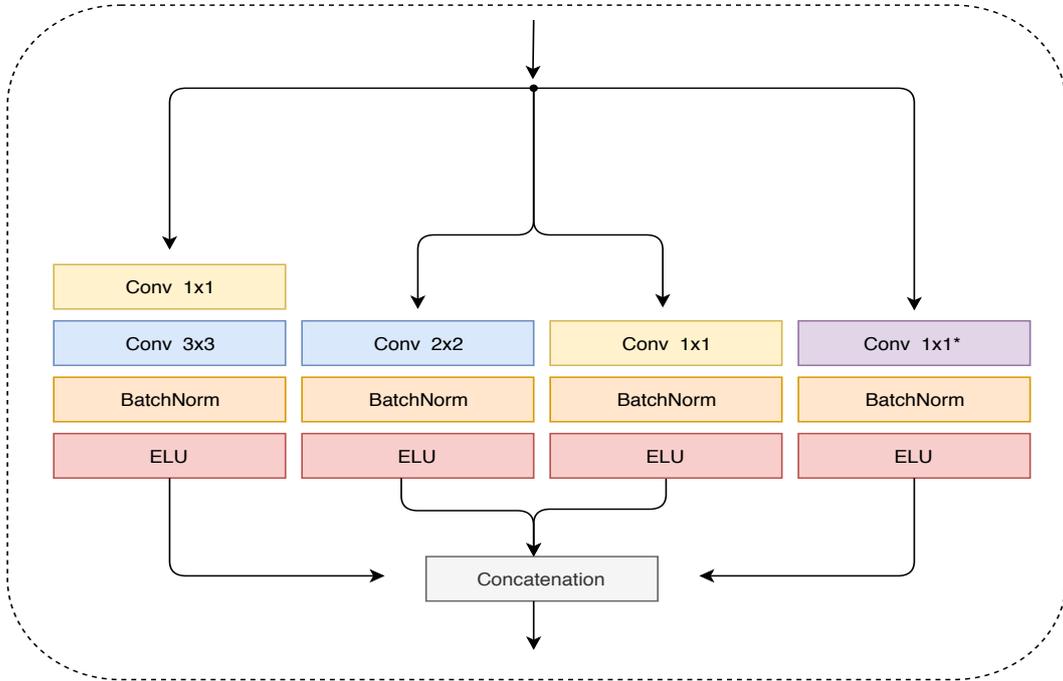


Figure 4.19: Detailed view of proposed Inception block used in the first two layer of the overall architecture mentioned in Fig.4.18.

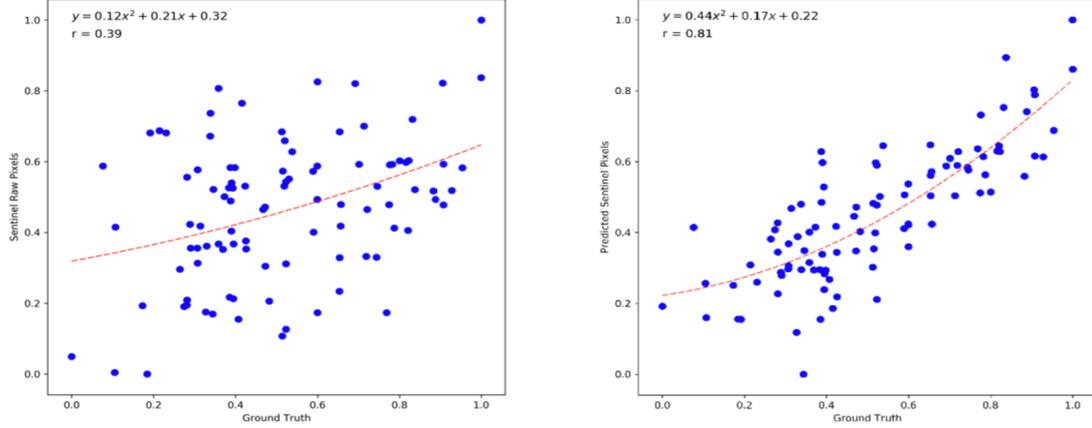


Figure 4.20: Scatter plot with regression equation and correlation coefficient (a) relation between Sentinel raw pixels and ground truth is presented, (b) coherence trend between predicted Sentinel pixels and ground truth.

the model size, offer a light-weight efficient solution to build a complex nonlinear map between satellite and UAV pixel information.

4.3.6 Experimental Discussion and Results

In this section we perform an experimental evaluation of the proposed methodology using data described in previous section, which further enlarged with a simple data augmentation technique. Firstly, an ablation study, reported in Tab. 4.6, has been performed in order to highlight the importance of the residual connection and consequently define the overall structure. Then, results on the test set are discussed and compared with the non-refined satellite pixels.

4.3.7 Data set Creation

We first processed raw data, in order to create a set of training satellite pixel samples $\mathbb{X} = \{\mathbf{X}_i\}$ with the related ground truth labels $\mathbb{Y} = \{\mathbf{y}_i\}$ coming from UAV images. First of all, sentinel data and down sampled drones acquisitions have been divided in temporal couples. Then, a zero padding operation have been

Table 4.6: Correlation of train set and test set for both network architectures.

	Proposed without Residual	Proposed with Residual
X-train r	0.74	0.85
X-test r	0.71	0.81

performed on both maps in order to obtain a rectangular shape dimension on both samples. Finally, working simultaneously on both pairs of raw data, patches of 3×3 have been extracted constructing a tri-dimensional tensor $\mathbf{X} \in \mathbb{R}^{i \times 3 \times 3}$ and $\mathbf{Y} \in \mathbb{R}^{i \times 1}$ where $\mathbf{X}_{i,:,:}$ are the training patches with $K = 3 \times 3$ and \mathbf{Y}_i are the ground truth. So, $\mathbf{X}_{i,:,:}$ is a training sample with its K neighbors where $\mathbf{X}_{i,1,1}$ is the sentinel pixel to be refined and \mathbf{Y}_i the UAV related sample. Moreover, in order to enlarge the number of available training examples, a simple data augmentation process have been applied; considering only one sample i and maintaining fixed the central sentinel pixel $\mathbf{X}_{i,1,1}$ and ground truth \mathbf{Y}_i , it is possible to generate a new sample rotating the K neighbors pixels. So, from a single sample i it is possible to produce $(K - 2)$ new samples for each original training data point. After the samples extraction from raw data, we adopted a simple pipeline of two steps to pre-process the data. A train test split with a percentage of 20% has been performed in order to divide the data set tensor X , with shape $(968, 3, 3)$, in a training and test set. Subsequently, the training set has been processed with the data augmentation technique just explained, producing a final X_{train} tensor with shape $(2520, 3, 3)$. It is worth to notice that no re-scaling was necessary due to the intrinsic nature of NDVI data.

4.3.8 Experimental setting

In this section, we examine the settings of the final network architecture. The basic RarefyNet, shown in Fig. 5.2, is the result of a careful design aimed at obtaining the best performance in terms of accuracy and computational cost. Indeed, the final model is a light weight model with 16,296 trainable parameters that can be easily trained with few samples without overfitting. Every inception block, as depicted in Fig. 4.19, has four different branches with different filter size f_n and dilatation rates.

First of all, the 1×1 convolution halved the number of feature maps in order to be processed by the following operation with a reduced computational cost. The first inception block produces for each branch 8 feature maps n_1 that are concatenated in an unique output tensor, after being pre-processed by a batch normalization layer and a ELU activation function. Equally, the second inception block produces $n_2 = 32$ filters for each branch that are concatenated in a final tensor that feeds the global average pooling layer. A fully connected layer with 32 and 1 output neurons processes the output tensor before the addition with the residual connection. Moreover, a dropout regularization technique with $p = 0.2$ is placed between the two fully connected layers in order to regularize the network and produce a much robust and reliable model. Finally, an output neuron, with a rectified activation function, closes the head of the network in order to compensate and mitigate the presence of possible biases.

In order to find the best training hyperparameters for the optimizer, we used 10%

of the training set to perform a random search evaluation, with few epochs, in order to select the most promising parameters. Then, after this first preliminary phase, the analysis has been focused only on the most promising hyperparameters value, fine tuning them with a grid search strategy. Also, for the optimization process, particular attention has been given to regularization using Adam with weight decay [129] with $\eta = 0.01$ and L2 norm as loss function.

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} m_t - \eta w_t \theta_t \quad (4.6)$$

This is a simple fix to the classic updating rule of Adam optimizer, but that has shown, in our case study, far better results than L2 regularization. In conclusion, we fed our model with 2520 training samples for 500 epochs with a batch size of 64 while cyclically varying the learning rate value, η with a cosine aneling strategy and reducing it every time the loss didn't improve for 10 subsequent epochs. All tests have been carried out with the TensorFlow framework on a workstation with 64 GB RAM, Intel Core i7-9700K CPU and a Nvidia 2080 Ti GPU.

4.3.9 Quantitative Results

Satellite set of pixels $\mathbb{X} = \{\mathbf{X}_i\}$ were divided into training (70%) validation (10%) and testing (20%). Coherence between the sentinel pixels \mathbb{X} and ground truth pixels \mathbb{Y} was evaluated using pearson correlation and regression model which can be seen in Fig. 4.20(a), which supports our conclusion in [4] that decametric satellite images are not well correlated with ground truth that describes real state of the vegetation over the vine yard. On the other hand, predicted sentinel pixels from our proposed CNN architecture (4.20 (b)), are found to be well correlated with the ground truth with 81% correlation, which is encouraging to use more frequently available satellite imagery in combination with few available UAV imagery for vineyard vegetation assessment.

4.3.10 Conclusions

In this work, a convolutional neural network (CNN) based approach was used to refine the moderate satellite images by using more detailed vegetation maps derived from centimetric UAV imagery over a vineyard. Predicted sentinel pixels and ground truth were found more coherent with the correlation of $r = 0.81$ than raw sentinel pixels. These out comes of our CNN based model encourage us to exploit the possibility to apply similar methodology to refine more satellite images by using few UAV images. In this way, monitoring with improved temporal resolution over the vineyards may be possible using more frequent satellite images in combination with few UAV images.

Chapter 5

Research contributions in Multi Temporal approaches for Agricultural Remote Sensing

5.1 Improvement in Land Cover and Crop Classification based on Temporal Features Learning

For productive land use management, agronomists and agricultural agencies must understand the use of current land and how to monitor changes over time. As the spatial and temporal resolution of images from globally available satellites (such as Sentinel-2) increases, we see new possibilities for using freely available multi-spectral optical images for research, with decametric spatial resolution and for more routine revisits for remote sensing applications of the field such as crop classification and land cover (LC&CC), agricultural and environmental monitoring, and management. Although the solutions that exist for cropland mapping can be divided based on their per-pixel and object basis, a challenge presents itself in cases of more agricultural crop classes and when these are considered at a larger scale. In this work, we developed an optimal deep learning framework for pixel-based LC&CC and implemented it based on Recurrent Neural Networks (RNN) alongside Convolutional Neural Networks (CNN) using multi-temporal sentinel-2 imagery of central north part of Italy, which's agricultural system is dominated by the diversity of different economic crop types. This proposed framework is also capable of learning time correlation of multiple images which allows a reduction on stages of manual feature engineering and also on modelling crops phenological stages. For the study, fifteen classes – including major agricultural crops – were considered. Other commonly used machine learning algorithms for comparison

were also tested, such as support vector machine SVM, random forest (RF), Kernel SVM and gradient boosting machine (XGBoost). Compared to existing mainstream methods, our proposed Pixel R-CNN, with an overall accuracy of 96.5%, showed considerable improvement. Thus, this study was able to demonstrate that the Pixel R-CNN based framework provides a very accurate system of assessing and employing time series data for multi-temporal classification tasks.

Part of the work discussed in this section 5.1 has been published in "Improvement in Land Cover and Crop Classification based on Temporal Features Learning from Sentinel-2 Data Using Recurrent-Convolutional Neural Network (R-CNN)" [139]. My contributions in this work as coauthor are as follows: Conceptualization, Data set preparations, Formal analysis, Investigation, Visualization, Writing—review editing.

In this work, we proposed a unique deep neural network architecture for LC&CC, which comprises of Recurrent Neural Network (RNN) that extracts temporal correlations from time series of sentinel-2 data in combination with Convolutional Neural Network (CNN) that analyzes and encapsulate the crops pattern through its filters. The remainder of this paper is organized as follows. Section 5.1.1 briefs about related work done for the LC&CC along with an overview of RNN and CNN. Section 5.1.2 provides an overview of the study area and raw data collected and exploited during the research. Section 5.1.3, 5.1.4, and 5.1.6 provides detailed information of proposed framework and the training strategies. Section 5.1.7 contains a complete description of the experimental settings, classification results 5.1.8 and discussion along with the comparison with previous state-of-the art results. Finally, Section 5.1.10 draws some conclusions.

5.1.1 Background and Related Work

With the global surge in population, there is an ever-increasing demand for agricultural productivity and precision in land cover and crop classification and their spatial distribution. It is vital that governments, policymakers, and farmers improve decision-making processes involved in the management of the agricultural practices and needs [66]. There is a growing necessity for useful cropland mapping and monitoring [138] to be developed, as needed for the essential role of agriculture in the management of our sustainable natural resources. Group on Earth Observations (GEO), with its Integrated Global Observing Strategy (IGOS), also emphasizes on an operational system for monitoring global land covers and agriculture by using remote sensing imagery.

The current systems in place for developing and maintain land cover and agricultural maps have also benefitted majorly from Satellite and Geographic Information System (GIS) [19]. Free sourced satellite data is one of the most used sources for mapping agricultural land and analyzing relevant indices about crop field conditions [238]. With recent advancements like the launch of Sentinel-2 – equipped

with multispectral imager capable of providing up to 10m per pixel spatial resolution with five days revisit time – there is more potential for remote sensing domain. Multi spectral time series data acquired from MODIS and LANDSAT have been widely used in many agricultural applications such as crop yield prediction [95], landcover and crop classification [192], leaf area index estimation [92], plant height estimation [121], vegetation variability assessment [191] and many more. The results can also be improved by using two different data sources simultaneously to extract more features. An example of this: Landsat-8 and Sentinel-1 used together for LC&CC [112].

There are different options available to choose supervised or unsupervised algorithms from, in order to map cropland, and these utilize mono or multi-temporal images [240, 242]. According to the current evidence, multitemporal images perform better than mono temporal mapping methods [66, 236]. For crop area estimation, it is also sufficient to attain imagery that can potentially be used for crucial phenological stages [59, 106]. [256] also showed that reduction in time series length affects the average accuracy of the classifier. Enhanced Vegetation Index, derived from 250 meters MODIS-Terra time series data, was used to establish crop patterns and to classify major crops like cotton, corn, and soybean in Brazil [9].

The feature space for vegetation assessment and monitoring was enriched by the exploitation and use of different vegetation indices (VIs) which are derived from various spectral bands [255, 231]. Crop classification makes use of a wide variety of common features of VIs such as normalized difference vegetation index (NDVI), normalized difference water index (NDWI), enhanced vegetation indexes (EVI). Other features used are statistical in nature, such as mean, standard deviation, and inertial moment, and also include textural features such grey-level co-occurrence matrix (GLCM). The accuracy of the algorithms involved can be increased by using ancillary data such as census data, road density or coverage, and elevation. However, all the features listed above – as well as other phenological metrics – make use of large volumes of data which may significantly increase computational complexity in contrast to the accuracy which then by relation has only improved slightly [131]. For this reason, [131] highlights the several feature selection methods that have been proposed.

In [76], various features have been derived from MODIS time series and best feature selection has been made using random forest algorithm.

LC&CC can also be classified as pixel based or object based. Blaschke also describes Object Based Image Analysis (OBIA) to demonstrate that satellite images can be segmented into homogenous image segments by utilizing high resolution sensors [17]. Many different object based classifications have been proposed that pertain to utilization of satellite imagery for production of crop maps [152, 127, 117].

- **Temporal feature representation:** LC&CC has been the subject of many

studies in the past, and a common classification approach that is adopted currently is using VIs time series (derived from remotely sensed imagery) for extracting temporal features and phenological metrics. There are also several thresholds and simple statistics based procedures which help in calculating the time of peak VI, Maximum VI, and other vegetation related metrics [227, 226]. These brought improvements in classification accuracy, as compared to when using VI features exclusively [201]. In addition to these simpler methods, more complex ones have also been adapted for extracting temporal features and addressing the vegetation phenology [196]. A set of functions [58] can also represent the time series of VI, and these include linear regression [56], curve fitting functions, and the Markov model [198]. [239] also exploits the Sigmoid function and it yielded improved results owing to not just the ease with which phenological features (for vegetation variability characterization [47]) but also its robustness. Nonetheless, even though the models mentioned above provide considerable alternatives and flexibilities in deployment for assessing vegetation dynamics in fields, there are some practical but very significant factors which make selecting these methods more difficult. These practical shortcomings are highlighted by factors depending, mostly, on external conditions, such as uncertain atmospheric conditions, intra-class variability, empirical seasonal patterns, among others. Due to this, there is a need for a proper approach that can take advantage of all the following information available from the time series of VI and can finally extract temporal patterns. Owing to the fact that the architecture proposed in our work is based on pixel based classification, the following subsections are of note as they provide relevant studies and descriptions.

- **Pixel based crops classification:** Khatami et. al. [108] highlights a detailed critical review of the modern supervised pixel-based methods used for land cover mapping. It was discovered that in terms of overall accuracy (OA) – with an OA of about 75% - the support vector machine (SVM) for mono temporal image classification. However, SVM is not only complicated; it is also demanding for time series multispectral data applications with larger area classifications. The second, with an OA of 74%, was Neural Networks (NN) based classifier. Random Forest (RF) bases classifiers is another common approach for remote sensing applications [65], though it should be noted that multiple features need to be derived and fed to the RF classifier for more effective output. One of the newest and most powerful concepts integrated into mapping is a branch of machine learning known as Deep Learning (DL). DL can be used to solve a wide range of problems like signal processing, computer vision, image processing, and natural language processing [116]. A different idea with DL is to break down complex problems into a series of smaller, more simple mappings, each defined by a different layer of the

model. In the end, these are composed hierarchically. Thus, it enables us to not only describe the mapping to output but also the representation involved as well. There are a plethora of modern models, frameworks, architecture, and benchmark databases of reference imagery available for this purpose to enable image classification domain.

