A Holistic Approach to Sustainable, Digital EU Agriculture, Forestry, Livestock and Rural Development based on Reconfigurable Aerial Enablers and Edge Artificial Intelligence-on-Demand Systems

CHAMELEON D4.3 CHAMELEON, Bundles, services v1

Work package	WP4: CHAMELEON Solution for multi-purpose on-boarding
Task	Task 4.3: CHAMELEON, Bundles, services for agriculture and rural area. Task 4.4: CHAMELEON, Bundles, services for livestock monitoring. Task 4.5: CHAMELEON, Bundles, services for forestry.
Authors	**Vaida Sirgedaitė-Šėžienė – LAMMC,** **Miguel Ángel Moreno Hidalgo – UCLM,** **David Sánchez Jiménez – USAL,** **Thanos Siouras – AiDEAS.**
Dissemination level	Public (PU)
Status	Final
Due Date	31/12/2023
Document date	14/06/2024
Version number	1.0

Funded by the European Union	Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the European Research Executive Agency can be held responsible for them.

Revision and history chart

Version	Date	Main author	Summary of changes
0.1	01/08/2023	LAMMC	Draft outline
0.2	21/11/2023	UCLM	Draft outline 2nd version
0.3	05/12/2023	USAL	Draft outline 3rd version
0.4	06/12/2023	AiDEAS	Draft outline 4th version
0.5	08/12/2023	LAMMC	Final draft
0.6	14/12/2023	SQD	Internal Review
1.0	18/12/2023	LAMMC	Final submitted version
2.0	13/06/2024	LAMMC	Revised version

Table of contents

Index of figures

Index of tables

List of abbreviations and acronyms

Abbreviation	Meaning
AGB	Above-ground biomass
AI	Artificial Intelligence
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer thermal images
BT	Brightness Temperature
CLMS	Copernicus Land Monitoring Service
CMBN	Concise multi-branch Network
CMYK	Subtractive color model, based on the CMY color model
CNN	Convolutional Neural Networks
CSAIL	Computer Science and Artificial Intelligence Laboratory
CSF	Cloth Simulation Filter
CSV	Comma-separated values
DSM	Digital surface model
DTM	Digital terrain model
ESA	European Space Agency
EVI	Enhanced Vegetation Index
FLOPs	Floating-point operations per second
FPN	Feature pyramid networks
FVC	Fractional Vegetation Cover
GCC	Green canopy cover
GCN	Graph convolutional network
GCP	Ground Control Point
GDAL	Translator library for raster and vector geospatial data formats
GEE	Google Earth Engine
GIS	Geographic Information System
GNDVI	Green Normalized Difference Vegetation Index
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPU	Graphics processing unit
GR	Greece
GSD	Ground Sampling Distance
ID	Personal Identification Number
IDW	Inverse distance weighting algorithm
JSON	JavaScript Object Notation
INS	An inertial navigation system
IoT	The Internet of Things
LAI	Leaf area index

LIDAR	Light detection and ranging
LMF	Local Maximum Filter
LPWAN	Low Power Wide Area Network
LSTM	Long Short-Term Memory Network
LTE	Long time Exposure
mAP	Mean Average Precision
MMPose	Pose estimation model
MS	Multispectral images
MS-TCN	Multi-Scale Temporal Convolutional Network
NASA	National Aeronautics and Space Administration
NBR	Normalized Burn Rate
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NGRDI	Normalized green-red difference index
NIR	Near-infrared
OD	Object Detection model
ODM	Open Drone Map
OpenCV	Open Computer Vision Library
PENet	Pose Estimation Network
PPK	Post-processed kinematic
R-CNN	Region-based Convolutional Neural Network
RDI	Regulated deficit irrigation
REDD+	Reducing Emissions from Deforestation and Forest Degradation
RGB	Red, green and blue
RGBT	Common channel order (Red, Green, Blue, Thermal)
RNN	Recurrent Neural Network
RoI	Selection of the region of interest
RS	Remote sensing
RTK	Real-time kinematic
SAHI	Slicing Aided Hyper Inference
SAVI	Soil Adjusted Vegetation Index
SRTM	Shuttle Radar Topography Mission
SW	South Western Slopes in Austria
SWIR	Short-wave infrared
TBGR	Common channel order (Thermal, Blue, Green, Red)
UAS	Unmanned Aerial Systems
UAV	Unmanned Air Vehicles
UNET	Neural Network Model
USGS	United States Geological Survey

V	Volume occupied by the crop
VARI	Visible atmospherically resistant index
Vis	Multispectral vegetation indices
VLOS	Visual Line of Sight
VTOL	Vertical take-off and landing autonomy flight
WebGIS	Web Geographical Information Systems
WebODM	Web Open Drone Map
WP2	Work package
YOLO	"You Only Look Once" Real-Time Object Detection system

1. EXECUTIVE SUMMARY

This Deliverable 4.3 presents the first version of the CHAMELEON bundles, services for agriculture and rural areas, livestock monitoring and forestry. The document outlines the main pillars of CHAMELEON operating ecosystem by identifying and analysing pilot cases' concerns. Three tasks (from W4) are addressed in this document:

- Task 4.3: CHAMELEON, Bundles, services for agriculture and rural area;
- Task 4.4: CHAMELEON, Bundles, services for livestock monitoring;
- Task 4.5: CHAMELEON, Bundles, services for forestry.

These tasks enable a Supportive CHAMELEON technical development of bundles and services for the agriculture and rural areas, livestock and forestry. Initially, Deliverable 4.3 provides the required systems analysis of identified different bundles related to agriculture and rural areas, forestry and livestock. Deliverable illustrates the functional and technical specifications of each service and bundle development steps:

- Relevance and problematic;
- UAV characteristics;
- Camera characteristics;
- Flight characteristics;
- Data processing;
- Results and outcomes.

Finally, these bundles will be utilizing the on-boarding tool and will be able to run in parallel achieving tailored per case environments.

2. INTRODUCTION

2.1. BACKGROUND

The application of drones in plant disease assessment offers efficient monitoring and detection capabilities for smart agriculture, forestry and livestock. Drones provide increased accessibility, improved coverage, and rapid data collection, enabling timely disease detection[1]. With advanced sensors (including digital, multispectral, thermal, and fluorescence sensors) and imaging techniques, drones can capture valuable data on plant and animal health indicators. These data can be processed using analytics and machine learning algorithms to identify disease patterns and assess severity. Integration of drones into plant disease and animal monitoring assessment systems allows for real-time monitoring, early detection, and targeted intervention[2]. Deep learning models require sophisticated and effective algorithms to address issues such as fluctuating illumination, growing diseases, occlusion, and shifting perspectives. Drones can contribute to sustainable farming practices, minimization of yield losses, reduced need for chemical treatments, and support for precision agriculture strategies[3]. Moreover, drones push the boundaries of forest cultivation and maintenance, redeeming wildlife and vegetation, while also monitoring risks to prevent the felling of trees[4]. With the adoption of drones, forestry can witness a high cascade of improvement and growth. Additionally, drones with thermal cameras are used to locate the herd in the pastures and move it or find the lost cattle in the mountains after storms. This helps reduce the time it takes to move cattle. Drones can help to inspect the health of cows/sheep.

This document outlines the main pillars of CHAMELEON operating culture by identifying and analysing pilot cases' concerns and the main challenges to be overcome in aerial platform development. Three tasks are addressed in this document:

- Task 4.3: CHAMELEON, Bundles, services for agriculture and rural area;
- Task 4.4: CHAMELEON, Bundles, services for livestock monitoring;
- Task 4.5: CHAMELEON, Bundles, services for forestry.

The main list of specified requirements from stakeholders was defined under WP2. Following the stakeholders' mapping, the key identified actors were invited to participate in the first CHAMELEON Stakeholders' Workshops. They were performed in 1. Avila (Spain), 2. Crete (Greece) and 3. Vienna (Austria) in local language, following online, physical and/or hybrid

---

1. Abbas, A.; Zhang, Z.; Zheng, H.; Alami, M.M.; Alrefaei, A.F.; Abbas, Q.; Naqvi, S.A.H.; Rao, M.J.; Mosa, W.F.A.; Abbas, Q.; et al. Drones in Plant Disease Assessment, Efficient Monitoring, and Detection: A Way Forward to Smart Agriculture. *Agronomy* **2023**, *13*, 1524. https://doi.org/10.3390/agronomy13061524

2. Shah, S.A.; Lakho, G.M.; Keerio, H.A.; Sattar, M.N.; Hussain, G.; Mehdi, M.; Vistro, R.B.; Mahmoud, E.A.; Elansary, H.O. Application of Drone Surveillance for Advance Agriculture Monitoring by Android Application Using Convolution Neural Network. *Agronomy* **2023**, *13*, 1764. https://doi.org/10.3390/agronomy13071764

3. Kumar, S.P., Subeesh, A., Jyoti, B., Mehta, C.R. (2023). Applications of Drones in Smart Agriculture. In: Pakeerathan, K. (eds) Smart Agriculture for Developing Nations. Advanced Technologies and Societal Change. Springer, Singapore. https://doi.org/10.1007/978-981-19-8738-0_3

4. Equinox's Drones. How Drone Technology is Becoming Essential for Forestry. Retrieved in November 2023 from https://www.equinoxsdrones.com/how-drone-technology-is-becoming-essential-for-forestry/

implementation. These three Workshops was organized in order to present CHAMELEON solution to main actors and end users, to exchange options and perceptions, to identify unforeseen risks or constraints, and to further understand the plant, plant-health, livestock, livestock-health and agri-environmental monitoring.

2.2. PURPOSE AND SCOPE

The scope of this deliverable is to create bundles of services for agriculture-rural areas, livestock, and forestry monitoring purposes, to solve problems related to these sectors by using UAV and AI technologies. The purpose of this deliverable is to provide the required systems analysis and workflow of the bundles related to agriculture and rural areas, forestry and livestock; and to describe the functional and technical specifications of each service and bundle development steps. Since each of the three service bundle has different goals and challenges, it is necessary to find a common strategy to cover all the bundles being developed. The bundles were categorized according to the Pilot use cases into the folowing: Greece (includes 3 bundles), Spain (includes 8 bundles), and Austria (includes 5 bundles).

2.2.1. GREECE

The bundles that will be deployed in the Greece mainly be related to the observation and monitoring of sheep and/or goat herds in the territory of Western Crete. The area is located in a small valley (above the mountainous region of Sfakia) which acts as the main base camp for most of the herders (including the owner of the GR Pilot herd). This area is characterized by isolated and difficult terrain and typical bush vegetation common to the island. The valley is surrounded by steep and barren mountain slopes with ancient trails known only to the local goat herders. Wind patterns change frequently in the region.

- **The main goals and expectations are:**

During the monitoring of the herd of animals, perform an analysis of animal behavior, identify possible health problems (temperature changes, posture disorders, etc.), assess the distribution of the herd in the monitored area (find lost animals), and implement a system to alert the herd owner about intruders or animals wandering the territory, using an automated mapping system that allows for quick and accurate location, which is especially valuable in emergency situations.

- **Complications**

Since the test area is located in a highly mountainous area, the area is subject to rapidly changing weather conditions, especially affecting the strength and speed of the wind. The size of the test region and the rapidly changing altitude affect successful and high-quality data collection. Correct selection of altitude can reduce the chance of errors, due to improperly collected data, too high altitude (>80 m) does not allow successful identification of animals, too low altitude (<30 m) disturbs animals and they no longer behave naturally.

- **Results/Outcomes:**

The implementation of the objectives of this service bundle will help to detect sick animals early, thus reducing the risk of them spreading a contagious disease to the rest of the herd.

