POLITECNICO DI TORINO

Master's Degree in ICT4SS

Master's Degree Thesis

Popillia Japonica Newman detection through remote sensing and AI computer vision

Candidate

Davide BRUSCO

Supervisors

Prof. Marco PIRAS

Ph.D. Elena BELCORE

June 2023

Summary

Food production is one of the most important human activities. Fighting against animals and vegetable weeds is necessary, as climate change and globalization allow for the diffusion of exotic pests and illnesses. ICTs could be an innovative solution for a sustainable struggle against these problems. The purpose of this thesis is to detect and estimate the number of *Popillia Japonica Neuman* in vineyards near Novara, Italy, which is a pest that has been causing problems to Piedmont wine producers since 2014. This research analyzes the vegetation indexes results obtained from satellite images and studies the pest and vine leaf spectrometric signatures in order to define which are the most useful electromagnetic bands to be exploited in the images data collection. Those are captured through UAS equipped with remote sensing sensors in the RGB and NIR bandwidths. A protocol is drawn up to collect useful data about the insect. These data are used to create two datasets that are exploited in two AI models, one for each band. The algorithm used belongs to the YOLO family and implements the semantic segmentation approach. In this specific domain, the result obtained is that remote sensing and artificial intelligence represent a solution for a quick and autonomous *Popillia Japonica Newman* detection.

Acknowledgements

La mia esperienza universitaria è durata quasi 6 anni, 5 anni e 7 mesi per l'esattezza. Un periodo lungo il quale ho avuto la fortuna di conoscere belle persone e anche la fortuna di incontrarne di meno piacevoli. Al primo anno di triennale eravamo in centinaia, all'ultimo di magistrale poche decine. Nello stesso periodo mi sono portato dietro vecchie amicizie, alcune si sono consolidate, altre si sono perse. Al di fuori della carriera universitaria, la vita è così, un susseguirsi di relazioni che non si sa mai se sono destinate a finire o a durare all'infinito.

In tutto questo vai e vieni però qualcuno silenziosamente seguiva i miei passi. Io sono un tipo orgoglioso, non chiedo quasi mai aiuto, però quando qualcuno si propone di sua spontanea volontà, allora accetto con piacere. Mamma e Papà sono stati così, silenziosi e latenti per la maggior parte del tempo, ma precisi, potenti e di conforto quando ne avevo bisogno. A voi due dico grazie.

Uno dei momenti più difficili di questo percorso sono state le 14 sessioni d'esame affrontate. Di sessioni ce ne sono tipicamente 3 l'anno: una tra gennaio e febbraio, una tra giugno e luglio e una a settembre; considerando che iniziavo ad entrare nel mood sessione un mese prima dell'inizio della stessa, di periodi tranquilli ne rimanevano ben pochi. Schopenhauer diceva: *"L'uomo è un pendolo che oscilla fra la noia e il dolore."*, adattandola allo studente universitario si può affermare che lo studente universitario è un pendolo che oscilla fra una sessione esami e la successiva. Durante la sessione esami ero intrattabile, sempre stressato e sul punto di fare la guerra. Un detto dice che *I veri amici si vedono nel momento del bisogno*, ma se si è nervosi 9 mesi l'anno allora questa frase, a mio avviso, perde un pò di valore. Io ho avuto la fortuna di essere circondato da amici e famigliari che sono stati comprensivi nei miei confronti sempre, giorno dopo giorno, per 5 anni e 7 mesi, un lavoro non da poco. A tutti voi dico grazie.

Quindi un brindisi ai miei genitori, per la loro costante e silenziosa presenza, ai mei famigliari, per il loro frenetico tifo nei miei confronti, ai miei vecchi amici, per aver rinnovato la loro presenza al mio fianco, ai miei nuovi amici, per aver portato freschezza e novità al mio percorso, alla mia ragazza, per aver cosparso di affetto le mie giornate, a voi tutti dico grazie.

Citando la celeberrima frase di un saggio: *"Davide, mi raccumand u risultatu!"*.

Table of Contents

List of Tables

List of Figures

Acronyms

GDP
: Gross Domestic Product

EU
: European Union

EEC
: European Economic community

LiDAR
: Light Detection And Ranging

NASA
: National Aeronautics and Space Administration

ESA
: European Space Agency

ENAC
: Ente Nazionale Aviazione Civile Italiana

EPPO
: European and Mediterranean Plant Protection Organization

GBIF
: Global Biodiveristy Information Facility

PAN
: Piano Azione Nazionale

PGI

Protected Geographical Indication

IGP

Indicazione Geografica Protetta

PDO

Protected Designation of Origin

DOP

Denominazione Origine Protetta

DOC

Denominazione Origine Controllata

DOCG

Denominazione Origine Controllata e Garantita

DIATI

Dipartimento Ingegneria dell'Ambiente, del Territorio e delle Infrastrutture

LLIN

Long-Lasting Insecticide-Treated Nets

ICT

Information Communication technology

PaaS

Platform as a Service

SaaS

System as a Service

GPU

Graphical Processing Unit

API

Application Programming Interface.

UAS
Unmanned Aereal System

APR
Aereomobile Pilotaggio Remoto

RGB
Red Green Blue

NIR
Near InfraRed

RTK
Real Time Kinematic

GPS
Global Positioning System

GNSS
Global Navigation Satellite System

VI
Vegetation Index

OES
Optical Emission Spectroscopy

NDVI
Normalized Difference Vegetation Index

SAVI
Soil-Adjusted Vegetation Index.

EVI
Enhanced Vegetation Index

AI
Artificial Intellegence

NN

Neural Network

CNN

Convolutional Neural Network

RCNN

Region-based Convolutional Neural Network

SVM

support vector machine

RPN

Region Proposal Network

YOLO

You Look Only Once

KNN

K Nearest Neighbours

NB

Naive Bayes

MSE

Mean Square Error

MAP

Mean Average Precision

Introduction

Since humans became sedentary, food production has been one of the most important human activities. From the beginning, fighting against animals and vegetable weeds was necessary. Climate change and globalization have resulted in the diffusion of exotic pests, such as *Popillia Japonica Newman* and *Asian Debug*, which have spread in Europe from Asia, or *Xylella fastidiosa*, native to South America. Those exotic pests attack the local agricultural system which is not able to fight and to eliminate them. ICT, *Information Communication Technology*, could provide innovative and sustainable solutions to address these problems. Actually, in the precision agricultural domain, ICTs are exploited to control smart connected machines, to monitor crops and as decision support software. In the north of Italy, the *Popillia Japonica Newman* was first discovered in 2014, and since then, research has been focused on finding sustainable solutions to control this invasive pest.

The objective of this thesis is to verify if remote sensing combined with AI, *Artificial Intelligent*, can identify and estimate the number of *Popillia Japonica Newman* infestation in a vine quickly and autonomously. The research consists on on-filed activities, for the data collection, and laboratory activities. All of those were possible to the support of the DIATI, *Dipartimento di Ingegneria dell'Ambiente del Territorio e delle Infrastrutture* of the Politecnico di Torino, and *Consorzio Tutela Nebbioli Alto Piemonte*, [1], directed by Andrea Fontana.

Remote sensing refers to the process of acquiring information about object or area from a distance, without having physical contact with it. This is typically done using sensors and instruments mounted on satellites, drones or rovers. On the other hand, spectral characterization, used as support to remote sensing, is done with spectrometers capable of detecting the spectral signatures of the interested objects.

The first part of this research involves quantifying *Popillia Japonica Newman* through satellite data, such as the images from the project Copernicus, from ESA, *European Space Agency*. The aim is to define if computing vegetation indexes using those data, it is possible to study the impact of the pest on crop fields and consequently quantify its presence.

The second part of this research involves the spectral characterization of insects and vine leaves as a preparatory study for the collection of UAS, *Unmanned Aereal System*, images. The aim is to exploit remote sensing instruments mounted onboard of them to collect multispectral data, which are processed according to their spectral signatures and used for pest quantification. The result of this section is the definition of a data collection protocol, specifically defined for the quantification of *Popillia Japonica Newman*.

In the third part, the collected data are labeled and used as input for an AI algorithm to understand which is the best computer vision approach and consequently to estimate the number of insects in the entire plantation.

Chapter 1

Overview of agricultural value, policies and market

1.1 Italian agriculture value

One of the Italian hallmarks in the world is the food, known for its quality and its long traditions. Food production and more generally agriculture and connected activities represent a milestone for the Italian economy [2]. To explain that, in 2019 the agriculture economy employed 5% of the active Italian population and it earned about 61.6 billion euros, this value increases significantly to 133.679 billion euros if the food industry is also considered, which is equal to 8.2% of the GDP, Gross Domestic Product, [3]. The primary production sector also plays an essential role from a social and cultural point of view. Since 1992 with EEC 2018/92 regulation, and then in 2006 with EEC 510/2006 regulation, the EU, European Union, aims to protect regional food products by introducing PGI, *Protected Geographical Indication*, IGP in Italian, and PDO, *Protected Designation of Origin*, DOP in Italian, certifications. The relative logos are reported respectively in Figure 1.1. Then in 2014, the Italian government applied the EU 1151/2012 European directive, useful to protect the names of specific products and to promote their unique characteristics linked to their geographical origin as well as traditional indication [4]. From that moment to 2021, Italy registered more than 800 food products, the higher number in the EU. Between them, there are 172 DOP and 140 IGP products and several DOC and DOCG, *Denominazione Origine Controllata e Garantita*, wines. These are shown divided by region in Table 1.1.

Figure 1.1: PGI and PDO certification logos.

Region	DOCG	DOC	Region	DOCG	DOC
Valle D'Aosta	0	1	Emilia Romagna	2	18
Piedmont	16	42	Abruzzo	1	8
Lombardia	5	27	Lazio	3	27
Liguria	0	8	Molise	0	4
Trentino Alto Adige	0	8	Campania	4	15
Friuli Venezia Giulia	4	10	Puglia	4	28
Veneto	14	28	Basilicata	1	4
Tuscany	11	41	Calabria	0	9
Marche	5	15	Sicilia	1	23
Unbria	2	13	Sardinia	1	17

Table 1.1: In Italy there are more than 400 quality certified wines, between them 74 are denominated DOCG and 341 DOC.

1.2 Italian agricultural economic trend and future challenges

According to [2], Italian agriculture in 2019 witnessed a sharp decline in the production, -12.1%, and value ,-17.5%, of wine, as well as a general reduction in fruit production ,-6.6%. However, the harvest of dried legumes increased by 8.7%, and olive oil production rose by 27.6%, due to a massive planting program aimed at countering the effects of *Xylella Fastidiosa* in recent years. The reasons behind these data are closely related to negative climate conditions and the spread of exotic pests and bacteria, such as the *Asian Debug*, the *Popillia Japonica Newman*, and the *Xylella Fastidiosa*, or actions taken to counter them.

Anthropogenic climate change exists and has a significant impact on agriculture and food systems in general. One of the main future challenges is to counter extreme weather events, such as late frosts, greater irregularity of rains, difficulties in water basin management, high soil disturbance, and variation in the sowing period, as reported by L. Mercalli in [5]. These events have caused a loss of 450 billion euros between 1980 and 2016 in Europe, as shown in Figure 1.2. Additionally, exotic weather conditions allow the spread of exotic and unknown parasites, mainly insects and bacteria, that the local environment cannot counter, resulting in a weakening of the food production system until a natural or, more often, a phytosanitary solution is found.

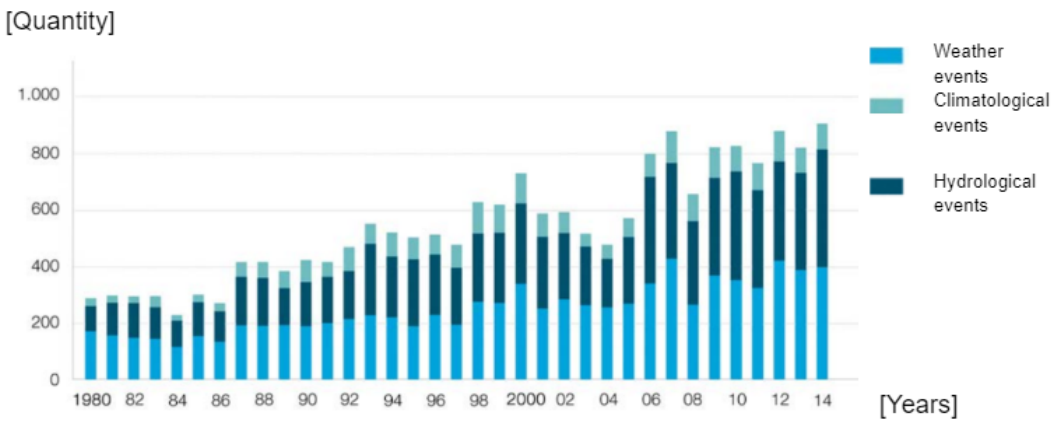

Figure 1.2: The plot shows clearly the increasing trend in the number of extreme weather events, between 1980 and 2014, [5].

Going deeper into unusual climate conditions, in Italy in 2018, climate change showed itself through a 1.5°C increase in the average minimum temperature and 1°C in the average maximum temperature, with long, dry winters and summers while the rainfalls are concentrated in short periods during the spring and the autumn. The forest's capability to absorb CO2 has decreased by about 42% also due to the increasing deforestation and soil cementification. The described context has a strong impact on agriculture in terms of trade, which is an index that takes into consideration final product prices and intermediate consumption prices. Between 2015 and 2019, considering the terms of trade reported in Figure 1.3, there was a a 20% decline in profits, basically due to the increasing cost of the intermediate consumption such as seeds, motive power, phytosanitary, insurance, irrigation waters, fertilisers and transportation, as shown in Figure 1.4. These findings are reported in [3]. An example of exotic bacteria, the *Xylella Fastidiosa*, is reported in Figure 1.5 and described in [6]. It is the cause of the *CoDiRO*, *Complesso del Distaccamento Rapido dell'Olivo*, that spread in the Puglia region, Italy.

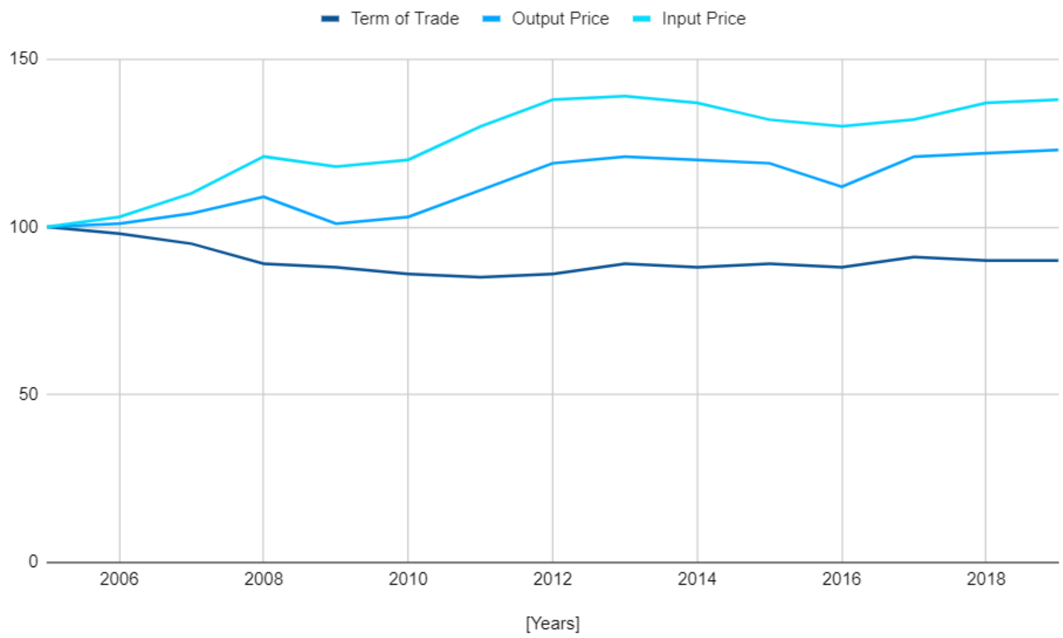

Figure 1.3: Terms of trade during the period between 2005 and 2019: The input price or intermediate consumption increase a lot, consequently also the output price is increased, [3].

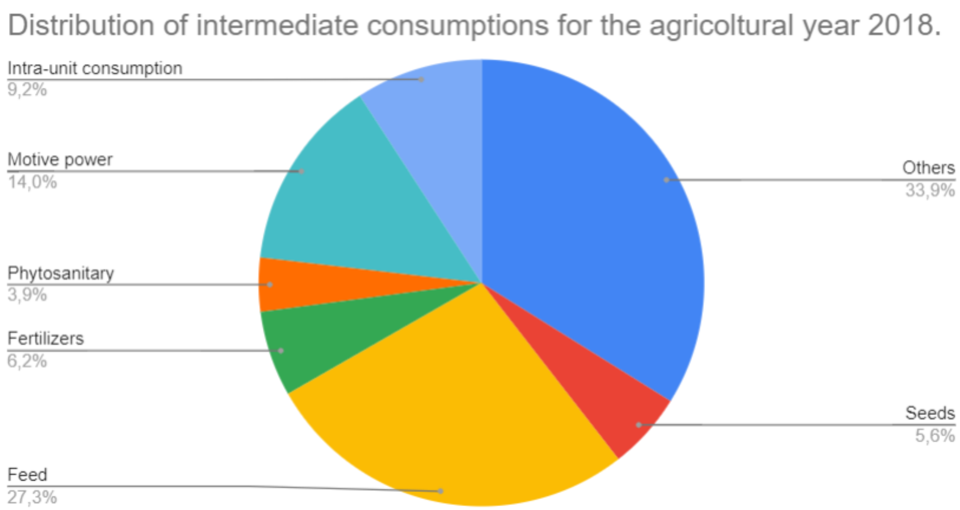

Figure 1.4: Representation of the intermediate costs in the food production sector, [3].

The olive trees show dry grey spots, as shown in Figure 1.6, and in a period long between some months to a few years, the tree died without any possibility of saving it. Around 6.5 million olive trees died in Puglia since the *Xylella* appeared in 2013, with significant consequences on the regional and national economy and on the society.

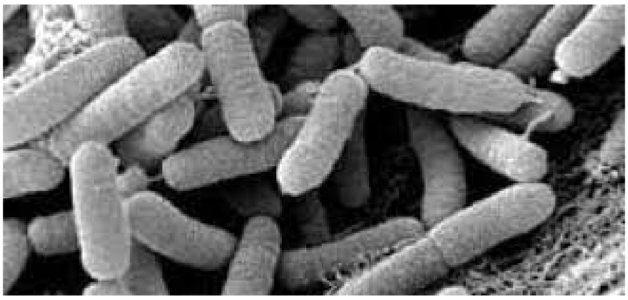

Figure 1.5: The bacteria Xylella Fastidiosa which is the cause of CoDiRO.

Figure 1.6: The effect of CoDiRO on olive trees, on them grey spots appear.