- **Neural networks (NNs):** Convolutional Neural Networks (CNNs) date back decades [115], emerging from the study of the brain’s visual cortex [87] and classical concepts of computer vision theory [212], [46]. Since the 1990s, these have been applied successfully in image classification [115].

Sequence data analysis is an important aspect in many domains, ranging from natural language processing, handwriting recognition, image captioning, to robot automation. In recent years, Recurrent Neural Networks (RNNs) have proven to be a fundamental tool for sequence learning [209], allowing to represent information from context window of hundreds of elements. Moreover, research community has over the years come up to different techniques to overcome difficulty of training over many time steps. Applications of CNNs, RNNs, and LSTM in the image classification domain are previously discussed in Chapter 3.

5.1.2 Study Area and Data

The study site near Carpi, Emilia-Romagna, situated in center-north part of Italy with central coordinates $44^{\circ}47'01''N$, $10^{\circ}59'37''E$ was considered for LC&CC shown in Figure 5.1. The Emilia-Romagna region is one of the most fertile plains of Italy. Area almost 2640 km^2 was considered which covers diverse crops land. The major crop fields in this region are Maize, Lucern, Barley, Wheat, and Vineyards. The yearly averaged temperature and precipitation are $14^{\circ}C$ and 843 mm for this region. Most of farmers practice single cropping in this area.

Table 5.1: Bands used in this study.

Bands used	Description	Central wavelength (μm)	Resolution (m)
Band 2	Blue	0.49	10
Band 3	Green	0.56	10
Band 4	Red	0.665	10
Band 8	Near infrared	0.705	10
NDVI	$(\text{Band}8 - \text{Band}4) / (\text{Band}8 + \text{Band}4)$	-	10

To know about the spatial distribution of crops, we deeply studied Land Use Cover Area frame statistical Survey (LUCAS) and extracted all the information

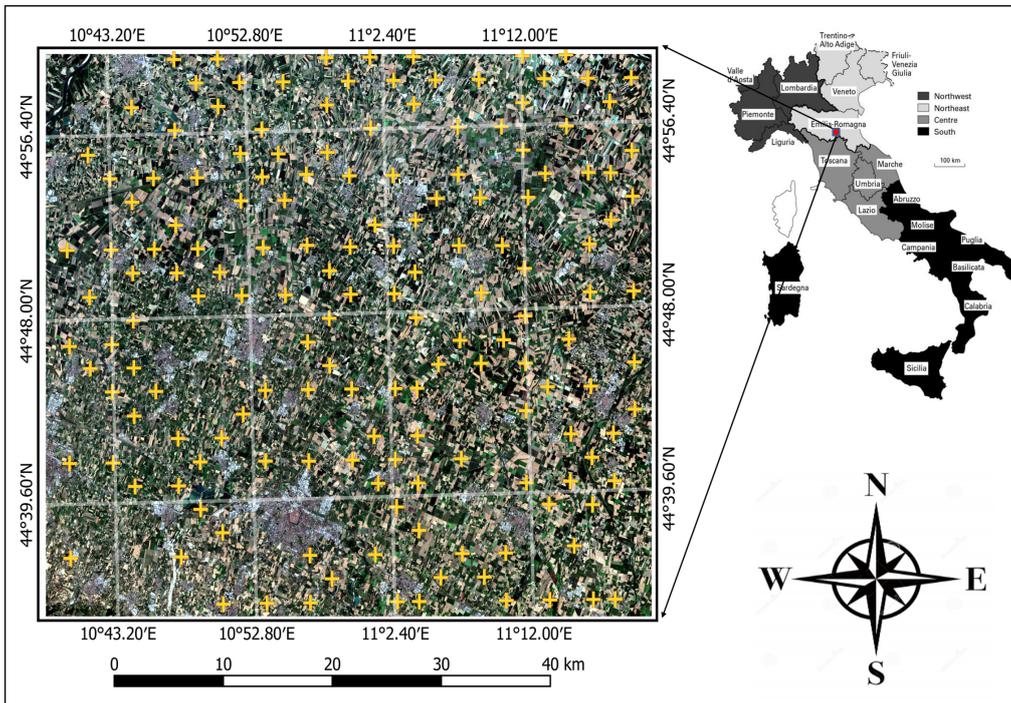


Figure 5.1: The study site is located in Carpi, region Emilia-Romagna is shown with the geo-coordinates (WGS84). RGB image composite derived from sentinel-2 imagery acquired in August-2015 is shown and the yellow marker showing geo-locations of ground truth land cover extracted from Land Use and Coverage Area frame Survey (LUCAS-2015).

we need for ground truth data. LUCAS was carried out by Eurostat to be able to monitor the agriculture, climate change, biodiversity, forest and water for almost all over the Europe [113]. The ground truth data was prepared using technical reference documents of LUCAS-2015. Quantum Geographic Information System (QGIS), open-source software for visualization, editing and analysis of geographical data, was used for importing spatial information of the crops and land cover types as well the geo-coordinates for the regions that were considered in the study segment. Pixels were selected manually by overlapping the LUCAS data and the images, which allowed a considerable amount of ground truth pixels to be extracted for training and testing the algorithm. Twin polar-orbiting satellites, part of the Sentinel-2 mission, launched by European Space Agency (ESA) in 2015 can be used for various application areas such as monitoring natural disasters, detecting land cover change changes, monitoring forest, and perhaps most importantly in agricultural management and monitoring.

It is equipped with multi-spectral optical sensors that captures 13 bands of different wavelengths. We used only high resolution bands that has 10meter/pixel

Table 5.2: Sentinel-2 data acquisition.

Date	Doy	Sensing Orbit #	Cloud pixel percentage
7/4/2015	185	22-Descending	0
8/3/2015	215	22-Descending	0.384
9/2/2015	245	22-Descending	4.795
9/12/2015	255	22-Descending	7.397
10/22/2015	295	22-Descending	7.606
2/19/2016	50	22-Descending	5.8
3/20/2016	80	22-Descending	19.866
4/29/2016	120	22-Descending	18.61
6/18/2016	170	22-Descending	15.52
7/18/2016	200	22-Descending	0

resolution as shown in Table 5.1. It has also short revisit time (10days at the equator and 5 days with twin satellites (Sentinel-2A, Sentinel-2B)). Its popularity in remote sensing is due to the full range of free data sources that are available and the key features that it possesses, such as free access to data products available at the ESA Sentinel Scientific Data Hub and with reasonable spatial resolutions (which is 10m for Red, Green, Blue and Near Infrared bands), reasonable spectral resolution, and high revisit times amongst others. In our study, we used ten multi-temporal sentinel-2 images reported in Table. 5.2, which are well co-registered from July-2015 to July-2016 with close to zero cloud coverage. We made the initial image selection based on cloudy pixel contribution at the granule level. The initial image selection was performed based on the cloudy pixel contribution at the granule level. This pre-screening was followed by further visual inspection of scenes and resulted in multi-temporal layer stack of 10 images. Sentinel Application Platform (SNAP) v5.0 along with sen2core v 2.5.1 were used to apply radiometric and geometric corrections to acquire Bottom of Atmosphere (BOA) Level 2A images from Top of Atmosphere (TOA) Level 1C. Further details about geometric, radiometric correction algorithms used in sen2cor can be found in [180]. Bands with 10 meter/pixel along with the derived Normalized Difference Vegetation Index (NDVI) were used for experiments as shown in Table 5.1.

5.1.3 Convolutional and Recurrent Neural Networks for pixel based crops classification

A single multi-temporal, multi-spectral pixel can be represented as a two-dimensional matrix $X^{(i)} \in \mathbb{R}^{t*b}$ where t and b are the number of time steps and spectral bands,

respectively. We aim to use the two dimensional matrix $X^{tocompute^{(i)}}$ a probability distribution $F(X^{(i)})$ consisting of K probabilities, with K equivalent to the number of classes. To achieve this, we propose a compact representation learning architecture composed of three main building blocks:

- **Time correlation representations** - this operation extracts temporal correlations from multi-spectral, temporal pixels $X^{(i)}$ exploiting a sequence-to-sequence recurrent neural network based on Long Short-Term Memory (LSTM) cells. A final Time-Distributed layer is used to compress and maintain a sequence like structure, preserving the multidimensionality nature of the data. In this way, it is possible to simultaneously take advantage of temporal and spectral correlations.

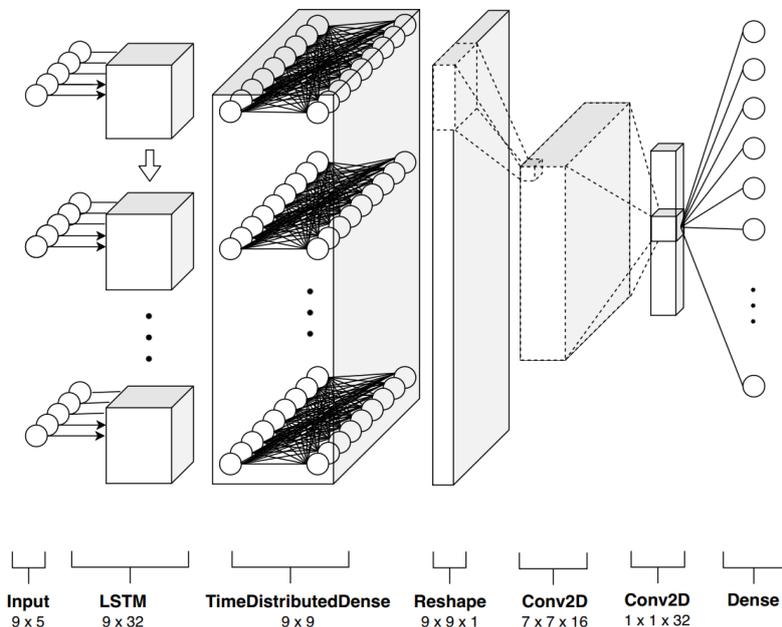


Figure 5.2: An overview of Pixel R-CNN model used for classification. Given a multi-temporal, multi-spectral input pixel $X^{(i)}$, the first layer of LSTM units extracts sequences of temporal patterns. A stack of convolutional layers hierarchically processes the temporal information.

In modern time series data analysis, RNNs are commonly used owing to outstanding results achieved in their use in many fields and applications over the years. Considering the simplest RNN, single-layered, (shown in Fig. 5.3), it has similarities to a Feedforward neural network, except it also has a connection going backwards. The single layer is not only fed an input $x^{(i)}$, it also receives $h^{(i)}$ (cell state) which is equal to the output neuron itself, $y^{(i)}$. In this way, at each time step t the recurrent layer receives its own input $x_t^{(i)}$ as

well as an output from the previous time steps. This is also made possible by a sort of intuitive ingrained memory that has an influence on all subsequent outputs. In this example, it is straightforward to compute a cell's output as shown in Eq. 5.1.

$$y_t^{(i)} = \phi(x_t^{(i)} \cdot W_x + y_{(t-1)}^{(i)} \cdot W_y + b) \quad (5.1)$$

where, in the context of this research, $x_t^{(i)} \in \mathbb{R}^{(1*b)}$ is a single time step of a pixel with n_{inputs} equal to the number of spectral bands b . $y_t^{(i)}$ and $y_{(t-1)}^{(i)}$ are the output of the layer at time t and $t - 1$, respectively, W_x and W_y are the weights matrices. It should be pointed out that although y_t and $x_t^{(i)}$ may be comprised of an arbitrary number of elements $n_{outputs}$, the representation Fig. 5.3 does not change. The depth dimension simply hides all the neurons. Unfortunately, the basic cell described above has many major limitations. The most significant one is that during the training stage, the loss function gradient fades away over time. Due to this, we adopted a more elaborated cell known as peephole LSTM unit for time correlation representation; see Fig. 5.4. This improves further on a concept proposed by Sepp Hochreiter and Jurgen Schmidhuber in 1997. The main concept behind this is to enable the network so that it learns what to use for the current state $h_{(t)}$, what to store in a long term state, what to throw away $c_{(t)}$, and to understand that, for the basic unit, $h_{(t)}$ and $y_{(t)}$ are equal. That is performed with simple element-wise multiplications working as "valves" for the fluxes of information. Fully Connected (FC) layers, receiving as input the current input state $x_{(t)}$ and the previous short-term memory term $h_{(t-1)}$, control the elements, V_1 , V_2 and

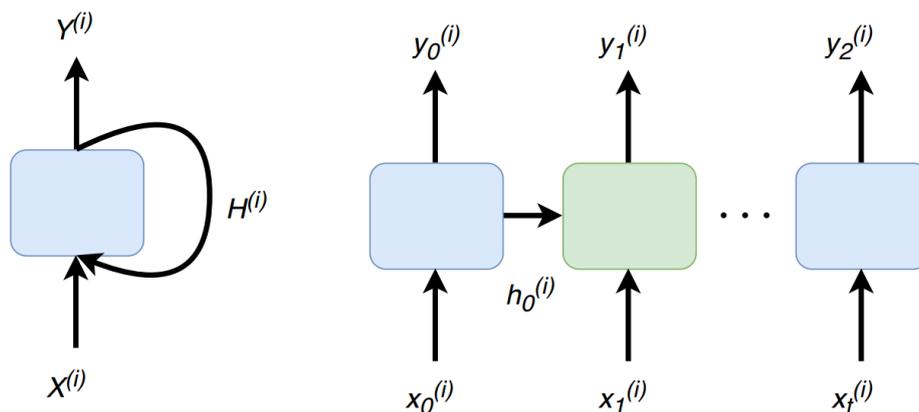


Figure 5.3: A recurrent layer and its unrolled through time representation. A multi-temporal, multi-spectral pixel $X^{(i)}$ is made by a sequence of time steps, $x_t^{(i)}$, that along the previous output $h^{(i)}$ feed the next iteration of the network.

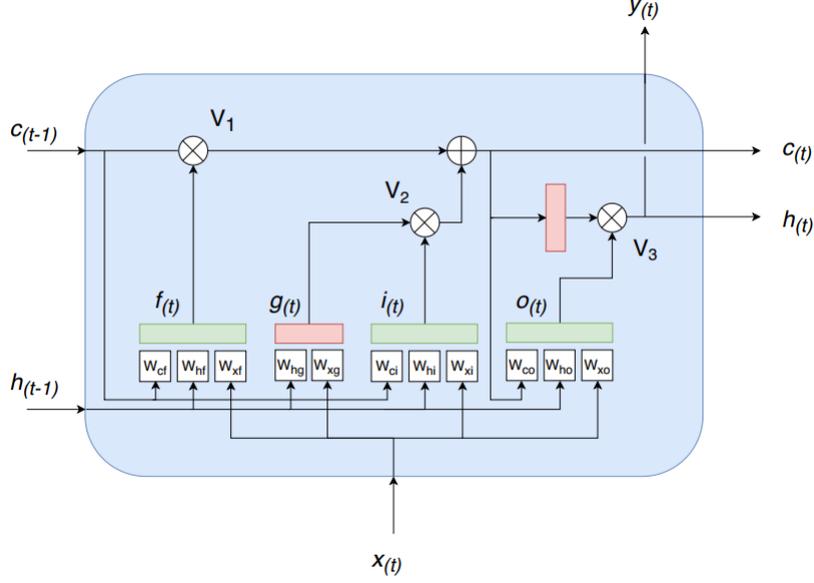


Figure 5.4: LSTM with peephole connections. A time step t of a multi-spectral pixel $x_t^{(i)}$ is processed by the memory cell which decides what to add and forget in the long-term state c_t and what discard for the present state $y_t^{(i)}$.

V_3 . Moreover, for the peephole LSTM cell, the previous long-term state c_{t-1} is added as an input to the FC of the forget gate, V_1 , and the input gate, V_2 . Lastly, the FC of the output gate is fed the current long-term state c_t as input. All “gates controllers” work with sigmoid as activation functions (green boxes), instead of relying on tanh ones to process the signals (red boxes). So, a peephole LSTM block’s signals can be summarized as follows: it has three signals for input and output; two for the standard input state $x(t)$ and one cell output $y(t)$. Here, the long and short term states are c and h respectively, and through its internal controllers and valves, the unit can provide these with useful information. Formally, as for the basic cell seen before, Eq (2). Eq (7). summarizes how to compute the cell’s long-term state, its short-term state, and its output at each time step for a single instance.

$$i(t) = \sigma(W_{ci}^T \cdot c_{t-1} + W_{hi}^T \cdot h_{t-1} + W_{xi}^T \cdot x_t^{(i)} + b_i) \quad (5.2)$$

$$f(t) = \sigma(W_{cf}^T \cdot c_{t-1} + W_{hf}^T \cdot h_{t-1} + W_{xf}^T \cdot x_t^{(i)} + b_f) \quad (5.3)$$

$$o(t) = \sigma(W_{co}^T \cdot c_t + W_{ho}^T \cdot h_{t-1} + W_{xo}^T \cdot x_t^{(i)} + b_o) \quad (5.4)$$

$$g(t) = \tanh(W_{hg}^T \cdot h_{t-1} + W_{xg}^T \cdot x_t^{(i)} + b_g) \quad (5.5)$$

$$c(t) = f(t) \otimes c_{t-1} + i(t) \otimes g(t) \quad (5.6)$$

$$y_{(t)} = h_{(t)} = o_{(t)} \otimes \tanh(c_{(t)}) \quad (5.7)$$

In summary, a first layer of LSTM is used to process multi-temporal and multi-spectral $X^{(I)}$ which yields a cumulative output $Y_{(lstm)}^{(i)}$. In the end of the process, a TimeDistributedDense layer is used to execute a Dense function, with the same weights, for all outputs attained over the time. This also preserves the multidimensionality of the processed data. Eq.(8).

$$F_{timeD}(F_{lstm}(X^{(i)})) = (W \cdot Y_{(lstm)}^{(i)} + B) \quad (5.8)$$

- **Temporal pattern extraction** - This comprises of a series of convolutional operations and rectifier activation functions, which map each of the elaborated spectral and temporal patterns, non-linearly onto higher dimensional representations. Thus, using a cascade of filters that extract meaningful features, RNN output temporal sequences are processed and finalized for the successive stage. Of these layers, the first one extracts a 2-dimensional tensor for each instance. After this, in the second operation, each of the tensors is mapped onto higher-dimensional space. This is executed by using two convolutional operations, built atop each other, that extract more abstract representations gradually by applying launched filters in a hierarchical order. The temporal patterns extraction, for the first convolutional layer is expressed, as F_{conv1} .

$$F_{conv1}(F_{timeD}(F_{lstm}^{(i)})) = \max(0, W_1 * Y_{(lstm)}^{(i)} + B_1) \quad (5.9)$$

where W_1 and B_1 represent filters and biases, respectively, and '*' is the convolutional operation. W_1 , contains n_1 filters with kernel dimension $f_1 \times f_1 \times c$, where f_1 is the spatial size of a filter and c is the number of input channels. The activation function for both of the layer units has been chosen as Rectified Linear Unit (ReLU), $\max(0, x)$, as common for CNNs. Conclusively, the process is summarized as such: the TimeDistributedDense layer's output $Y_{timeD}^{(i)}$ feeds a stack of convolutional networks which gradually reduce the first two dimensions, and slowly work on extracting the higher level representations and thus generating high-dimensional arrays. In addition, with the n_1 and n_2 filters common across all units, in order to execute the same operation with a likely-sized dense fully connected layers, we would need a far greater number of parameters and computational power. The synergy of CNN and RNN presents an optimal and more efficient method of elaborating the overall temporal pattern.