During the scanning of the territory, if a "rogue sheep" is identified, which is in a difficult situation (injured) or is separated from the herd, it will be possible to provide help in time, thus saving the animal in the herd. The installation of a warning system will allow the herdsman/owner to identify the visit of stray persons or animals to the territory.

2.2.2. SPAIN

Tietar Valley (Avila, Spain) is a wooded area surrounded by mountains, where the natural vegetation characteristic of this place grows. However, due to the mountainous and rocky terrain, some places do not have access roads. Also, the area is close to cities, so wooded natural areas are intertwined with urban vegetation, which can affect the spread of fires.

- **The main goals and expectations:**

Determining high fire risk areas and wildfire outbreaks. The data gathered and collected by the UAV is used for fire prevention purposes. Identifying the vegetation most associated with high fire risks. Determining the connections between wild vegetation and urban greenery, and prevention in order to prevent the spread of fires between wild and urban areas, by offering evacuation routes, extinguishing means, the most effective places to stop fires, etc. Identification and quick response to fires that have already started, providing a forecast of fire spread, as well as the best ways to reach the fire, for firefighters.

Remote monitoring of animals. During remote monitoring, the aim is to collect data about the herd, ensuring the discipline and calmness of the animals, but ensuring a clear and high-quality image, so that the data is suitable for use for the following purposes. collected data will help to reveal possible animal health problems and perform behavioral analysis due to possible attacks by wild animals. The use of thermal cameras will be adapted to determine the state of health of animals by measuring changes in their body temperature and will allow to determine the visit of other unwanted animals or persons to the pastures, thus protecting the animals in the herd more quickly. An automated mapping system will allow you to quickly identify animals and their location in the pasture, which is especially valuable in emergency situations.

Analysis of crop development conditions. Using UAVs to monitor crop growth and identify potential causes of development failure in time. early detection of water deficit in plants and quick implementation of management measures in time to prevent crop loss. Determining the characteristics of the soil of small areas in order to create optimal conditions for the uniform growth of crops in the entire supervised area.

- **Complications:**

Possible complications of animal herd management due to the size of the fields, since the cattle live freely, depending on the location of the area and woody vegetation, it can be difficult to assess and precisely determine the location of the animals, and possible problems. Therefore, it is better to use closed areas so that the animals cannot get lost. There are no problems with timely detection of fires or determining the state of vegetation, but there are possible obstacles due to licenses and permits to fly drones in some areas.

- **Results/Outcomes:**

The implementation of the goals of this package of services will help to detect sick, injured or stray animals in the territory early, which will allow herd owners to quickly react to the situation. Identifying the optimal conditions for crops and quickly detecting a decrease in rainfall or any of the nutrients required for growth will allow farmers to select appropriate cropland management strategies. The use of drones for fire prevention will allow you to prevent or respond properly to wildfires that have just started, thus preventing the experience of large losses due to all-consuming fires.

2.2.3. AUSTRIA

Kleines Rodeltal (SW slope) is a small forest in Upper Austria. The forest is located in a valley, on its river slopes. A small river flows through the valley with a small settlement. The location is safe for test flights and due to suitable vegetation conditions, it is favorable to test how the system works. The forest is conifers with deciduous trees that have started to spread since 2000. In the same region, there are also vineyards, which are located in a mountainous area but are completely privately owned. There are no restrictions around, so drones can work without special permits.

- **The main goals and expectations:**

During the monitoring of the forest, the UAVs will be used in order to monitor and carry out an inventory of woody vegetation. At the same time, determining whether there are no fallen or fallen trees in the stands, which could pose a danger to people in the area. It is also expected to monitor the health status of trees in order to maintain a healthy and sustainable forest ecosystem. For that purpose, the aim is to develop a system for early detection of damage by bark beetles. For the preventive protection of forests against fires, soil and plant moisture levels are assessed in order to provide timely notification of increased fire risk in some regions. Monitoring in vineyards. The goal is to provide remote monitoring of damage after strong environmental factors such as storms, hail, etc. As in forests, vineyards monitor soil moisture and the degree of wilting of plants, so that the harvest is not lost, and irrigation systems are used in time.

- **Complications:**

Due to the hilly terrain, multiple serial flights will be required to cover the area as VLOS flight is mandatory.

- **Results/Outcomes:**

The created system allows timely identification of threats arising in forests due to extreme environmental factors, such as natural meteorological phenomena (wind, drought) or damage by bark beetles. It will allow a proper response and ensure high-quality and sustainable use of forests. Also, the system will benefit not only the forest but also agriculture, especially vineyards. In which the influence of environmental factors will be monitored and when the warning mechanisms are activated, the situation will be responded to in time.

2.3. APPROACH

This deliverable gives the approach of how specific issues related to agriculture, livestock and forestry will be solved. In this document every bundle is described step by step from requirements to programs used for the development of each bundle. Detailed description of development of each bundle for agriculture, forestry and livestock monitoring is provided according to their action level:

- The relevance and problematic of the bundle section will provide a clearer view of the problematic of each bundle of agriculture, livestock and forest. The aim of this information is to provide an elusive concept of specific problematic and to define practical and especially social applicability. It is related to the solution of any problem, which requires the prior identification of the relevant elements from which a solution can be constructed.
- Required UAV characteristics will be selected by bundle developers, regarding the specific aim of each bundle, the UAV characteristics. Under CHAMELEON project will utilize a three-layered approach in the monitoring of plants, animals and forests by exploiting two types of unmanned aerial vehicles (Figure 2.3.1):
  - Low altitude drones;
  - Medium-high altitude, fixed wing UAVS;
  - Very High altitude, Satellite imaging.

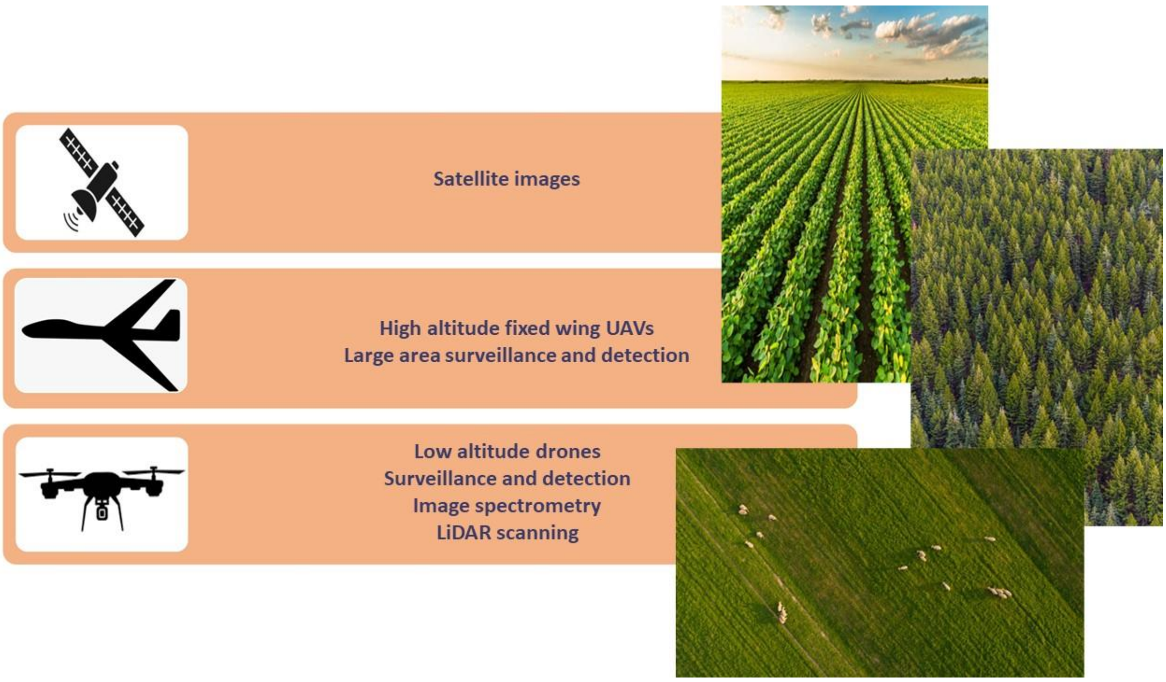


Figure 2.3.1: CHAMELEON three layered approach

- Required camera characteristics section will provide the camera specifications needed to run bundles. For small areas the RGB imaging or LIDAR scanning using drones could be used. Meanwhile, for medium and/or large area c the same cameras using fixed-wing UAVs or satellite imaging could be used. In order for the user to properly collect the necessary data to solve the problem, the requirements of the camera must be clearly defined. The best thing to do is to provide the minimum requirements for the camera/photos required, which would solve the problem.
- The required flight characteristics section will provide brief information about the specific requirements for each bundle. The imaging plan depends on such characteristics as:
    a. selected bundles (forestry/livestock/agriculture);
    b. selected area type (forested area, dense agricultural crops, animal distribution, etc.;
    c. area size and area complexity (square, rectangular, a complex figure);
    d. the overlap of photo or video material, which depends on the type of area that is being displayed and will determine the speed at which images are taken.
- Overlap and flight height must be adjusted depending on the selected area. In most cases, it is recommended to acquire images with a regular grid pattern. The recommended overlap is a minimum of 75% forward overlap (with respect to the direction of flight) and a minimum of 60% lateral overlap (between runways). To ensure data collection of the desired quality, the camera must be kept at a constant height and position throughout the selected area.

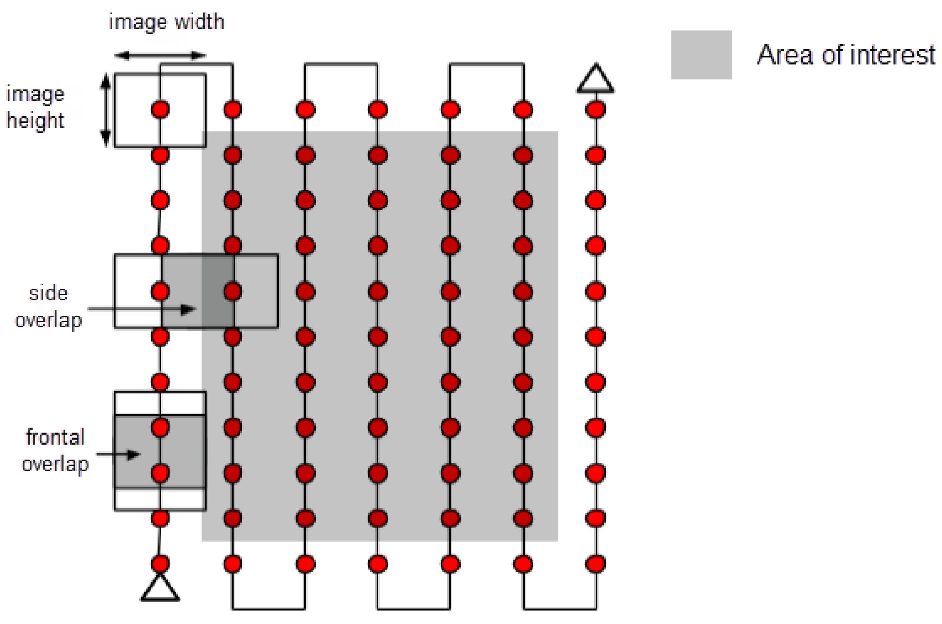


Figure 2.3.2: Flight execution scheme[5]

[5] PIX4D. Image acquisition - PIX4Dmapper. Retrieved in November 2023 from https://support.pix4d.com/hc/en-us/articles/115002471546-Image-acquisition-PIX4Dmapper

Example: A minimum of 80% front and side overlap is recommended for flat agricultural fields. In areas of forest, dense vegetation, it is recommended to increase the overlap to at least 85% front and side overlap and fly higher to facilitate the detection of overlapping image similarities. Thermal imaging projects require at least 90% front and side overlap. For projects with multiple flights, the different flights must coincide and the conditions (sun direction, weather conditions, unbuilt new buildings, etc.) should be similar. All the flight plans described below should be executed automatically with cross-planning programs that provide either single geolocated photographs or an orthophotomosaic of the entire area at once.