In [7], L. Meadows reports that a future challenge will be the growing population. Around 10 billion people are expected to be living on Earth in 2050, and the quantity of food needs to be increased to meet this demand. However, the loss of soil fertility, pollution, climate change, and the spread of pests and bacteria represent strong obstacles to the sustainability of food production. In this scenario, smart and precision agriculture, and more generally, sustainable agriculture, could be a solution to these challenges. In particular, new ICTs provide real support in the farmer decision-making process, which could help to reduce harvest losses.

1.3 Precision farming spreading

According to G. Orefice in [8], the number of investments in smart agriculture has increased significantly in recent years, with a global turnover of 15 billion euros and around 750 new startups entering the market. In Italy the precision agriculture business has grown from 540 million euros in 2020 to 1.6 billion euros in 2021. Of the total investments, 47% are devoted to purchasing smart connected machines, while 53% are used for monitoring and controlling crops through sensor systems. Decision support software and system deployment have also seen a substantial increase, with 26% of farmers opting for these solutions. In 2021, 60% of farmers used at least one ICT solution, a 4% increase from the previous year, and 40% used at least two. These trends indicate that ICTs will play an increasingly crucial role in the agricultural industry in the years to come.

In this context, the massive amounts of data collected by Industry 4.0 systems play a crucial role. While some of these data are used for managing warehouses and food production, processing and analyzing them to generate trends, models, and AI software is more challenging.

Examples of new technologies that are transforming agriculture include UAS, *Unmanned Aereal Systems*, commonly known as drones, and ground rovers that are supported by satellite images and positioning systems. These tools are used to improve phytosanitary shedding, crop disease and pest detection. Weather stations are also employed to describe microclimates and vegetation health status, while autonomous machinery can work without farmer monitoring thanks to precise satellite positioning. Figure 1.7 shows an example of a precision farming drone.

The combination of precision farming and AI represents a major innovation in the primary sector, and could be a key solution to counteract future agricultural challenges.

Figure 1.7: A few researchers of the Politecnico di Torino using UAS to detect the *Popillia Japonica Newman* in vineyards.

Chapter 2

ICT technologies for precision farming

EPI-AGRI, [9], is the productivity and sustainability agriculture European partnership and it was established in 2012 to achieve the European objective of the Agenda 2030, [10]. It defines precision farming as "*a management approach that focuses on near real-time observation, measurement, and responses to variability in crops, fields and animals. It can help increase crop yields and animal performance, reduce costs, including labour costs, and optimize processes inputs. All of these can help increase profitability. At the same time, precision farming can increase worker safety and reduce the environmental impacts of agriculture and farming practices, thus contributing to the sustainability of agricultural production*".

The precision farming tools are ICTs such as remote sensing, UAS, satellites, positioning systems, on-field sensors which generate a huge amount of data. These tools have high technological requirements, such as internet coverage, high computational power or data center and require technicians.

2.1 Remote sensing

Remote sensing is the acquisition of information about an object or phenomenon without physical contact with it, but based on energy source to illuminate the target. This energy is in the form of electromagnetic radiation. The remote sensing instruments are divided into active and passive sensors. The active ones emit signals, such as microwaves, and then measure the reflected or backscattered energy to derive information about the target. Examples of active remote sensing sensors include radar and LiDAR systems. On the other hand, passive sensors detect natural or ambient electromagnetic radiation, such as sunlight or thermal radiation, emitted by the target, without emitting any radiation of their own, example in

Figure 2.1. *A* represents the sun, while *B* the sunlight radiation that is reflected by the Earth's surface, *C*. The electromagnetic radiation is captured by the satellite *D*. Thanks to base stations and other satellites, *E*, visual data, such as false colour maps, are generated, *F* and *G*.

Passive sensors are highly sensitive to the spectral characteristics of the target and are therefore used extensively in applications such as land-use mapping, agricultural monitoring, and forest inventory, [11]. The detected response generates a specific spectrometric signature, which is a result of the interaction between electromagnetic radiation and matter as a function of the wavelength or frequency of the radiation. The spectrometric signature characterizes each type of matter and is defined in one or more bands, [11].

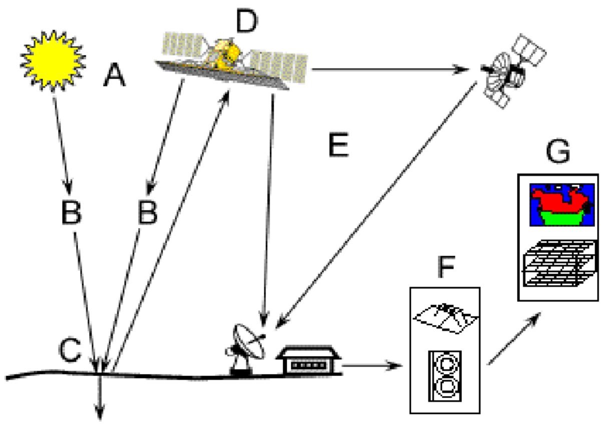

Figure 2.1: A schema of a passive remote sensing instrument: the satellite.

Examples of passive sensors include radiometers and cameras, an example is shown in Figure 2.2, cameras can be multispectral or hyperspectral and can be mounted on board satellites, drones or rovers, or alternatively, carried by hand. A camera is defined as multispectral if the number of narrow spectral bands that the sensor is able to detect is between 3 and 10, while a hyperspectral camera can sense hundreds or even thousands of very narrow spectral bands, covering a much wider range of the electromagnetic spectrum than a multispectral camera, a graphical representation is reported in Figure 2.3.

Cameras and generally remote sensing device have their own spectral resolution which describes the ability of a sensor to define fine wavelength intervals. The finer the spectral resolution, the narrower the wavelength range for a particular channel or band. Table 2.1 reports the bands commonly used in remote sensing. NIR, *Near-infrared Radiation*, which falls between 700 nm and 1000 nm, is absorbed by vegetation chlorophyll, other pigments, and water content in leaves. NIR

Figure 2.2: In a typical multispectral or hyperspectral camera, each camera lens is able to cover a specific range of the electromagnetic spectrum.

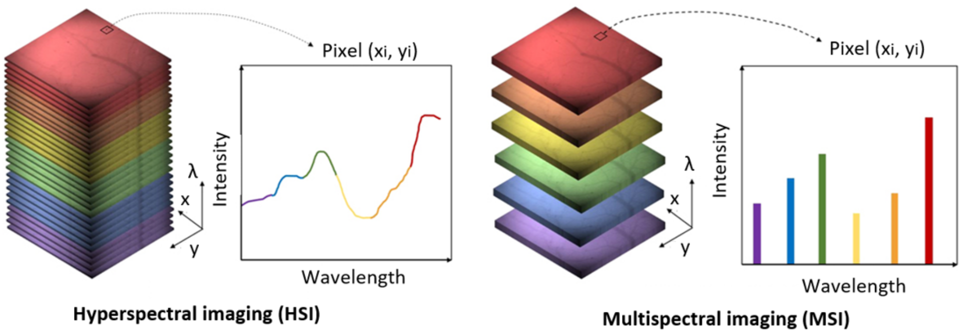

Figure 2.3: Comparing between multispectral and hyperspectral approach.

radiation is used to determine plant health, leaf area index, and water stress. It is also used in mineral exploration to detect hydroxyl groups in clay minerals. Red-Edge is a specific range of wavelengths between 680 nm and 750 nm that is used in remote sensing for vegetation analysis. This spectral range is particularly sensitive to changes in chlorophyll content and is used to estimate plant biomass, leaf area index, and other vegetation parameters. Human-visible radiation is the type of radiation used in photography to capture photos and video. UV, *Ultraviolet radiation*, which has wavelengths between 10 nm and 400 nm, is used in remote sensing for atmospheric monitoring and aerosol detection. It is also used in fluorescence spectroscopy to detect the presence of organic compounds and minerals with fluorescence properties.

The choice of bandwidth depends on the specific remote sensing application and the information that needs to be extracted from the target surface. Different bandwidths can provide unique information about the target surface, and combining multiple bandwidths can enhance the accuracy and reliability of remote sensing data analysis.

Band	Range(nm)
NIR	700-1000
Red-Edge	680-750
Human visible	400-700
UV	10-400

Table 2.1: The most common bands used in remote sensing in various applications.

Finally the radiometric correction is very important, which is a process used to correct errors and variations in the brightness values of digital images captured by remote sensing instruments. The aim of radiometric correction is to convert the raw digital numbers recorded by the sensor into meaningful physical units of radiance or reflectance. Radiometric correction is essential for many applications, such as land cover mapping, vegetation analysis, and environmental monitoring. By removing the effects of atmospheric attenuation and sensor calibration, radiometric correction enables the comparison of images acquired at different times, under different weather conditions, or by different sensors, thus facilitating long-term studies and trend analysis, [11].

2.2 Vegetation indexes

A VI, *Vegetation Index*, is a spectral imaging transformation of two or more image bands designed to enhance the contribution of vegetation properties, allowing reliable spatial and temporal inter-comparisons of terrestrial photosynthetic activity and canopy structural variations. The values of the image bands are obtained by measuring the reflectance of vegetation to solar light in several electromagnetic spectrum bands. Figure 2.4 shows the qualitative electromagnetic responses of a dead, stressed, and healthy leaf.

In the following list, some VIs are presented:

1. NDVI stands for *Normalized Difference Vegetation Index*. It is defined as the vigour index and indicates the health of vegetation through the measurement of photosynthesising biomass. Its mathematical definition is reported in Equation 2.1. The more the value is low, the lower are photosynthesising abilities of the vegetation cover. NIR and Red are the surface reflectances of the object to the solar light in the respective bands.

$$NDVI = (NIR - Red)/(NIR + Red) \tag{2.1}$$

2. SAVI instead is the *Soil-Adjusted Vegetation Index*, which is similar to the NDVI index but with a bare soil correction factor. It works well with low

Figure 2.4: Electromagnetic response of a leaf in the interested bands used to vegetation indexes computation.

vegetation cover and it allows the comparison of a large variety of fields. It is defined mathematically in Equation 2.2. NIR and Red are the surface reflectances of the object to the solar light in the respective bands. L is the canopy background adjustment, and it is equal to 0.5.

$$SAVI = (1 + L)(NIR - Red)/(NIR + Red) \tag{2.2}$$

3. The EVI, *Enhanced Vegetation Index*, is defined by Equation 2.3 and it can be considered as an optimized VI with improved sensitivity in high biomass regions by a de-coupling of the canopy background signal and a reduction in atmosphere influences. NIR and Red are the surface reflectances of the object to the solar light in the respective bands. L is the canopy background adjustment, and it is equal to 1, C1 and C2, respectively equal to 6 and 7.5, which are the coefficients of the aerosol resistance term. The blue is the aerosol influences correction factor in the red band, while G is a gain factor equal to 2.5.

$$EVI = G \times (NIR - Red)/(NIR + C1 \times Red - C2 \times Blue + L) \tag{2.3}$$

2.3 Remote sensing sensor supports

2.3.1 Satellites

Since the 1960s, satellites have been widely used for communication, navigation, military purposes, and more recently, remote sensing. This technology allows for the collection of data that are available for scientific research and public use. One major advantage of satellite technology in precision farming is the ability to collect a large amount of data about the field, covering a wide area in a short time. Satellites can also provide various weather data and help with machinery positioning. However, bad weather conditions like dense clouds and fog can limit these opportunities. Actually for remote sensing purposes, the main active constellations are:

1. *Landsat* program by NASA is the first space program devoted to data acquisition through remote sensing;
2. Sentinel2 and *Corine Land Cover* belong to the Copernicus program by ESA, they are specifically developed for image acquisition through remote sensing techniques;
3. *Terra and Aqua* is operated by NASA, it provides remote sensing data for studying the Earth's climate, atmosphere, land, and oceans;
4. SPOT is operated by Airbus Defence and Space, it provides high-resolution imagery for various applications including agriculture, forestry, and defense.
5. *RapidEye* is operated by Planet Labs, it provides multispectral imagery for various applications including agriculture, mapping, and disaster management.

The Landsat program by NASA, *National Aeronautics and Space Administration*, is the longest-running civil satellite program. It was launched in 1975, and the most recent launch was Landsat 9 on September 27, 2021. Only Landsat 8 and Landsat 9 are currently active. This system has eight spectral bands with spatial resolution ranging from 15 to 60 meters and a temporal resolution of 16 days. It has different applications in agriculture, cartography, geology, forestry, and surveillance.

In addition, in 2015 the ESA launched the Sentinel-2 mission as part of the Copernicus program. It comprises a constellation of two polar-orbiting satellites, Figure 2.5, which aim to monitor variability in land surface conditions. The swath width is 290 km, and the revisits time is between 10 to 5 days at the equator with 2 satellites under cloud-free conditions and between 2 to 3 days at mid-latitudes in the same conditions.

The main differences between the constellations are the spatial and temporal resolution. Landsat satellites have about 30 meters of spatial resolution and a temporal resolution ranging from 8 to 16 days, which may not be enough during certain growing seasons. In contrast, Sentinel-2 satellites have a 10-meter spatial resolution and a temporal resolution ranging from 3 to 5 days. It's important to note that some data may be lost due to bad weather conditions, which must be taken into consideration.

2.3.2 Unmanned Aerial System

UAS is a vehicle that is operated without a human pilot on board, it typically consists of three main components: the UAS or drone, a ground-based control

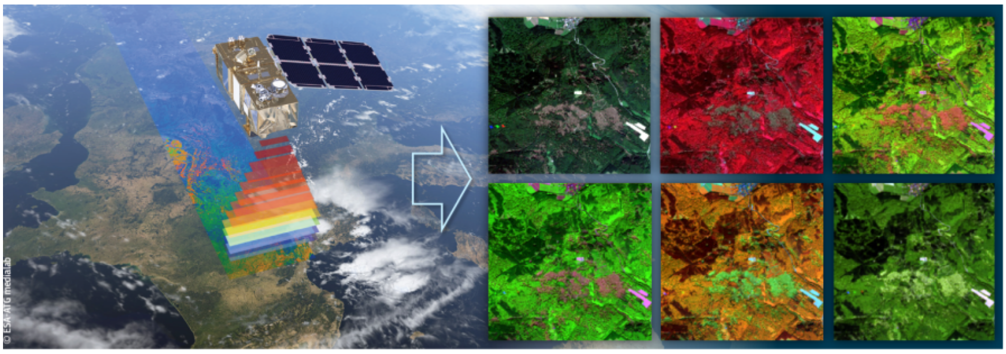

Figure 2.5: Examples of multispectral images acquired by Sentinel-2, the same image captured by the satellite processed in different ways gives false colour maps

station, and a communication link between the two. They can be equipped with remote sensing sensors as shown in the first drone in Figure 2.7. Remote sensing is one of the more common applications connected with UAS technology. On top of them, high-resolution multispectral or hyperspectral cameras can be carried, which are able to collect huge amounts of images with high resolution. This feature is enhanced by the possibility to plan and schedule flights in order to obtain a repeatable acquisition protocol. The most advanced drones can carry out the mission without human intervention. This feature allows farmers to automate and deepen the study of a field portion that may require specific activity. These systems are often connected to an automatic drone charging station that enables the industrial drone to land, take off, recharge the batteries, and download and upload data.

2.4 ICT technologies supporting remote sensing

2.4.1 GNSS positioning

GPS, *Global Positioning System*, GALILEO, GLONASS and BeiDou are GNSS, *Global Navigation Satellite System* providing positioning and timing information worldwide. They are provided respectively by United States government, EU, Russian Federation and Republic of China. RTK, *Real-Time Kinematic* instead is a high precision positioning technique used to determine the location of a receiver with centimetre-level accuracy in real-time, it exploits the GNSS systems. RTK positioning relies on a technique called carrier-phase differential GPS, which involves using two receivers, a base station and a rover, to calculate highly precise positioning information. The base station receives signals from satellites and calculates its precise position using carrier-phase measurements, which are extremely accurate but prone to errors. The rover also receives signals from satellites, and the difference between the carrier-phase measurements at the base station and the rover is used to calculate the precise position of the rover in real time. RTK positioning systems can also provide additional information, such as heading and velocity, to further enhance the accuracy of navigation and control applications. RTK positioning is commonly used in applications that require high accuracy, such as land surveying and precision agriculture, exploiting UAS. In Figure 2.6, a UAS equipped with RTK is reported.

Figure 2.6: The RTK positioning service is provided by the circular element on the top of the UAS.

2.4.2 Optical emission spectroscopy

Optical emission spectroscopy , OES, is a remote sensing technique that is based on the detection of an object's reflectance, which is the intensity of radiation emitted by an object when it is hit by solar or artificial light. The term reflectance refers to the electromagnetic radiation intensity reflected by an object subject to incident radiation. In contrast, transmittance refers to the object's ability to be penetrated by the incident radiation, and absorbance refers to the electromagnetic radiation intensity absorbed by an object subject to incident radiation. The sum of these parameters is always 1, and they together completely describe an object's electromagnetic signature.

A spectrometer is generally used to perform this kind of analysis. This term refers to several tools for measuring electromagnetic radiation. The basic principle involves exploiting wave interference to decompose radiation into its wavelengths and measuring them through a photodiode.

2.5 Others ICTs exploited by precision farming

UASs are an establish technology, their first appearance was in the early twentieth century. Then, they experienced rapid development during the two world wars and again with transistor engineering in the sixties. UASs have been affordable to common users since 2006, but it was only in the last few years that the EU Commission began working on producing a common European directive for both recreational and working use.

Currently, different UAS typologies can be used for several purposes, such as remote sensing, precision farming, surveillance, and search and rescue. In the near future, drones might be used for the autonomous delivery of drugs and small packages, as well as flying taxis for people. Example of UASs exploited in smart farming is reported in Figure 2.7: on the left a multispectral camera-equipped drone, which is flexible and easy to use and generally able to cover surface lower than one kilometer square, then a fixed-wing drone in the middle, with a high precision positioning system which is able to fly very long distances, and finally a quadcopter spraying fertilizers or pesticides on the right, able to shoulder payloads up to 30 kilograms. The main potential benefits of using UAS in precision farming

Figure 2.7: Examples of UAS types.

is to have the possibility to cover large areas quickly and efficiently, saving farmers time and labor cost, to sense and mapping the crop field improving monitoring activities, to harvest and to spraying pesticides or fertilizers. On the other hand the drawbacks are high initial investment, technical limitations, such as battery life and weather condition, and regulatory restrictions.

In the agricultural field, the EU issued the CE 128/09 directive in 2009 [12], which defines some boundaries in the use of phytosanitary substances through UAS. In Italy, this directive was adopted earlier, first with Decreto Legislativo 150/2012 and then with PAN, *Piano Azione Nazionale*, in 2014. This policy introduced strong limitations on the use of drones in smart agriculture. Article 14 states: "*Aerial spraying of pesticides has the potential to cause significant adverse impacts on human health and the environment, particularly from spray drift. Therefore, aerial spraying should generally be prohibited with derogations possible where it represents clear advantages in terms of reduced impacts on human health and the* *environment in comparison with other spraying methods, or where there are no viable alternatives, provided that the best available technology to reduce drift is used.*" Nowadays, an effective and comprehensive international directive that regulates the use of flying drones for both working and recreational use is still missing.