- **Classification** - this final operation maps the feature space with a probability distribution $F(X^{(i)})$ with K different probabilities, where K , as previously stated, is equal to number of classes. At the end of the process, $Y_{conv2}^{(i)}$ the extracted feature vectors are mapped onto a probability distribution of K

probabilities, with K being set equal to the number of the classes. This is done by performing a weighted sum and consequently a SOFTMax activation function.

$$\hat{p}_k = \sigma(s(x))_k = \frac{\exp s_k(x)}{\sum_{j=1}^K \exp s_j(x)} \text{ for } j = 1.., K \quad (5.10)$$

where $s(x) = W^T \cdot Y_{conv2}^{(i)} + B$ is a vector containing the scores of each class for the input $Y_{conv2}^{(i)}$. During the training process, the Weights W and Bias B are learned in a system that allows for the classification of the high dimensional space arrays into K different classes. The estimated probability of an extracted feature vector $Y_{conv2}^{(i)}$ belonging to class K , given the scores of each class for the particular instance, is denoted by \hat{p}_k .

The complete list of operations makes up a compact, lightweight architecture that can map, non-linearly, the multi-temporal information with its intrinsic nature, and attain substantially better results as compared to the previous state-of-the-art solutions. The present architecture was inspired by mental imagery studies of the human brain [141], with the images being a form of inner neural representation. Also, the combined practical use of RNN and CNN allows knowledge representation to be distributed across the whole model thus allowing one of the most significant characteristics of deep learning known as distributed learning to be exploited directly. This model, dubbed Pixel R-CNN, is depicted in Fig. 5.2.

5.1.4 Network Training

The overall mapping function F can be learned through the estimation of network parameters Θ of the three different model parts. This can be achieved by reducing the loss between the pixel class prediction $F(X^{(i)})$ and the respective ground truth $y^{(i)}$. Thus, provided with true class sets $\{y_i\}$ and the respective data set with n pixel samples $\{X_i\}$, the categorical cross-entropy can be used as the loss function:

$$J(\Theta) = -1/n \sum_{i=1}^n \sum_{k=1}^K y_k^{(i)} \log(\hat{p}_k^{(i)}) \quad (5.11)$$

where $y_k^{(i)}$ cancels all classes loss except for the true one. Using AMSGrad optimizer [172], which is an adaptive learning rate method modifying the basic ADAM optimizer algorithm [109], we can also minimize Equation 5.11. The overall algorithm update rule without the debiasing step is:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (5.12)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (5.13)$$

$$\hat{v}_t = \max(\hat{v}_{t-1}, v_t) \quad (5.14)$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} m_t \quad (5.15)$$

The exponential decay of the gradient and the gradient squared are represented by eq. 5.13 and eq. 5.14, respectively. Instead, with the Eq. 5.15., keeping a higher v_t term results in a much smaller learning rate, η , fixing the exponential moving average and preventing to converge to a sub-optimal point of the cost function. We used a technique called cosine aneling to vary the learning rate value between certain boundary values in a cyclical fashion [204]. It is possible to attain this value through a preliminary training procedure, whilst increasing the learning rate linearly and simultaneously observing the loss function value. In the end, we insert our only regularization methodology, known as “Dropout” [207], in the time representation stage between the LSTM and the Time-Distributed layer. By this simple adjustment, we can train more robust temporal patterns – which are also noise-resilient - in the extraction stage. Without the hassle of relying on temporal activations, we can use the CNN and achieve vastly improved abstraction of the generated representations that distribute knowledge across all the available units.

5.1.5 Data set preparation

We processed raw data to create a set of \mathbf{n} pixel samples $\mathbb{X} = \{\mathbf{X}_i\}$ with related ground truth labels $\mathbb{Y} = \{y_i\}$. Then, we applied Principal Component Analysis (PCA) – one of the most popular dimensionality reduction algorithms – to do a visual inspection of the data set and to project the training set onto a lower tridimensional hyperplane. Lastly, a discussion of the quantitative and qualitative results accompanied by a detailed description of the architecture settings is also presented. In order to find the devised architecture, the sample pixels need to be extracted from the raw data and then reordered. So, the first RNN stage entails that data points are collected in a slice of time series as needed. In order to achieve successive pre-processing pipeline, we separated the labeled pixels from the raw data and divided them into chunks of data to form a tri-dimensional tensor $\mathbf{X} \in \mathbb{R}^{i \times t \times b}$. In Fig. 5.5, a visual representation of the data set tensor \mathbf{X} generation, where fixing the first dimension $\mathbf{X}_{i,:}$, there are the individual pixel samples $X^{(i)}$ with $t = 9$ time steps and $b = 5$ spectral bands. However, it should be noted that the number of time bands and steps are chosen arbitrarily based on the availability of raw data.

We pre-processed the data by adopting a simple two-step pipeline. Stratified sample was applied to the data set tensor \mathbf{X} to divide it with the shape (92116, 9, 5), in a training and test set. This is an important step, because of the natural unbalanced number of instances per class in the data set Table . 5.3, and it preserves the same percentage in the two sets. After selecting a split percentage for the training of 60%, we obtained two tensors \mathbf{X}_{train} and \mathbf{X}_{test} with shape (55270, 9, 5) and (36846, 9, 5), respectively. We also adopted standard scaling $(x - \mu)/\sigma$, a

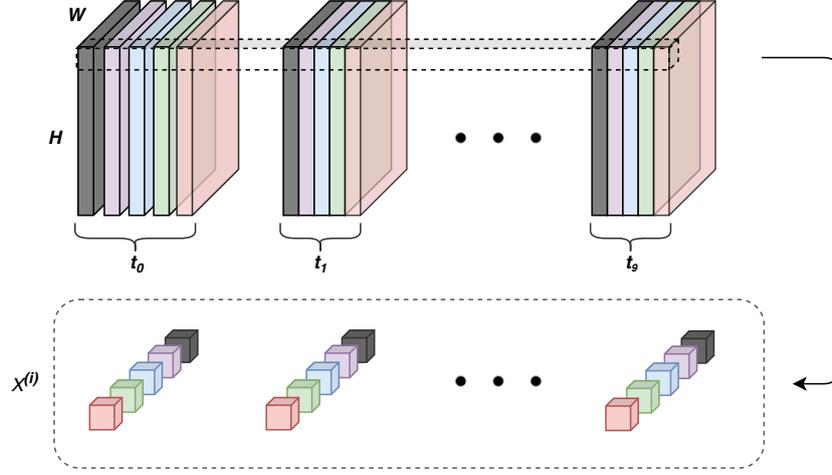


Figure 5.5: Overview of the tensor $\mathbf{X} \in \mathbb{R}^{i \times t \times b}$ generation. The first dimension i represents the collected instances $X^{(i)}$, the second t the different time steps and b the number of spectral bands.

Table 5.3: Land cover types contribution in the reference data.

Class	Pixels	Percentage
Tomatoes	3020	3.20%
Artificials	9343	10.14%
Trees	7384	8.01%
Rye	4382	4.75%
Wheat	12826	13.92%
Soya	5836	6.33%
Apple	849	0.92%
Peer	495	0.53%
Temp Grass	1744	1.89%
Water	2451	2.66%
Lucern	17942	19.47%
Drum Wheat	1188	1.28%
Vineyard	6110	6.63%
Barley	2549	2.76%
Maize	15997	17.37%
Total	92116	100%

common practice to facilitate the training, to normalize the two sets of data points.

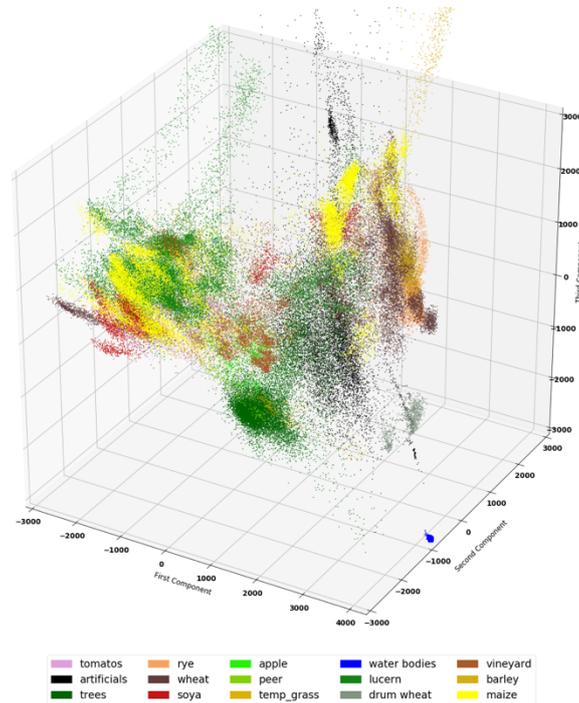


Figure 5.6: Visual representation of the data points projected in the tri-dimensional space using PCA. The three principal components taken into account preserve 64.5% of the original data set variance.

5.1.6 Dataset Visualization

For exploring and visualizing the generated points, we used the Principle Component Analysis (PCA), which reduces the high dimensionality of the data set. For this operation, we took the features of our data points to be components t and b . By applying Singular Value Decomposition (SVD) and consequently selecting the first three principal components, $W_d = (c_1, c_2, c_3)$, we were able to plot the different classes in a tri-dimensional space, achieving a visual representation of the projected data points. Fig. 5.6 shows the projected data points as plotted in a tri dimensional space. It's worth mentioning how much intra-class variance is present, except for the water bodies. Majority of the classes lied on more than one hyperplane, which also accounts for the difficulty of the task underhand and the studied data set. It was also possible to perform an analysis of the explained variance ratio varying the number of dimensions. In Fig. 5.7 it should be noted that on getting closer to higher components, the explained variance trends stop growing fast. This can be considered due to the intrinsic dimensionality of the data set. Thus consequently, it can be assumed reasonably that the reduction of the number of time steps would not greatly affect the overall results.

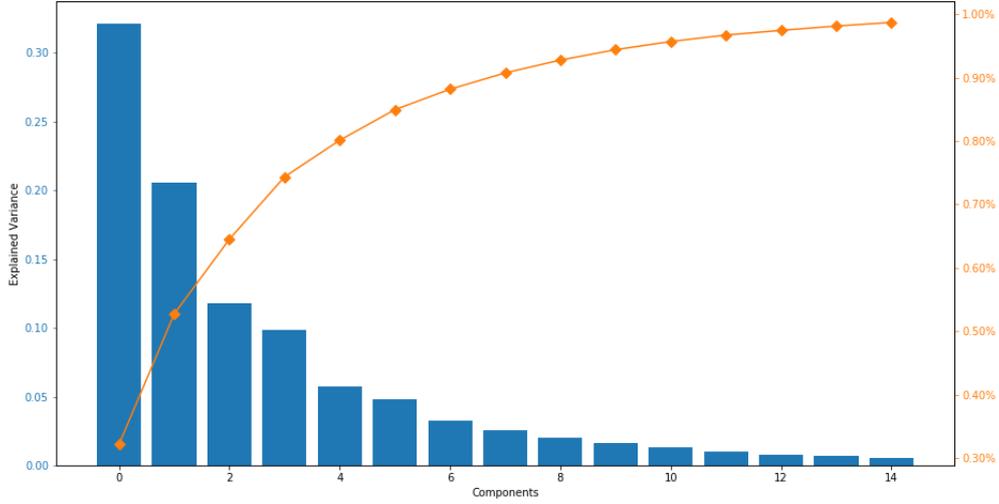


Figure 5.7: Pareto Chart of the explained variance as a function of the number of components.

5.1.7 Experimental settings

We will examine the final network architecture settings in this section. The basic Pixel R-CNN framework – showcased in Fig 5.2 – was obtained as a result of careful consideration and design aimed at obtaining optimal performances in terms of both accuracy and computation. The final model, equipped with 30, 936 trainable parameters (less than 1 MB) is lightweight, fast, and more accurate than the present high level modern solutions available. Following the suggested approach, we used only an RNN layer with 32 output units for peephole LSTM cell randomly turned off, having a probability of $p = 0.2$ with a Dropout regularization operation. The peephole LSTM showed an improvement for all experiments, particularly in overall accuracy with around 0.8% over the standard LSTM cells. Then, the Time Distributed Dense transforms $Y_{(lstm)}^{(i)}$ in a 9×9 square matrix that feed a stack of two CNN layers with a number of features $n_1 = 16$ and $n_2 = 32$, respectively. A one-dimensional array is produced with the first layer as a filter size of $f_1 = 3$ and the second as $f_2 = 7$. Finally, $Y_{(conv2)}^{(i)}$ is mapped to the probability of the $K = 15$ different classes using a fully connected layer with a SoftMax activation function. We adopted ReLU as activations functions for the layers, except for the final one. We used 10% of the training set for performing random search evaluations, with a few epochs, to select the most promising parameters. This was in a bid to find the best training hyperparameters for the optimizer. After this preliminary phase, we shifted the analysis to be focused solely on the most promising hyperparameters' value, fine tuning them with a grid search strategy. So, for the AMSGrad optimizer we set $\beta_1 = 0.86$, $\beta_2 = 0.98$ and $\epsilon = 10^{-9}$. Then, finally, while observing the loss function, to estimate the initial value of this vital hyperparameter, we linearly increased

the learning rate η as was previously shown with a preliminary procedure. So conclusively, while cyclically varying the learning rate value using a cosine aneling strategy, we fed our model more than 62000 samples for 150 epochs with a batch size of 128. The testing was done with the TensorFlow framework operating on a workstation with the following specifications: an Intel Core i7-9700K CPU, with 64GB RAM, and an NVIDIA 2080Ti GPU.

5.1.8 Classification Results

The classifier performance was evaluated with user’s accuracy (UA), overall accuracy (OA), producer’s accuracy (PA), and the most common metric for classification tasks – shown in confusion matrix see Table 5.4 – the kappa coefficient (K) [112, 40, 42, 203]. The overall performance of our proposed Pixel R-CNN architecture, indicated by overall accuracy (OA), is calculated as the ratio between correctly classified total number of pixels and total ground truth pixels for all classes. The pixels that were classified correctly for each class are denoted as the diagonal elements of the matrix. We calculated the individual class accuracy by dividing the number of correctly classified pixels in each category by the total number of pixels in each corresponding row, referred to as the User’s accuracy, and corresponding columns, referred to as Producer’s accuracy. PA indicates the probability that a certain crop type on the ground is classified as such. UA represents the probability

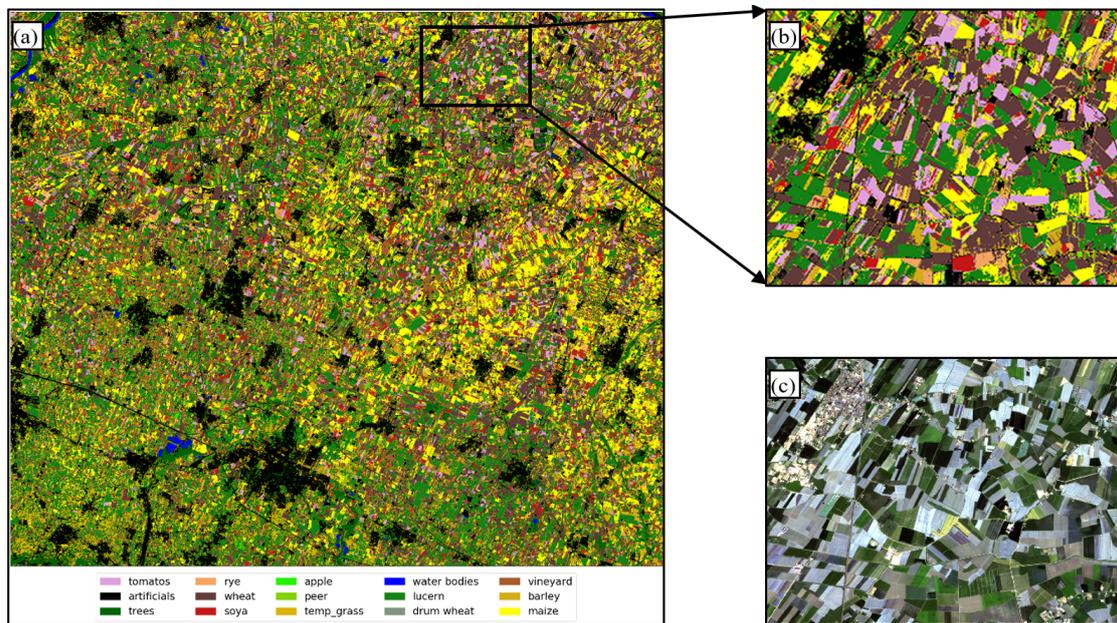


Figure 5.8: (a). Final classified map using Pixel R-CNN, (b). zoomed in region of the classified map, and (c). Raw Sentinel-2 RGB composite of the zoomed region.

that a pixel classified in given class actually belongs to that class. We were able to

achieve OA=96.5% and Kappa=0.914 with 15 number of classes with our proposed pixel-based R-CNN method. This was for a diverse and large-scale area, and thus it exhibited many improvements as compared to the other mainstream methods. The highest UAs were for water-bodies and trees, with 99.1% and 99.3% respectively. This is mainly because of the intra-class variability and minor NIR band reflectance changes over time, although it was learnt quickly by our Pixel R-CNN framework. Most of the classes, including the significant crop types such as Wheat, Lucern, Maize, Vineyard, Soya, Rye, and Barley, had more than 95% UA. The worst class was Grassland, classified with PA = 65% and UA = 63%. The primary confusion of grassland class was with Lucern and Vineyard. On the other hand, Artificial Class, which is comprised of roads, buildings, and urban areas, and represents pixel reflectance with a mixed nature, was detected accurately with UA = 97% and PA = 99%.