- The processing of the collected data is described in detail or clearly graphically presented. A structured scheme is presented from the acquisition of data from UAVs to the presentation of processed data for deep learning or application of algorithms. Programs or their packages that are used in the initial data processing process are presented. The extraction, labeling, and representation of the required structures for simulation and algorithm application are presented. The objectives of applying models and algorithms are described. The required classification of objects or actions selected for bundle execution or other necessary sorting method is presented.
- The results/outcomes section should provide information about what is expected and what will be obtained after performing all the processes with the obtained data. Depending on the executed bundle, it must be indicated in which format the required information will be presented to the end user or customer.

2.4. RELATION WITH OTHER DELIVERABLES AND WORK PACKAGE'S

The direct dependence of CHAMELEON WP is presented in Figure 2.4.1. This deliverable is the outcome of WP4, which is tightly supported by WP2, specifically by the T2.1 Stakeholders use cases, requirements, and workshop, and T2.3 Understanding plant, plant-health, livestock-heath, and agri-environmental monitoring requirements. Deliverable has close relation with the other internal one tasks of the WP4 – T4.1 Plug-n-play platform – the task that serve the technological capabilities for the bundles and services deployment; and T4.6 ADSS visualization and user interface – the task that will ensure relevant and simple interface for final bundles users. Inextricably, this deliverable is in close relation with tasks of WP3, which serves UAV Core platform (T3.1) for bundles execution and provides CHAMELEON Open payload toolset (T3.3). Moreover, Deliverable is closely related with WP6, while the bundles are adapted and will be finally tested and applied on the Pilot use cases areas. Finally, D4.3 is a part of WP5 CHAMELEON innovation platform.

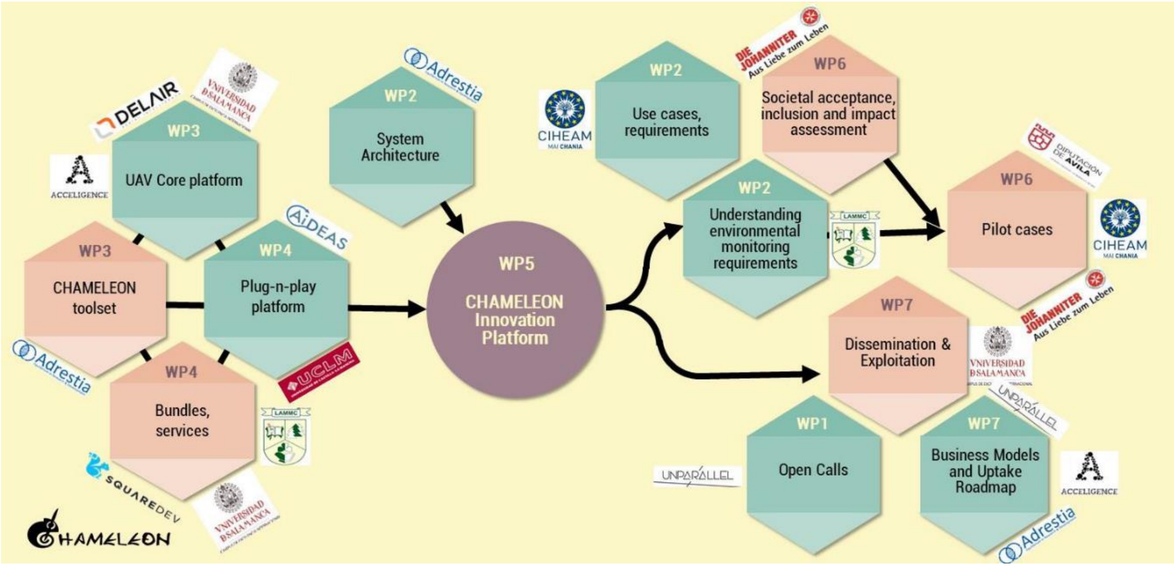


Figure 2.4.1: CHAMELEON WP relations

The main list of the services and bundles, which will be implemented in the CHAMELEON project WP4 were defined during workshops of stakeholders in Avila (Spain), Crete (Greece) and Vienna (Austria). These workshops were planned and organized under WP2 MAICh, JOAFG, USAL and Ávila Diputación together with WP leaders. The main outcomes were provided in Deliverable 2.1, which was uploaded to the European Commission system.

Figure 2.4.2: CHAMELEON Workshops of stakeholder

Two main pillars - pasture and livestock underlined for the livestock management in Cretan mountains, involving the monitoring, control and improvement of grazing fields, focused on biodiversity monitoring, fertilization and sowing of grazing fields. A hybrid workshop performed by the MAICh team referred to livestock monitoring and management in Crete, Greece. Essential feedback gained from the local and regional authorities, livestock owners' association, producers, and other key actors, which was compiled with the investigated technical specifications of the proposed CHAMELEON solutions (Figure 2.4.3).

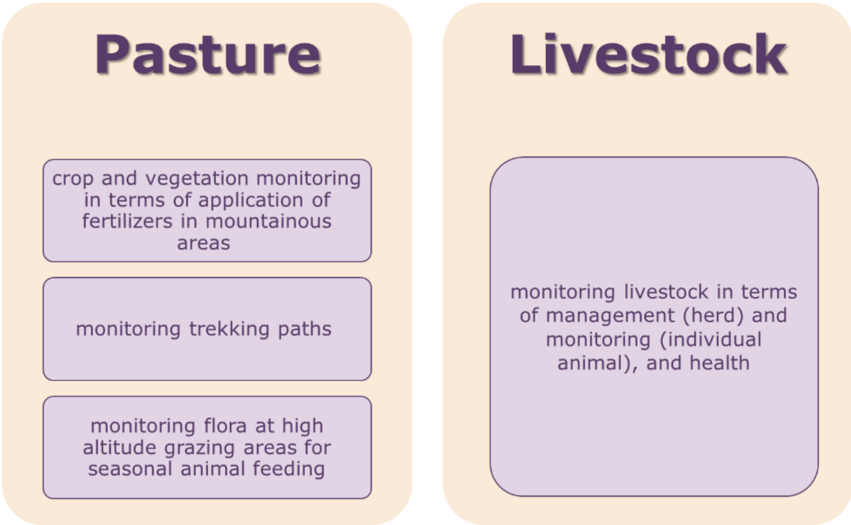


Figure 2.4.3: The main pillars in Crete, Greece

Surveys and personal interviews as workshops were performed by JOAFG in order to collect more information for forest monitoring of potential dangers, and vineyards from Austrian stakeholder. This way two main pillars of the Austrian pilot case were defined – forest and vineyard (Figure 2.4.4).

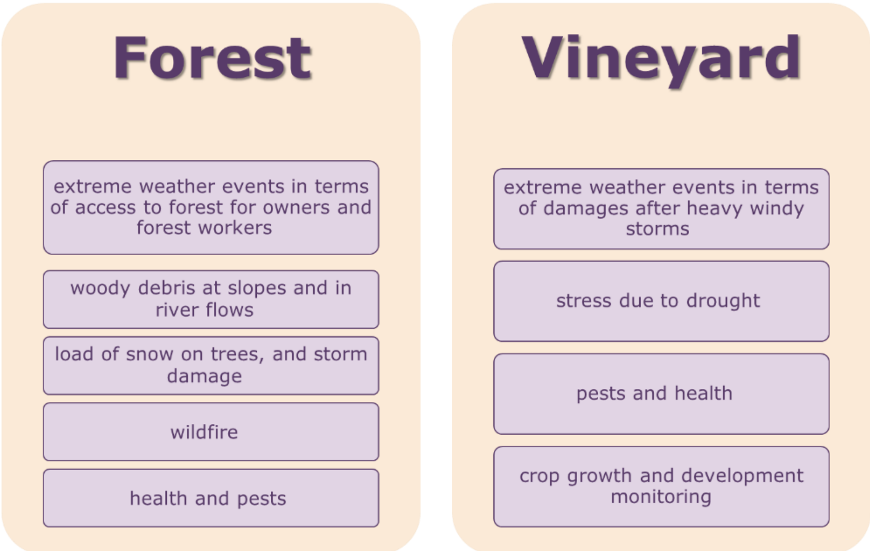


Figure 2.4.4: The main pillars in Viena, Austria

A workshop performed by USAL and Ávila Diputación referring to forest fire defense plans for rural areas, livestock, and vineyards of local stakeholders (mainly regional and local authorities, and livestock owners' associations, forest, and vineyard end users). During the workshop, four main pillars of this use case were defined – forest, vineyard, livestock, and soil, which are referred to as main concerns, presented in the Figure 2.4.5.

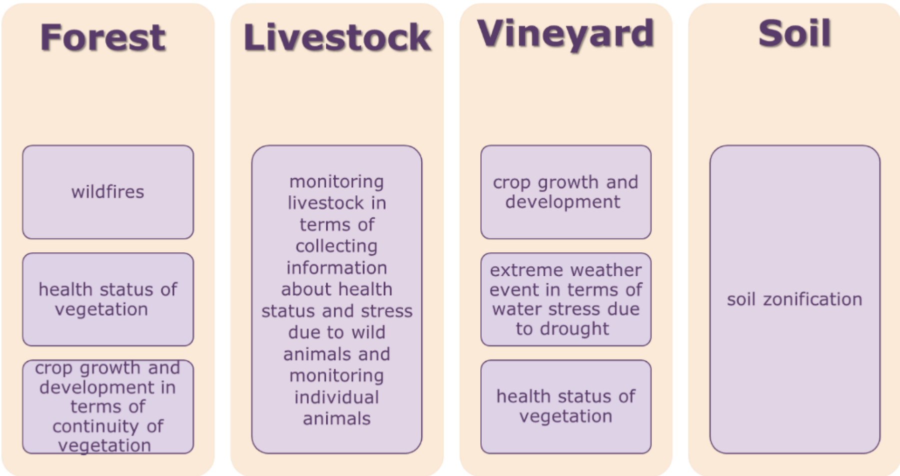


Figure 2.4.5: The main pillars in Avila, Spain

The conceptual, functional, and implementation view, together with the deployment and use case view of the bundle development were provided in Deliverable 2.2, which was uploaded in the European Commission system. The content of this deliverable is related with the outcomes of WP2.

Table 2.4.1: The conceptual view represents the main categorised groups of CHAMELEON stakeholders.

Group of Stakeholders	Functions
Drone owners	Unmanned Air Vehicles (UAVs) provided by drone owners will be used to perform the monitoring and management procedures of livestock, agriculture and forestry.
Bundle developers	Bundles Developers construct the logic of their developed bundles that will be implemented on the reconfigurable drones of CHAMELEON. The Cloud infrastructure will be the main point between the CHAMELEON ecosystem and the developers.
Livestock associations	It is beneficiary of the bundles regarding livestock monitoring and management, and the end users of the CHAMELEON.
Forest owners	The second group of the end users of the CHAMELEON will be focusing on the monitoring and management of the forestry bundles.
Vineyard owners	Bundles focusing on vineyard monitoring and management will be utilised by vineyard owners. They are another end-user and beneficiary of the CHAMELEON platform.
Authorities	Authorities, such as firefighters, municipalities, and civil aviation will provide the law regulations in order to ensure lawful CHAMELEON activities.