In Italy, the size of a drone can vary from a few hundred grams, which only requires registration with ENAC, *Ente Nazionale Aviazione Civile Italiana*, to a few kilograms, which require a complete license.

Satellites and flying drones are limited in their ability to collect data from the upper portions of plants and areas covered by leaves. To overcome these limitations, rovers are used as autonomous ground vehicles to collect data, such as images or other measurements. Rovers can navigate around the field and move in close proximity to the plants, allowing for more detailed and accurate data collection. Figure 2.8 shows an example of a rover used for agricultural applications.

Figure 2.8: A small rover moving in a cornfield.

Widely spread in the large farms of the Po Valley, France, Germany, and the USA are autonomous agricultural machinery tracked by high-quality positioning systems. A few examples are shown in Figure 2.9. Thanks to the positioning system, the vehicles can customize the amount of fertilizer, water, and pesticides specifically for each portion of the field. In addition to that, high-tech sensors can be installed on these machines to collect data about the ground and plant health status.

Figure 2.9: Autonomous machinery used without worker monitoring.

On-field sensors are an established technology used for several decades in all sectors of agriculture and livestock. Nowadays, integrated systems based on networks of sensors and devices are available on the market that collaborate to generate a complete and in-depth vision of the fields. Alternatively, punctual solutions can measure the values of a restricted area. An example is the weather station, which can sense climate parameters such as air temperature and humidity, wind speed and direction, and rain volumes, as shown in Figure 2.10. More complex systems can measure the leaf coverage, terrain humidity, and detect pests or illnesses using cameras and AI image segmentation algorithms.

Figure 2.10: Example of weather station.

2.6 Algorithms

The described systems collect a large amount of data that needs to be processed and analyzed to extract useful information to manage the entire agricultural chain. First of all, management software are quite common. They help the farmer in warehouse management, planning working activities, and controlling and maintaining machinery.

In this scenario, AI and ML *machine learning* approaches have a primary role. In fact, thanks to them, generating trends, models, and predictions are possible. Examples of algorithms exploited in smart agriculture are:

1. NN, *Neural network*, for data prediction: CNN, *Convolutional Neural Networks*, are often used to learn time-series trends end to predict future values of temperature and humidity;

2. Computer vision algorithms: R-CNN, *Region proposal Convolutional Neural Network*, mask R-CNN and YOLO, *You Look Only Once* are image and

semantic segmentation algorithms that can identify determined objects in an input image, this class of algorithms are exploited to detect pests and illnesses;

3. Remote sensing algorithms and vegetation index: from the computation of the vegetation indexes is easy to create specific maps for optimizing the irrigation process and phytosanitary and fertilisers distribution process;

4. Mathematical models correlate the weather condition with plants' growth levels.

In the last few years, the concept of blockchain has also affected the agricultural sector. It allows us to trace the agri-food supply chain and monitor food production, transformation, and distribution. The goal is to certify the origin of the product, mainly IGP, DOP, DOC, and DOCG ones, making the customer aware of the entire supply chain and the high-quality value products.

2.6.1 Computer vision algorithms in precision farming

Computer vision is an interdisciplinary field that deals with how computers can interpret and understand digital images or videos. It involves developing algorithms and techniques that allow computers to analyze, process, and extract meaningful information from visual data. This field combines elements from computer science, mathematics, physics, and biology to create artificial intelligence systems that can perform tasks such as image and object recognition, image segmentation, motion analysis, and 3D reconstruction.

In precision farming, often the background is complex, and the objects to detect are small and colored in the same way as the surrounding vegetation. Consequently, computer vision techniques can be difficult to implement. It was necessary to study and evolve algorithms for almost ten years to obtain efficient AI software that can detect precise objects in these scenarios.

Image segmentation is a ML method based on digital image breaking down into subgroups called image segments. An example is shown in Figure 2.11. The goal is to reduce the image complexity to make further processing possible. In the object detection a ML model processes the segments and gives as output the bounding boxes identifying specific labels. In Figure 2.12, a schema is shown about the logical steps of object detection, [13].

The evolution of image segmentation has led to the development of instance segmentation and semantic segmentation. Instance segmentation is a special form of image segmentation that deals with detecting instances of objects and demarcating their boundaries. It groups pixels in a semantically meaningful way, as shown in the example in Figure 2.13. In fact, both image segmentation and instance segmentation are performed with deep learning techniques, which are a class of ML methods based on NNs. The CNN and the RCNN belong to this family, [13].

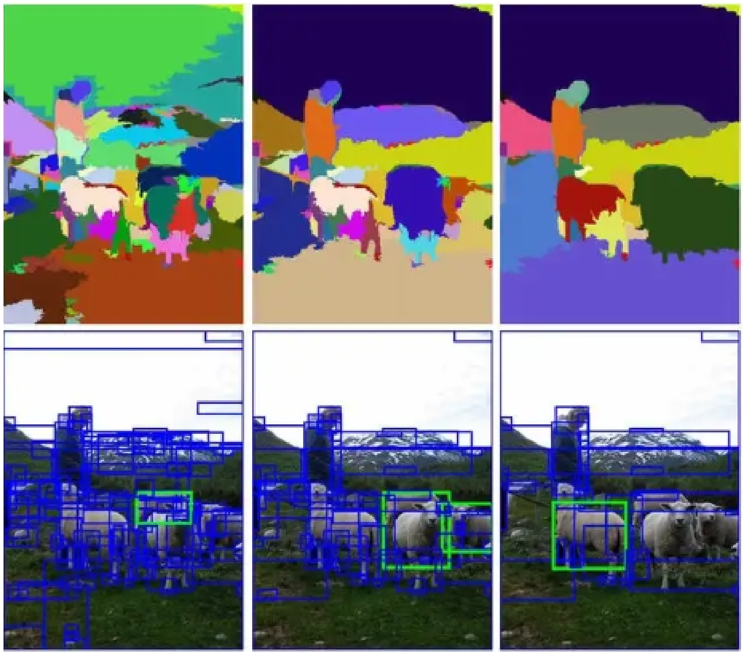

Figure 2.11: On the top row the image is divided into subgroups, called segments or regions, based on the pixel analysis. On the bottom row instead, the bounding boxes are obtained from the region's proposal.

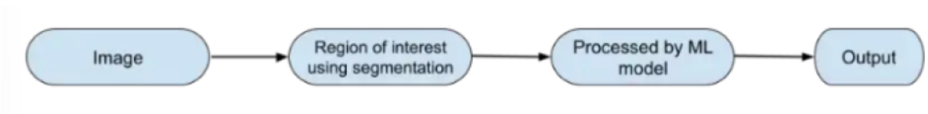

Figure 2.12: In this schema, the input is an image while the output is the searched objects. In order to obtain that, the first step is the detection of regions of interest through image segmentation, then those are processed by a ML model.

2.6.2 From image segmentation to instance segmentation

One of the first effective approaches to the problem was proposed in 2014 by R. Girshick et al [14]. They developed a scalable detection algorithm that improved the success rate from 30.0% to 53.3%. This method was based on CNN, and then on a pre-trained supervised ML classification model. The resulting algorithm was called R-CNN. In the following years, the algorithm was significantly improved by increasing the number of independent region proposals obtained from CNN to 2000 and by exploiting better classification algorithms, such as the *Support Vector Machine*, SVM.

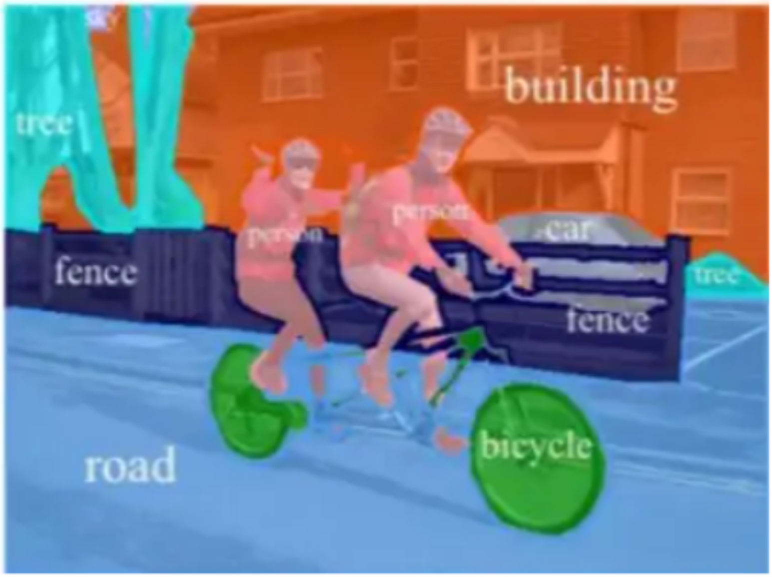

Figure 2.13: In the instance segmentation the subgroups are obtained considering not only the single features of a pixel but considering its role in the object detected.

The R-CNN algorithm had two obvious pitfalls: multi-stage training and test stage resource requirements, and the limitation on the input image size. To fix the first issue, He et al. [15] proposed *Fast R-CNN* in 2015, which achieved significantly higher performance in less time and with fewer resources. In the same year, the fixed-size input requirement was improved by He et al [16] with the *Spatial Pyramid Pooling*, a pooling layer in the CNN that does not require a fixed-size input image.

Another improvement in the R-CNN family algorithms was the *Faster R-CNN*. It was able to reduce the region proposal stage from 2 seconds to 10 milliseconds per image, which was the bottleneck of the system at the time, thanks to the *Region Proposal Network*, RPN, algorithm. The RPN is a fully convolutional network that simultaneously predicts object bounds and objectness scores at each position. Ren et al. proposed this solution in 2015, [17].

The *Mask R-CNN* extends the Faster R-CNN by adding a branch for predicting segmentation masks on each region of interest. Like its predecessors, the Mask R-CNN is based on two steps, the RPN and the classifier, as reported in [18].

Finally, YOLO was presented for the first time in 2016 [19]. It is currently the best, fastest, and most accurate computer vision algorithm to perform instance segmentation. Its main feature is that it predicts multiple bounding boxes and class probabilities using features from the entire image to predict each bounding box.

2.6.3 YOLO family algorithm

As described in the previous section, the YOLO family is a class of deep learning algorithms used for computer vision, exploiting the instance segmentation process for object detection, [20]. The main steps are:

1. YOLO takes the input image and splits it into square grids;
2. On each grid, image classification and localization are applied, with only one object per grid;
3. Each grid is responsible to detect the object for which its centre falls into it;
4. For each grid, some bounding boxes with confidence scores are predicted.

The confidence scores reflect how confident the model is about an object contained in a bounding box.

The YOLO network design is inspired by the *GoogLeNet* model for image classification, Figure 2.14, which is a variant of the *Inception Network*, an important milestone in the development of CNN classifiers through deep learning [21]. Its network has 24 convolutional layers used as features extractors, followed by 2 fully connected layers used to predict the coordinates of bounding boxes. Convolutional

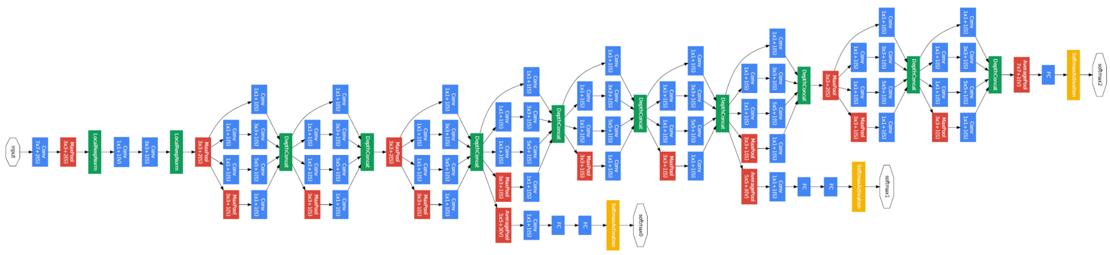

Figure 2.14: *GoogLeNet* network for image classification, a variant of *Inception Network*: the blue boxes are convolutional layers and fully connected layers, the red ones are max-pooling layers, the yellows are softmax activation layers.

layers are pre-trained with the *ImageNet* image database, [22], at half the resolution and then double the resolution for detection. YOLO uses a linear activation function for the final layer and leaky ReLU for all other layers.

YOLO might have some limitations as it imposes a maximum of 2 bounding box predictions for each grid cell, which limits the number of nearby objects that the model can predict. Additionally, it struggles with small objects that appear in groups, such as flocks of birds.

Other variants include *Fast YOLO*, which uses 9 convolutional layers instead of 24, and *YOLOv2* and *YOLOv3*. *YOLOv2* uses 5 max-pooling layers and a softmax layer instead of the *Darknet architecture*. It includes batch normalization, which leads to significant improvements in convergence and eliminates other forms of regularization, and drops out layers without overfitting. Overfitting occurs when a model is trained too well on the training data, to the point that it starts to fit the noise in the data rather than the underlying patterns. Consequently, it performs very well on the training data but generalizes poorly to unseen data. It can occur when a model becomes too complex and captures both the signal and the noise in the training data, or when there are too few training examples, [23]. Additionally, it runs a *K-Means clustering* algorithm on the bounding boxes to get priors or anchors, which are a set of predefined bounding boxes of a certain height and width used to capture the scale and aspect ratio of specific object classes to be predicted. This is useful to detect multiple objects, objects with different scales, and overlapping objects, improving the speed and efficiency by avoiding the need to scan the image using a sliding window.

Finally, *YOLOv3* uses 9 anchors, employs logistic regression instead of the softmax function, and uses *Darknet-53* with 53 convolutional layers.

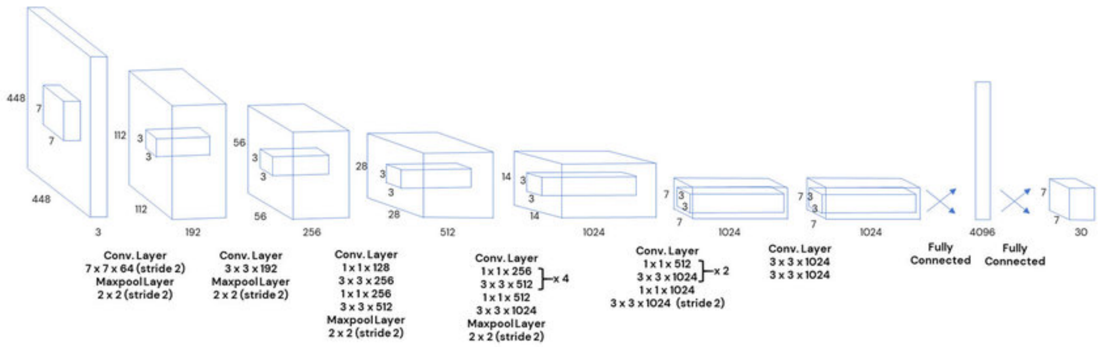

Figure 2.15: *YOLOv3* network schema.

2.6.4 Image and semantic segmentation in pest detection

According to [24], between 2010 and 2020, China, the USA, Europe, and Brazil published 980 studies about computer vision in precision farming. In 78.3% of these studies, the goal was to identify crop pests, with only 3 papers targeting beneficial insects, bee species, and predatory species. Thanks to the increasing trend of deep learning methods, such as CNNs, it is now possible to identify up to 28 classes of insects in entomology observations.

The evolution of pest image segmentation algorithms follows the evolution of computer vision methods from image segmentation to semantic segmentation. As described by Kasinathan et al., of several ML algorithms analyzed, CNNs achieved more than 90% precision with 24 classes on insects, while other classifiers such as

NN, SVM, KNN, *k-nearest neighbors*, and NB, *naive Bayes*, achieved classification accuracy rates between 30% and 75%.

Later, Li et al. [25] compared the most effective deep learning methods in semantic segmentation. They compared the Faster R-CNN, Mask R-CNN, and YOLOv5 algorithms, and all obtained classification accuracy rates higher than 97%. The comparison was performed across several databases: for the *Baidu AI insect* database characterized by simple background images, YOLOv5 was the best method, while for the *IP102* database with complex backgrounds and abundant categories, the optimal solution was Faster and Mask R-CNN.

With the spread of Platform as a Service, PaaS, systems in recent years, the common goal is to provide mobile applications for pest recognition based on cloud computing. An example is presented by Esmail Karar et al. [26], who provided a mobile application able to identify pests such as Aphids, Cicadellidae, Flax Budworm, Flea Beetles, and Red Spider with an accuracy of 99%, exploiting the Faster R-CNN algorithm.

Chapter 3

Case study description

The object of this thesis is detecting and quantifying the presence of *Popillia Japonica Neuman* that have attacked a vineyard. This is part of a more general goal, which is to monitor a field and provide a hint to the farmer on how to act in advance in case the pests start to spread. The way used to achieve this aim is through advanced technologies, such as remote sensing, satellites, drones, and AI algorithms. The research was performed near Novara in Piedmont, Italy, thanks to the technical support of the DIATI department of the Politecnico di Torino and the collaboration with the *Consorzio Tutela Nebbioli Alto Piemonte*.

3.1 *Popillia Japonica* in Italy

The *Popillia Japonica* is known to the EPPO, *European and Mediterranean Plant Protection Organization*, and it was reported in the CE 2000/29 directive by the EU. This insect is also present in the central-eastern United States, Canada, Central Europe, Japan, and some Indian areas. In figures 3.1, 3.2, and 3.3, some plots, reported by [27], show the occurrence of the insect in terms of month, year, and geographical area.

The Piedmont region website [28] reported that in July 2014, in the *Parco Naturale del Ticino*, the beetle *Popillia Japonica* was found for the first time. This insect is native to Japan. Since its first detection, many monitoring, containment, and prevention activities have started. To fight the insect sustainably, farmers and phytosanitary authorities have used traps, treatments, and biological insecticides based on entomopathogenic nematodes. However, these actions have failed, and the pest is present in Piedmont. It attacks all of the eastern region. The area is shown in Figure 3.4.

Consequently, the Italian government issued the D.M. 22 Gennaio 2018 about some emergency measures to avoid the spread of the pest within the national borders.

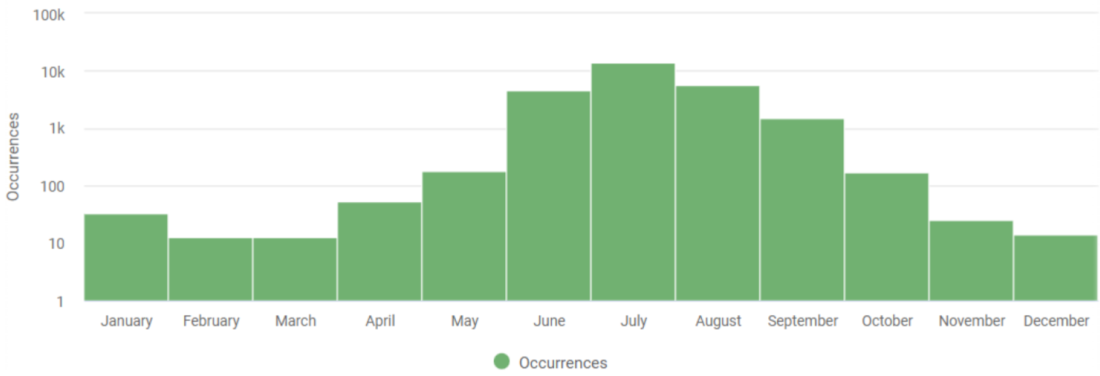

Figure 3.1: As shown in this plot, the population increases by the end of May and strongly in June and July. During August, September and October the population decreases.