Table 5.4: Obtained confusion matrix.

Ground Truth	Classified Classes																Total PA	
	TM	AR	TR	RY	WH	SY	AP	PR	GL	WT	LN	DW	VY	BL	MZ			
Tomatoes (TM)	1096	0	0	0	4	11	0	0	0	0	0	0	0	0	0	0	1111	98%
Artificial (AR)	0	3752	8	1	2	0	2	1	9	9	12	2	6	0	4	3808	99%	
Trees (TR)	0	31	2967	1	0	0	0	3	10	0	17	0	2	0	0	3031	98%	
Rye (RY)	0	1	0	1960	25	0	0	0	0	0	0	0	0	5	0	1991	98%	
Wheat (WH)	38	7	0	221	4981	6	0	0	10	0	14	1	2	38	42	5360	93%	
Soya (SY)	3	0	0	0	3	1226	0	0	0	0	11	0	3	0	41	1287	95%	
Apple (AP)	0	0	0	0	0	0	142	0	0	0	2	0	21	0	0	165	86%	
Pear (PR)	0	0	11	0	0	0	27	124	0	0	0	0	6	0	0	168	73%	
Grassland (GL)	0	39	3	7	0	1	0	0	239	0	72	0	3	0	4	368	65%	
Water (WT)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	906	100%	
Lucern (LN)	0	0	0	2	0	2	0	0	48	0	7250	0	26	0	10	7338	98%	
Durum.Wheat (W)	0	4	0	0	0	0	0	0	2	0	0	0	0	0	0	328	98%	
Vineyard (VY)	11	7	4	4	11	1	50	1	21	0	93	0	2139	0	7	2349	91%	
Barley (BL)	0	1	0	2	24	0	0	0	1	0	16	1	6	0	817	96%		
Maize (MZ)	17	14	0	0	10	24	0	3	10	0	0	0	0	0	7689	7790	99%	
Total	1165	3856	2993	2198	5060	1271	221	132	350	915	7488	326	2214	860	7797			
UA	94%	97%	99%	89%	98%	96%	64%	93%	68%	99%	96%	99%	96%	95%	98%			

Table 5.5: Comparison with recent studies.

Study	Details			
	Sensor	Features	Classifier	Accuracy
Our	Sentinel-2	BOA Reflectances	Pixel R-CNN	96.50%
Rußwurm and Körner [188], 2018	Sentinel-2	TOA Reflectances	Recurrent Encoders	90%
Skakun et al. [203], 2016	Radarsat-2 + Landsat-8	Optical+SAR	NN and MLPs	90%
Conrad et al. [42], 2014	RapidEye	Vegetation Indices	RF and OBIA	86%
Violo et al. [224], 2018	Sentinel-2	Optical	RF	91-95%
Hao et al. [76], 2015	MODIS	Stat + phenological	RF	89%
J.M. Peña-Barragán [159], 2011	ASTER	Vegetation Indices	OBIA+DT	79%

For class Apple, obtained PA was 86% while UA = 68% which shows that 86% of the ground truth pixels were identified as Apple but only 68% of the of the pixels classified as Apple in the classification were actually belongs to class Apple. Some Pixels (see Table. 5.4) belongs to Peer and Vineyard were mistakenly classified as Apple.

We also compared our approach with other recent multi-temporal classifications. We have shown the final classified map, with an example of the zoomed part and the actual RGB image, in Fig. 5.8. To our knowledge, there isn't a multi-temporal benchmark dataset that compares the classification approaches on an equal footing. This is because some of the data sets available online for classification don't have ground truth of other land cover types such as Water bodies, Artificial land (build-up) etc. Various dependencies have to be considered such as the extent of the considered study area, the number of classes to be evaluated, and the number of evaluated ground truth samples and this makes a direct quantitative comparison very difficult. However, we provided a well related comparison of the study domain by their applied approaches, number of considered classes, used sensors and achieved overall accuracy in Table. 5.5. Hao et al. [76], achieved 89% OAA by using RF classifier on the extracted phenological features from MODIS time series data. It was determined that, provided the data's temporal resolution is sufficient, it is possible to achieve good classification accuracies with handcrafted classification algorithms and features. Still, the MODIS sensor data is not apt for areas of large homogenous regions because of its low spatial resolution (500m). By using high spatial resolution data from RapidEye sensor, Conrad et al [42] achieved 90% OA for the nine considered classes. [203], shows features from optical and SAR were extracted and used by the committee of neural network of multilayer perceptrons to classify a diverse agriculture region. Recurrent encoders were employed in [188] to classify a large area for the 17 considered classes. High spatial resolution (10m) Sentinel 2-data was used and achieved 90% of OAA, thus demonstrating successfully that recurrent encoders are useful in capturing the temporal information of spectral features, which leads to higher accuracies. Using the RF classifier, though nine classes being considered, and the Sentinel-2 data, Voulo et al [224] achieved a maximum 95% classification accuracy.

5.1.9 Non deep learning classifiers

We also tested four more traditional classifiers on the same dataset for comparison: Support Vector Machine (SVM), Kernal SVM, Random Forest (RF), and XGBoost. All of these are quite considered as baseline models for classification tasks [53] and are quite well known for high performance. SVM is capable of a non-linear classification, making use of the kernel functions by separating hyperplanes. A widely used RF classifier is an ensemble of decision trees based on bagging approach [197]. XGBoost is built on a gradient of framework of decision trees, something

that garnered a lot of attention in the machine learning community. Both have been used extensively in remote sensing applications. [131, 76, 65]. Each classifier involves hyper parameters which need to be tuned at the time of classification model development.

Table 5.6: Comparison of Pixel R-CNN with non deep learning classifiers.

Model	Parameters	OA
SVM	C: 0.01, 0.1, 1 , 10, 100, 1000 Kernel: linear	79.50%
Kernel SVM	C: 0.01, 0.1, 1 , 10, 100, 1000 Kernel: rbf Gamma: 0.1 , 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8	76.20%
Random Forest	n_estimators: 10, 20, 100, 200, 500 max_depth: 5 , 10, 15, 30 min_samples_split: 3, 5 , 10, 15, 30 min_samples_leaf: 1, 3, 5 , 10	77.90%
XGBoost	learning_rate: 0.01 , 0.02, 0.05, 0.1 gamma: 0.05, 0.1 , 0.5, 1 max_depth: 3, 7 , 9, 20, 25 min_child_weight: 1, 5, 7, 9 subsamples: 0.5, 0.7 , 1 colsample_bytree: 0.5 , 0.7, 1 reg_lambda: 0.01, 0.1, 1 reg_alpha: 0, 0.1, 0.5, 1	77.60%
Pixel R-CNN	Mentioned in experimental settings	96.50%

We performed a “random search” approach to optimize the major hyperparameters. The bold letters in Table. 5.6 denote the best values of hyperparameters, as selected based on classification accuracies achieved for the validation set. Table. 5.6 also reports other details about the hyperparameters and the achieved overall accuracy (OA) for all the other traditional classifiers that were tried out (SVM, Kernel SVM, RF, and XGBoost). The highest OA out of these non-deep learning classifiers was for SVM, which had OA = 79.6%. It was followed closely by RF, Kernel SVM, and XGBoost, which had 77.5%, 76.8%, and 77.2% respectively. However, our proposed Pixel R-CNN based classifier attained an OA score of 96.5%, which is far better than the non-deep learning classifiers. Indeed, handling large data sets and learning temporal and spectral correlations is a challenge for the traditional non-deep learning classifiers. With the advent of deep learning models in remote sensing domain, there is increased flexibility which means that temporal features

can be exploited in such a way that the amount of information gained is increased, thus yielding better and more reliable results for classification tasks.

5.1.10 Conclusion and future works

In this work, we developed a unique deep learning architecture with Recurrent and Convolutional Neural Network called Pixel R-CNN to perform Land Cover and Crop Classification by using multitemporal decametric sentinel-2 imagery of central north part of Italy. Our framework promises significantly better results as compared to mainstream methods, attaining a 96.5% OA with Kappa = 0.914 for 15 number of classes. Besides, we also tested the other non-deep learning classifiers such as SVM, RF, SVM Kernel, and XGBoost, all widely in use. In comparing the results with our own proposed classifier and indeed, the comparison revealed our method to be much more effective, more so when the classification accuracy can be increased through the temporal feature extraction. One significant advantage of our architecture is that by learning the time correlation of multiple images, it is capable of performing automated feature extraction. This, thus, reduces the modelling crops' phenological stages and the manual feature engineering involved. Nonetheless, the method is not exclusively limited to classification tasks and can be extended to other applications that need the extraction of temporal features.

5.2 Assessment of Biophysical Parameters of Crops using Sentinel-2 Imagery

For maize crop, examining biophysical parameters such as canopy height and above-ground biomass are the important agro-ecological indicators that are used in describing the crop growth, photosynthetic efficiency and carbon stock. Remote sensing is a broadly used approach and proved to be the most relevant source of information that can be used to observe vegetative conditions over the agricultural fields. In this work, sentinel-2 multispectral imagery is used to determine spectral vegetation indices covering the different maize growth period by using some visible bands, as well as the near-infrared band. The relationship was established and analyzed between in-field measurements of maize biophysical parameters (such as the height of the canopy and above ground biomass) and the derived spectral vegetation indices from remotely sensed imagery using simple linear regression and Pearson correlation to exploit the possibility of using satellite imagery for estimation of crop biophysical parameters.

Part of the work described in this section. 5.2 has been previously published in "Analyzing relationship between maize height and spectral indices derived from remotely sensed multi-spectral imagery" [105].

5.2.1 Introduction and Background

Crop monitoring is an essential element in precision agriculture since it is used to assess crop conditions over time, to foretell irrigation time, and predict total yield. Besides, understanding crop growth model and its relevant biophysical parameters such as height and biomass are essential in crop monitoring to achieve better results in agricultural productions [89, 35, 165]. Canopy height (Hcanopy) and above-ground biomass (AGB) of crops are two essential biophysical parameters that can be used to evaluate the growth of the crop for the study of precision agriculture. The Hcanopy and AGB of crops also present important indications on the light use efficiency, and carbon stocks in agricultural ecosystems [118], that can be associated with the yield production and their environmental impact. A cost-effective and quantitative estimation of crop biomass is useful in risk management, policy-making, and decision-making for the national and global agricultural-related businesses [15]. There are several methods of obtaining biophysical variables, the direct one is to acquire the values of crop height and biomass is to measure the plant height by tape and pluck plants and record the measurements of biomass by using a conventional method adapted by the agronomists mentioned in [118]. Hence, it is hectic, labour-intensive and time-consuming activity for agronomic experts and ecologist to perform long-term measurements of crops spanning over large areas.

From the last two decades, remote sensing has been emerged as the most valuable resource in the domain of precision agriculture due to its capability to produce information about the condition of the crops over regional and global scales. Remote sensing is a widely adopted approach for agricultural monitoring purposes. It can be used to measure various spectral vegetation indices and estimate biophysical variables over extensive croplands. Vertical plant structure is the foremost factor responsible for the spectral reflectance of vegetative canopies in the visible and near-infrared (NIR) spectral bands, and studies show that there are considerable correlations between biophysical parameters and spectral vegetation indices (VIs) using several combinations of visible, NIR, and shortwave reflectance [25, 26]. Nevertheless, in the past studies, not much work has been carried out about crops biophysical parameter estimation using satellite imagery due to low spatial and temporal resolution.

In [175], images obtained from low-resolution satellite image time series (with spatial resolution from 250m to 1000m) have been used for agricultural land monitoring and management for over three decades. Although it can provide some relevant information for vegetation monitoring but also has some drawbacks due to the big pixel dimensions, which leads to biased values of vegetative indicators. In monitoring crops, the development and progress of high spatial and temporal resolution satellite sensors provide new prospects for the researchers to exploit this finely detailed data for numerous applications. Recent satellite such as Sentinel-1 and -2, SPOT5, QuickBird, WorldView-1 and -2 can provide high-resolution imagery with a fine spectral and temporal resolution can be used to extract more precise and reliable information in vegetation monitoring.

Several studies [176, 150, 79] confirmed the concept of estimating maize crop biophysical parameters such as height and biomass using UAV imagery. However, UAV imagery shows some limitations such as difficulty in sensors calibration, strict flights campaign following crop phenological cycles, labour extensive. Availability of free moderate resolution satellite imagery, such as Sentinel-2, can be exploited for estimation of the crops biophysical parameters. Accordingly, the focus of this study is to compare the several vegetation indices derived from Sentinel-2 multi-temporal imagery with maize biophysical parameter parameters Hcanopy and AGB. Based on the analysis in this study, we offer some propositions on how to make full use of the moderate spatial resolution images obtained from sentinel-2 in agricultural monitoring such as height and biomass estimation of maize crop.

5.2.2 Study Area and Data Set

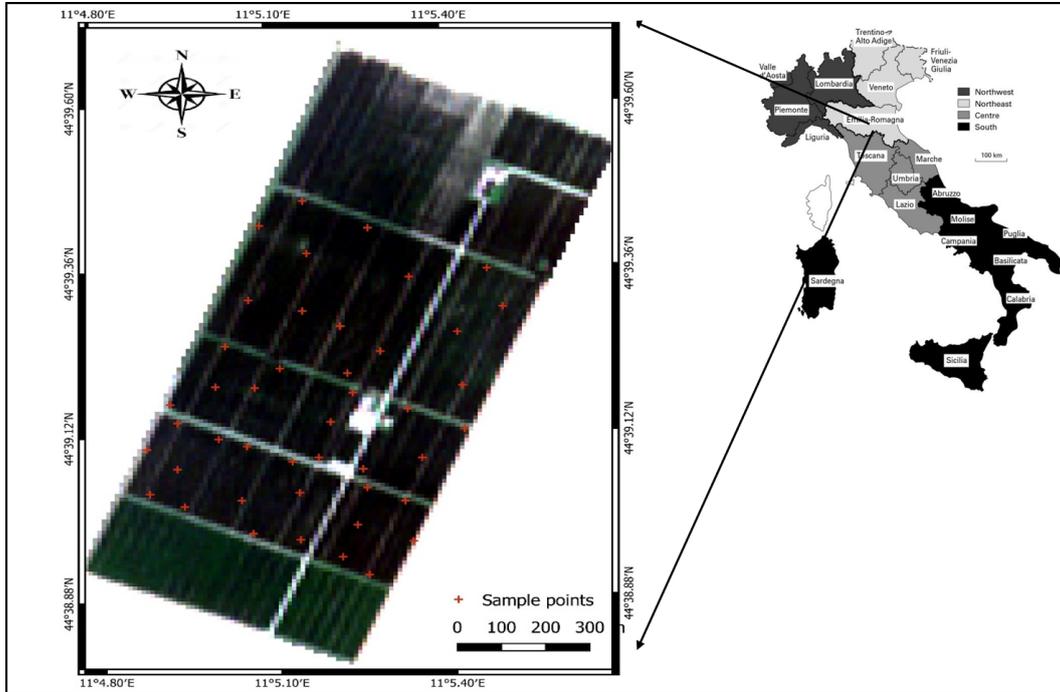


Figure 5.9: The study site of city carpi, region emilia-romagna is shown with the geo-coordinates (WGS84). RGB image composite derived from sentinel-2 imagery acquired in August 2018 is shown and the red marker showing location points where physical observations were taken.

In this study, several maize crop fields shown in Figure 5.9 are chosen from the carpi, emilia-romagna region situated in the north part of Italy with the central coordinates $44^{\circ}39.24'N$, $11^{\circ}5.25'E$. The emilia-romagna region is one of the most fertile plains of Italy. Maize is one of the major crop type in this area, usually sowed in late May, flowering near early August, and harvested late September. The mean annual temperature is $14.2\text{ }^{\circ}\text{C}$ with the coldest weather in late December and January and the hottest in July. The average precipitation during the maize phenological cycle is 89 mm . Sentinel-2 consists of twin polar-orbiting satellite (sentinel-2A and sentinel-2B) launched by European Space Agency (ESA) in 2015 and 2017 respectively and are used in several application areas such as land cover detection and classification, natural disaster monitoring, forest monitoring and in agricultural monitoring and management [194]. It is equipped with multi-spectral optical sensors which captures 13 bands of different wavelengths shown in Table 5.7. It has become more popular due to fact that it offers various key features such as, free data availability at reasonable spatial resolution up to 10m (for Red, Green,

Blue and Near Infrared bands), short revisit time and has good spectral resolution among other available free data sources.

Table 5.7: Spectral bands of Sentinel-2.