These groups play the crucial role for CHAMELEON system components such as: CHAMELEON Cloud, CHAMELEON Edge, and The Physical layer of CHAMELEON. Moreover, these groups of stakeholders of CHAMELEON will conduct and assess the effectiveness of the main activities proposed in the CHAMELEON project. An overview of the functional elements of the CHAMELEON system, including their respective responsibilities and interactions depends on the categories of the bundles – agriculture and rural areas, livestock monitoring and forestry (Figure 2.4.6).

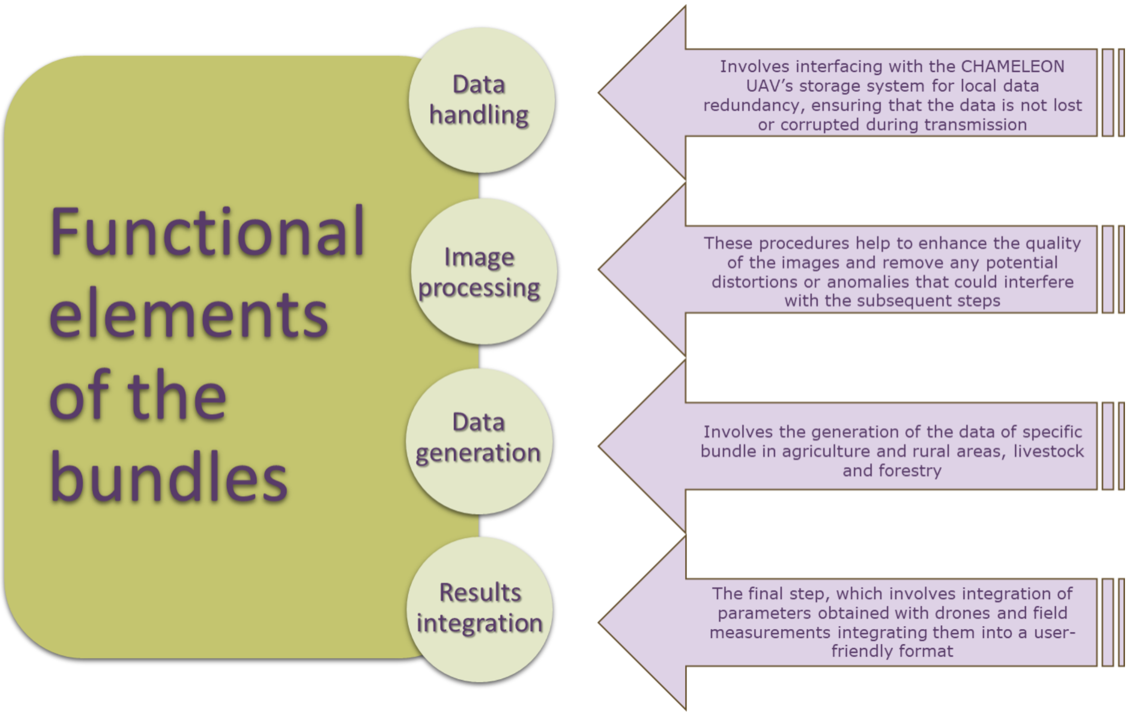


Figure 2.4.6: The functionalities of the bundles in CHAMELEON system

All bundles will be integrated into the CHAMELEON platform during the implementation stage, which consists of a detailed view of how the individual parts of each bundle will be integrated into the overall system architecture. The CHAMELEON system has main following requirements in order to be deployed: Software Deployment (represents the software deployment requirements (e.g., cloud microservices, containerized bundles etc.); Hardware Requirements (represents requirements for UAVs, Gateway, and Cloud platform); Network Requirements (represents the network connections requirements by the different layers CHAMELEON); High-level Deployment Diagram (represents the first version of the CHAMELEON Architecture). Main elements of the CHAMELEON ecosystem as well as interactions between these different components are documented through several end-to-end operations:

Figure 2.4.7: Main elements of the CHAMELEON ecosystem

Moreover, the prioritization of the Agriculture, Livestock and Forestry bundles were done under WP6 and provided in the Deliverable 6.1. During the WP6 Workshop the most important bundles were selected, which will be implemented during the CHAMELEON project under WP4. The selected bundles are provided in Table 2.4.2.

Table 2.4.2: List of bundles selected to developed in the CHAMELEON project.

Pilot Case	Group	Developer	Business Use Case	Bundle
Austria	**Forest**	**USAL**	Crop and vegetation monitoring	Vegetation monitoring and census
	Forest	**USAL**	Extreme weather event and drought	Large woody debris on rivers
	Forest	**LAMMC**	Health and pests	Health status of vegetation (mainly bark beetle), game browsing, ground cover, and fungal growth
	Vineyard	**UCLM**	Crop and vegetation monitoring	Crop growth and development monitoring
	Vineyard	**UCLM**	Extreme weather event and drought	Vineyard water stress due to drought.
Greece	**Livestock**	**AiDEAS**	Monitoring livestock	Livestock management (herd) and monitoring (individual animal)
	Livestock	**AiDEAS**	Monitoring livestock	Animals' health
	Pasture	**UCLM**	Crop and vegetation monitoring	Monitoring flora at high-altitude grazing areas for seasonal animal feeding
Spain	**Forest**	**USAL**	Crop and vegetation monitoring	Continuity of vegetation
	Forest	**USAL**	Wildfire	Characterization of Wildland-urban interface.
	Forest	**USAL**	Wildfire	Hot spot identification at the beginning of wildfire
	Livestock	**USAL**	Monitoring livestock	Collecting information about health status and stress (wild animals)

	Livestock	**AiDEAS**	Monitoring livestock	Cow lameness detection
	Vineyard	**UCLM**	Crop and vegetation monitoring	Crop growth and development monitoring.
	Vineyard	**UCLM**	Extreme weather event and drought	Vineyard water stress due to drought.
	Vineyard	**UCLM**	Soil	Soil zonification

Moreover, the list of the bundles with priority "medium" was selected and prepared for submission for the Open calls' procedures under WP1, which are planned during the implementation stage CHAMELEON project. The list of selected bundles for Open calls are provided in Table 2.4.3.

Table 2.4.3: List of bundles selected to developed in the Open Calls

Pilot Case	Group	Business Use Case	Bundle
Austria	**Forest**	Extreme weather event and drought	Woody debris on forests slopes
Austria	**Forest**	Extreme weather event and drought	Load of snow on trees (depth)
Greece	**Pasture**	Crop and vegetation monitoring	Application of fertilizers in inaccessible grazing areas of high altitude
Spain	**Livestock**	Monitoring livestock	Monitoring livestock/individual animal/virtual fences
Spain	**Forest**	Health and pests	Early detection of health status in forest (pest and dry vegetation)

Some of the bundles were eliminated as medium or low priority. Additionally, some bundles have the same implementation procedures and the same purpose. Therefore, in order to avoid overlapping information, some bundles have been combined. The list of merged or/and eliminated bundles are provided in Table 2.4.4.

Table 2.4.4: List of bundles, which were merge/eliminated from the CHAMELEON project.

Pilot Case	Group	Business Use Case	Bundle
Austria	**Forest**	Extreme weather event and drought	Access to forest (paths, roads)
Austria	**Vineyard**	Extreme weather event and drought	Vineyard damage evaluation due to heavy wind storms
Austria	**Vineyard**	Health and pests	Health status of vineyards and early detection of pest and fungal infestations (e.g., *Grapevine cicada*, *Esca fungus* and *Phytoplasmosis*)
Austria	**Vineyard**	Health and pests	Vineyard damage evaluation due to wild animals (game browsing)
Austria	**Vineyard**	Health and pests	Health status of vegetation, game browsing, ground cover and fungal growth
Greece	**Pasture**	Crop and vegetation monitoring	Engraving and monitoring of trekking paths (tourism)
Spain	**Forest**	Health and pests	Early detection of health status in forest (pest and dry vegetation)

Spain	**Vineyard**	Health and pests	Health status of vineyard and early detection of pest (e.g. *Bactrocera oleae* or *Dryocosmus kuriphilus*)

The development of each bundle has been split into tasks and subtasks according to the activities. Gantt chart of each bundle is provided in the Deliverable 6.1 in order to responsibly follow the advances of each bundle's associated work.

3. BUNDLES, SERVICES FOR AGRICULTURE AND RURAL AREAS

The Bundles (Business use case) for agriculture and rural areas of each pilot use case are actions to be conducted on each Pilot with a specific purpose and indicator. The pilot use cases, and the Bundles for agriculture and rural areas will be developed in CHAMELEON project, according to the conceptualisation and use cases definition of Deliverable 2.1. The CHAMELEON solution for agriculture and rural areas will be validated and demonstrated under three relevant pilot use cases: i) Spain (Avila), ii) Greece (Crete), and iii) Austria (Vienna) (Table 3).

Table 3: List of bundles for agriculture and rural areas developed in the CHAMELEON project.

Pilot use case	Bundles (Business use case)
Spain	CROP GROWTH AND DEVELOPMENT MONITORING
	VYNEYARD WATER STRESS DUE TO DROUGHT
	SOIL ZONIFICATION
Greece	MONITORING FLORA AT HIGH-ALTITUDE GRAZING AREAS FOR SEASONAL ANIMAL FEEDING
Austria	CROP GROWTH AND DEVELOPMENT MONITORING
	VINEYARD WATER STRESS DUE TO DROUGHT

Different bundles consist of different action steps of the workflow, which was provided in the Deliverable 6.1. Detailed description of development of each bundle for agriculture and rural areas is provided according to their action level: Relevance and problematic; UAV characteristics; Camera characteristics; Flight characteristics; Data processing; AI deep learning process/algorithms; Results/Outcomes.

3.1 CROP GROWTH AND DEVELOPMENT MONITORING

RELEVANCE AND PROBLEMATIC

Crop growth and development monitoring provides valuable information for precision agriculture practices. By analysing crop health indicators, farmers can tailor their management strategies to specific areas of the field, optimizing the use of fertilizers, pesticides, and water resources. This leads to more efficient and sustainable farming practices, reducing costs and environmental impact. Differentiated crop development in a plot is affected by a series of factors that includes: a) soil differences, b) individual plants differences, c) micro-climate conditions, and d) farmers labours, among many others. Analysing these factors independently would miss the interaction between factors. The strategy of using the crop as a sensor, through the estimation of the vigour of the plants, would capture the interaction between all these factors. It will allow to determine differentiated zones in the plot where the inputs (water, fertilizers, and others) application can be also differentiated.

The main parameters that describe the vigour of the plants are the green canopy cover (GCC) and the leaf area index (LAI). In fact, there are many studies that determine the relationship between those two variables because it is much easier to determine GCC than LAI. It is important to emphasize at this point that, estimations of plant vigour based on NDVI calculation, is an indirect measurement of the GCC, which has turned into a useful and highly used procedure in case low-resolution products are obtained. In orchards, the volume occupied by the canopy is also an important parameter to estimate when monitoring crop vigour.

Multispectral cameras compared with conventional RGB cameras, has the main advantage of acquiring information in part of the spectra that is of special interest for agriculture, such as near-infrared and red-edge, which improves radiometry characterization. However, when aiming for geometric characterization of the plants (GCC and volume characterization), the use of multispectral cameras may present a disadvantage due to their lower spatial resolution. Also, the cost of multispectral cameras is much higher than conventional RGB cameras, which in some cases can jeopardize its applicability.

Final users demand easy to use products with the tools they have available, in this case RGB or multispectral sensors. To supply a solution to this demand, in this bundle we developed different methods for crop growth and monitoring, based on both RGB and multispectral information, which can be utilized independently (based on user demand) or jointly to obtain synergies between these products.