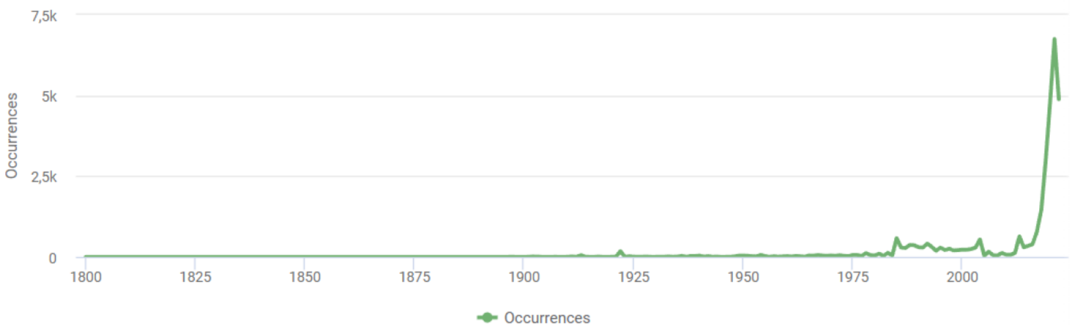

Figure 3.2: The number of occurrences per year increases exponentially since the beginning of the twenty-first century.

In this document, some risk areas were defined, and guidelines were provided on how to control and monitor larvae and adults, eventually using phytosanitary treatments. The government also allocated economic aid for farmers.

3.2 The Novara province vineyards

The Novara province is located in the north east of Piedmont, as shown in Figure 3.5, and it is bordered by the Ticino and Sesia rivers. Physically, the region is characterized by hilly areas covered by vineyards and flat areas of the Po valley. The elevation ranges from 150 to 300 meters above sea level.

The research was performed in Ghemme and Briona, Figure 3.6, between the beginning of June 2022 to the end of July 2022. They are known for the production

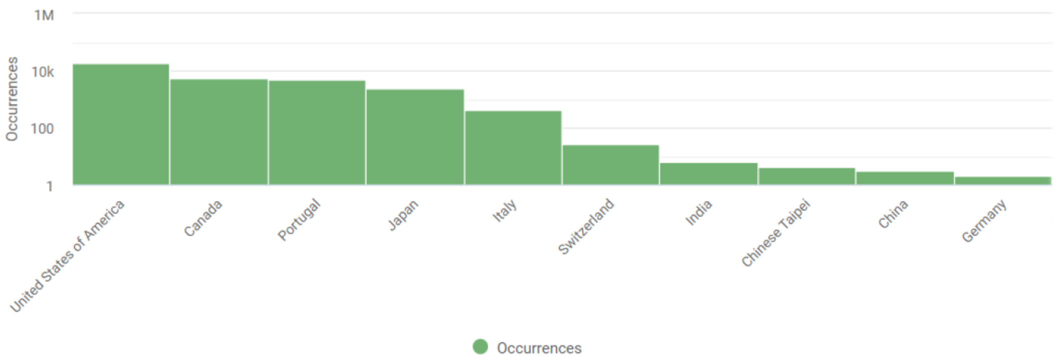

Figure 3.3: The geographical spread of the insects: the USA is the country with the higher number of occurrences, then Canada, Portugal and Japan have a similar pests diffusion, then Italy has several attacked sites between 100 and 1000.

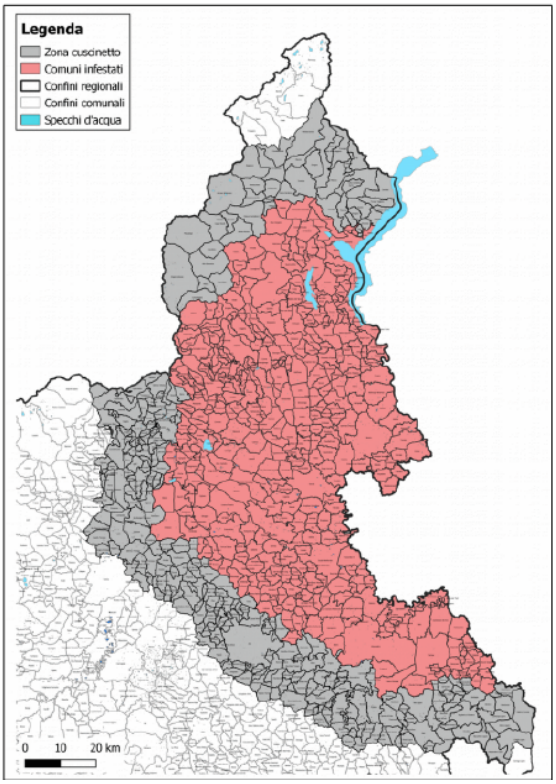

Figure 3.4: Maps obtained by ATTO DD 872/A1703B/2021 from the Piedmont government, updating the delimited area for *Popillia Japonica* presence.

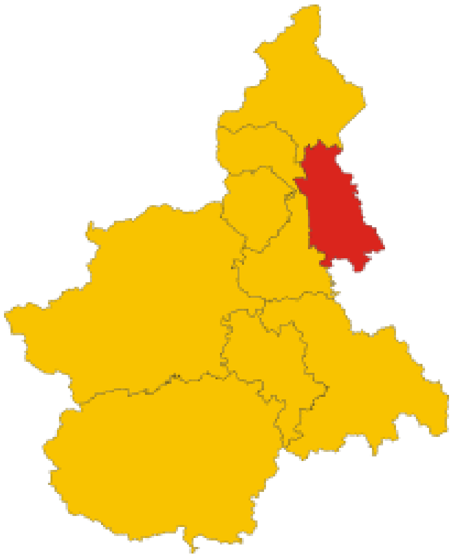

Figure 3.5: Piedmont region, in red the Novara province.

of *Ghemme DOCG* wine, which is obtained from *Nebbiolo* grapes, 85% and from *Vespoli* or *Uva Rara* grapes, 15%. The wine has a minimum alcohol volume of 12.0%, an intense garnet red color, and a tannic and particularly structured taste.

The Ghemme vineyards are located along slopes or on the top of low hills,

Figure 3.6: A Landscape of Ghemme city vineyards.

which sometimes end with wide highlands. Woods and lawns are widely spread, contributing to a humid micro-climate that provides a great environment for the growth of *Popillia Japonica*. The weather is characterized by warm summers, between 17°C and 31°C, and very cold winters, between -3°C and 10°C. The wettest months are April, May, October, and November, and the sky is mostly cloudy. Despite this, the climate is changing due to global warming. In 2022, the rains diminished and were concentrated in a few days at the end of June. Additionally, the average temperature increased in spring and summer. In these specific weather conditions, the spread of the analyzed pest was concentrated in a short period after the rain, thanks to a high level of humidity and warm temperature.

3.3 *Popillia Japonica* entomology

As previously described in Section 1.2, exotic pests are one of the most invasive challenges that farmers have to deal with in the near future. In order to develop an effective ICT solution to counteract the spread of these insects, studying their entomology and phenology is necessary. From this research, all the technical steps and studies required to achieve the goal will be derived.

In Figure 3.7, a larvae and an adult are shown. The insect appears in green and brown on the head, and green and coppery on the back. Finally, it has tufts on the backside that allow it to be distinguished from other beetles, like *Anomala Vitis*, *Mimela Junii*, *Cetonia Aurata* and *Anisoplia Agricola*, shown in Figure 3.7. It is between 8 and 12 millimetres long and between 5 and 7 millimetres wide. The larvae are about 30 millimetres long, and their color changes between white and light brown.

It is a pest, and adult specimens can attack different crops such as fruit trees, vines, hazel trees, as well as some cereals and vegetables. An example of its impact on a vine is shown in Figure 3.8. They feed on leaves and fruits, reducing the plant's photosynthetic capacity and ultimately causing a reduction in production. On the other hand, the larvae, which are born in the soil, can harm turf and pastures by feeding on their roots. The *Popillia Japonica* sleeps during the night and awakens with the heat of the sun. At sunrise, it tends to stay stationary. Even though stressed, it starts to fly when the surrounding temperature reaches a certain value. This is the moment in which treatment and data collection can be done.

3.4 *Popillia Japonica* in the literature

In the literature, several proposed solutions aim to avoid the spread and negative effects of *Popillia Japonica*. Some solutions adopt a mechanical approach, such as traps, while others exploit bacteria to counter the pest's diffusion. The common goal is to avoid the use of insecticides that can have negative effects on other insects, crops, and the soil. For example, in their work [29], Pinero and Dudenhoeffer proposed a novel mass trapping system that captured the *Popillia Japonica* in elderberry and blueberry orchards in Missouri, USA. Over 3 years, 10.3 million

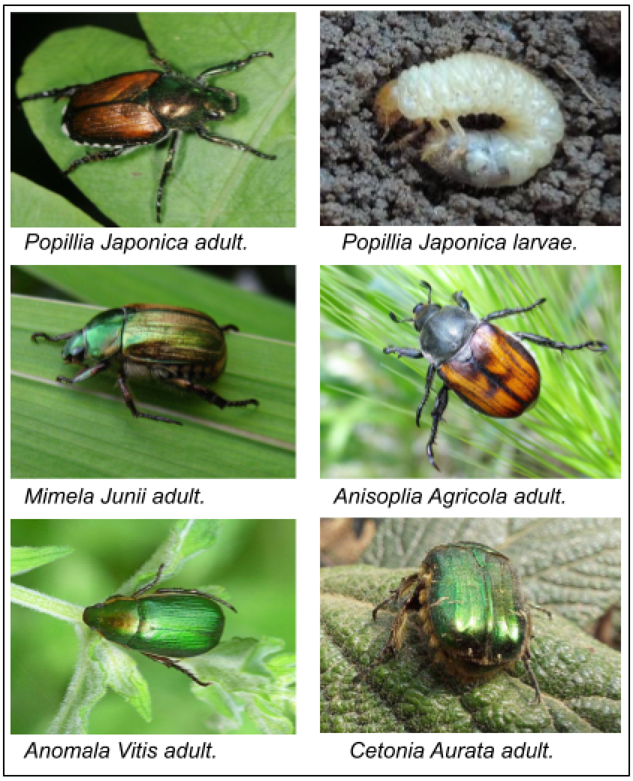

Figure 3.7: Images of the *Popillia Japonica Newman* adult and larvae and other similar insects.

Figure 3.8: In the picture is clear the presence of some insects and the loss of lushness of the leaves.

insects were collected, reducing the damage. In contrast, Marianelli et al. proposed an integrated pest management approach to improve pest control and monitoring in [30]. LLIN, or *Long-Lasting Insecticide-treated Nets*, are used as an attract-and-kill strategy. L. Marianelli developed different nets characterized by different insecticides, and the results showed that the viability of the Japanese beetle was strongly affected by the exposure to LLINs. After 24 hours, all the treated individuals were affected, and often 100% of the pests were killed, even after a short exposure.

An alternative approach is presented in [31], the work of Carl T. Redmond et al. where the pests are fought with target-selective biopesticides. The *Bacillus Thuringiensis Galleriae* has a strong effect on the spread of the *Popillia Japonica*, reducing the insect's leaves feeding for a period between 3 to 14 days. However, it fails to control the pest in the turfgrass soils.

Improving monitoring and control techniques is another common goal against the spread of *Popillia Japonica.* In [32] of Bruce A. Hungate et al., hydrogen isotope records are used to evaluate the geographic position of the insects. The hydrogen isotope precipitation, specific to a particular area, can be transferred to the insects and then analyzed to define the time and the place from which that generation of pests came. This study is useful for verifying the efficacy of insect control and monitoring measurements.

Chapter 4

ICT solutions and methodologies

4.1 ICTs solutions

The thesis approach is multi-scale, considering satellites and UASs, and multi-resolution, considering several remote sensing sensors, in order to test a wide quantity of possible solutions and identify the best one for the proposed case. The research is composed by three main steps, reported by the schema in Figure 4.1.

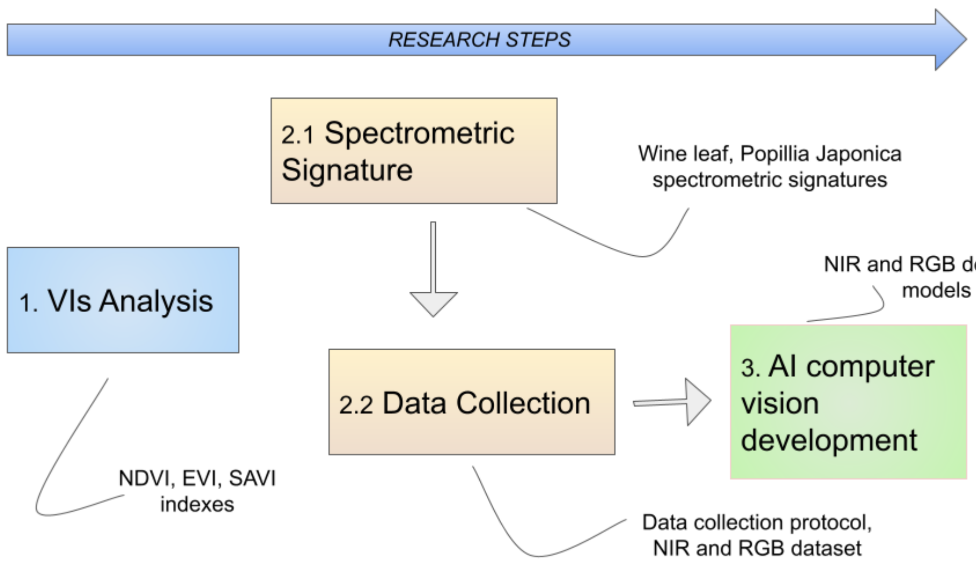

Figure 4.1: Research steps schema.

The *step 1* is the VIs analysis, that gives as result the charts of the NDVI, EVI, SAVI curves. The *step 2* is composed by the spectrometric analysis and the data collection. It gives as results the wine leaf and *Popillia Japonica* spectrometric signatures, the data collection protocol and the NIR and RGB datasets. Finally during *step 3* the AI computer vision detection models are developed.

4.1.1 VIs analysis

The first approach to achieve the thesis goal, the detection and counting of *Popillia Japonica* in the vineyard, is done with satellites, particularly with the data collected by the *Sentinel-2* mission belonging to the ESA *Copernicus* program. Those data does not require the radiometric correction, which is previously applied on them.

The data is available for public and private use through several online services and APIs such as *Google Earth Engine* [33]. It is a SaaS, *Software as a Service*, provided by Google that allows writing scripts and managing a multi-petabyte catalog of satellite imagery and geospatial datasets with planetary-scale analysis capabilities. It provides an online JavaScript dashboard and many Python and JavaScript APIs, as shown in Figure 4.2.

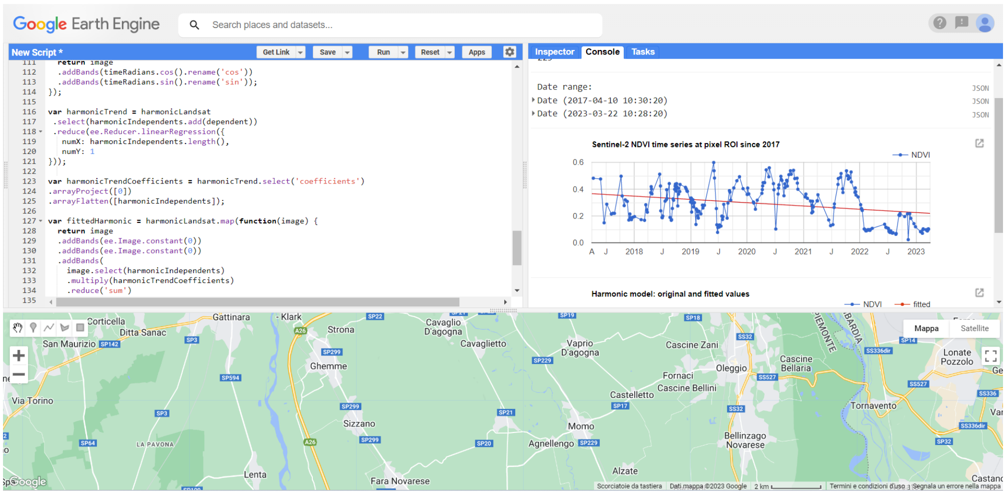

Figure 4.2: The *Google Earth Engine* dashboard, it present a terminal for the output, a map to visualize the results and IDE for coding.

4.1.2 Spectrometric signatures and data collection

The second research step is the vine leaf and *Popillia Japonica* spectral characterizations in order to identify the best practise in the remote sensing data collection.

This analysis is performed with the *Ocean Insights HR spectrometer*, [34]. The active instrument is composed of a halogen lamp that generates a light beam using a collimator. It is then transmitted through an optical fiber to the analyzed object, which finally reflects the light and generates its spectrometric signature detected by the instrument. In Figure 4.3, some steps of the experiment are shown.

Figure 4.3: The miniature spectrometer components are visible: on the left the halogen lamp and the spectrometer, the blue cable is the optical fibre and the black parallelepiped is the collimator. The white circle is used to calibrate the instruments. On the right the spectrometric signature acquisition phase on the *Popillia Japonica*.

After that there is the data collection. Between the beginning of June 2022 and the end of July 2022, a significant amount of data was collected through UAS equipped with multispectral and RGB cameras in several missions carried out in Ghemme and Briona, as detailed in Section 3.2, the locations are shown in Figure 4.4. The basic features of each mission are reported in Table 3.1.

Figure 4.4: Novara city and Ghemme and Briona vineyards geographical position.