Spectral Bands	Central Wavelength (μm)	Resolution (m)
Band 1 (B1) – Coastal	0.443	60
Band 2 Blue	0.49	10
Band 3 Green	0.56	10
Band 4 Red	0.665	10
Band 5 (B5)– Vegetation Red Edge	0.705	20
Band 6 (B6)– Vegetation Red Edge	0.74	20
Band 7 (B7)– Vegetation Red Edge	0.783	20
Band 8 NIR	0.842	10
Band 8A (B8A)– Narrow NIR	0.865	20
Band 9 (B9)– Water vapor	0.945	60
Band 10 (B10)– SWIR	1.375	60
Band 11 (B11)– SWIR	1.61	20
Band 12 (B12)– SWIR	2.19	20

5.2.3 Methodology

In Figure.5.10, overall methodology used in this work is presented. Spectral bands from multi temporal data products were extracted and atmospheric correction was performed to all bands using SNAP tool box provided by ESA.

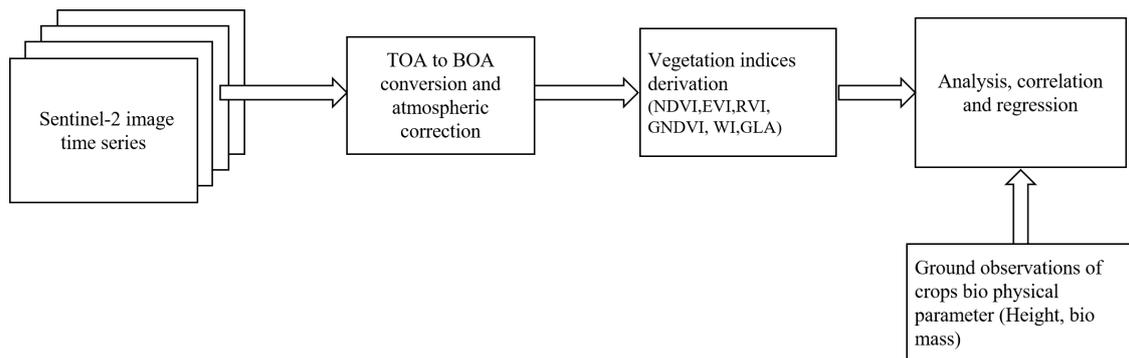


Figure 5.10: Adopted methodology for comparing spectral vegetation indices (VIs) derived from satellite imagery with biophysical parameters of maize crop..

5.2.4 Vegetation Indexes

Various vegetation indices reported in Table.5.8 were derived from multi spectral sentinel-2 imagery. These VIs were calculated as various combination of RGB and NIR digital numbers associated with their corresponding bands.

Table 5.8: Vegetation Indices Formulation

Vegetation Indices	Formula	Source
GLA	$(2G - R - B)/(2G + R + B)$	
GNDVI	$(G - R)/(G + R)$	
RVI	NIR/R	
NDVI	$(NIR - R)/(NIR + R)$	
EVI	$G^*(NIR - R)/(NIR + C_1 * R - C_2 * B + L)$	
WI	$(G - B)/(R - G)$	

Where NIR , R , G and B are atmospherically corrected surface reflectance. C_1 , C_2 are the coefficients of the aerosol resistance term. G^* is the gain factor for EVI. Coefficients adopted for sentinel-2 imagery are: $G^* = 2.5$, $C_1 = 6$, $C_2 = 7.5$ and $L = 1$. Most of these are the indicators of greenness of the vegetation crop that can be compared with the biophysical variables of crop. In [67, 233], similar VIs were derived and used for the estimation of leaf area content (LAI), chlorophyll content and AGB. These VIs have the ability to describe the spectral variability of vegetation canopies in space. RVI is one of the strong indicator that can provide spectral variability. Sentinel-2 imagery were acquired (from late May till end of September) in order to analyze the crop pattern in accordance with the maize crop phenological period mentioned in Figure. 5.11. The Statistical values reported in Table. 5.9 were derived by only those sentinel-2 pixels which were spatially and temporally overlapped with the ground physical measurements. It was noted that standard deviation of GNDVI and GLA were quite low as compared to other VIs. EVI and RVI were found to be good descriptors of spatial variability in the study fields.

5.2.5 Field Measurements

Considering the maize crop cycle shown in Figure. 5.11, four field measurements were conducted on different dates in the selected 43 field plots mentioned in Figure. 5.9. These dates are subjectively chosen in accordance with the satellite imagery acquisition dates in order to have reliable comparison for the similar environmental and crop growth conditions. Each field plot has a large area, almost flat terrain, and under uniform growing conditions. The geographic position (latitude and longitude) of all the plots were recorded using a hand-held GPS. In each plot, the heights of four representative maize plants were measured using a tape. Canopy hieght (H_{canopy})

for each plot was calculated as the average height (m) of the measured maize plants. Plant density ($D, plant/m^2$) and AGB was determined according to conventional method [118]. Statistical measure derived from physical measurements taken on stage T3 (tasseling and flowering) are reported in Table. 5.10.

Growth Stages	Emergence	Vegetative				Productive		Maturity
Maize development								
	Stage 0	Stage 1&2	Stage 3&4	Stage 5	Stage 6	Stage 7		Stage 8
	Germination	Leaf Development	Stem Elongation	Heading	Flowering, enthesis	Tasseling	Fruit development	Ripening
Satellite imagery Acquisition	26 th May	6 th June	29 th June	17 th July	2 nd August	12 th August	4 th September	20 th September
Ground Measurements	-	4 th June (T1)	-	18 th July (T2)	-	11 th August (T3)	-	17 th September (T4)

Figure 5.11: Maize growth development stages with satellite data acquisition dates along with field crop height and biomass measurements.

Table 5.9: Basic statistics of the plot-wise spectral vegetation indices (VIs)

Spectral vegetation index	Min	Max	Mean	SD
NDVI	0.292	0.75	0.544	0.124
RVI	3.02	6.771	5.103	0.872
EVI	0.284	0.873	0.611	0.163
GNDVI	0.053	0.184	0.123	0.034
WI	-0.254	0.391	0.134	0.151
GLA	0.014	0.089	0.051	0.019

5.2.6 Results Assessment

In this section, relative comparison between spectral VIs and maize biophysical parameters are analyzed. Pearson correlation between all the VIs mentioned in Table. 5.11 and maize bio- physical measurements (H_{canopy} and AGB) are calculated for the four different growth stages from T1-T4 shown in Figure. 5.11. It was noted that the correlation values change significantly with time. Results showed that, at growth stage T3 (Tasseling and flowering), high correlation were found between physically measured maize parameters (H_{canopy} and AGB) and VIs derived from

Table 5.10: Basic statistics of field measurements.

	Hcanopy (m)	D (plant/m ²)	AGB (kg/m ²)
Min	1.58	4.35	0.79
Max	3.64	9.7	2.63
Mean	2.46	6.01	1.65
SD	0.47	1.27	0.48

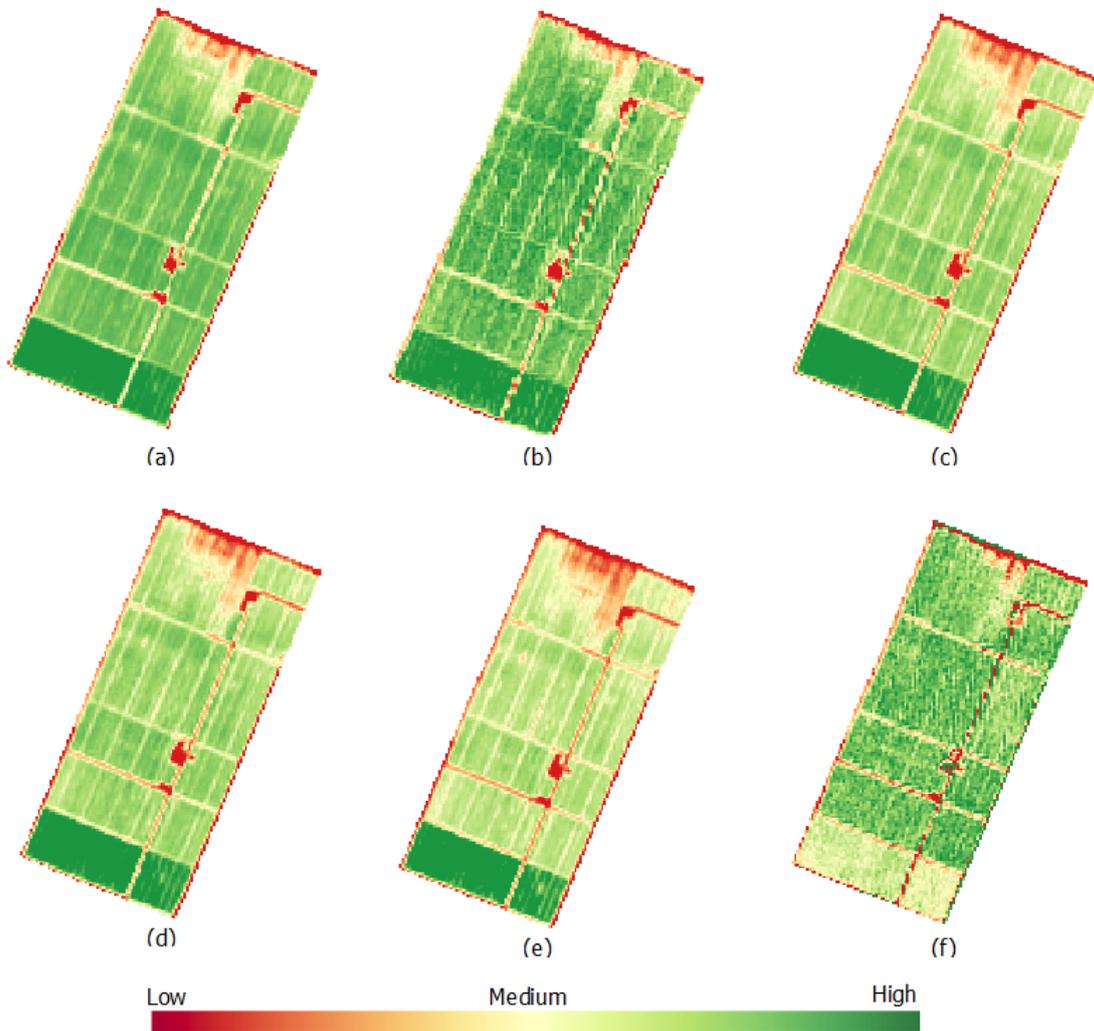


Figure 5.12: False color representation of Vegetation indices (a) GLA, (b) GNDVI, (c) RVI, (d) NDVI, (e) EVI, (f) WI derived from sentinel-2 imagery for the study area.

satellite imagery as compared to other stages.

Nevertheless, spectral VIs showed less correlation to *AGB* as compared to *Hcanopy*.

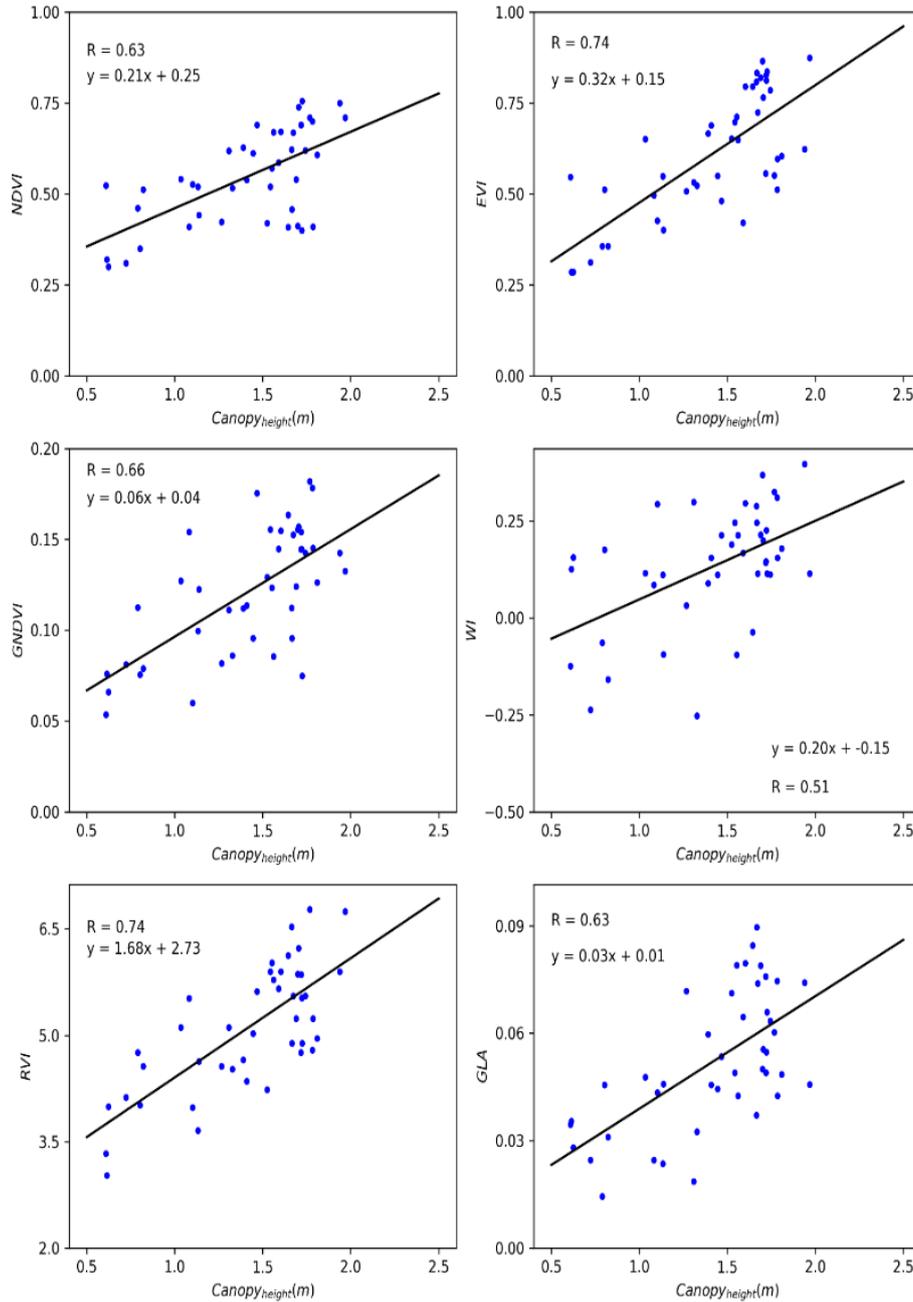


Figure 5.13: Relationship between field-measured canopy height $H_{canopy}(m)$ and VIs (NDVI, GNDVI, RVI, EVI, WI and GLA) derived from satellite imagery acquired on maize growth stage T3.

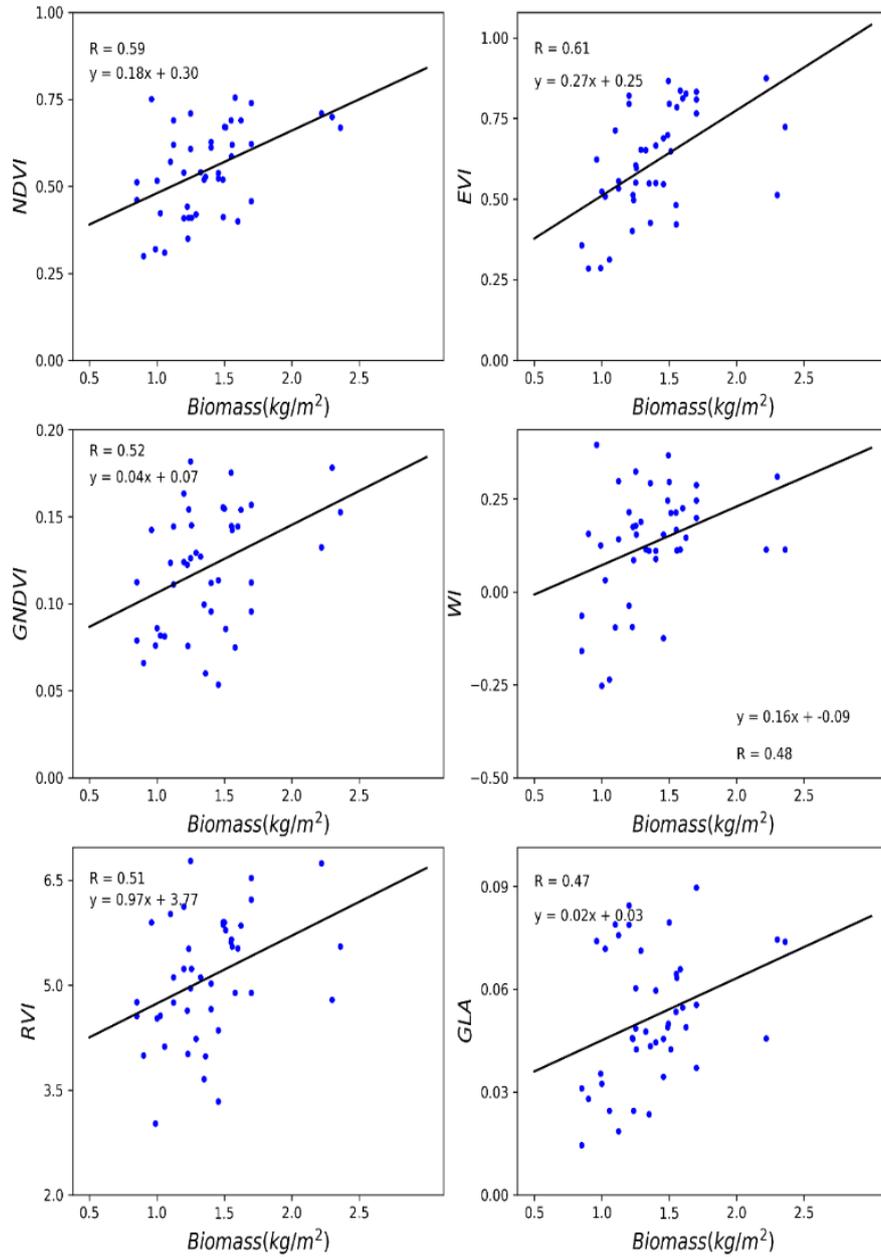


Figure 5.14: Relationship between field-measured biomass (AGB kg/m²) and VIs (NDVI, GNDVI, RVI, EVI, WI and GLA) derived from satellite imagery acquired on maize growth stage T3 .