The obtained parameters are the followings:

- Green canopy cover (GCC)
- Volume occupied by the crop (V)
- Multispectral vegetation indices (VIs), such as NDVI, SAVI, among others.

Final user will decide which parameter will utilize to analyse crop growth and development monitoring.

UAV CHARACTERISTICS

To help in the bundle development process, information has been acquired with the PAFyC-UCLM equipment before the own Chameleon drones are operative. In this case, a DJI Matrix600 Pro drone was utilized, with the following specifications.

Table 3.1.1: Matrix 600 Pro UAV characteristics

Technical Specifications	Requirements
Max. MTOW	15.5 kg
Wing type	Hexacopter
Battery	LiPo 6S 4500 mAh, 22.2 V
Mounted sensors	Sony Alpha6000. RGB Micasense MX. Multispectral Flir Tau2. Thermal All sensors are mounted at the same time on a gymbal

Operational requirements	
Flight time (Max.)	20 minutes
Cruising speed	5 m/s (max 65 km/h)
Radio link range	Up to 3.5 km
Wind resistance	8 m/s
Ground sampling Distance (GSD)	1-8 cm depending on flight height and sensor
Relative orthomosaic/3D model accuracy	1-3x GSD
Absolute horizontal/vertical accuracy (w/GCPs)	Down to 3 cm / 5 cm
Absolute horizontal/vertical accuracy (no GCPs)	1-5 m
Automatic 3D flight planning	Yes

CAMERA CHARACTERISTICS

To obtain the required products the following sensors, included in Chameleon project, can be used:

RGB camera such as SONY Alpha 6000, for obtaining green canopy cover and volume.

Multispectral camera, such as Micasense RedEdge MX, for obtaining vegetation indices.

Figure 3.1.1: Cameras mounted on DJI M600 Pro

FLIGHT CHARACTERISTICS

A typical photogrammetric flight planning should be implemented. After selection of the region of interes (RoI), the main parameters to consider for each type of product (RGB or multispectral) are the followings:

- Ground sample distance (GSD)
  - RGB products: max 2 cm
  - Multispectral products: max. 5 cm.
- Overlapping and sidelapping:
  - RGB products: 80% (forward overlap) 50% (side overlap)

Multispectral products: 80% (forward overlap) 50% (side overlap), , which ensures a high quality of the final geomatic products. Final user could decrease these numbers but with a risk of reducing final product quality.n any case, the maximum flight height of 120 m will not be exceeded as well as any other restriction commanded by law.

In case of vineyards, it is preferable to perform the flight perpendicular to the vine's rows, to better characterize the whole plant. Flight execution will be autonomous based on a typical flight planning procedure. Oblique images could improve the vines characterization, but flight planning and execution will be more complex and it would make difficult a final productive tool. After evaluation of the bundles results, it is concluded that oblique images do not contribute to improve the final results.

Figure 3.1.2: Example of flight route for a photogrammetric flight planning covering a plot of 7,5 ha.

DATA PROCESSING

The proposed methodology is based on the geometric and radiometric characterization of each individual vine in the plot. Geomatic products were obtained using version 2.0.1 of the Agisoft Metashape Professional software (Agisoft LLC, St. Petersburg, Russia). This software allows the generation of dense point clouds, digital terrain models (DTMs), and orthomosaics, from the aerial images acquired by UAVs. Also, it allows a first treatment of these data, such as point cloud classification. Fig. 3.1.3. shows a flowchart that summarizes the proposed methodology.

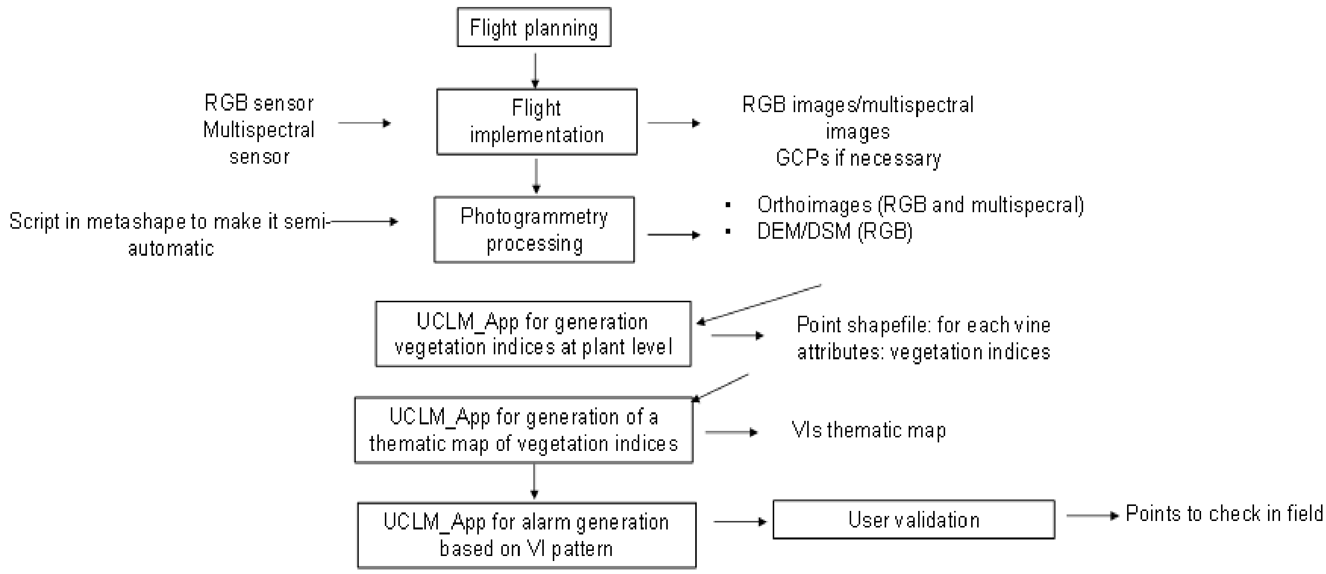


Figure 3.1.3: Flowchart of the proposed methodology

Step 0. Photogrammetric processing

Photogrammetry processing was performed with Agisoft Metashape Professional software (Agisoft LLC, St. Petersburg, Russia) version 2.0.1. A software in Python was implemented to automatize the whole process without the need of human intervention. A two options software was implemented: 1) case 1, in which RTK or PPK systems are available in the drone and, therefore, the coordinates of the centre of the images can accurately obtained; and 2) case 2, in which GCPs should be placed. In this last case, the coordinates of the images should be generated with a manual step in Metashape by locating the GCPs in the images and performing the firsts steps manually. After aligning images and generating the accurate coordinates of the images, these can be saved in a .csv file and continue the whole process automatically.

After generation of the dense point cloud, the classified ground points tool in Metashape is utilized to perform the segmentation between vegetation and ground, which is a critical step to determine the final values of GCC and volume occupied by the canopy. The segmentation is performed exclusively using geometry parameters, although radiometry approach could be also implemented. In this case, in which the disposal of the vegetation is on trellis, it is not necessary to complicate the processing flow with a radiometry approach. The parameters that best perform to classify regular vines mounted on trellis are incorporated in the developed bundles after calibration.

With the classified point cloud as ground and vegetation, a digital terrain model (DTM) is generated by considering point classified as ground. Also, a digital surface model (DSM) I generated with all the points (ground and vegetation).

Summarizing, the following geomatic products are generated for each flight: a) classified point cloud (ground and vegetation), b) orthoimage (RGB and/or multispectral), c) digital terrain model, considering ground points, and d) digital surface model, considering ground and vegetation points.

To evaluate the different vegetation parameters at plant level, a set of frames that defines the area occupied by each plant is previously defined. Different approaches to define these frames can be applied, such as detecting each vine in an orthoimage, measuring each vine with a GNSS-RTK system, or automatically detect each vine with a geometry approach (to be implemented on chameleon project). The automatic detection demands final user modification/validation of the obtained solution.

Each frame is a rectangle with a side corresponding to the distance between plant rows and the other corresponding to the distance between plants in a row. The centroid of the rectangle is coincident with the location of the vine trunk. Figure 3.1.4 show the disposal of all the frames corresponding to each vine of the plot.

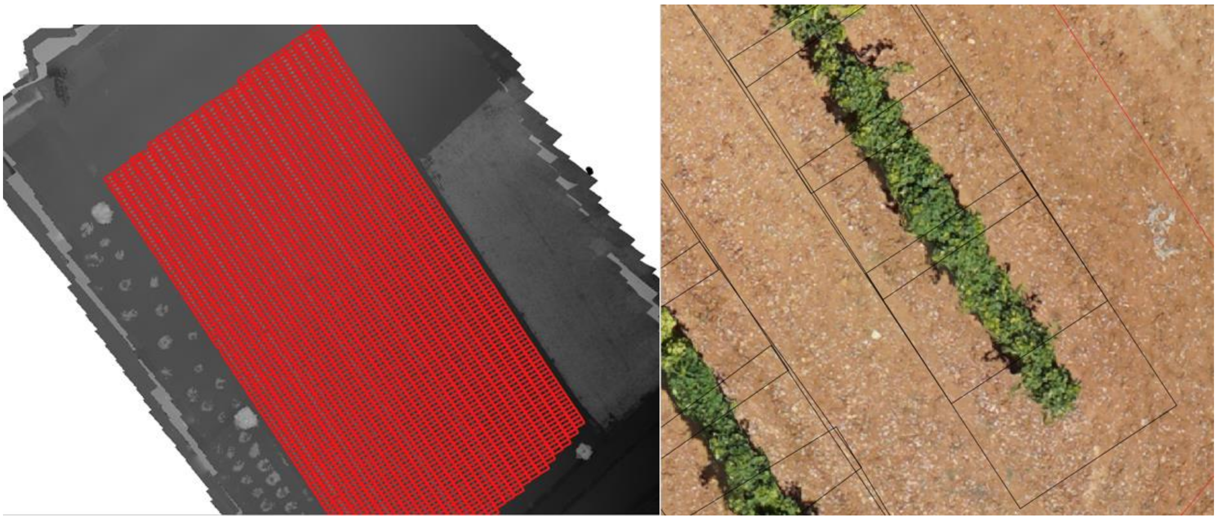


Figure 3.1.4: Location of each frame corresponding to each vine in the plot.

Step 1.1. Determining crop parameters based on RGB products.

To determine GCC and V, the difference between DSM and DTM is obtained. The result is the elevation of the vegetation in each pixel. If the elevation of the vegetation in each pixel is multiplied by the area of each pixel, the volume occupied by the vegetation is obtained. It is important to highlight that the volume obtained with this methodology is an orthogonal projection of the vegetation to the soil. This overestimates the volume occupied by the plant because it does not consider that there is a part of the plant without vegetation in the lower part of the plant. If necessary, this can be corrected by measuring in some plants the distance between the soil and the plant. We do not consider it to be necessary because this overestimation is considered for every plant in the plot and could be considered as an offset in any model that relates with parameter with any other variable of interest. Also, it will not affect plot zoning applications.

The presence of vegetation is obtained by calculating the green canopy cover (GCC), when the difference between DSM and DTM exceeds a limit. This limit has been calibrated and included in the bundle development.

With the GCC and the volume occupied by each plant, referred to each frame, a mapping of the whole plot for any of these two parameters is generated and utilized for different precision viticulture applications.

Step 1.2. Determining vegetation indices (VIs) based on multispectral products.

In this case, from the multispectral orthoimages different VIs are calculated. There will be two options:

- Calculate VIs for the whole vine frame, including soil and plant. In this case, in which only around a 15% of the soil is covered by vegetation, the soil effect is high. It integrates the vegetation response (reflectivity) and the GCC. Thus, it is highly affected by the vigour of the plants.
- Calculate VIs for only the vegetation inside the frame, that requires a vegetation segmentation. In this case, this segmentation is performed using radiometric characteristics of the vegetation, shadows establishing different thresholds in the different bands calibrated specifically for vineyard.