The *DJI Matrice 300*, shown in Figure 4.5, is an industrial UAS capable of carrying various sensors such as RGB, thermal, and multispectral cameras. More details can be found on the website [35]. In this thesis, this drone is equipped with a high-resolution multispectral camera and RTK positioning service. This

Date	Time	Location	Camera Support	Camera	Data
3/06/2022	Early morning	Briona (NO) Ghemme (NO)	DJI Matrice 300	Multispectral	Photos
15/06/2022	Early morning	Loc. Pianazze Briona (NO)	6X Sensors Sentera	Multispectral	Photos
17/06/2022	Early morning	Briona (NO) Ghemme (NO)	6X Sensors Sentera	Multispectral	Photos
21/06/2022	Early morning	Loc. Pianazze Briona (NO)	DJI P4 / Autel Nano EVO	Multispectral /HD camera	Photos /video
30/06/2022	Early morning	Loc. Pianazze Briona (NO)	Autel Nano EVO	HD camera	Photos /video

Table 4.1: Table reporting the characteristics of each mission between the beginning of June 2022 to the end of July 2022, in Ghemme (NO) and Briona (NO).

sensor typology is also used in two other solutions: the *DJI P4 Multispectral*, a UAS equipped with RTK positioning service, and the *Sentera 6X Sensors*, a hand carried camera. The first one is shown in Figure 4.6, and more details can be found on the website [36], while the second one is presented in Figure 4.7 and more details can be found on the website [37]. Finally, during the missions, the *Autel Evo Nano* is also used, which is a 249-gram drone with a high-resolution RGB camera capable of capturing both photos and videos. More details can be found on the website [38].

All the data were captured in early morning, all the missions were performed in June and the acquisitions were done in similar light conditions. It is possible to conclude that no radiometric correction is required.

4.1.3 AI computer vision development

The data collected during the missions are then processed and exploited as training and testing images for the semantic segmentation algorithm, which has the goal of identifying and counting the *Popillia Japonica* presented in the vineyard. The computer vision algorithm identified is *YOLOv3*, which is available on GitHub [39]. Its neural network is composed of 52 convolutional layers with different filter values, ranging from 32 to 1024, and different activation functions such as mish, linear, and leaky, as well as some max pool layers. The anchors are 9.

AI algorithms require high computational power, so in order to train the computer vision algorithm, it was necessary to subscribe to *Google Colab Pro*, a PaaS, *Platform as a Service*, that offers an online software development platform with strong computational power through GPU *Graphical Processing Unit*, [40], it is shown in Figure 4.8.

Given the semantic segmentation algorithm and the cleaned dataset, the main

Figure 4.5: The *DJI Matrice 300* used during the 3 June 2022 mission, fitted with high technological equipment it is able to fly for more than 50 minutes, more than 10 kilometres far away and carrying a payload of a few kilograms.

Figure 4.6: The *DJI P4 Multispectral* used during the 21 June 2022 mission, fitted with multispectral equipment it is a high performance UAS with about 30 minutes fly time and 7 kilometer of transmission distance, [36].

steps needed to obtain a working computer vision solution are reported below:

1. Dataset labeling and subdivision into training and testing datasets;

2. Algorithm training and parameter tuning;

3. Algorithm testing and performance metrics computation.

Dataset labeling means creating files with the label and the bounding box coordinates for each pest in each photo. There are many software tools to do this, and each of them provides a different format for the information detected, based

Figure 4.7: The *Sentera 6X Sensors*, used during the 15 and 17 June 2022 missions, is a hand-carried sensor, it is possible to sense closer to the studied object, [37].

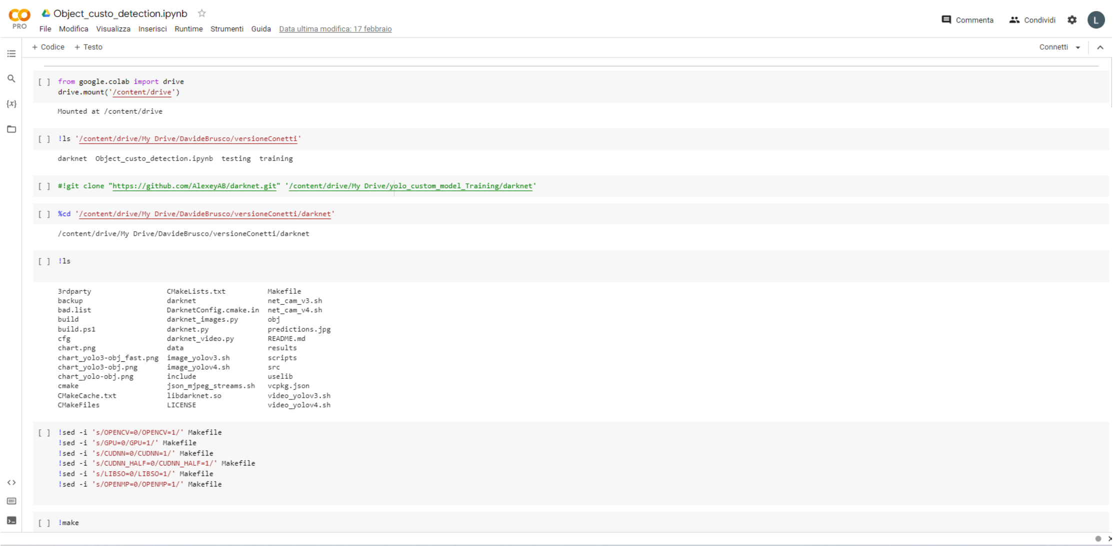

Figure 4.8: The *Google Colab Pro* notebook dashboard.

on different detection algorithms used. In order to use the chosen algorithm, a Python GitHub application is used [41]. The layout is reported in Figure 4.9.

The training phase is a fundamental step used to train the network and indeed to compute its weights, that are the parameters that connect the neurons in one layer to the neurons in the next layer. These weights represent the strength of the connections between the neurons, and they are adjusted during the training process to optimize the network's performance. The weights are initialized randomly before training and are learned through the backpropagation algorithm. By adjusting the

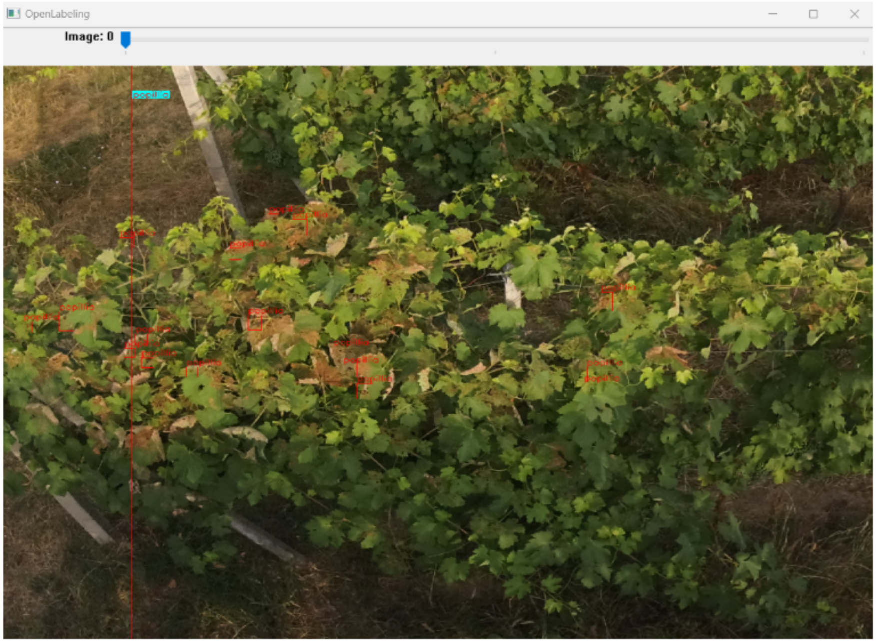

Figure 4.9: Git Hub application layout used to label the images. In each red square is contained an insect.

weights, the neural network is able to learn and to make predictions for the input data it is trained on, and generalize to new input data, [23].

The main performance metric used is a loss function which represents the difference between the predicted output and the actual output for a given input example during the training phase. The loss function value represents the error of the network's predictions and is used to update the network's weights in order to minimize the error. There are many different loss functions that can be used in NN, each with different properties and intended uses. For example, some common loss functions include MSE, *Mean Squared Error*, binary cross-entropy and categorical cross-entropy. The choice of loss function depends on the type of problem being solved and the desired characteristics of the network's output, [23].

4.2 Methodologies

4.2.1 VIs analysis

In this section, how the *Sentinel-2* constellation is used for computing VIs is showed, with the final goal of understanding if it is possible to recognize a field hit by *Popillia Japonica*. Algorithm 1 reports the developed script launched with *Google*

Earth Engine, common to all the VIs computation. Then, specific VI formulas are shown in Algorithms 2, 3, and 4. The script inputs are:

1. The interested region's geographic coordinates, centered at 45°32'36" Nord and 8°29'45" East;

2. The cloud filter values, in this specific case equal to 10, used to filter images covered by clouds;

3. The start date, January 2018, and end date, June 2022, of the analysis.

The region of analysis is a polygon, as shown in Figure 4.4, obtained by a list of coordinates that define the vertices of the figure, Line 2. The collection is then filtered considering the cloud filter values and the period of analysis, Line 4. Finally, the desired VI is computed, Line 6, which is exploited to generate a regression model. The complete used codes are reported in Appendix A.

Algorithm 1 General VI computation.

▷ Inputs list of geographic coordinates, cloud filter values, start date, end date
region = instance of a polygon defined by the list of geographic coordinates
dataset = instance of image collection recovered by *Copernicus/S2_SR*
$dataset_{filtered}$ = filtering images considering cloud filter values, start data and end date
$VI = expression_{VI}$
$dataset_{new} = dataset_{filtered} + VI$
model = VI regression modelling through time independent variable
▷ Plotting the $dataset_{new}$ and the *model*

Algorithm 2 NDVI expression function.

$NIR = image.('B8')$
$Red = image.('B4')$
return $(NIR - Red)/(NIR + Red)$

Algorithm 3 SAVI expression function.

$NIR = image.('B8')$
$Red = image.('B4')$
return $(NIR - Red)/(NIR + Red + 0.5) * 1.5$

Algorithm 4 EVI expression function.

$NIR = image.('B8')$
$Red = image.('B4')$
$Blue = image.('B2')$
return $(NIR - Red)/(NIR + 6 * Red - 7.5 * Blue + 1) * 2.5$

4.2.2 Spectrometric signatures

In the following section, how the *Ocean Insight HD spectrometer* is used to obtain spectrometric signatures of vine leaves and pests is discussed. The obtained results will be used to determine whether a multispectral camera can promptly distinguish between the pest and the vine when scanning the vineyard.

Theoretically, vegetation exhibits a strong reflection in the NIR bandwidth as it is not useful in the plant's growth cycle, and it helps to avoid excessive warming and chlorophyll evaporation. In contrast, it absorbs the incident radiation in the visible red band, which is used in chlorophyll photosynthesis. However, literature lacks specific articles about the spectrometric sign of a beetle obtained with incident light. Nevertheless, we expect a low reflection in the NIR bandwidth due to the general warming of insects caused by sunlight.

The instrument is configured as reported in Table 4.2. It needs to be calibrated with a white circular sample, as shown in Figure 4.3. This practice is used to eliminate noise generated by the electronic parts of the instrument. After calibration, the light beam is used to measure four spectrometric signatures: on the vine leaves, on the head and body of the pest, and on the whole insect.

Parameter	Value
Sensed Bandwidth	from 350 nm to 1000 nm
Resolution	1.33 nm
Frequency Scan	4500 scans per second

Table 4.2: The *Ocean Insight HD spectrometer* input parameters.

4.2.3 Data collection

For the thesis purpose, an existing dataset containing *Popilla Japonica* images is firstly searched. One supplied by GBIF, *Global Biodiversity Information Facility*, [27], with more than 23,000 images was found but those are not fit with the research goal, because the insects are contextualized in different environments and they are too much zoomed. Some examples in Figure 4.10.

Figure 4.10: 3 GBIF datasets' images, the background change for every photo and the insect size and resolution are not comparable with the ones obtained with drones.

Due to the fact that the dataset is not suitable for the thesis purposes, it is necessary to create a custom dataset. As described in Section 4.1.2 and in Table 3.1, lots of images and videos were collected.

During the first mission on June 3, 2022, some test images were collected using the *DJI Matrice 300* fitted with high-performance equipment, such as a multispectral camera and RTK positioning system. This UAS allows the user to plan at priori the flight that the drone will perform autonomously. The time required is between 20 to 30 minutes to capture several photos of the target vineyard.

During the missions dated on June 15 and 17, 2022, the images were collected with *Sentera 6X Sensors*, which allows for the capture of high-resolution photos close to the wine rows. The difficulties faced are reported below:

1. The sensor has a 1.5 m focus distance to the scanned object and the wine rows' width is between 1.8 m and 2 m;

2. The sensor needs a battery supply and a device on which the control and management dashboard is launched. At least 2 people are required.

In the mission on June 21, 2022, the collection tests were performed with a *DJI P4 Multispectral* armed with a multispectral camera and with the *Autel Nano EVO*. The multispectral drone allowed for covering a wide area in less time than the *Sentera 6X Sensors* with only one UAS pilot. On the other hand, the second drone is small, fast, and manageable, used to collect both images and video.

During the mission on June 30, 2022, several tests were done with the *Autel Nano EVO* with the purpose of defining the flying protocol. In order to do that, several parameters such as flying height, the camera inclination, the flying speed, and the flying direction were tested.

The collected data needs to be classified and pre-processed before being used in a precision farming algorithm or service. Firstly, all the wrong photos are deleted, and the video is cut in those parts that are useless, such as landing and take-off phases. Finally, from the described missions, the following were captured:

1. About 150 photos and 15 minutes of video from the *Autel Nano EVO*;
2. About 100 photos from the *DJI P4 Multispectral*;
3. About 400 photos from the *Sentera 6X Sensors*;
4. About 200 from the *DJI Matrice 300*.

The data is classified by date and location, Ghemme or Briona, then considering the type of sensors or UAS and finally by format, photo, or video. In the end the images are labelled exploiting the tool reported in Figure 4.9.

4.2.4 AI computer vision development

As described in Section 4.1.1, the algorithm used for detecting the *Popillia Japonica* in the vineyards is the *YOLOv3* with the *Darknet network*.

The framework used is plug and play: the images need to be uploaded and the training phase launched, tuning the network hyperparameters and setting some configuration files. Finally the testing phase is run. This framework needs a *Google Drive* connection and a specific folders structure for storing the results and cloning the *Darknet network* git repository, it is shown in Figure 4.11.

Table 4.3 reports the main hyperparameters for the neural network. The batch size is set to 64, while the maximum batch size is 6000. A batch is a subset of the training dataset that is processed at once during the training process. By using batches, the training process can be more efficient in terms of computational resources and memory usage, especially for large datasets. Batches allow the model to update its parameters after processing a subset of the data, rather than waiting until the entire dataset has been processed. This framework is capable of changing the batch size considering the system complexity and the computational resources available, [23].

The learning rate, on the other hand, controls the step size at each iteration while moving toward a minimum of a loss function during the training process. It determines how quickly or slowly the model learns the optimal weights and biases for the problem. If the learning rate is too high, the model may overshoot the optimal weights and biases and fail to converge to the minimum of the loss function. On the other hand, if the learning rate is too low, the model may take too long to converge, and the training process may be slow. In this case, the learning rate is set low due to the small size of the training dataset, [23].

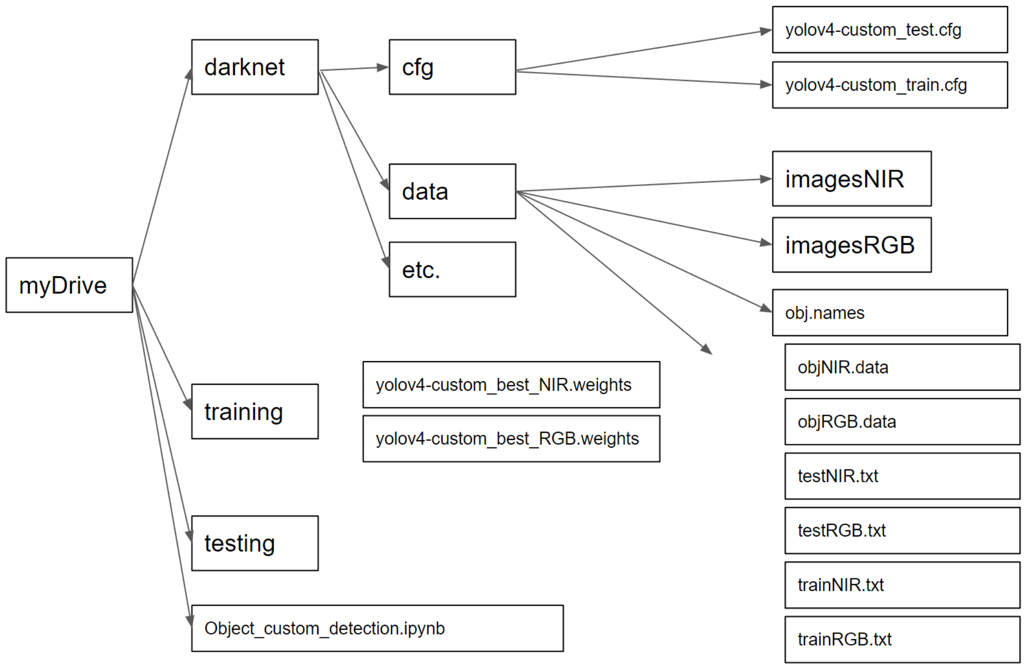

Figure 4.11: Schema of the project folders structure: the notebook file is the main of the project, in the *darknet* folder there are all the NN configuration files and settings, such as the searched label and the paths for the training and testing datasets, while in the training and testing folders, there are respectively the weights files and the NN results.

Parameter	Value
Batch	64
Max Batches	6000
Learning Rate	0.001
Steps	4800,5400

Table 4.3: NN input hyperparameters, those values are set considering the instruction of the used framework and the results of the training phase.

The algorithm exploits the transfer learning technique for which some pre-trained weights are loaded in order to speed up and simplified the training process. In this thesis, two training sessions were performed using RGB images and NIR images, and two weight files were obtained. Each training session required approximately 14 hours of working time.

The goal is to identify which network guarantees better performance. The MSE loss function is used as the performance metric, defined as the average of the squared differences between the predicted output and the actual output over a set of examples N. The relative equation is reported in Equation 4.1.

$$MSE = \frac{1}{N} \cdot \sum_{i=1}^{N} (y_{predicted_i} - y_i)^2 \tag{4.1}$$

The testing phase consists of two steps: in the first step, the trained NNs are tested with random RGB and NIR images to generally understand the NN performances, while in the second step, the testing campaign is extended. During the test phase, the algorithm provides the input image with the identified pests labeled and for each detection, a confidence percentage is provided, which represents how reliable the result is. For each output image, the number of insects detected is also counted, considering only the number of identified pests with a confidence percentage greater than 80%.

Chapter 5

Results and discussion

The application of the *Popillia Japonica* case study resulted in the collection of a large amount of data, which are analyzed in the following chapter to obtain several protocols for identifying and counting pests in a vineyard.

5.1 VIs analysis results

As described in Section 4.2.1, various VIs were computed from the collected data, and the results are shown in Figure 5.1. Upon examining these curves, it is apparent they are quite similar, indicating homogenous vegetation coverage. As there is no bare soil present, the SAVI correction factor is not useful and the EVI correction factor for canopy background adjustment does not improve the results. A period in which there are a few differences is from mid 2022 until early 2023 where the NDVI serie presents lower values. Those were dry months with a few rains and snowfalls, the health of the vegetation coverage is low and the SAVI and the EVI indexes are less prone to that.