VIs such as RVI and EVI were found better descriptor for the Hcanopy with maximum correlation of 0.74, which shows that, VIs derived from high resolution imagery such as sentinel-2, can be used to estimate maize crop height over the large crop fields.

Moreover, scatter plots along with linear regression fitting lines for VIs and biophysical parameters (H_{canopy} and AGB) are shown in Figure. 5.13 and Figure. 5.14 respectively, which revealed that EVI and RVI are more sensitive to H_{canopy} than the AGB .

Table 5.11: Pearson correlation between vegetation indices derived and height and biomass field measurements for four different maize growth stages.

Pearson Correlation	T1	T2	T3	T4
$R_{NDVI-Hcanopy}$	0.55	0.58	0.63	0.57
$R_{NDVI-AGB}$	0.58	0.59	0.59	0.51
$R_{GNDVI-Hcanopy}$	0.34	0.42	0.66	0.43
$R_{GNDVI-AGB}$	0.57	0.63	0.52	0.63
$R_{RVI-Hcanopy}$	0.61	0.64	0.74	0.65
$R_{RVI-AGB}$	0.59	0.55	0.51	0.58
$R_{EVI-Hcanopy}$	0.53	0.61	0.74	0.66
$R_{EVI-AGB}$	0.57	0.62	0.61	0.55
$R_{WI-Hcanopy}$	0.36	0.41	0.51	0.38
R_{WI-AGB}	0.47	0.59	0.48	0.51
$R_{GLA-Hcanopy}$	0.36	0.52	0.63	0.51
$R_{GLA-AGB}$	0.47	0.47	0.47	0.5

Chapter 6

Overall Conclusions and Future Work

Starting with the aim to take full advantage of advancements in the remote sensing technology in terms of spectral, spatial and temporal resolution of remote sensing data and to develop methodologies to address agricultural applications for improving the efficacy of practices used in farming system. This thesis presents a collection of work done as a part of doctoral studies in remote sensing using time series data analysis for several agricultural applications. To be more specific, this study is comprised of several research tasks and are divided in to two parts. First part is devoted for the tasks related to precision viticulture. In first task, a detailed analysis and comparison of multispectral imagery of vineyards, provided by decametric resolution satellite and low altitude UAV platforms, is presented. The effectiveness of the Sentinel-2 imagery and the high-resolution UAV aerial images was evaluated by considering the well-known relation between the NDVI and vineyard vigour. The results show how, in the case of crops where the inter-row surfaces involve a relevant portion of the cropland, such as vineyards, the radiometric information acquired by decametric resolution satellite platforms has difficulties in properly evaluating crop status and variability. In these situations, the vigour of the vineyard could often be in discord with the inter-row areas (e.g., grass, plants for pest and disease integrated control, or soil), leading to biased vineyard vigour assessments from decametric resolution imagery, such as the Sentinel-2 imagery. This was proved by a detailed analysis of the radiometric unbundled contribution of different elements within the cropland, performed by defining three different NDVI indices from the high-resolution UAV imagery, considering: (i) the whole cropland surface; (ii) only the crop canopy pixels; and (iii) only the inter-row terrain pixels. In this study, the NDVI maps derived from the satellite imagery were found not to be in accordance with the in-field crop vigour assessment as confirmed by ANOVA test and the obtained Pearson correlations coefficients values which ranges from 0.29 to 0.55 for three considered parcels. In addition, the satellite-based NDVI maps

were found to be more related to the NDVI maps computed by the high-resolution UAV imagery, considering only the pixels representing inter-row surfaces. As a validation, a new type of NDVI map from the UAV imagery, generated by considering only the pixels representing the vine canopies, was defined. The effectiveness of this last type of map in describing the observed vineyard vigour was found to be relevant which is backed by ANOVA tests and the obtained Pearson correlation coefficients ranges from 0.58 to 0.64 for three vineyard parcels.

Another research problem discussed in this thesis is an automatic coverage path planning for UGV for hilly vineyard environment. UAV imagery is used to obtain a Digital Terrain Model (DTM) and Digital Surface Model (DSM), which are then used in the mask generation to extract the vine rows. The experimental results show that the work as a whole presents significant contribution in coverage path planning for UGV in the challenging environment like hilly vineyards that can be useful for the farmers to manage agricultural tasks.

As concluded earlier in the work of variability assessment of vineyard that NDVI maps derived from sentinel-2 imagery were not found to be coherent with the reference data. So another research aim was to make use of the freely available satellite data for vineyard vegetation assessment. To achieve this, refinement of satellite-based NDVI maps is performed for vineyard assessment using high resolution UAV imagery. In this work, a convolutional neural network (CNN) based approach was used to refine the moderate satellite images by using more detailed vegetation maps derived from centimetric UAV imagery. Refined NDVI map and ground truth were found more coherent raw sentinel pixels which was confirmed by the Pearson correlation coefficient $R = 0.81$. The synergy formed between satellite and UAV can be more exploited to take full advantages of both platforms and the same approach can be further extended to other crop types with the similar row formation.

In the second part of the thesis, multi temporal remote sensing data analysis was performed only considering sentinel-2 satellite platform . A novel and optimal deep learning framework for pixel-based Land Cover and Crop Classification is developed and implemented based on Recurrent Neural Networks (RNN) in combination with Convolutional Neural Networks (CNN) using multi temporal sentinel-2 imagery of central north part of Italy, which has diverse agricultural system dominated by economic crop types. Proposed framework is capable of automated features extraction by learning time correlation of multiple images, which reduces manual feature engineering and modelling crops phenological stages. Fifteen classes, including major agricultural crops were considered in this study. Other widely used traditional machine learning algorithms such as support vector machine SVM, random forest (RF), Kernel SVM and gradient boosting machine also called XGBoost were also tested for comparison. The overall accuracy (OAA) achieved by our proposed Pixel R-CNN is 96.5% which shows considerable improvements from the other mainstream methods. The future direction of this research could be the testing of same framework on the other regions. There is still need of creating a global benchmark

data set to test the performance of different classifiers at equal footings.

Estimating biophysical parameters of crops is the very essential for managing agricultural practices. To this end, spectral vegetation indices were derived from high resolution satellite (sentinel-2) imagery and compared with the field measured biophysical pair of variables maize Canopy Height (Hcanopy) and Above Ground Biomass (AGB). Derived vegetation indices were found well correlated with the vertical structure such as height during the particular growth period (tasseling and flowering stage) of maize. It was also noted that considerable correlation was found between AGB and EVI but it was not consistent for all the crop growth stages. The results showed that VIs such as EVI and RVI can be used to estimate Hcanopy. This work also offers some suggestion on how to make use of the freely available sentinel-2 satellite imagery to estimate the crop physical structure.

Considering the addressed remote sensing application in the field of precision agriculture, there are some needs and suggestions for future research work. For example, past archives of satellite-based remote sensing data at moderate to high spatial resolution and ordinary spectral resolution should be combined with real-time remote sensing data at high spatial and spectral resolution in order to improve the decision-making process in precision agriculture. Similarly, more spectral information should continue to be exploited to develop new spectral vegetation indices that concurrently provide an accurate assessment of multiple crop features such as (LAI, biomass, etc.) and stresses (e.g. water and N; weeds and insects, etc.)

Bibliography

- [1] Accenture. *Evolution of digital agriculture*. :<https://www.accenture.com/au-en/insight-accenture-digital-agriculture-solutions>. (Accessed on 24/12/2019).
- [2] Agisoft. *PhotoScan Software*. :<https://www.agisoft.com/>. (Accessed on 24/11/2018).
- [3] Irene Aicardi et al. “Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach”. In: *Journal of Cultural Heritage* (2018).
- [4] Johanna Albetis et al. “Detection of Flavescence dorée grapevine disease using unmanned aerial vehicle (UAV) multispectral imagery”. In: *Remote Sensing* 9.4 (2017), p. 308.
- [5] Md Zahangir Alom et al. “The history began from alexnet: A comprehensive survey on deep learning approaches”. In: *arXiv preprint arXiv:1803.01164* (2018).
- [6] J Angerer et al. “Use of satellite-derived data to improve biophysical model output: An example from Southern Kenya”. In: (2001).
- [7] Charles E Antoniak. “Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems”. In: *The annals of statistics* (1974), pp. 1152–1174.
- [8] Jaume Arnó Satorra et al. “Precision viticulture. Research topics, challenges and opportunities in site-specific vineyard management”. In: *Spanish Journal of Agricultural Research*, 2009, vol. 7, núm. 4, p. 779-790 (2009).
- [9] Damien Arvor et al. “Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil”. In: *International Journal of Remote Sensing* 32.22 (2011), pp. 7847–7871.
- [10] Nicolas Audebert, Bertrand Le Saux, and Sébastien Lefèvre. “Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 140 (2018), pp. 20–32.

- [11] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. “Neural machine translation by jointly learning to align and translate”. In: *arXiv preprint arXiv:1409.0473* (2014).
- [12] Pedro Ballester and Ricardo Matsumura Araujo. “On the performance of GoogLeNet and AlexNet applied to sketches”. In: *Thirtieth AAAI Conference on Artificial Intelligence*. 2016.
- [13] A Bannari et al. “A review of vegetation indices”. In: *Remote sensing reviews* 13.1-2 (1995), pp. 95–120.
- [14] M Pir Bavaghar. “Deforestation modelling using logistic regression and GIS”. In: *Journal of Forest Science* 61.5 (2015), pp. 193–199.
- [15] Inbal Becker-Reshef et al. “A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data”. In: *Remote Sensing of Environment* 114.6 (2010), pp. 1312–1323.
- [16] Mariana Belgiu and Lucian Drăguț. “Random forest in remote sensing: A review of applications and future directions”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 114 (2016), pp. 24–31.
- [17] Thomas Blaschke. “Object based image analysis for remote sensing”. In: *ISPRS journal of photogrammetry and remote sensing* 65.1 (2010), pp. 2–16.
- [18] E Borgogno-Mondino et al. “A comparison between multispectral aerial and satellite imagery in precision viticulture”. In: *Precision agriculture* 19.2 (2018), pp. 195–217.
- [19] Claire Boryan et al. “Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program”. In: *Geocarto International* 26.5 (2011), pp. 341–358.
- [20] RGV Bramley and RP Hamilton. “Understanding variability in winegrape production systems: 1. Within vineyard variation in yield over several vintages”. In: *Australian Journal of Grape and Wine Research* 10.1 (2004), pp. 32–45.
- [21] *Calibrated Reflectance Panel (Micasense)*. : <https://www.micasense.com/accessories/#!/Calibrated-Reflectance-Panel>. (Accessed on 13/11/2018).
- [22] JB Campbell. *Introduction to Remote Sensing, S. 6*. 2002.
- [23] Gustau Camps-Valls and Lorenzo Bruzzone. *Kernel methods for remote sensing data analysis*. John Wiley & Sons, 2009.
- [24] Maureen Caudill. “Neural networks primer, part VIII”. In: *AI Expert* 4.8 (1989), pp. 61–67.

- [25] Jing M Chen and Josef Cihlar. “Retrieving leaf area index of boreal conifer forests using Landsat TM images”. In: *Remote sensing of Environment* 55.2 (1996), pp. 153–162.
- [26] Jing M Chen et al. “Leaf area index of boreal forests: Theory, techniques, and measurements”. In: *Journal of Geophysical Research: Atmospheres* 102.D24 (1997), pp. 29429–29443.
- [27] Liang-Chieh Chen et al. “Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs”. In: *IEEE transactions on pattern analysis and machine intelligence* 40.4 (2017), pp. 834–848.
- [28] Gong Cheng et al. “Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.8 (2015), pp. 4238–4249.
- [29] Gong Cheng et al. “Multi-class geospatial object detection and geographic image classification based on collection of part detectors”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 98 (2014), pp. 119–132.
- [30] Anil M Cheriyyadat. “Unsupervised feature learning for aerial scene classification”. In: *IEEE Transactions on Geoscience and Remote Sensing* 52.1 (2013), pp. 439–451.
- [31] Jan K Chorowski et al. “Attention-based models for speech recognition”. In: *Advances in neural information processing systems*. 2015, pp. 577–585.
- [32] Junyoung Chung et al. “Empirical evaluation of gated recurrent neural networks on sequence modeling”. In: *arXiv preprint arXiv:1412.3555* (2014).
- [33] Junyoung Chung et al. “Gated feedback recurrent neural networks”. In: *International Conference on Machine Learning*. 2015, pp. 2067–2075.
- [34] Alberto Cina et al. “Network Real Time Kinematic (NRTK) Positioning–Description, Architectures and Performances”. In: *Satellite Positioning-Methods, Models and Applications*. InTech, 2015.
- [35] EA Cloutis et al. “Airborne multi-spectral monitoring of agricultural crop status: effect of time of year, crop type and crop condition parameter”. In: *Remote Sensing* 17.13 (1996), pp. 2579–2601.
- [36] Ronan Collobert and Jason Weston. “A unified architecture for natural language processing: Deep neural networks with multitask learning”. In: *Proceedings of the 25th international conference on Machine learning*. ACM. 2008, pp. 160–167.
- [37] Lorenzo Comba, Paolo Gay, and Davide Ricauda Aimonino. “Robot ensembles for grafting herbaceous crops”. In: *Biosystems engineering* 146 (2016), pp. 227–239.

- [38] Lorenzo Comba et al. “Unsupervised detection of vineyards by 3D point-cloud UAV photogrammetry for precision agriculture”. In: *Computers and Electronics in Agriculture* 155 (2018), pp. 84–95.
- [39] Lorenzo Comba et al. “Vineyard detection from unmanned aerial systems images”. In: *computers and Electronics in Agriculture* 114 (2015), pp. 78–87.
- [40] Russell G Congalton. “A review of assessing the accuracy of classifications of remotely sensed data”. In: *Remote sensing of environment* 37.1 (1991), pp. 35–46.
- [41] Jerome T Connor, R Douglas Martin, and Les E Atlas. “Recurrent neural networks and robust time series prediction”. In: *IEEE transactions on neural networks* 5.2 (1994), pp. 240–254.
- [42] Christopher Conrad et al. “Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of Uzbekistan using multi-temporal RapidEye images”. In: *Computers and Electronics in Agriculture* 103 (2014), pp. 63–74.
- [43] Brett A Cruden, Dinesh Prabhu, and Ramon Martinez. “Absolute radiation measurement in venus and mars entry conditions”. In: *Journal of Spacecraft and Rockets* 49.6 (2012), pp. 1069–1079.
- [44] TW Cui et al. “Assessing and refining the satellite-derived massive green macro-algal coverage in the Yellow Sea with high resolution images”. In: *ISPRS journal of photogrammetry and remote sensing* 144 (2018), pp. 315–324.
- [45] Jean Pierre Da Costa et al. “Delineation of vine parcels by segmentation of high resolution remote sensed images”. In: *Precision Agriculture* 8.1-2 (2007), pp. 95–110.
- [46] N Dalal and Bill Triggs. “Histograms of Oriented Gradients for Human Detection [C] IEEE Computer Society Conference on Computer Vision and Pattern Recognition”. In: *Washington: IEEE, 2005* 886 (2005), p. 893.
- [47] Matthew P Dannenberg et al. “Empirical evidence of El Niño–Southern Oscillation influence on land surface phenology and productivity in the western United States”. In: *Remote Sensing of Environment* 159 (2015), pp. 167–180.
- [48] Jia Deng et al. “Imagenet: A large-scale hierarchical image database”. In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee. 2009, pp. 248–255.
- [49] SZ Dobrowski, Susan L Ustin, and JA Wolpert. “Remote estimation of vine canopy density in vertically shoot-positioned vineyards: determining optimal vegetation indices”. In: *Australian Journal of Grape and Wine Research* 8.2 (2002), pp. 117–125.

- [50] Juan Du. “Understanding of Object Detection Based on CNN Family and YOLO”. In: *Journal of Physics: Conference Series*. Vol. 1004. 1. IOP Publishing. 2018, p. 012029.
- [51] Jorge R Ducati, Magno G Bombassaro, and Jandyra MG Fachel. “Classifying vineyards from satellite images: A case study on burgundy’s côte d’or”. In: *Oeno One* 48.4 (2014), pp. 247–260.
- [52] Jane Elith, John R Leathwick, and Trevor Hastie. “A working guide to boosted regression trees”. In: *Journal of Animal Ecology* 77.4 (2008), pp. 802–813.
- [53] Manuel Fernández-Delgado et al. “Do we need hundreds of classifiers to solve real world classification problems?” In: *The Journal of Machine Learning Research* 15.1 (2014), pp. 3133–3181.
- [54] FAO Food. “Agriculture Organization of the United Nations.(2017). The future of food and agriculture”. In: *Trends and challenges*. Rome ().
- [55] Sigfredo Fuentes et al. “Automated estimation of leaf area index from grapevine canopies using cover photography, video and computational analysis methods”. In: *Australian Journal of Grape and Wine Research* 20.3 (2014), pp. 465–473.
- [56] Chris Funk and Michael E Budde. “Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe”. In: *Remote Sensing of Environment* 113.1 (2009), pp. 115–125.
- [57] Enric Galceran and Marc Carreras. “A survey on coverage path planning for robotics”. In: *Robotics and Autonomous systems* 61.12 (2013), pp. 1258–1276.
- [58] Gillian L Galford et al. “Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil”. In: *Remote sensing of environment* 112.2 (2008), pp. 576–587.
- [59] J Gallego et al. “Best practices for crop area estimation with remote sensing”. In: *Ispira: Joint Research Center* (2008).
- [60] John A Gamon et al. “Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types”. In: *Ecological Applications* 5.1 (1995), pp. 28–41.
- [61] AN Ganeshamurthy et al. “Biomass distribution and development of allometric equations for non-destructive estimation of carbon sequestration in grafted mango trees”. In: *Journal of Agricultural Science* 8.8 (2016), p. 201.
- [62] Qishuo Gao, Samsung Lim, and Xiuping Jia. “Hyperspectral image classification using convolutional neural networks and multiple feature learning”. In: *Remote Sensing* 10.2 (2018), p. 299.