Step 2. Maps generation

After calculating the vegetation parameters for each vine (one value per vine) a map will be generated that shows the spatial distribution of the plant-based parameter in the whole plot. To do so, a kriging processing of the data is performed. This results in maps for each of the selected parameters that are useful for visual crop growth and monitoring process.

Step 3. Alarm generation

Generation of GCC, V or VIs by themself are useful products to perform crop growth and monitoring. However, an alarm system is also implemented based on two aspects:

- Differences of GCC, V or VIs in a flight date (which plants are more or less developed). The alarm will highlight those vines with lower (threshold managed by user) development state.
- Differences of these parameters between one flight date and the next (which plants have a higher or lower growth rate. The alarm will highlight those vines with lower (threshold managed by user) growth rate.

RESULTS / OUTCOMES

Some of the results obtained after applying the bundle on a vineyard plot are shown in the following figures.

Figure 3.1.5: NDVI map of individual vines, considering soil and plant

This result is a huge improvement compared with the generation of classical NDVI (or other VIs) since it represents the area of each individual vine, and not the response of each pixel of the plot.

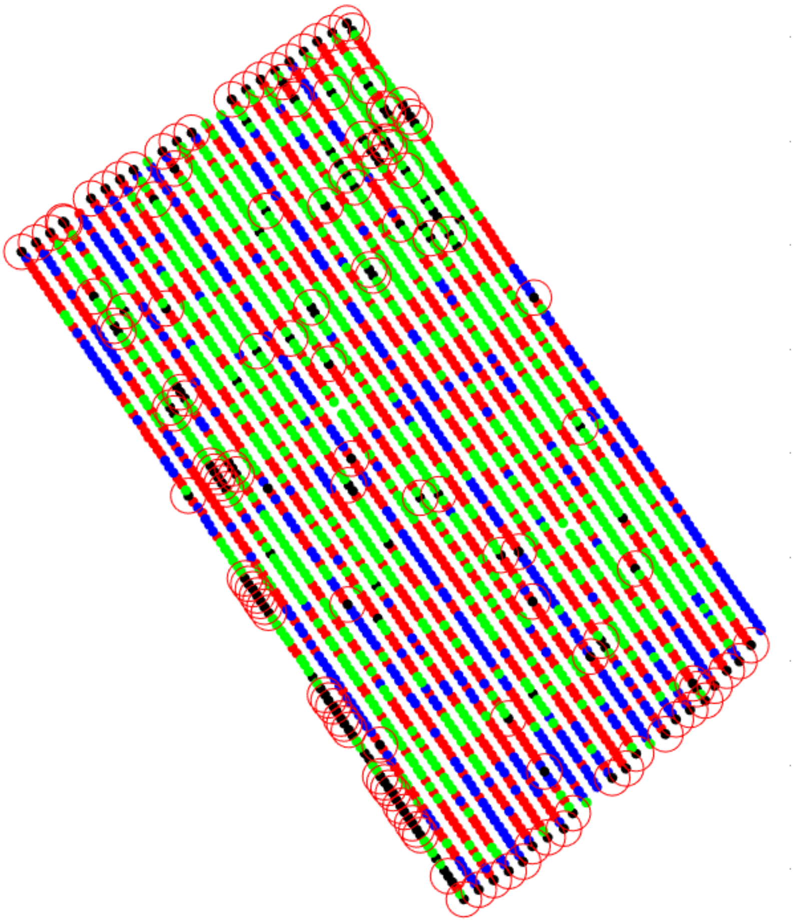


Figure 3.1.6: Detection of the 10% of vines with lower development in a flight date (red circle) and clustering of vines based on crop development (4 classes)

AI DEEP LEARNING PROCESS / ALGHORITHMS

An attempt to localize individual vines using AI algorithm will be tested in a collaboration of UCLM and AIDEAS, with a high risk due to de complexity of the action. In case it does not work with a high accuracy, the above explained approaches will be implemented.

- Aim and challenges

The objective of this undertaking is to oversee a vineyard, focusing on the detection and quantification of vine plants through the application of object detection, a widely recognized computer vision technique.

Challenges: The task of monitoring a vineyard and employing object detection for the identification and counting of vine plants involves many challenges. One significant challenge involves the inherent variability in the appearance of vine plants due to factors such as growth stage, lighting conditions, and seasonal changes. Adapting the object detection model to accurately identify vines under these diverse circumstances is a complex task. Another challenge is posed by occlusions within the vineyard environment. Vines may be obscured by foliage, structures, or other plants, making it difficult for the object detection algorithm to precisely delineate individual plants. This requires the model to exhibit robustness in handling partial visibility and overlapping instances. Moreover, the scale of vineyards can be extensive, leading to large datasets that demand efficient processing and computational resources. Balancing the need for high-resolution imagery with the computational constraints of real-time monitoring is a challenge that needs careful consideration.

The dynamic nature of vine growth introduces further complexity. Continuous changes in the appearance and layout of vines over time necessitate a model capable of adapting and learning from evolving patterns, requiring not only accurate object detection but also temporal awareness. Additionally, environmental factors such as varying lighting conditions, weather changes, and the presence of shadows can impact the performance of the object detection model. Ensuring the model's robustness to these environmental variations is crucial for its reliable deployment in real-world vineyard monitoring scenarios. In summary, the challenges encompass addressing variations in vine appearance, handling occlusions, managing large-scale datasets, adapting to dynamic growth patterns, and ensuring resilience to environmental factors. Overcoming these challenges is essential for the successful implementation of an effective and accurate vineyard monitoring system using object detection techniques.

- Proposed methodology

Vine detection requires a series of pre-processing techniques and accurate prediction models. To accomplish this, we follow a sequence of steps to process and analyse the data effectively. These stages encompass handling the dataset, implementing image preprocessing, applying various annotation approaches, employing augmentation techniques to enhance generalization, utilizing object detection methods, and conducting a thorough validation process.

  - **Dataset**

UCLM supplied AiDEAS with an orthoimage captured by a drone at a precise moment during the day. This image encompasses a vast and meticulously arranged vineyard, housing a total of 2639 vine plants. Additionally, a second image was acquired from an alternative perspective within a neighbouring vineyard.

Figure 3.1.7: A vineyard from a vertical view

In the below depiction, the plant roots are showcased in an enlarged format, aiding in the differentiation between shadows and the particular features of interest evident in the image above.

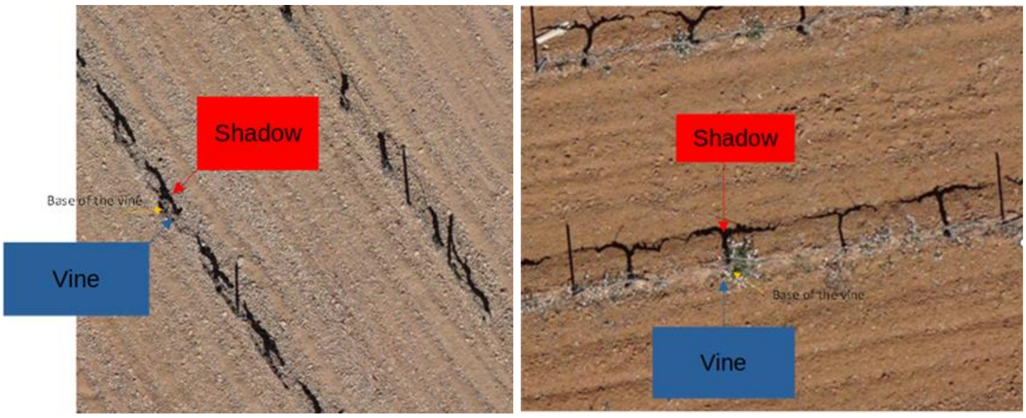


Figure 3.1.8: A vineyard in zoom view

- **Image pre-processing**

In response to the inherent challenges of the project, we undertook the development of three distinct approaches to address and overcome the complexities associated with this formidable task.

1st Approach of image pre-processing

UCLM, utilizing a software, supplied us with individual images isolating each vineyard root. These images encompass the complete root structure, along with details of the surrounding shade, wires, and soil. Therefore, in one of our approaches we used these 2639 images to annotate the vines and then train our model.

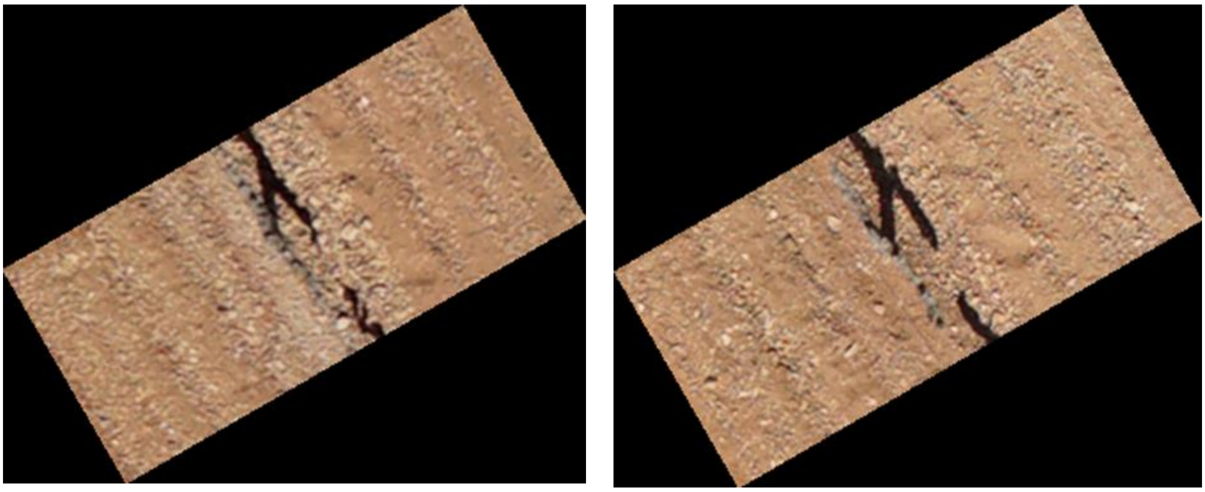


Figure 3.1.9: Isolated vine plants.

2nd Approach of image pre-processing

In this particular approach, we implemented a 60-degree rotation of our images.

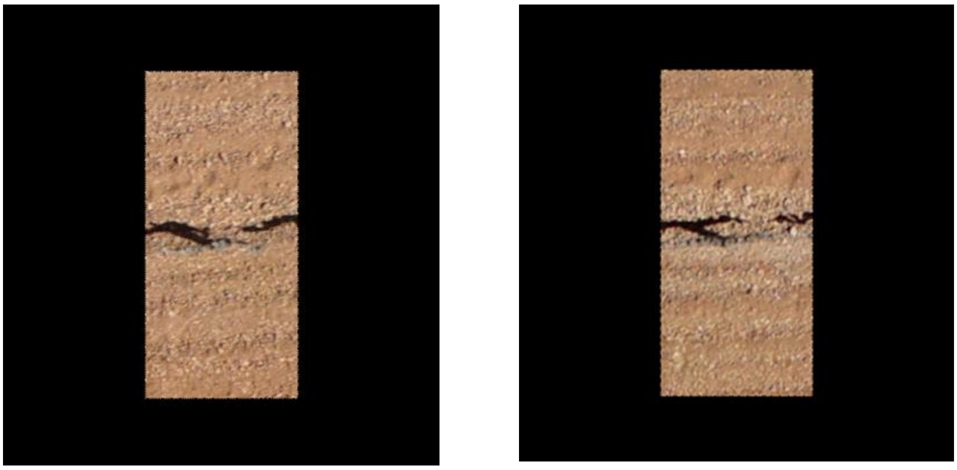


Figure 3.1.10: Rotated isolated vine plants.