However, the plot reported in Figure 5.2 shows the cycle of photosynthetic biomass levels, which correspond to changing seasons. From July to November, the period during which the vineyard grows grapes, the photosynthetic biomass level is higher.

While the *Popillia Japonica* has a significant impact on the reduction of leaf coverage, as shown in Figure 5.3, its effect on the photosynthetic biomass level is not detectable from satellites, as the measurements are also influenced by grass or other plants close to the vine. Therefore, this ICT solution cannot be used for this specific purpose. Alternatively, UAS can be considered as an alternative solution, as they can be driven closer to the vine rows, thus increasing image resolution.

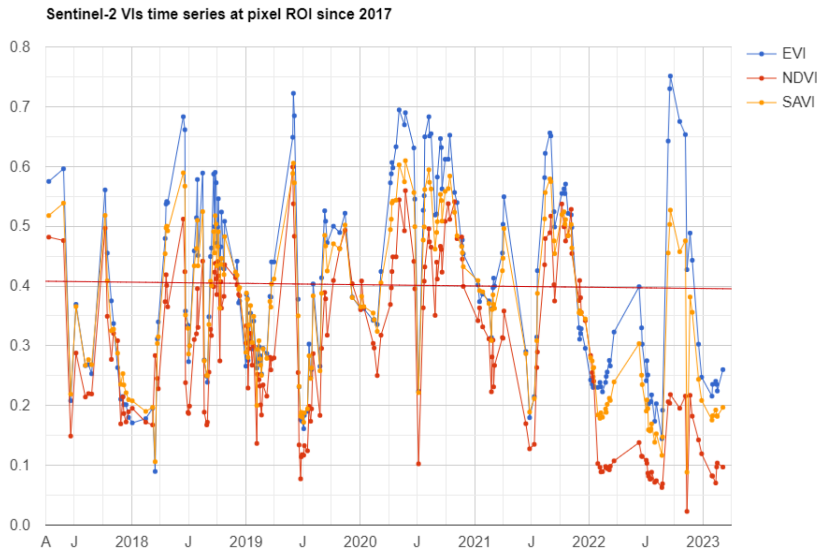

Figure 5.1: In blue the EVI serie, in red the NDVI serie and in yellow the SAVI serie.

5.2 Spectrometric signature resutls

As described in Section 4.2.2, the spectrometric signatures are used to determine if wine leaves and pests are clearly distinguishable in the NIR bandwidth. The results obtained through the spectrometer are shown below. In Figure 5.4, the vine leaf signature is shown, which has a small peak around 550 nm corresponding to the green color and a larger peak between 700 nm and 900 nm. The leaf also has a high reflectance at the red edge and NIR bands.

Regarding Figure 5.5, the pest's iridescent head exhibits low reflectance in the NIR band, indicating that it absorbs most of the light. The resulting spectrometric signature is flat and low, except for a slight increase in the infrared bandwidth, possibly due to the insect's heating. Similarly, the pest's body shows a high absorption factor, but with increasing reflectance between 700 nm and 1000 nm, which defines its red and brown color, as shown in Figure 5.6. Overall, the spectrometric signature of the pest is a combination of the head and the body analysis, resulting in a low-medium reflectance between 700 nm and 1000 nm, as

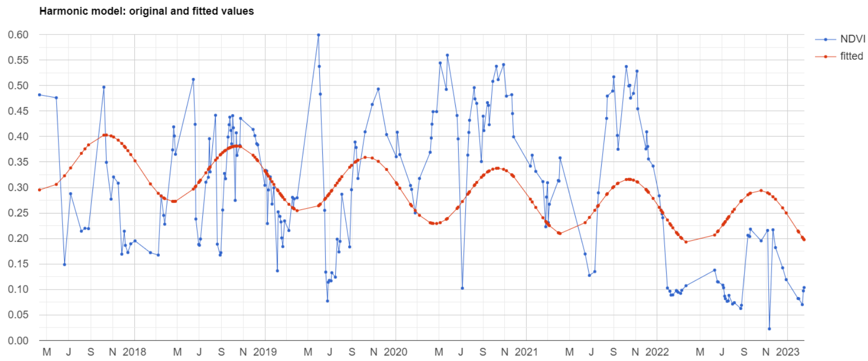

Figure 5.2: In blue the NDVI serie while in red the armonic model derived from the computed values. The NDVI values follow a cyclical pattern meaning the changing seasons.

Figure 5.3: An example of leaves eaten by the *Popillia Japonica* in a vine row.

depicted in Figure 5.7. In the collected images, the insect is very little, only a few pixels represent it, consequently, those represent an averaging between the head and the body spectrometric behaviours. In addition to that, as shown in Figure 5.8, the pests look like black points in the NIR band, this is justified by low reflectance

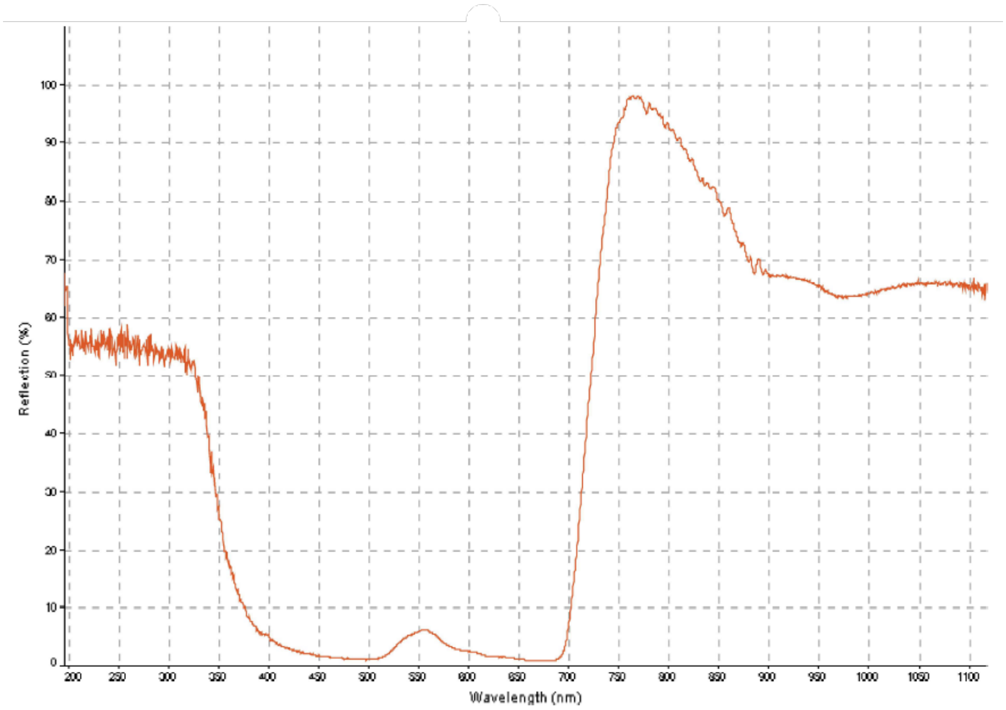

Figure 5.4: Reflectance diagram of a vine leaf.

at these frequencies.
Finally in the NIR band plants and insects have different spectrometric behaviours, so working in that range can help pest detection. In the visible range instead, that difference is not so visible and detecting the *Popillia Japonica* between leaves is more difficult. These differences are visible in Figure 5.8 where the same pictures in both bandwidths are shown.

5.3 Data collection results

As described in Section 4.2.3, in order to collect useful that through UAS equipped with remote sensing sensors, several devices are tested during the missions presented in Table 3.1.

The *DJI Matrice 300* solution did not allow to obtain useful images because the studied object is too far from lens camera and the pest are not detected as exposed in Figure 5.9. The data collected with the *Sentera 6X Sensors* have enough image resolution as shown in Figure 5.10. Anyway it is difficult to cover long distances, such as a wine field, walking and carrying on such kind of sensor. The most promising solution is to exploit the *DJI P4 Multispectral*. This UAS is littler than the *DJI Matrice 300*, that allows to fly closer to the vine rows, it is equipped

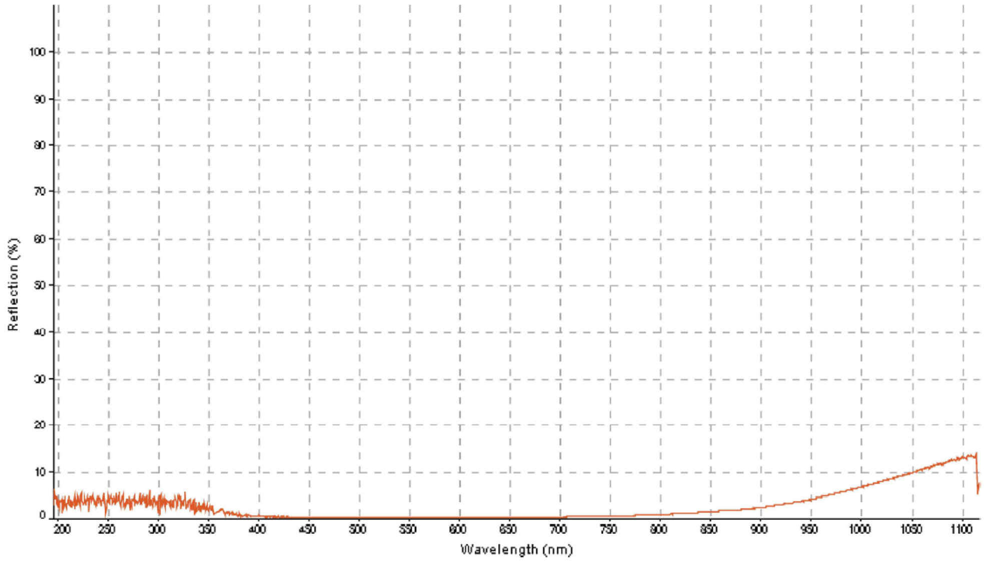

Figure 5.5: Reflectance diagram of the head of the *Popillia Japonica*.

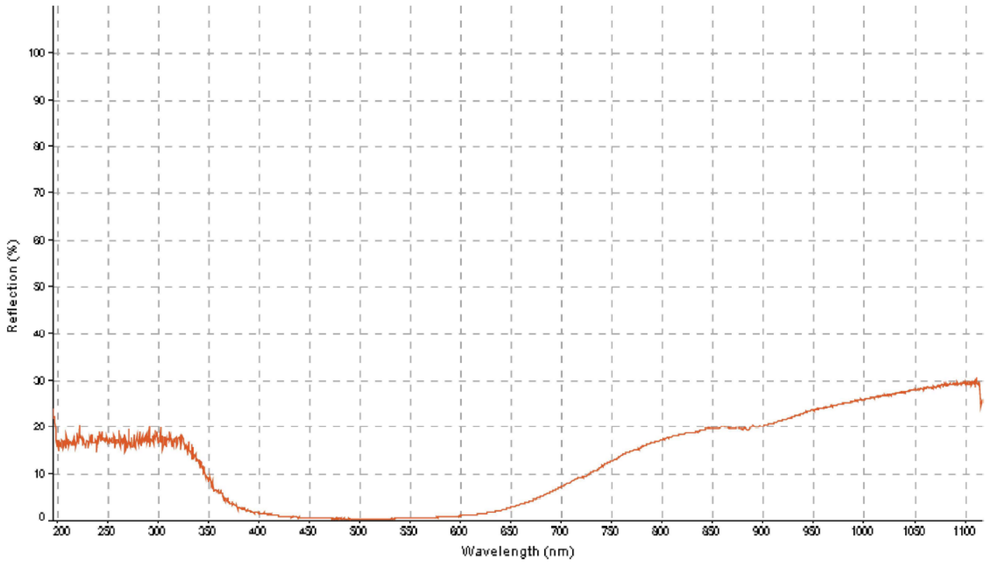

Figure 5.6: Reflectance diagram of the body of the *Popillia Japonica*.

with a high resolution multispectral camera and it allows to cover long distances, 5.11.

The *Autel Nano EVO* is equipped with an RGB camera and can capture both photos and videos. It was specifically used to define the best flying protocol for ensuring correct data collection. Figure 5.12 shows a photo acquired using this

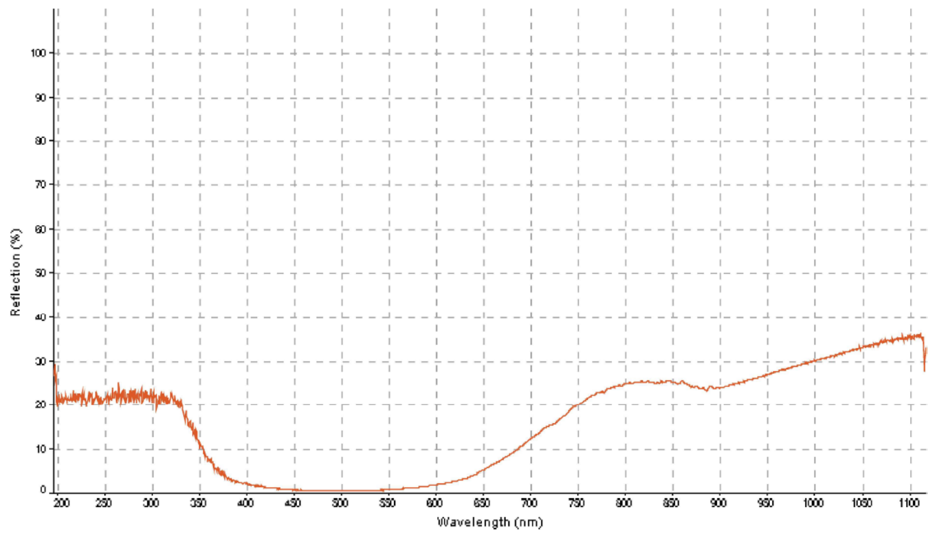

Figure 5.7: Total body *Popillia Japonica* reflectance diagram.

Figure 5.8: On the left the RGB picture, on the right the same picture in NIR bandwidth.

camera, and the features of the defined protocol are reported in Table 5.1. The optimal flying height should be close enough to the vine row to capture the pests in the picture, while also avoiding turbulence produced by the propeller that moves the leaves. Additionally, the flying direction and camera inclination should provide images containing both the top and side of the trees. Finally, the acquisition speed should be adjusted to avoid out-of-focus photos. Although the flying protocol and the NIR band, the images appear quite hard, as well, to be analysed by sight, in fact, the pests are small and the background has different colours and shared areas.

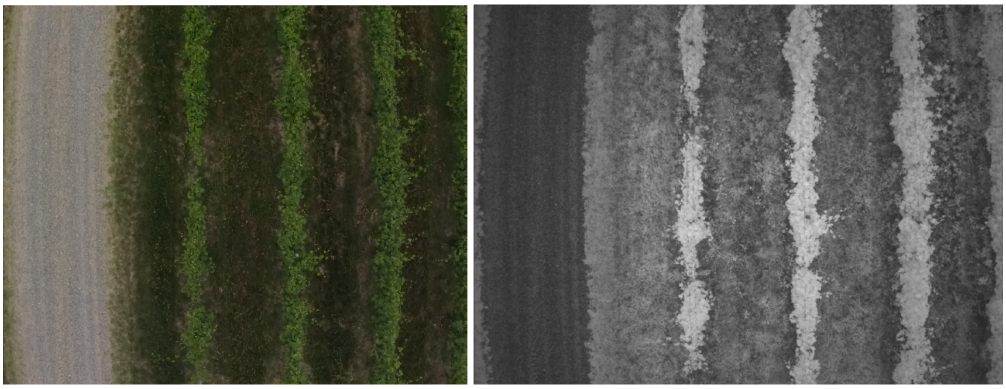

Figure 5.9: On the left an RGB image and on the right a NIR image captured with *DJI Matrice 300*.

Figure 5.10: Images captured With the *Sentera 6X Sensors*.

Figure 5.11: Results obtained with *DJI P4 Multispectral*.

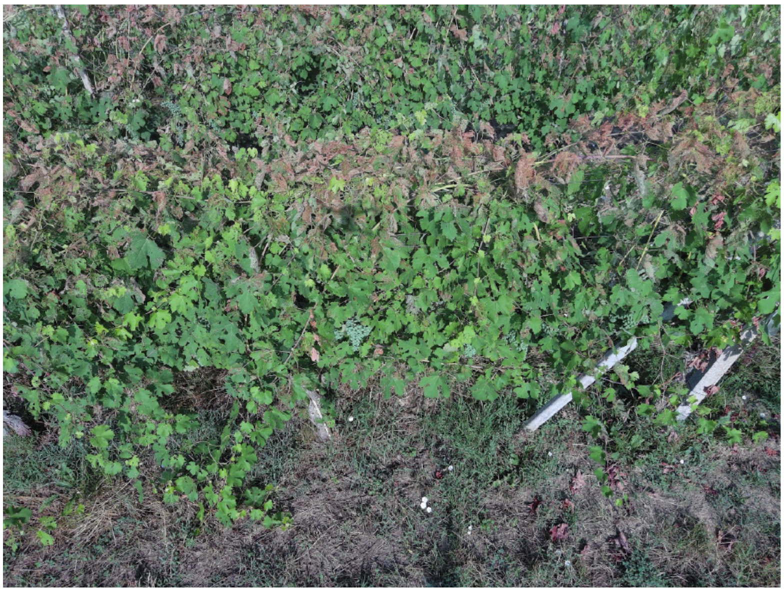

Figure 5.12: This is an example of an image collected following the protocol defined during the missions

Parameter	Value
Flying height	2 m from the wine row top
Flying direction	Perpendicular to the wine row
Camera inclination	45°
Flying speed	0.5-2.5 m/s
Acquisition method	Photos and video
Resolution	FullHD/Ultra HD and 2.7k

Table 5.1: Flying protocol features to collect as best as possible images for the dataset creation.

5.4 AI computer vision development results

5.4.1 Labelling phase

The background in the collected images is complex, as shown in Figure 5.11. Consequently, visually recognizing each *Popillia Japonica* present in both RGB and NIR images is quite difficult, due to the color similarity between the insect and the leaves. However, in NIR images, the labeling is slightly easier because of the strong color difference between insects and leaves. On the other hand, strong light absorption is not an exclusive pest feature.

Labeling is essential, as "garbage in produces garbage out", meaning that providing incorrect input to a system results in incorrect output. To improve the labeling process quality, an automatic system can be used to exploit pre-trained NNs to identify the desired object. Ultimately, two datasets of 72 images each were labeled for the training phase, and four groups of 15 images were labeled for the testing phase.

5.4.2 Training phase

As described in Section 4.2.4, two training phases of the *YOLOv3* algorithm were performed, one on the RGB dataset and one on the NIR one. The interim results for the RGB dataset are reported in Table 5.2 and Figure 5.13, while the NIR results are presented in Table 5.3 and Figures 5.14.