- [63] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. “Learning to forget: Continual prediction with LSTM”. In: (1999).
- [64] Ross Girshick. “Fast r-cnn”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 1440–1448.
- [65] Pall Oskar Gislason, Jon Atli Benediktsson, and Johannes R Sveinsson. “Random forests for land cover classification”. In: *Pattern Recognition Letters* 27.4 (2006), pp. 294–300.
- [66] Cristina Gomez, Joanne C White, and Michael A Wulder. “Optical remotely sensed time series data for land cover classification: A review”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 116 (2016), pp. 55–72.
- [67] Alemu Gonsamo. “Leaf area index retrieval using gap fractions obtained from high resolution satellite data: Comparisons of approaches, scales and atmospheric effects”. In: *International Journal of Applied Earth Observation and Geoinformation* 12.4 (2010), pp. 233–248.
- [68] Laura Gow. “A Land Surface Temperature Model-data Differencing Approach to Quantifying Subsurface Water Used by Vegetation: Application in the Condamine Region, South-eastern Queensland”. PhD thesis. University of Queensland, School of Geography, Planning and Environmental ..., 2017.
- [69] J Grace et al. “Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence?” In: *Global Change Biology* 13.7 (2007), pp. 1484–1497.
- [70] E Grilli, F Menna, and F Remondino. “A review of point clouds segmentation and classification algorithms”. In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 42 (2017), p. 339.
- [71] Daniel Guidici and Matthew L Clark. “One-Dimensional convolutional neural network land-cover classification of multi-seasonal hyperspectral imagery in the San Francisco Bay Area, California”. In: *Remote Sensing* 9.6 (2017), p. 629.
- [72] Driss Haboudane et al. “Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture”. In: *Remote sensing of environment* 81.2-3 (2002), pp. 416–426.
- [73] Andrew Hall, John Louis, and David Lamb. “A method for extracting detailed information from high resolution multispectral images of vineyards”. In: *Proceedings of the 6th International Conference on Geocomputation*. Cite-seer. 2001, pp. 24–26.

- [74] Andrew Hall et al. “Optical remote sensing applications in viticulture—a review”. In: *Australian journal of grape and wine research* 8.1 (2002), pp. 36–47.
- [75] Ibrahim A Hameed, Dionysis Bochtis, and Claus A Sørensen. “An optimized field coverage planning approach for navigation of agricultural robots in fields involving obstacle areas”. In: *International Journal of Advanced Robotic Systems* 10.5 (2013), p. 231.
- [76] Pengyu Hao et al. “Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA”. In: *Remote Sensing* 7.5 (2015), pp. 5347–5369.
- [77] Peter E Hart, Nils J Nilsson, and Bertram Raphael. “A formal basis for the heuristic determination of minimum cost paths”. In: *IEEE transactions on Systems Science and Cybernetics* 4.2 (1968), pp. 100–107.
- [78] John A Hartigan and Manchek A Wong. “Algorithm AS 136: A k-means clustering algorithm”. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1 (1979), pp. 100–108.
- [79] Steve Harwin, Arko Lucieer, and Jon Osborn. “The impact of the calibration method on the accuracy of point clouds derived using unmanned aerial vehicle multi-view stereopsis”. In: *Remote Sensing* 7.9 (2015), pp. 11933–11953.
- [80] Trevor Hastie, Robert Tibshirani, and Martin Wainwright. *Statistical learning with sparsity: the lasso and generalizations*. Chapman and Hall/CRC, 2015.
- [81] Kaiming He et al. “Deep residual learning for image recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.
- [82] Mingzhu He et al. “Regional crop gross primary productivity and yield estimation using fused landsat-MODIS data”. In: *Remote Sensing* 10.3 (2018), p. 372.
- [83] Geoffrey E Hinton and Ruslan R Salakhutdinov. “Reducing the dimensionality of data with neural networks”. In: *science* 313.5786 (2006), pp. 504–507.
- [84] Jacinta Holloway and Kerrie Mengersen. “Statistical machine learning methods and remote sensing for sustainable development goals: A review”. In: *Remote Sensing* 10.9 (2018), p. 1365.
- [85] Fan Hu et al. “Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery”. In: *Remote Sensing* 7.11 (2015), pp. 14680–14707.

- [86] Wei Hu et al. “Deep convolutional neural networks for hyperspectral image classification”. In: *Journal of Sensors* 2015 (2015).
- [87] David H Hubel and Torsten N Wiesel. “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex”. In: *The Journal of physiology* 160.1 (1962), pp. 106–154.
- [88] Alfredo Huete. “Huete, AR A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment”. In: *Remote sensing of environment* 25 (1988), pp. 295–309.
- [89] E Raymond Hunt et al. “Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring”. In: *Remote Sensing* 2.1 (2010), pp. 290–305.
- [90] Tomohiro Ishii et al. “Detection by classification of buildings in multispectral satellite imagery”. In: *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE. 2016, pp. 3344–3349.
- [91] Sylvain Jay et al. “Exploiting the centimeter resolution of UAV multispectral imagery to improve remote-sensing estimates of canopy structure and biochemistry in sugar beet crops”. In: *Remote Sensing of Environment* 231 (2019), p. 110898.
- [92] Huaan Jin et al. “Improvement of spatially and temporally continuous crop leaf area index by integration of CERES-Maize model and MODIS data”. In: *European journal of agronomy* 78 (2016), pp. 1–12.
- [93] Xiuliang Jin et al. “Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery”. In: *Remote Sensing of Environment* 198 (2017), pp. 105–114.
- [94] Zhenong Jin et al. “Crop model-and satellite imagery-based recommendation tool for variable rate N fertilizer application for the US Corn system”. In: *Precision agriculture* 18.5 (2017), pp. 779–800.
- [95] David M Johnson. “A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products”. In: *International Journal of Applied Earth Observation and Geoinformation* 52 (2016), pp. 65–81.
- [96] Lee F Johnson. “Temporal stability of an NDVI-LAI relationship in a Napa Valley vineyard”. In: *Australian Journal of Grape and Wine Research* 9.2 (2003), pp. 96–101.
- [97] LF Johnson et al. “Mapping vineyard leaf area with multispectral satellite imagery”. In: *Computers and electronics in agriculture* 38.1 (2003), pp. 33–44.

- [98] LF Johnson et al. “Remote sensing of vineyard management zones: Implications for wine quality”. In: *Applied Engineering in Agriculture* 17.4 (2001), p. 557.
- [99] Carl F Jordan. “Derivation of leaf-area index from quality of light on the forest floor”. In: *Ecology* 50.4 (1969), pp. 663–666.
- [100] Olav Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 2006.
- [101] Michael Kampffmeyer, Arnt-Borre Salberg, and Robert Jenssen. “Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2016, pp. 1–9.
- [102] Arnon Karnieli et al. “Use of NDVI and land surface temperature for drought assessment: Merits and limitations”. In: *Journal of climate* 23.3 (2010), pp. 618–633.
- [103] Yoram J Kaufman and Claudia Sendra. “Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery”. In: *International Journal of Remote Sensing* 9.8 (1988), pp. 1357–1381.
- [104] Aleem Khaliq, Vittorio Mazzia, and Marcello Chiaberge. “Refining satellite imagery by using UAV imagery for vineyard environment: A CNN Based approach”. In: *2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)*. IEEE. 2019, pp. 25–29.
- [105] Aleem Khaliq, Maria Angela Musci, and Marcello Chiaberge. “Analyzing relationship between maize height and spectral indices derived from remotely sensed multispectral imagery”. In: *2018 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. IEEE. 2018, pp. 1–5.
- [106] Aleem Khaliq, Leonardo Peroni, and Marcello Chiaberge. “Land cover and crop classification using multitemporal Sentinel-2 images based on crops phenological cycle”. In: *2018 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*. IEEE. 2018, pp. 1–5.
- [107] Aleem Khaliq et al. “Comparison of Satellite and UAV-Based Multispectral Imagery for Vineyard Variability Assessment”. In: *Remote Sensing* 11.4 (2019), p. 436.
- [108] Reza Khatami, Giorgos Mountrakis, and Stephen V Stehman. “A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research”. In: *Remote Sensing of Environment* 177 (2016), pp. 89–100.
- [109] D Kingma and J Ba. “Adam: a method for stochastic optimization (2014). arXiv preprint”. In: *arXiv preprint arXiv:1412.6980* ().

- [110] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “Imagenet classification with deep convolutional neural networks”. In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.
- [111] Nataliia Kussul et al. “Deep learning classification of land cover and crop types using remote sensing data”. In: *IEEE Geoscience and Remote Sensing Letters* 14.5 (2017), pp. 778–782.
- [112] Nataliia Kussul et al. “Parcel-based crop classification in ukraine using landsat-8 data and sentinel-1A data”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9.6 (2016), pp. 2500–2508.
- [113] *Land Use and Coverage Area frame Survey (LUCAS) Details*. <https://ec.europa.eu/eurostat/web/lucas>. Accessed: 2019-01-14.
- [114] S Lanjeri, J Melia, and D Segarra. “A multi-temporal masking classification method for vineyard monitoring in central Spain”. In: *International Journal of Remote Sensing* 22.16 (2001), pp. 3167–3186.
- [115] Yann LeCun et al. “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324.
- [116] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *nature* 521.7553 (2015), pp. 436–444.
- [117] Qingting Li et al. “Object-based crop classification with Landsat-MODIS enhanced time-series data”. In: *Remote Sensing* 7.12 (2015), pp. 16091–16107.
- [118] Wang Li et al. “Airborne LiDAR technique for estimating biomass components of maize: A case study in Zhangye City, Northwest China”. In: *Ecological indicators* 57 (2015), pp. 486–496.
- [119] Ying Li, Haokui Zhang, and Qiang Shen. “Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network”. In: *Remote Sensing* 9.1 (2017), p. 67.
- [120] Shunlin Liang. *Quantitative remote sensing of land surfaces*. Vol. 30. John Wiley & Sons, 2005.
- [121] Muhammad Usman Liaqat et al. “Evaluation of MODIS and Landsat multi-band vegetation indices used for wheat yield estimation in irrigated Indus Basin”. In: *Computers and electronics in agriculture* 138 (2017), pp. 39–47.
- [122] Bee Lim et al. “Enhanced deep residual networks for single image super-resolution”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017, pp. 136–144.
- [123] Tsung-Yi Lin et al. “Focal loss for dense object detection”. In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2980–2988.

- [124] Tsung-Yi Lin et al. “Microsoft coco: Common objects in context”. In: *European conference on computer vision*. Springer. 2014, pp. 740–755.
- [125] Qingshan Liu et al. “Learning multiscale deep features for high-resolution satellite image scene classification”. In: *IEEE Transactions on Geoscience and Remote Sensing* 56.1 (2017), pp. 117–126.
- [126] Xianming Liu et al. “Feedback neural network for weakly supervised geo-semantic segmentation”. In: *arXiv preprint arXiv:1612.02766* (2016).
- [127] John A Long et al. “Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and random forest”. In: *GIScience & Remote Sensing* 50.4 (2013), pp. 418–436.
- [128] Francisca López-Granados et al. “Early season weed mapping in sunflower using UAV technology: variability of herbicide treatment maps against weed thresholds”. In: *Precision Agriculture* 17.2 (2016), pp. 183–199.
- [129] Ilya Loshchilov and Frank Hutter. “Fixing weight decay regularization in adam”. In: *arXiv preprint arXiv:1711.05101* (2017).
- [130] Marine Louargant et al. “Unsupervised classification algorithm for early weed detection in row-crops by combining spatial and spectral information”. In: *Remote Sensing* 10.5 (2018), p. 761.
- [131] F Löw et al. “Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines”. In: *ISPRS journal of photogrammetry and remote sensing* 85 (2013), pp. 102–119.
- [132] Francois PS Luus et al. “Multiview deep learning for land-use classification”. In: *IEEE Geoscience and Remote Sensing Letters* 12.12 (2015), pp. 2448–2452.
- [133] Haobo Lyu et al. “Long-term annual mapping of four cities on different continents by applying a deep information learning method to landsat data”. In: *Remote Sensing* 10.3 (2018), p. 471.
- [134] Haobo Lyu, Hui Lu, and Lichao Mou. “Learning a transferable change rule from a recurrent neural network for land cover change detection”. In: *Remote Sensing* 8.6 (2016), p. 506.
- [135] Emmanuel Maggiori et al. “High-resolution aerial image labeling with convolutional neural networks”. In: *IEEE Transactions on Geoscience and Remote Sensing* 55.12 (2017), pp. 7092–7103.
- [136] Troy S Magney, Jan UH Eitel, and Lee A Vierling. “Mapping wheat nitrogen uptake from RapidEye vegetation indices”. In: *Precision Agriculture* 18.4 (2017), pp. 429–451.

- [137] Dimitrios Marmanis et al. “Deep learning earth observation classification using ImageNet pretrained networks”. In: *IEEE Geoscience and Remote Sensing Letters* 13.1 (2015), pp. 105–109.
- [138] Nicolas Matton et al. “An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series”. In: *Remote Sensing* 7.10 (2015), pp. 13208–13232.
- [139] Vittorio Mazzia, Aleem Khaliq, and Marcello Chiaberge. “Improvement in Land Cover and Crop Classification based on Temporal Features Learning from Sentinel-2 Data Using Recurrent-Convolutional Neural Network (R-CNN)”. In: *Applied Sciences* 10.1 (2020), p. 238.
- [140] Ronald E McRoberts et al. “Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery”. In: *Remote Sensing of Environment* 111.4 (2007), pp. 466–480.
- [141] Emmanuel Mellet et al. “Functional anatomy of spatial mental imagery generated from verbal instructions”. In: *Journal of Neuroscience* 16.20 (1996), pp. 6504–6512.
- [142] V Mnih et al. *Human-level control through deep reinforcement learning*. *Nature*. 2015.
- [143] Andrea Bordone Molini et al. “DeepSUM: Deep neural network for Super-resolution of Unregistered Multitemporal images”. In: *arXiv preprint arXiv:1907.06490* (2019).
- [144] István Monostori et al. “Relationship between SPAD value and grain yield can be affected by cultivar, environment and soil nitrogen content in wheat”. In: *Euphytica* 211.1 (2016), pp. 103–112.
- [145] FJ Montero et al. “Assessment of vine development according to available water resources by using remote sensing in La Mancha, Spain”. In: *Agricultural Water Management* 40.2-3 (1999), pp. 363–375.
- [146] Seyed-Mohsen Moosavi-Dezfooli et al. “Universal adversarial perturbations”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1765–1773.
- [147] M Susan Moran, Yoshio Inoue, and EM Barnes. “Opportunities and limitations for image-based remote sensing in precision crop management”. In: *Remote sensing of Environment* 61.3 (1997), pp. 319–346.
- [148] Lichao Mou, Pedram Ghamisi, and Xiao Xiang Zhu. “Unsupervised spectral-spatial feature learning via deep residual Conv-Deconv network for hyperspectral image classification”. In: *IEEE Transactions on Geoscience and Remote Sensing* 56.1 (2017), pp. 391–406.

- [149] Lichao Mou and Xiao Xiang Zhu. “RiFCN: Recurrent network in fully convolutional network for semantic segmentation of high resolution remote sensing images”. In: *arXiv preprint arXiv:1805.02091* (2018).
- [150] David J Mulla. “Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps”. In: *Biosystems engineering* 114.4 (2013), pp. 358–371.
- [151] Anh Nguyen, Jason Yosinski, and Jeff Clune. “Deep neural networks are easily fooled: High confidence predictions for unrecognizable images”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 427–436.
- [152] Antonio Novelli et al. “Performance evaluation of object based greenhouse detection from Sentinel-2 MSI and Landsat 8 OLI data: A case study from Almería (Spain)”. In: *International journal of applied earth observation and geoinformation* 52 (2016), pp. 403–411.
- [153] Timo Oksanen and Arto Visala. “Coverage path planning algorithms for agricultural field machines”. In: *Journal of Field Robotics* 26.8 (2009), pp. 651–668.
- [154] Aaron van den Oord, Nal Kalchbrenner, and Koray Kavukcuoglu. “Pixel recurrent neural networks”. In: *arXiv preprint arXiv:1601.06759* (2016).
- [155] John Richard Otukei and Thomas Blaschke. “Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms”. In: *International Journal of Applied Earth Observation and Geoinformation* 12 (2010), S27–S31.
- [156] Gonzalo Pajares. “Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs)”. In: *Photogrammetric Engineering & Remote Sensing* 81.4 (2015), pp. 281–330.
- [157] Federico Pallottino et al. “Science mapping approach to analyze the research evolution on precision agriculture: world, EU and Italian situation”. In: *Precision Agriculture* 19.6 (2018), pp. 1011–1026.
- [158] Omkar M Parkhi, Andrea Vedaldi, and Andrew Zisserman. “Deep face recognition”. In: (2015).
- [159] José M Peña-Barragán et al. “Object-based crop identification using multiple vegetation indices, textural features and crop phenology”. In: *Remote Sensing of Environment* 115.6 (2011), pp. 1301–1316.
- [160] José Manuel Peña et al. “Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images”. In: *PloS one* 8.10 (2013), e77151.