3rd Approach of image pre-processing

In this approach, we conducted segmentation of our initial image, resulting in the generation of 377 individual images.

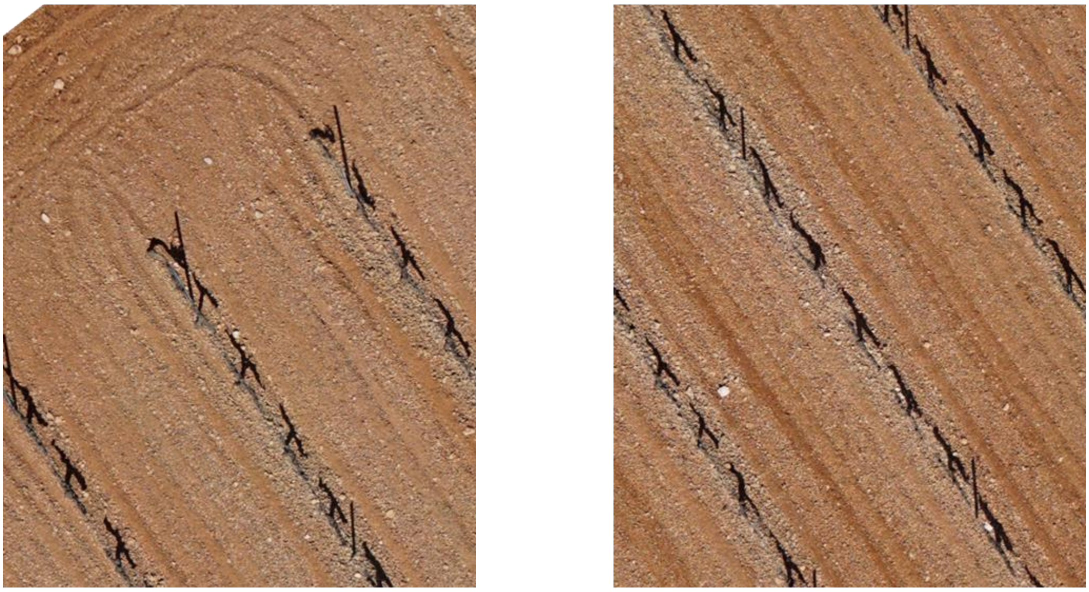


Figure 3.1.11: Isolation of vine plants in clusters

– **Annotation approaches**

Due to the diverse image preprocessing approaches, we integrated three distinct methods for annotating images. Two of these methods were implemented using Python code, while the third was accomplished through the use of the free software, LabelMe.

Boundary Box in the whole image

In this particular approach, we devised a manually crafted code to execute automatic annotation on the outline of each image.

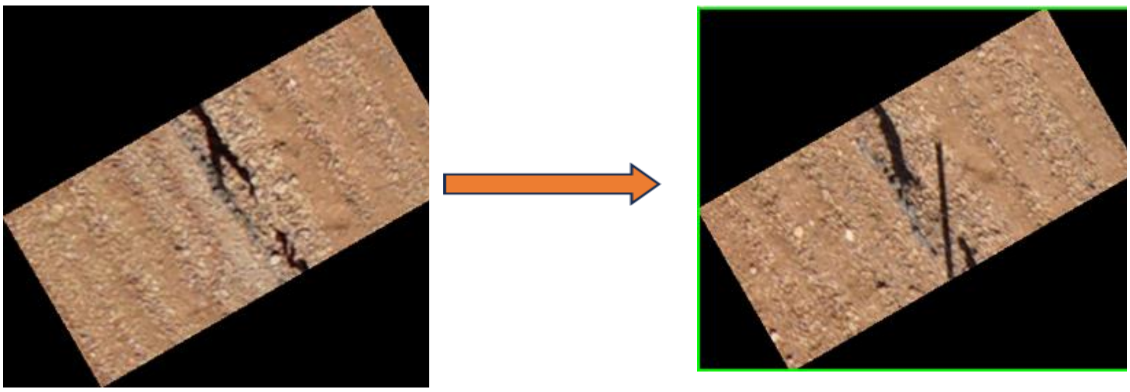


Figure 3.1.12: Example of boundary box in the whole image.

Boundary Box in the rotated image

For images that were previously rotated by 60 degrees, we designed a code to eliminate the black regions and exclusively annotate the ground containing the vine plants.

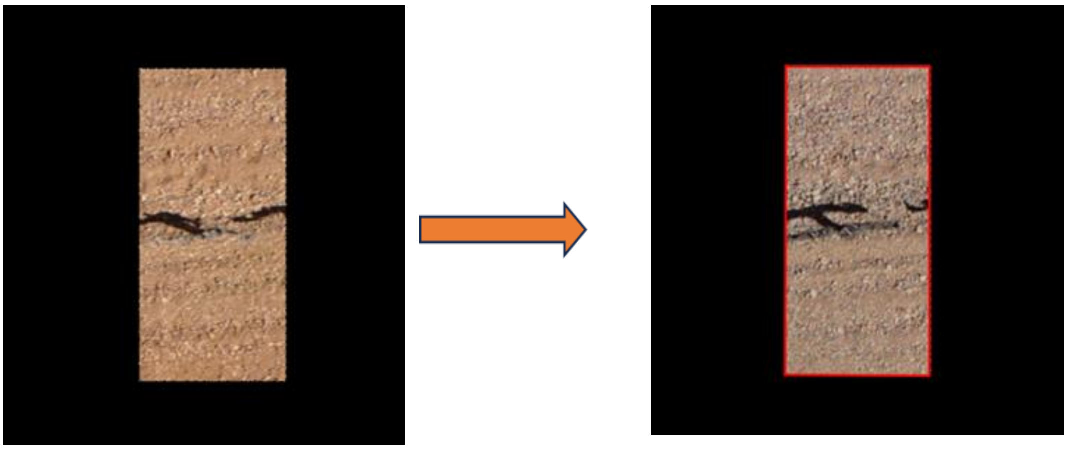


Figure 3.1.13: Example of boundary box in the rotated image.

Boundary Boxes in each cluster image

In order to construct the boundary boxes in each vine plant of each cluster image (annotation of these images) the LabelMe tool was employed. LabelMe is a widely utilized open-source software tool in the domain of computer vision and machine learning for image annotation and segmentation tasks. Developed by the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), LabelMe provides an intuitive graphical user interface that enables users to annotate images by drawing bounding boxes around objects and labelling them with corresponding categories. LabelMe also supports semantic segmentation annotations, enabling pixel-level labelling for a more granular understanding of object boundaries.

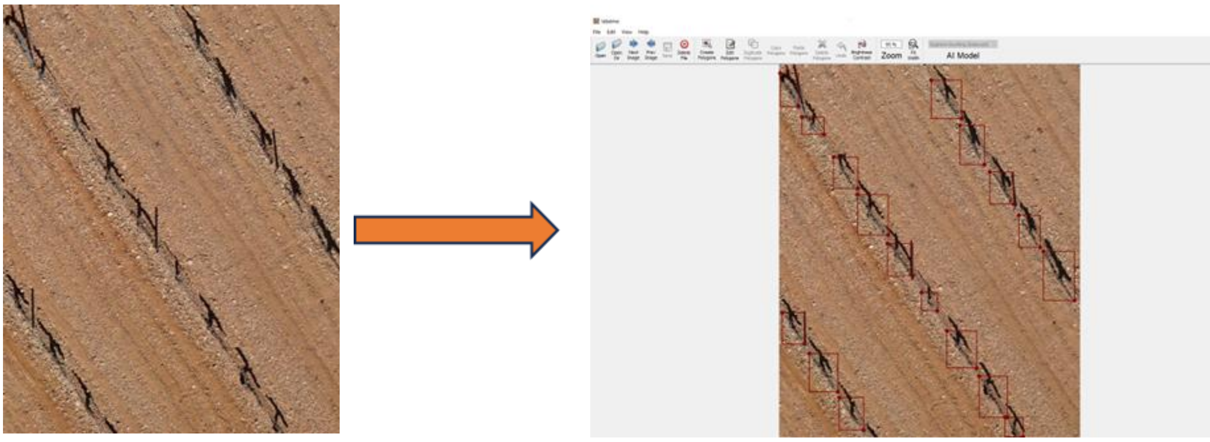


Figure 3.1.14: Label me annotation.

– **Augmentation**

Data augmentation offers several advantages in the context of training deep learning models, particularly for computer vision tasks. By applying diverse transformations to the training dataset, augmentation enriches the dataset and contributes to the overall improvement of model performance. One primary advantage of data augmentation is enhanced model generalization. Through the introduction of various transformations such as flipping, rotation, scaling, and cropping, the model learns to recognize objects in different orientations, sizes, and spatial configurations. This increased adaptability translates to better performance when faced with unseen data during deployment. Moreover, data augmentation acts as a regularization technique. The augmented training samples introduce randomness and variability, preventing the model from memorizing specific patterns in the training data. This regularization helps mitigate overfitting, ensuring that the model does not become overly specialized to the training set and can generalize effectively to new, unseen examples. Additionally, data augmentation contributes to improving the model's robustness to noise and variations in real-world data. Techniques such as adding noise, applying blurs, and adjusting brightness and contrast help the model become more resilient to environmental factors, ensuring its effectiveness in diverse and unpredictable conditions.

In summary, the advantages of data augmentation include enhanced generalization, regularization to prevent overfitting, expanded dataset size, and improved robustness to real-world variations. By incorporating these diverse examples during training, the model becomes more versatile and capable of handling a wide range of scenarios, ultimately leading to better overall performance and reliability in practical applications. Below are presented in detail the augmentation techniques applied in our dataset.

HorizontalFlip (p=0.5): Horizontal flipping is a fundamental augmentation technique that transforms images by mirroring them along the horizontal axis with a 50% probability. This operation is instrumental in training models to recognize objects regardless of their orientation. By presenting the model with both the original and flipped versions of an image, it learns to identify features invariant to horizontal reflections, enhancing its ability to generalize across diverse scenarios.

VerticalFlip (p=0.5): Similar to horizontal flipping, vertical flipping involves mirroring images along the vertical axis with a 50% probability. This augmentation technique contributes to the robustness of a model by exposing it to variations in object perspectives. It aids in the development of models capable of recognizing objects from different angles, a crucial aspect for real-world applications.

RandomRotate90 (p=0.5): The RandomRotate90 augmentation introduces randomness by rotating images randomly by 90 degrees with a 50% probability. This operation helps the model learn to recognize objects in various orientations, preventing it from relying too heavily on specific angles. By incorporating rotational variations during training, the model becomes more adaptable to diverse real-world scenarios.

Rotate (limit=45, p=0.5): Rotate augmentation goes a step further by randomly rotating images by up to ±45 degrees with a 50% probability. This fine-tuned rotation enables the model to handle images with subtle variations in object alignment, making it more resilient to different spatial configurations and improving its overall robustness.

RandomScale (scale_limit=0.1, p=0.5): Random scaling introduces diversity by randomly resizing images by up to 10% with a 50% probability. This augmentation helps the model generalize better to objects of varying sizes, ensuring it remains effective across a spectrum of scales commonly encountered in real-world scenarios.

RandomCrop (width=450, height=450, p=0.5): The RandomCrop technique involves randomly cropping images to a size of 450x450 pixels with a 50% probability. This operation encourages the model to focus on important regions, improving its ability to recognize objects regardless of their placement within the image.