Iteration	Loss	Delta Loss	Time left(h)
600	0.252		11.15
1700	0.126	1.15E-04	12.58
2900	0.073	4.38E-05	5.95
3600	0.065	1.09E-05	4.64
4800	0.044	1.82E-05	2.77
5900	0.043	9.09E-08	0.39
6000	0.043	7.00E-06	0.19

Table 5.2: Interim results of the RGB training phase.

Iteration	Loss	Delta Loss	Time left(h)
1000	12.981		10.07
3000	12.695	1.43E-04	8.38
3600	10.078	4.36E-03	5.36
4200	8.827	2.09E-03	2.47
4800	8.644	3.05E-04	1.83
5400	6.645	3.33E-03	0.41
6000	8.412	-2.95E-03	0.05

Table 5.3: Interim results of the NIR training phase.

In the RGB training, the loss function decreases exponentially from about 0.3 to less than 0.05. During this phase, the *YOLOv3* algorithm learns very fast and

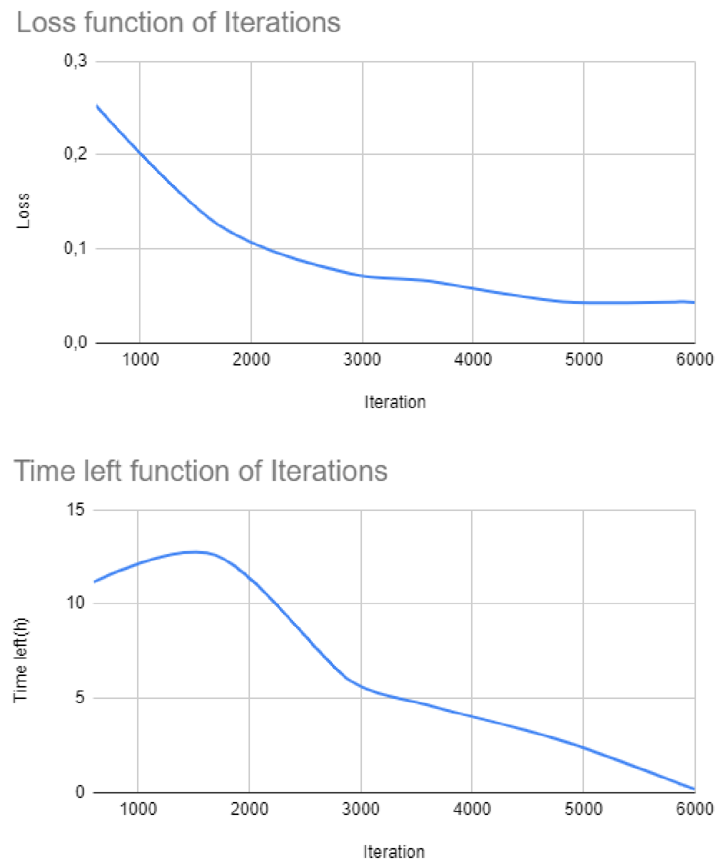

Figure 5.13: Plots on the RGB training phase: above the loss metric function of the number of iteration, below the evolution of the estimated time needed to complete the training.

precisely from the training images. However, this could be a case of overfitting, which is a common problem in machine learning. It can occur when a model becomes too complex and captures both the signal and the noise in the training data, or when there are too few training examples. Both of these situations could be linked to this case.
On the other hand, in the NIR training, the loss function is one to two orders of magnitude higher, and it has a linear behavior. There is no overfitting problem, but the testing accuracy could be low.

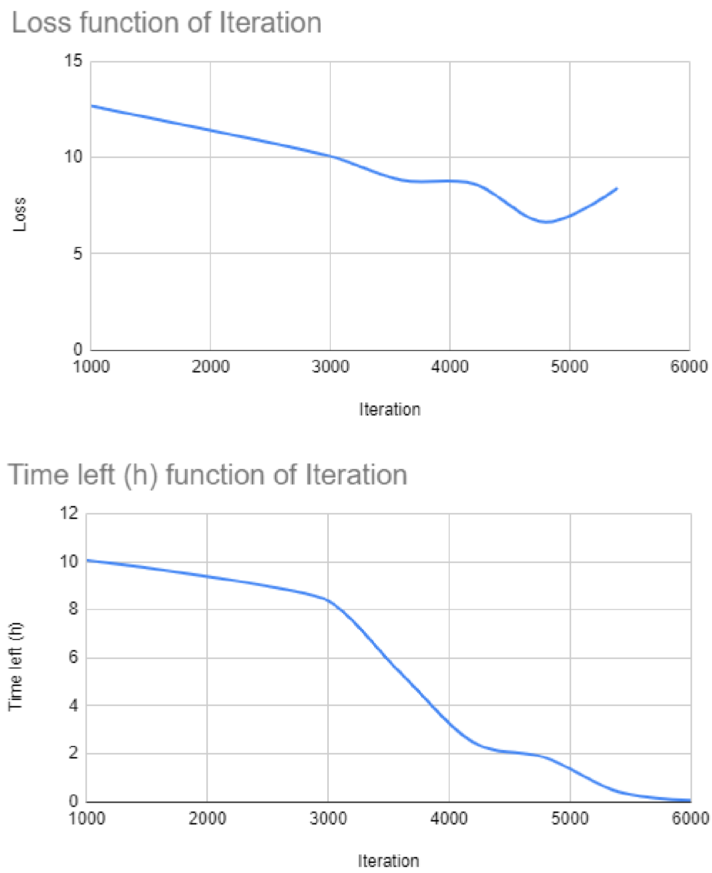

Figure 5.14: Plots on the NIR training phase: above the loss metric function of the number of iteration, below the evolution of the estimated time needed to complete the training.

5.4.3 Testing phase

The testing phase is composed of two steps. In the first step, the general behavior of the trained models is analyzed with different images, while in the second step, we studied only the most promising cases. Table 5.4 reports the tests done during this phase and the variables taken into account.

Considering the RGB tests, as expected, the NN trained with RGB images is not able to detect *Popillia Japonica* in new data. As shown in Figure 5.15, only one insect is detected in a picture recovered by *Sentera 6X Sensors*.

In Figure 5.16, there are no pests, as no insect is recognized by the algorithm, and there are no false positive cases.

On the other hand, the results for the NIR tests are much better, as shown in

Device	Resolution	Size	Band	Pests present	Pests detected
Sentera 6X	5184X3888	2.7 MB	RGB	Severals	1
Sentera 6X	5184X3888	2.6 MB	RGB	None	0
Sentera 6X	.png convertion		NIR	Severals	Severals
DJI P4	1600X1300	856 KB	RGB	Severals	0
DJI P4	1600X1300	742 KB	RGB	None	0
DJI P4	1600X1300	742 KB	NIR	Severals	Severals

Table 5.4: In table are reported the tests performed during the first step: the NIR NN model guarantee better performance.

Figure 5.17, where *YOLOv3* is able to detect several insects in an input test image, which belongs to the same dataset used for training and it shares some features with it. In the end with the same dataset size the NIR trained NN has obtained

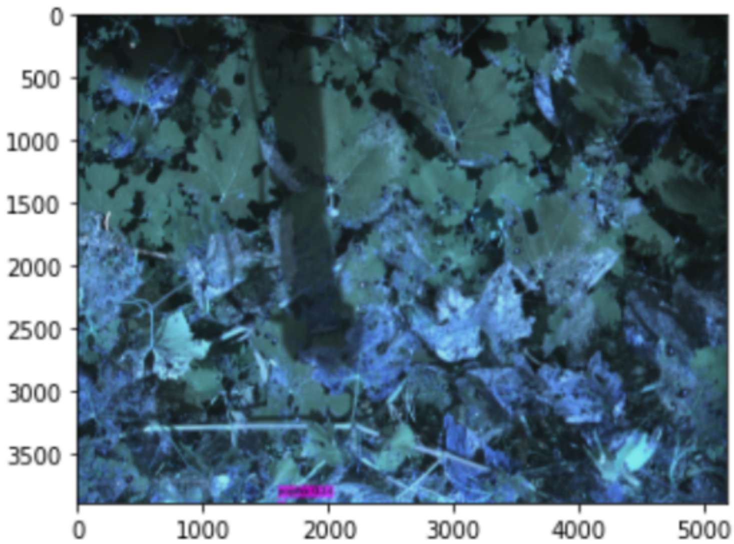

Figure 5.15: The RGB trained NN is able to catch only one insect in the picture, small violet square on the bottom.

much better results than the RGB trained one. Thanks to the complex background and small dataset used, the RGB trained NN is proned to overfitting. The issue is partially fixed with NIR trained NN where the environment is simplified thanks to the lower amount of information carried out by the single NIR band.

In the second step the tests were performed on 4 groups of 15 images each:

1. Train-like NIR images, recovered with *DJI P4 Multispectral*, dated 21 June

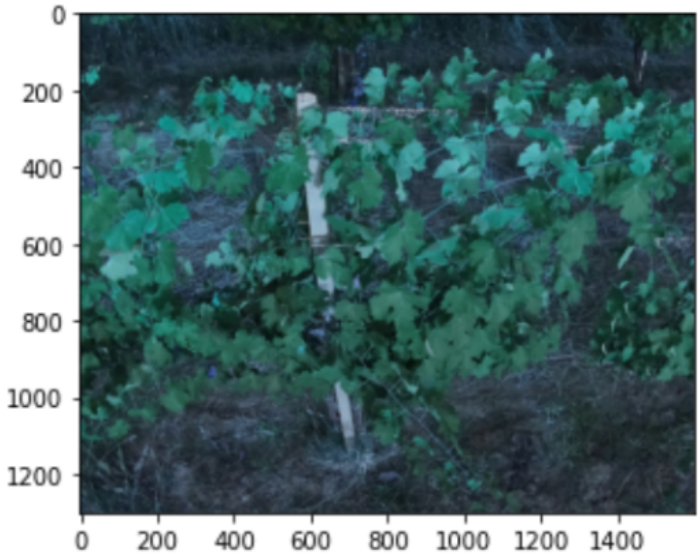

Figure 5.16: Image in which there are not false positive, no insect are present in the picture.

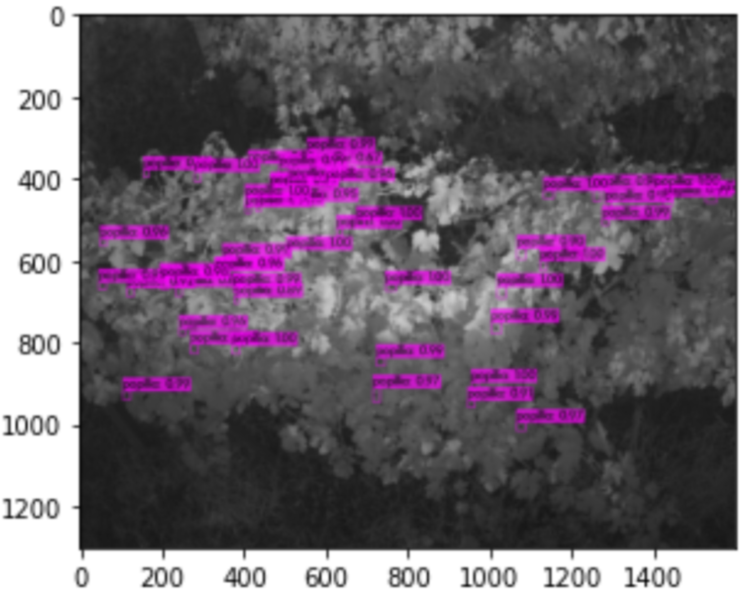

Figure 5.17: The NN detects several pests, represented by the violet squares.

2022, in Briona (NO);

2. *Sentera 6X Sensors* NIR images, dated 15 June 2022, Briona (NO);

3. *Sentera 6X Sensors* NIR images, dated 17 June 2022, Ghemme (NO);
4. *Sentera 6X Sensors* NIR images, dated 17 June 2022, Briona (NO);

In Table 5.5 the percentage of insects identified for each group are reported. A threshold of 80% confidence value is applied on the detected pests.

Dataset	Size	Labelled pests	Detected pests	Detected pests >80%	Perf. tot. (%)	Perf. >80% (%)
Train-like	15	329	170	159	52	48
15/06/2022	15	290	128	54	44	19
17/06/2022 Ghemme	15	703	155	5	22	8
17/06/2022 Briona	15	955	62	15	7	2

Table 5.5: In table are reported the performance recovered from the application of the NIR NN model on each testing dataset.

Analyzing the *Train-like* dataset, example in Figure 5.18, about half of the insects are detected and the 93% of the insect gain a score higher than 80%, as shown in Figure 5.19. This dataset images are similar to the ones used in the training phase.

Next, for the *15/06/2022* dataset, as shown in Figure 5.20, although the total number of detected pests is similar to the previous case, the percentage of images with 80% or higher confidence drops from 93% to 36%. This is due to a change in the background, making it more difficult for the NN to identify the *Popillia Japonica*. Finally, when the NN is applied to the *17/06/2022 Ghemme* and *17/06/2022 Briona* datasets, as shown in Figure 5.21, it exhibits similar behavior to group 2. The confidence score distribution is shown in Figure 5.22. For groups 2, 3, and 4, the percentage of labels with 80% or higher confidence drops below the threshold value, indicating an increase in false positive probability.

5.4.4 Conclusions

In Section 2.6.3, it was noted that YOLO family algorithms limit the number of nearby objects that the model can predict to a maximum of 2 bounding boxes for each grid cell. In the *17/06/2022 Ghemme* and *17/06/2022 Briona* datasets, the number of pests increased significantly, including the relative aggregates, which presented the worst cases for the algorithm, and thus led to lower performance. However, despite this, in the literature, the YOLO algorithm is successfully used in this scenario, achieving results higher than 95% with careful training.

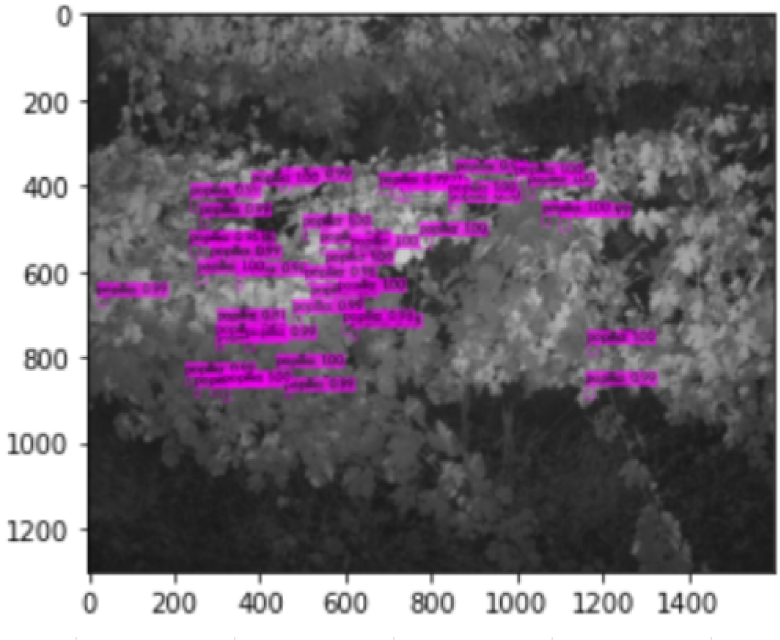

Figure 5.18: Image example of group 1.

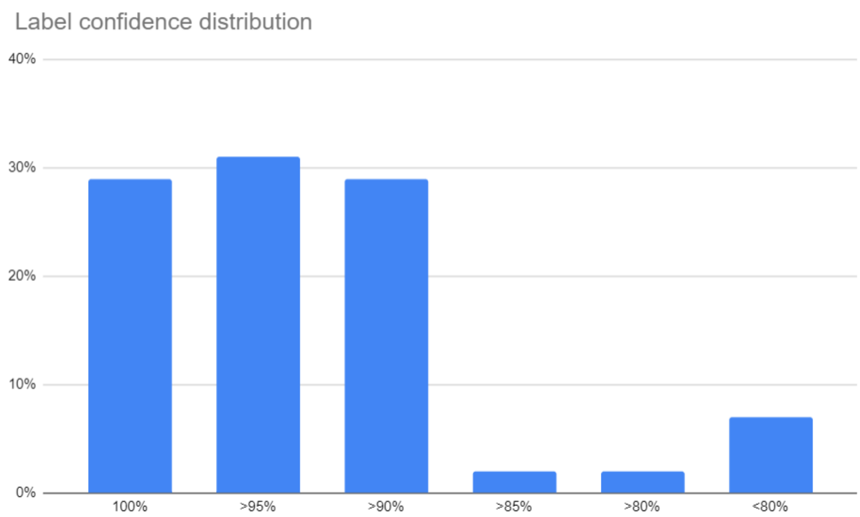

Figure 5.19: Group 1 label confidence distribution.

As previously described, the RGB trained NN is prone to overfitting due to the complex background and small dataset size, and the hand-made labelling can be challenging and inaccurate. Furthermore, the NIR-trained NN has difficulties generalizing to the background of leaves and vegetation. Therefore, to have an

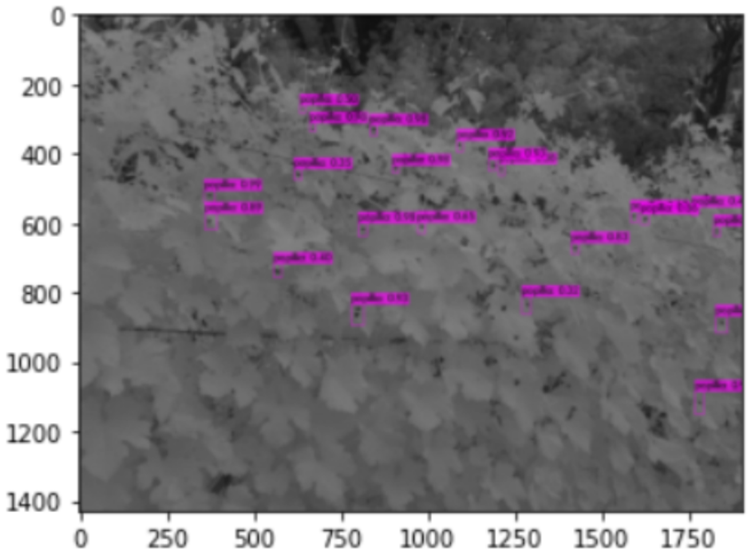

Figure 5.20: Image example of group 2.