- [161] José Peña et al. “Object-based image classification of summer crops with machine learning methods”. In: *Remote Sensing* 6.6 (2014), pp. 5019–5041.
- [162] Otávio AB Penatti, Keiller Nogueira, and Jefersson A Dos Santos. “Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?” In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2015, pp. 44–51.
- [163] Maria Pérez-Ortiz et al. “A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method”. In: *Applied Soft Computing* 37 (2015), pp. 533–544.
- [164] Isabel Pôças et al. “Satellite-based evapotranspiration of a super-intensive olive orchard: Application of METRIC algorithms”. In: *Biosystems Engineering* 128 (2014), pp. 69–81.
- [165] Violeta Poenaru et al. “Multi-temporal multi-spectral and radar remote sensing for agricultural monitoring in the Braila Plain”. In: *Agriculture and Agricultural Science Procedia* 6 (2015), pp. 506–516.
- [166] Daniela Poli and Thierry Toutin. “Review of developments in geometric modelling for high resolution satellite pushbroom sensors”. In: *The Photogrammetric Record* 27.137 (2012), pp. 58–73.
- [167] Anup K Prasad et al. “Crop yield estimation model for Iowa using remote sensing and surface parameters”. In: *International Journal of Applied Earth Observation and Geoinformation* 8.1 (2006), pp. 26–33.
- [168] Jacopo Primicerio et al. “Individual plant definition and missing plant characterization in vineyards from high-resolution UAV imagery”. In: *European Journal of Remote Sensing* 50.1 (2017), pp. 179–186.
- [169] J Primicerio et al. “NDVI-based vigour maps production using automatic detection of vine rows in ultra-high resolution aerial images”. In: *Precision agriculture’15*. Wageningen Academic Publishers, 2015, pp. 693–712.
- [170] Z Quan, Z Xianfeng, and J Miao. “Eco-environment variable estimation from remote sensed data and eco-environment assessment: models and system”. In: *Acta Botanica Sinica* 47 (2011), pp. 1073–1080.
- [171] Mirco Ravanelli and Yoshua Bengio. “Speaker recognition from raw waveform with sincnet”. In: *2018 IEEE Spoken Language Technology Workshop (SLT)*. IEEE, 2018, pp. 1021–1028.
- [172] Sashank J Reddi, Satyen Kale, and Sanjiv Kumar. “On the convergence of adam and beyond”. In: *arXiv preprint arXiv:1904.09237* (2019).
- [173] D Sushma Reddy and P Rama Chandra Prasad. “Prediction of vegetation dynamics using NDVI time series data and LSTM”. In: *Modeling Earth Systems and Environment* 4.1 (2018), pp. 409–419.

- [174] Aliny Aparecida dos Reis et al. “Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods”. In: *New Zealand Journal of Forestry Science* 48.1 (2018), p. 1.
- [175] Felix Rembold et al. “Using low resolution satellite imagery for yield prediction and yield anomaly detection”. In: *Remote Sensing* 5.4 (2013), pp. 1704–1733.
- [176] Clara Rey-Caramés et al. “Using RPAS multi-spectral imagery to characterise vigour, leaf development, yield components and berry composition variability within a vineyard”. In: *Remote Sensing* 7.11 (2015), pp. 14458–14481.
- [177] José F Reyes, Christian Correa, and Javier Zuniga. “Reliability of different color spaces to estimate nitrogen SPAD values in maize”. In: *Computers and Electronics in Agriculture* 143 (2017), pp. 14–22.
- [178] Md Nasim Reza et al. “Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images”. In: *Biosystems engineering* 177 (2019), pp. 109–121.
- [179] Arthur J Richardson and CL Wiegand. “Distinguishing vegetation from soil background information”. In: *Photogrammetric engineering and remote sensing* 43.12 (1977), pp. 1541–1552.
- [180] R Richter, J Louis, and U Müller-Wilm. “Sentinel-2 MSI—level 2A products algorithm theoretical basis document”. In: *European Space Agency, (Special Publication) ESA SP* 49.0 (2012), pp. 1–72.
- [181] Rudolf Richter et al. “Correction of cirrus effects in Sentinel-2 type of imagery”. In: *International journal of remote sensing* 32.10 (2011), pp. 2931–2941.
- [182] Paola Rizzoli et al. “Relative height error analysis of TanDEM-X elevation data”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 73 (2012), pp. 30–38.
- [183] Nathaniel Robinson et al. “A dynamic Landsat derived normalized difference vegetation index (NDVI) product for the conterminous United States”. In: *Remote Sensing* 9.8 (2017), p. 863.
- [184] Adriana Romero, Carlo Gatta, and Gustau Camps-Valls. “Unsupervised deep feature extraction for remote sensing image classification”. In: *IEEE Transactions on Geoscience and Remote Sensing* 54.3 (2015), pp. 1349–1362.
- [185] O Ronneberger, P Fischer, and TU-net Brox. “Convolutional networks for biomedical image segmentation”. In: *Paper presented at: International Conference on Medical Image Computing and Computer-Assisted Intervention 2015*.

- [186] JW Rouse et al. “Monitoring vegetation systems in the Great Plains with ERTS”. In: *NASA special publication 351* (1974), p. 309.
- [187] Joshua D Rudd, Gary T Roberson, and John J Classen. “Application of satellite, unmanned aircraft system, and ground-based sensor data for precision agriculture: A review”. In: *2017 ASABE Annual International Meeting*. American Society of Agricultural and Biological Engineers. 2017, p. 1.
- [188] Marc Rußwurm and Marco Körner. “Multi-temporal land cover classification with sequential recurrent encoders”. In: *ISPRS International Journal of Geo-Information 7.4* (2018), p. 129.
- [189] Michael Schmidt et al. “A framework for large-area mapping of past and present cropping activity using seasonal Landsat images and time series metrics”. In: *Remote Sensing 8.4* (2016), p. 312.
- [190] Kathryn A Semmens et al. “Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach”. In: *Remote Sensing of Environment 185* (2016), pp. 155–170.
- [191] Cornelius Senf et al. “A Bayesian hierarchical model for estimating spatial and temporal variation in vegetation phenology from Landsat time series”. In: *Remote sensing of environment 194* (2017), pp. 155–160.
- [192] Cornelius Senf et al. “Mapping land cover in complex Mediterranean landscapes using Landsat: Improved classification accuracies from integrating multi-seasonal and synthetic imagery”. In: *Remote Sensing of Environment 156* (2015), pp. 527–536.
- [193] J Senthilnath et al. “Application of UAV imaging platform for vegetation analysis based on spectral-spatial methods”. In: *Computers and Electronics in Agriculture 140* (2017), pp. 8–24.
- [194] *Sentinel-2 User Handbook*. :https://earth.esa.int/documents/247904/685211/Sentinel-2_User_Handbook. (Accessed on 01/10/2018).
- [195] Yang Shao and Ross S Lunetta. “Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points”. In: *ISPRS Journal of Photogrammetry and Remote Sensing 70* (2012), pp. 78–87.
- [196] Yang Shao et al. “An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data”. In: *Remote Sensing of Environment 174* (2016), pp. 258–265.
- [197] Di Shi and Xiaojun Yang. “An assessment of algorithmic parameters affecting image classification accuracy by random forests”. In: *Photogrammetric Engineering & Remote Sensing 82.6* (2016), pp. 407–417.

- [198] Sofia Siachalou, Giorgos Mallinis, and Maria Tsakiri-Strati. “A hidden Markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data”. In: *Remote Sensing* 7.4 (2015), pp. 3633–3650.
- [199] Oriana Silvestroni et al. “Canopy management strategies to control yield and grape composition of Montepulciano grapevines”. In: *Australian journal of grape and wine research* 25.1 (2019), pp. 30–42.
- [200] Élizabeth Simms and Heather Ward. “Multisensor NDVI-based monitoring of the tundra-taiga interface (Mealy Mountains, Labrador, Canada)”. In: *Remote Sensing* 5.3 (2013), pp. 1066–1090.
- [201] Vincent Simonneaux et al. “The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco”. In: *International Journal of Remote Sensing* 29.1 (2008), pp. 95–116.
- [202] SistemaPiemonte. *Carta dei paesaggi agrari e forestali 1:250.000*. URL: <http://www.sistemapiemonte.it/eXoRisorse/dwd/servizi/Agricoltura/ServiziGeografici/note.pdf>.
- [203] Sergii Skakun et al. “Efficiency assessment of multitemporal C-band Radarsat-2 intensity and Landsat-8 surface reflectance satellite imagery for crop classification in Ukraine”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9.8 (2015), pp. 3712–3719.
- [204] Leslie N Smith. “Cyclical learning rates for training neural networks”. In: *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2017, pp. 464–472.
- [205] Jianqiang Song et al. “Pinot Noir wine composition from different vine vigour zones classified by remote imaging technology”. In: *Food chemistry* 153 (2014), pp. 52–59.
- [206] Ravi P Sripada et al. “Aerial color infrared photography for determining early in-season nitrogen requirements in corn”. In: *Agronomy Journal* 98.4 (2006), pp. 968–977.
- [207] Nitish Srivastava et al. “Dropout: a simple way to prevent neural networks from overfitting”. In: *The journal of machine learning research* 15.1 (2014), pp. 1929–1958.
- [208] Liang Sun et al. “Daily mapping of 30 m LAI and NDVI for grape yield prediction in California vineyards”. In: *Remote Sensing* 9.4 (2017), p. 317.
- [209] I Sutskever, O Vinyals, and QV Le. “Sequence to sequence learning with neural networks”. In: *Advances in NIPS* (2014).

- [210] Christian Szegedy et al. “Going deeper with convolutions”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9.
- [211] Christian Szegedy et al. “Inception-v4, inception-resnet and the impact of residual connections on learning”. In: *Thirty-first AAAI conference on artificial intelligence*. 2017.
- [212] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- [213] Brian W Szuster, Qi Chen, and Michael Borger. “A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones”. In: *Applied Geography* 31.2 (2011), pp. 525–532.
- [214] A Tagarakis et al. “Management zones delineation using fuzzy clustering techniques in grapevines”. In: *Precision Agriculture* 14.1 (2013), pp. 18–39.
- [215] Ying Tai, Jian Yang, and Xiaoming Liu. “Image super-resolution via deep recursive residual network”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 3147–3155.
- [216] Jianbin Tao et al. “A study of a Gaussian mixture model for urban land-cover mapping based on VHR remote sensing imagery”. In: *International Journal of Remote Sensing* 37.1 (2016), pp. 1–13.
- [217] Kenichi Tatsumi et al. “Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data”. In: *Computers and Electronics in Agriculture* 115 (2015), pp. 171–179.
- [218] Charles Toth and Grzegorz Józków. “Remote sensing platforms and sensors: A survey”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 115 (2016), pp. 22–36.
- [219] Johan Uddling et al. “Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings”. In: *Photosynthesis research* 91.1 (2007), pp. 37–46.
- [220] Babawuro Usman. “Satellite Imagery Land Cover Classification using K-Means Clustering Algorithm Computer Vision for Environmental Information Extraction”. In: *Elixir International Journal of Computer Science and Engineering* 63 (2013), pp. 18671–18675.
- [221] Gary E Varvel, James S Schepers, and Dennis D Francis. “Ability for in-season correction of nitrogen deficiency in corn using chlorophyll meters”. In: *Soil Science Society of America Journal* 61.4 (1997), pp. 1233–1239.
- [222] Jay M Ver Hoef and Hailemariam Temesgen. “A comparison of the spatial linear model to nearest neighbor (k-NN) methods for forestry applications”. In: *PloS one* 8.3 (2013), e59129.

- [223] C Vivalda et al. “Forest wildfire risk mapping and the influence of the weather and geo-morphological input data”. In: June 2017, pp. 171–171. DOI: [10.1201/9781315210469-145](https://doi.org/10.1201/9781315210469-145).
- [224] Francesco Vuolo et al. “How much does multi-temporal Sentinel-2 data improve crop type classification?” In: *International journal of applied earth observation and geoinformation* 72 (2018), pp. 122–130.
- [225] Kiri Wagstaff et al. “Constrained k-means clustering with background knowledge”. In: *ICML*. Vol. 1. 2001, pp. 577–584.
- [226] JJ Walker, KM De Beurs, and GM Henebry. “Land surface phenology along urban to rural gradients in the US Great Plains”. In: *Remote Sensing of Environment* 165 (2015), pp. 42–52.
- [227] JJ Walker, KM De Beurs, and RH Wynne. “Dryland vegetation phenology across an elevation gradient in Arizona, USA, investigated with fused MODIS and Landsat data”. In: *Remote Sensing of Environment* 144 (2014), pp. 85–97.
- [228] Xiaoqing Wan et al. “Stacked sparse autoencoder in hyperspectral data classification using spectral-spatial, higher order statistics and multifractal spectrum features”. In: *Infrared Physics & Technology* 86 (2017), pp. 77–89.
- [229] Limin Wang et al. “Places205-vggnet models for scene recognition”. In: *arXiv preprint arXiv:1508.01667* (2015).
- [230] Qiang Wang et al. “Learning attentions: residual attentional siamese network for high performance online visual tracking”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 4854–4863.
- [231] Brian D Wardlow and Stephen L Egbert. “A comparison of MODIS 250-m EVI and NDVI data for crop mapping: a case study for southwest Kansas”. In: *International Journal of Remote Sensing* 31.3 (2010), pp. 805–830.
- [232] Tim Wheeler and Joachim Von Braun. “Climate change impacts on global food security”. In: *Science* 341.6145 (2013), pp. 508–513.
- [233] Chaoyang Wu et al. “Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation”. In: *Agricultural and forest meteorology* 148.8-9 (2008), pp. 1230–1241.
- [234] Guorong Wu et al. “Unsupervised deep feature learning for deformable registration of MR brain images”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2013, pp. 649–656.
- [235] Kuo-Ping Wu and Sheng-De Wang. “Choosing the kernel parameters for support vector machines by the inter-cluster distance in the feature space”. In: *Pattern Recognition* 42.5 (2009), pp. 710–717.

- [236] Jingge Xiao et al. “Land Cover Classification Using Features Generated From Annual Time-Series Landsat Data”. In: *IEEE Geoscience and Remote Sensing Letters* 15.5 (2018), pp. 739–743.
- [237] Z. Xiao et al. “Use of General Regression Neural Networks for Generating the GLASS Leaf Area Index Product From Time-Series MODIS Surface Reflectance”. In: *IEEE Transactions on Geoscience and Remote Sensing* 52.1 (Jan. 2014), pp. 209–223. ISSN: 1558-0644. DOI: [10.1109/TGRS.2013.2237780](https://doi.org/10.1109/TGRS.2013.2237780).
- [238] Yichun Xie, Zongyao Sha, and Mei Yu. “Remote sensing imagery in vegetation mapping: a review”. In: *Journal of plant ecology* 1.1 (2008), pp. 9–23.
- [239] Qinchuan Xin et al. “Modeling grassland spring onset across the Western United States using climate variables and MODIS-derived phenology metrics”. In: *Remote Sensing of Environment* 161 (2015), pp. 63–77.
- [240] Jun Xiong et al. “Automated cropland mapping of continental Africa using Google Earth Engine cloud computing”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 126 (2017), pp. 225–244.
- [241] Jinru Xue and Baofeng Su. “Significant remote sensing vegetation indices: A review of developments and applications”. In: *Journal of Sensors* 2017 (2017).
- [242] L Yan and DP Roy. “Improved time series land cover classification by missing-observation-adaptive nonlinear dimensionality reduction”. In: *Remote Sensing of Environment* 158 (2015), pp. 478–491.
- [243] Guijun Yang et al. “Inversion of a radiative transfer model for estimating forest LAI from multisource and multiangular optical remote sensing data”. In: *IEEE Transactions on Geoscience and Remote Sensing* 49.3 (2010), pp. 988–1000.
- [244] Ming-Der Yang, Su-Chin Chen, and Hui Ping Tsai. “A Long-Term Vegetation Recovery Estimation for Mt. Jou-Jou Using Multi-Date SPOT 1, 2, and 4 Images”. In: *Remote Sensing* 9.9 (2017), p. 893.
- [245] Simon X Yang and Chaomin Luo. “A neural network approach to complete coverage path planning”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 34.1 (2004), pp. 718–724.
- [246] Qiangqiang Yuan et al. “Deep learning in environmental remote sensing: Achievements and challenges”. In: *Remote Sensing of Environment* 241 (2020), p. 111716.
- [247] Jun Yue et al. “Spectral–spatial classification of hyperspectral images using deep convolutional neural networks”. In: *Remote Sensing Letters* 6.6 (2015), pp. 468–477.

- [248] Amir R Zamir et al. “Feedback networks”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 1308–1317.
- [249] Wojciech Zaremba, Ilya Sutskever, and Oriol Vinyals. “Recurrent neural network regularization”. In: *arXiv preprint arXiv:1409.2329* (2014).
- [250] Chunhua Zhang and John M Kovacs. “The application of small unmanned aerial systems for precision agriculture: a review”. In: *Precision agriculture* 13.6 (2012), pp. 693–712.
- [251] Wangjian Zhang et al. “Boosted regression tree model-based assessment of the impacts of meteorological drivers of hand, foot and mouth disease in Guangdong, China”. In: *Science of the Total Environment* 553 (2016), pp. 366–371.
- [252] X Zhang et al. “Evaluating the fraction of vegetation cover based on NDVI spatial scale correction model”. In: *International Journal of Remote Sensing* 27.24 (2006), pp. 5359–5372.
- [253] Wenzhi Zhao and Shihong Du. “Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach”. In: *IEEE Transactions on Geoscience and Remote Sensing* 54.8 (2016), pp. 4544–4554.
- [254] Xinwei Zheng et al. “Automatic annotation of satellite images via multi-feature joint sparse coding with spatial relation constraint”. In: *IEEE Geoscience and Remote Sensing Letters* 10.4 (2012), pp. 652–656.
- [255] Liheng Zhong, Peng Gong, and Greg S Biging. “Phenology-based crop classification algorithm and its implications on agricultural water use assessments in California’s Central Valley”. In: *Photogrammetric Engineering & Remote Sensing* 78.8 (2012), pp. 799–813.
- [256] Fuqun Zhou, Aining Zhang, and Lawrence Townley-Smith. “A data mining approach for evaluation of optimal time-series of MODIS data for land cover mapping at a regional level”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 84 (2013), pp. 114–129.
- [257] Peicheng Zhou et al. “Weakly supervised target detection in remote sensing images based on transferred deep features and negative bootstrapping”. In: *Multidimensional Systems and Signal Processing* 27.4 (2016), pp. 925–944.
- [258] Jurgen Zoto et al. “Automatic Path Planning for Unmanned Ground Vehicle Using UAV Imagery”. In: *International Conference on Robotics in Alpe-Adria Danube Region*. Springer. 2019, pp. 223–230.
- [259] Qin Zou et al. “Deep learning based feature selection for remote sensing scene classification”. In: *IEEE Geoscience and Remote Sensing Letters* 12.11 (2015), pp. 2321–2325.

This Ph.D. thesis has been typeset by means of the T_EX-system facilities. The typesetting engine was pdfL^AT_EX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete T_EX-system installation.