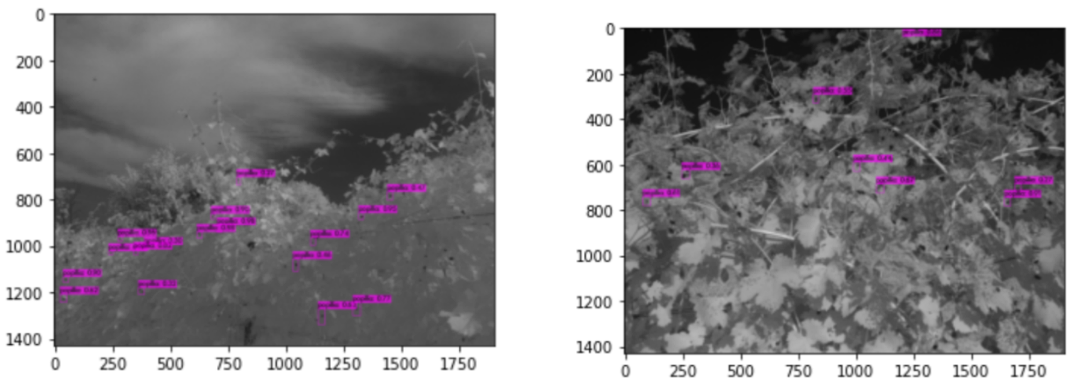

Figure 5.21: Images example of group 3 and 4.

efficient NN training, it is recommended to:

1. Exploit automatic labelling software in order to have more accurate results;

2. It is a good practise, often found in literature, using at least 500 training pictures. If necessary, turn to data augmentation techniques such as flipping, rotating, or scaling images, which can increase the effective size of the dataset and improve the model's performance with a smaller number of original images;

3. Collect photos from different points of view of the vine rows, possibly including small parts of the sky, in order to change the background and reduce the

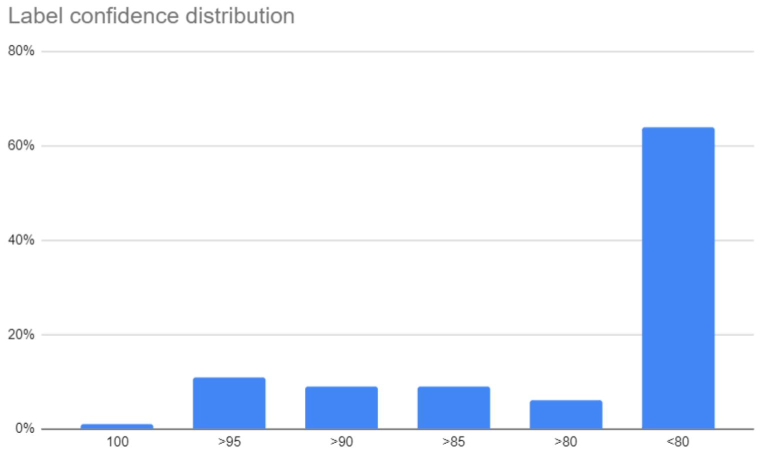

Figure 5.22: Group 2 label confidence score distribution.

probability of overfitting.

In addition, to address overfitting, regularization techniques such as dropout, weight decay, or early stopping can be applied. These techniques help the model to better generalize to new, unseen data and improve its overall performance.

Finally, considering the results of the testing phase, it is possible to conclude that the *YOLOv3* algorithm can be used for insect detection, but adjustments and tips need to be taken into account to have efficient software. NIR band allows for extremely high performance compared to the RGB band, making it easier and faster to achieve optimal NN training.

Chapter 6

Conclusion

ICTs represent a set of innovative solutions that support the farmer's decision-making process and activity management, in particular, they are an effective response to exotic pest fight. The objective of this thesis was to verify if remote sensing combined with AI computer vision can identify and estimate the number of pests in a vineyard attacked by *Popillia Japonica Newman*. To achieve this goal, a VIs analysisis is firslty performed, then through a spectrometer the wine leaf and the pest spectrometric signatures are obtained. Finally UASs are used to collect datasets which are exploited in the AI computer vision algorithm.

From the tests performed it is derived that the VIs analysis cannot identify the effects of the *Popillia Japonica* in a vineyard, due to the low image resolution and large scale. On the other hand, UASs represent an effective way to collect NIR and RGB datasets, consequently a specific data collection protocol is defined. At the same time , the spectral signatures of the wine leaf and the insect were derived and the strong differences between them were exploited to identify and to count the pests with the YOLOv3 algorithm. In particular the NIR model performs better than the RGB model which is affected by overfitting.

This research proves that the described technologies can be exploited to achieve the thesis purposes and provides a list of tips and protocols for data collection and processing, with the aim of speeding up future improvements. From this study, the prominent role of data collection methodologies is apparent: the devices used, the sensors exploited, the acquisition time, and the flying characteristics need to be carefully defined. Equally important is knowledge of the studied environment, the vineyard in this thesis, including which insects could be found in a vine field, which agricultural treatments are done, and the entomological features of the studied pest.

Future developments could be addressed to the improvements of the labeling techniques and AI computer vision performance. For instance, in this work, the RGB and NIR images are analyzed separately without considering the possibility of mixing red, green, blue, and NIR channels. It is supposed that mixing the NIR and the Red channels could fix the non-exclusive *Popillia Japonica Newman* spectrometric behavior problem and improve labeling techniques. Also, data pre-processing can help achieve this goal by cleaning wrong or useless data, which is mandatory to achieve high-quality results. Regarding the computer vision algorithm, expanding and diversifying the data collection could improve the testing confidence score. Moreover, the analysis of the distances between the insects might allow the identification of pest agglomerations and the estimation of the relative component numbers.

The potential of this project is very high. *Popillia Japonica Newman detection through remote sensing and AI computer vision* could be supported by charging stations that enable the UAS to land, takeoff, recharge the batteries, and upload the collected data to a cloud, creating a completely autonomous pest detection system.

Finally, what is expected in the near future is that ICTs will support all activities in the agriculture sector with the goal of improving the quality and quantity of agricultural products while taking into account environmental, social, and economic sustainability.

Appendix A

Popillia Japonica Satellite Case Study

```
// Ghemme (NO) field geographical coordinates
var roi = ee.Geometry.Point([8.490, 45.545])
var list_coordinates = [[8.48395586013794, 45.54224663397568],
    [8.500628471374512, 45.54224663397568],
    [8.500628471374512, 45.55037651450808],
    [8.48395586013794, 45.55037651450808]]
var region = ee.Geometry.Polygon(list_coordinates)
var timeStartField= 'system:time_start';
var timeEndField= 'system:time_end';
// cloud image filtering value
var CLOUD_FILTER = 10

var sentinel2 = ee.ImageCollection('COPERNICUS/S2_SR').filterBounds(
    region)
                //.filterDate(timeStartField, timeEndField)
                .select('B1','B2','B3','B4','B5','B6','B7','B8','
    B8A','B9','B11','B12')
                .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE',
    CLOUD_FILTER));
Map.centerObject(roi, 10);

var addVariables = function(image) {
  var dateStart =ee.Date(image.get(timeStartField));
  var dateEnd =ee.Date(image.get(timeEndField));
  var years = dateStart.difference(ee.Date.fromYMD(1980,1,1), 'year')
    ;
  var NDVI = image.normalizedDifference(['B5', 'B4']).rename('NDVI');
  var EVI = image.expression(
```

```
     '2.5 * ((NIR - RED) / (NIR + 6 * RED - 7.5 * BLUE + 1))', {
         'NIR': image.select('B8').divide(10000),
         'RED': image.select('B4').divide(10000),
         'BLUE': image.select('B2').divide(10000)
     }).rename("EVI");
   var SAVI = image.expression(
     '((NIR - Red) / (NIR + Red + 0.5)) * 1.5', {
         'NIR': image.select('B8').divide(10000),
         'Red': image.select('B4').divide(10000)
       }).rename('SAVI');

   return image
   .addBands(NDVI).float()
   .addBands(EVI).float()
   .addBands(SAVI)
   .addBands(ee.Image(years).rename('t').float())
   .addBands(ee.Image.constant(1));
};

//adding bands for charts creation
var filteredSentinel2 = sentinel2
   .filterBounds(roi)
   .map(addVariables);

//creating a image collection
var sudyarea = sentinel2
    .filterBounds(region)
var final = ee.ImageCollection(sudyarea.map(addVariables));


// printing some statistics
var count = filteredSentinel2.size();
print('Count: ', count);
var range = filteredSentinel2.reduceColumns(ee.Reducer.minMax(), ["
    system:time_start"])//["2022-05-01"])
print('Date range: ', ee.Date(range.get('min')), ee.Date(range.get('
    max')))


// priting charts: in order to change from to NDVI to SAVI or EVI
    just change
// .select('NDVI') with .select('SAVI') or .select('EVI')
var sentinel2chart = ui.Chart.image.series(filteredSentinel2.select('
    NDVI'), roi)
.setChartType('ScatterChart')
.setOptions({
   title: 'Sentinel-2 NDVI time series at pixel ROI since 2017',
   trendlines:{0: {
     color: 'cc0000'
```

```
    }},
    lineWidth: 1,
    pointSize: 3,
});

print(sentinel2chart);

//linear model, in order to change from to NDVI to SAVI or EVI just
   change
// String('NDVI') with String('SAVI') or String('EVI')
var independents = ee.List(['constant', 't']);
var dependent = ee.String('NDVI');

//calculating  a linear trend by "join" the dep variable ((VI)) to
   the indep one ((time)) and creating a linear regression
var trend = filteredSentinel2.select(independents.add(dependent))
.reduce(ee.Reducer.linearRegression(independents.length(), 1));

//creating a linear trend map on VI changes across time
//  Map.addLayer(trend, {}, 'trend array image');

var coefficients = trend.select('coefficients')
.arrayProject([0])
.arrayFlatten([independents]);

var harmonicIndependents = ee.List(['constant', 't', 'cos', 'sin']);

var harmonicSentinel = filteredSentinel2.map(function(image) {
var timeRadians = image.select('t').multiply(2*Math.PI);
  return image
  .addBands(timeRadians.cos().rename('cos'))
  .addBands(timeRadians.sin().rename('sin'));
});

var harmonicTrend = harmonicSentinel
 .select(harmonicIndependents.add(dependent))
 .reduce(ee.Reducer.linearRegression({
   numX: harmonicIndependents.length(),
   numY: 1
}));

var harmonicTrendCoefficients = harmonicTrend.select('coefficients')
.arrayProject([0])
.arrayFlatten([harmonicIndependents]);

var fittedHarmonic = harmonicSentinel.map(function(image) {
  return image
  .addBands(ee.Image.constant(0))
  .addBands(ee.Image.constant(0))
```

```
  .addBands(
    image.select(harmonicIndependents)
    .multiply(harmonicTrendCoefficients)
    .reduce('sum')
    .rename('fitted'));
});

// In order to change from NDVI to SAVI or EVI just change 'NDVI'
   with 'SAVI' or 'EVI'
print(ui.Chart.image.series(
  fittedHarmonic.select(['fitted','NDVI']), roi, ee.Reducer.mean(),
   30)
  .setSeriesNames(['NDVI', 'fitted'])
  .setOptions({
    title: 'Harmonic model: original and fitted values',
    lineWidth: 1,
    pointSize: 3,

}));
```

Bibliography

[1] «Consorzio Tutela Nebbioli Alto Piemonte. Visited in 24/03/2023». In: (). URL: https://www.consnebbiolialtop.it/ (cit. on p. 1).

[2] ISTAT. «Andamento Economia Agricola». In: (2019) (cit. on pp. 3, 4).

[3] CREA. «Consiglio per ricerca in agricoltura e l'analisi dell'economia agraria». In: (2019) (cit. on pp. 3, 5, 6).

[4] EU. «Geographical indications and quality schemes explained. Visited on 24/03/2023». In: (). URL: https://agriculture.ec.europa.eu/farming/geographical-indications-andquality-%20schemes/geographical-indications-and-quality-schemes-explaineden (cit. on p. 3).

[5] A. Buffa G. Ricciardi L. Mercalli. «Cambiamenti climatici e sistemi agroalimentari». In: (2017) (cit. on p. 5).

[6] Riccardo Grasso. «300 milioni per il problema della Xylella: è polemica. Ma come agisce questo batterio?» In: *Il Superuovo* (2019) (cit. on p. 5).

[7] H. Meadows J. Randers W. Behrens L. Meadows. «. The limits of growth». In: (1972) (cit. on p. 7).

[8] Giorgio dell'Orefice. «Agricoltura 4.0, investimenti a 1,6 miliardi nel 2021(+23%). Visited on 24/03/2023». In: *Il Sole 24 Ore* (2022) (cit. on p. 7).

[9] EU. «EPI-AGRI. Visited on 24/03/2023». In: (). URL: https://ec.europa.eu/eip/agriculture/en/node (cit. on p. 9).

[10] ONU. «Take action for sustainable development goals. Visited on 24/03/2023». In: (). URL: https://www.un.org/sustainabledevelopment/ (cit. on p. 9).

[11] Natural Resources Canada. «Foundamentals of Remote Sensing». In: () (cit. on pp. 10, 12).

[12] EU. «EPI-AGRI Precision Farming. Visited on 24/03/2023». In: (). URL: https://ec.europa.eu/eip/agriculture/en/digitising-agriculture/developingdigital-%20technologies/precision-farming-0 (cit. on p. 17).

[13] J. Donahue T. Darrell R. B. Girshick and J. Malik. «Rich feature hierarchies for accurate object detection and semantic segmentation». In: *CoRR* (2013) (cit. on p. 20).

[14] Medium. «A brief history of CNNs in image segmentation: from R-CNN to Mask R-CNN. Visited on 24/03/2023». In: (2017). URL: `https://medium.com/athelas/a-brief-history-of-cnnsin-image-segmentation-from-r-cnn-to-mask-r-cnn-34ea83205de4` (cit. on p. 21).

[15] R. Girshick. «Fast R-CNN». In: *CoRR* (2015) (cit. on p. 22).

[16] K. Zhang X. Ren S. Sun J. He. «Spatial pyramid pooling in deep convolutional networks for visual recognition». In: *Computer Vision - ECCV 2014* (2014) (cit. on p. 22).

[17] S. He K. Girshick R. Sun J. Ren. «Faster r-cnn: Towards real-time object detection with region proposal networks». In: *CoRR* (2015) (cit. on p. 22).

[18] Georgia Gkioxari Piotr Dollár Ross B. Girshick Kaiming He. «Mask R-CNN». In: *CoRR* (2017) (cit. on p. 22).

[19] Santosh Kumar Divvala. Ross B. Girshick Ali Farhadi Joseph Redmon. «You only look once: Unified, real-time object detection». In: *CoRR* (2015) (cit. on p. 22).

[20] Medium. «Real time object detection with YOLO, YOLO2 and YOLO3. Visited on 24/03/2023». In: *https://jonathan-hui.medium.com/real-time-object-detection-withyolo-yolov2-28b1b93e2088* (2018) (cit. on p. 23).

[21] Wei Liu Yangqing Jia Pierre Sermanet Scott E. Reed Dragomir Anguelov Dumitru Erhan Vincent Vanhoucke Andrew Rabinovich Christian Szegedy. «Going deeper with convolutions». In: *CoRR* (2014) (cit. on p. 23).

[22] Standford University Princeton University. «Image Net». In: (2015) (cit. on p. 23).

[23] Kevin P. Murphy. «Machine Learning. A probabilistic Perspective». In: (2017) (cit. on pp. 24, 39, 43).

[24] J. Grundy H. Parry A. Dorin D. C. Amarathunga. «Methods of insect image capture and classification: A systematic literature review». In: *Smart Agricultural Technology* (2021) (cit. on p. 24).

[25] Wei Li, Tengfei Zhu, Xiaoyu Li, Jianzhang Dong, and Jun Liu. «Recommending advanced deep learning models for efficient insect pest detection». In: *Agriculture* (2022) (cit. on p. 25).

[26] Mohamed Esmail Karar, Fahad Alsunaydi, Sultan Albusaymi, and Sultan Alotaibi. «A new mobile application of agricultural pests recognition using deep learning in cloud computing system». In: *Alexandria Engineering Journal* (2021) (cit. on p. 25).

[27] GBIF. «Popillia Japonica Newman, 1838 in GBIF secretariat. GBIF backbone taxonomy. Visited on 24/3/2023». In: (2021). URL: https://www.gbif.org/species/4425774 (cit. on pp. 26, 41).

[28] Regione Piemonte. «Lotte Obbligatorie al Coleottero Scarabeide del Giappone (Popillia Japonica Newman). Visited on 24/03/2023». In: (). URL: https://www.regione.piemonte.it/web/temi/agricoltura/servizi-fitosanitaripan/lotte-obbligatorie-coleottero-scarabeide-giappone-popillia-japonica-newman (cit. on p. 26).

[29] Dudenhoeffer AP. Piñero JC. «Mass trapping designs for organic control of the japanese beetle, popillia japonica(coleoptera: Scarabaeidae)». In: *Pest Manag Sci* (2018) (cit. on p. 30).

[30] Paoli F. Sabbatini Peverieri G Benvenuti C Barzanti GP Bosio G Venanzio D Giacometto E Roversi PF. Marianelli L. «Long-lasting insecticide-treated nets: A new integrated pest management approach for popillia japonica (coleoptera: Scarabaeidae)». In: *Integr Environ Assess Manag.* (2019) (cit. on p. 32).

[31] Wallis L Geis M Williamson RC Potter DA. Redmond CT. «Strengths and limitations of bacillus thuringiensis galleriae for managing japanese beetle (Popillia Japonica) adults and grubs with caveats for cross-order activity to monarch butterfly (danaus plexippus) larvae». In: *Pest Manag Sci* (2020) (cit. on p. 32).

[32] Kearns DN. Ogle K Caron M Marks JC Rogg HW. Hungate BA. «Hydrogen isotopes as a sentinel of biological invasion by the japanese beetle, Popillia Japonica (Newman)». In: *PLoS One* (2016) (cit. on p. 32).

[33] Google. «Google Earth Engine. Visited on 24/03/2023». In: (). URL: https://earthengine.google.com/ (cit. on p. 34).

[34] Ocean Insight. «Ocean Insight HD Spectrometer. Visited on 24/03/2023». In: (). URL: https://www.oceaninsight.com/products/spectrometers/ (cit. on p. 35).

[35] DJI. «DJI Matrice 300. Visited on 24/03/2023». In: (). URL: https://www.dji.com/it/matrice-300 (cit. on p. 35).

[36] DJI. «DJI P4 Multispectral RTK. Visited on 24/03/2023». In: (). URL: https://www.dji.com/it/p4-multispectral?site=brandsitefrom=insitesearch (cit. on pp. 36, 37).

[37] Sentera. «Sentera 6X Sensor. Visited on 24/03/2023». In: (). URL: https://sentera.com/products/fieldcapture/sensors/6x/ (cit. on pp. 36, 38).

[38] Autel Robotics. «Autel Evo Nano. Visited on 24/03/2023». In: (). URL: https://www.autelrobotics.com/productdetail/evo-nano-series-drones.html (cit. on p. 36).

[39] Medium. «Train a custom yolov4 object detector (using google colab). Visited on 24/03/2023». In: (2021). URL: https://medium.com/analytics-vidhya/train-a-custom-yolov4-objectdetector-using-google-colab-61a659d4868 (cit. on p. 36).

[40] Google. «Google Colab Pro. Visited on 24/03/2023». In: (). URL: https://colab.research.google.com/signup (cit. on p. 36).

[41] R. Ventura M. Veloso J. Cartucho. «Robust object recognition through symbiotic deep learning in mobile robots». In: (2018) (cit. on p. 38).