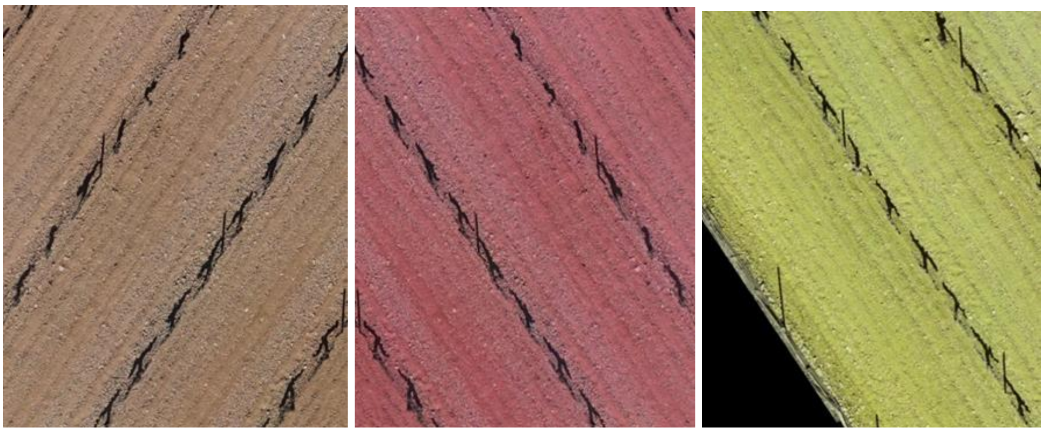
GaussNoise (p=0.2): GaussNoise augmentation introduces Gaussian noise to images with a 20% probability. This random addition of noise during training helps the model become more robust to variations in pixel intensity, making it less sensitive to minor fluctuations in input data and enhancing its generalization capabilities.

GaussianBlur (blur_limit=3, p=0.5): GaussianBlur applies a Gaussian blur to images with a limit of 3 and a 50% probability. This augmentation technique smoothens images, reducing sensitivity to high-frequency noise. By incorporating blur variations during training, the model becomes more resilient to distortions and improves its overall performance.

RandomBrightnessContrast (p=0.2): RandomBrightnessContrast augments images by applying random changes to brightness and contrast with a 20% probability. This variation in intensity during training enhances the model's adaptability to diverse lighting conditions, ensuring it performs well under different brightness and contrast levels.

RGBShift (p=0.2): RGBShift augments images by randomly shifting the values of the red, green, and blue channels with a 20% probability. This operation helps the model learn to recognize objects under different color distributions, contributing to its color invariance and improving its performance across a range of color variations in real-world scenarios.

Here, we observe the implementation of several of the previously mentioned techniques in the vineyard settings.



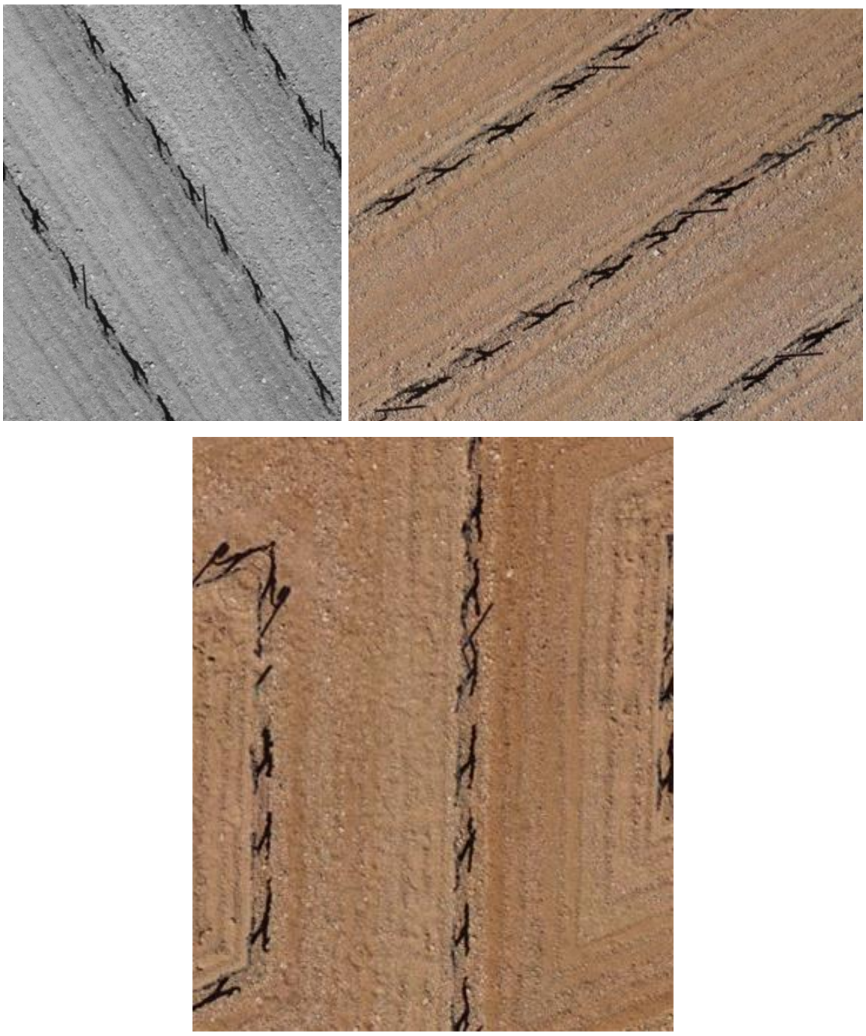


Figure 3.1.15: Augmentation techniques in vineyards

- **Object detection with Yolov8**

Object detection in vineyards

Since 2010, an increasing number of farmers have embraced the utilization of unmanned aerial vehicles (UAVs) for obtaining remote sensing data. These UAVs offer high spatial-temporal resolution, allowing farmers to assess the condition of their crops and monitor changes in their fields more effectively. Among the various imaging sensors employed, such as multispectral and RGB cameras, the latter are widely preferred in vineyards. They play a crucial role in characterizing the vegetative development of the canopy and identifying missing vines along the rows. In 2023, Salvatore et al. diverse methodologies aimed at identifying and pinpointing individual vines within a commercial vineyard. The focus is on angled RGB images captured during the winter dormant period when the canopy is devoid of leaves. Specifically, they accurately combined photogrammetric techniques and spatial analysis tools[6]. In another work, for the detection of missing vine plants by using RGB images which were acquired by UAV

[6] Di Gennaro, S.F., Vannini, G.L., Berton, A., Dainelli, R., Toscano, P. and Matese, A., 2023. Missing Plant Detection in Vineyards Using UAV Angled RGB Imagery Acquired in Dormant Period. Drones, 7(6), p.349.

platform they applied 2.5D and 3D approaches[7]. Additionally, an automated technique for identifying grapevine trunks by leveraging 3D point cloud data obtained from UAV sources was proposed. The methodology in this workplace emphasizes discerning critical geometric parameters, ensuring the inclusion of every plant within the 3D model. The primary objective is to facilitate the detection of missing plants through a thorough analysis of the spatial characteristics[8].

Object detection with Yolov8

Yolov8, a state-of-the-art object detection model, has demonstrated significant prowess in the field of computer vision, particularly in the task of plant detection in vineyards. Yolov8, an evolution of the YOLO (You Only Look Once) architecture, excels in real-time object detection, offering a balance between precision and computational efficiency. In the context of vineyard applications, Yolov8's ability to simultaneously detect multiple objects in a single pass is particularly advantageous. The model's deep neural network architecture enables it to capture intricate features and spatial relationships within the vineyard imagery, allowing for precise identification of individual plants.

The Yolov8 model's robustness is further enhanced by its capacity to handle variations in plant appearances and sizes, crucial attributes in the context of diverse vineyard landscapes. The model's anchor-based approach aids in accurately localizing and classifying plants, even under challenging conditions such as varying illumination and occlusions. The efficiency of Yolov8 in processing high-resolution images ensures that it can effectively cover large vineyard areas, making it well-suited for comprehensive plant detection tasks.

Moreover, Yolov8's adaptability and ease of integration with different hardware platforms make it a practical choice for deployment in real-world vineyard monitoring systems. Its capability to identify and track plants in real-time allows for timely and informed decision-making in precision agriculture applications. In conclusion, Yolov8 stands out as a powerful solution for plant detection in vineyards, offering a compelling combination of accuracy, efficiency, and versatility that aligns with the demands of modern agricultural practices.

Building on the strengths outlined above, we applied YoloV8 specifically for vine recognition in this task. This technology's precision in detecting and tracking plants aligns perfectly with our objective of enhancing the efficiency and effectiveness of vine management practices.

---

7. Di Gennaro, S.F. and Matese, A., 2020. Evaluation of novel precision viticulture tool for canopy biomass estimation and missing plant detection based on 2.5 D and 3D approaches using RGB images acquired by UAV platform. Plant Methods, 16, pp.1-12.

8. Jurado, J.M., Pádua, L., Feito, F.R. and Sousa, J.J., 2020. Automatic grapevine trunk detection on UAV-based point cloud. Remote Sensing, 12(18), p.3043.

Validation metrics

The evaluation criteria used to assess the performance of the vine detection model, which employs YOLOv8, include three primary metrics: Precision, Recall, and Mean Average Precision (mAP). After consulting with the end-users, it was determined that an acceptable level of mAP is greater than 0.80.

Table 3.1.2: Distribution of metrics.

Metric	Description
Precision	Measures the accuracy of positive predictions
Recall	Assesses the model's ability to identify all relevant instances
mAP	Evaluates the model's precision and recall at different thresholds

These metrics are crucial for understanding both the accuracy and the generalization capabilities of our model. To facilitate a comprehensive assessment, we have divided our dataset into three parts: 70% is used for the training set, 20% for the validation set, and 10% for the testing set. This division ensures a balanced and effective approach for both training and evaluating the model's performance across various data segments.

- Results

The results section showcases outcomes in the form of labels and predictions, illustrating the model's performance in detecting vines across three distinct approaches. Detailed information, including accuracy metrics and the model's efficacy under diverse conditions, is presented to provide a comprehensive evaluation of its performance.

– **Results of the rotated Images approach**

In this approach, we have adopted an innovative approach where boundary boxes are rotated vertically. The details of the deployed approach are given in the table below.

Table 3.1.3: Parameters of training.

Dataset	
Image pre-processing	60-degree rotation of images
Number of images	2639
Learning	
OD model	Yolov8 Nano version
Hyper-parameters	Epochs: 100, Batch size: 16, Image size: 640, Optimizer: auto, Weight decay: 0.0005
Validation	
Data split	70 train set/20 validation set/10 test set
Validation metrics	mAP, precision, recall

The results of this method are presented in a series of figures. Figure 3.1.16 displays a collection of images used during the training phase. Each image is marked with a boundary box, highlighting the vine, its shadow, and the adjacent ground area. These images offer a clear view of how the algorithm interprets and learns from the training data.

Figure 3.1.16: Batch images of Approach 1 training.

Figure 3.1.17 The figure illustrates the metrics obtained from the training process over each epoch. The main focus is on the mean Average Precision (mAP), a crucial metric for drawing conclusions. The graph demonstrates that the highest mAP score achieved during the 100 epochs is 0.995. These results come from the same dataset, which was internally split into training, validation, and testing sets. Data augmentation was also employed to ensure some level of generalization. While the mAP score is very high, it does not entirely reflect the model's reliability. Future experiments in D4.7 will include cross-dataset validation to more precisely evaluate the generalization to other datasets collected under partially different conditions.

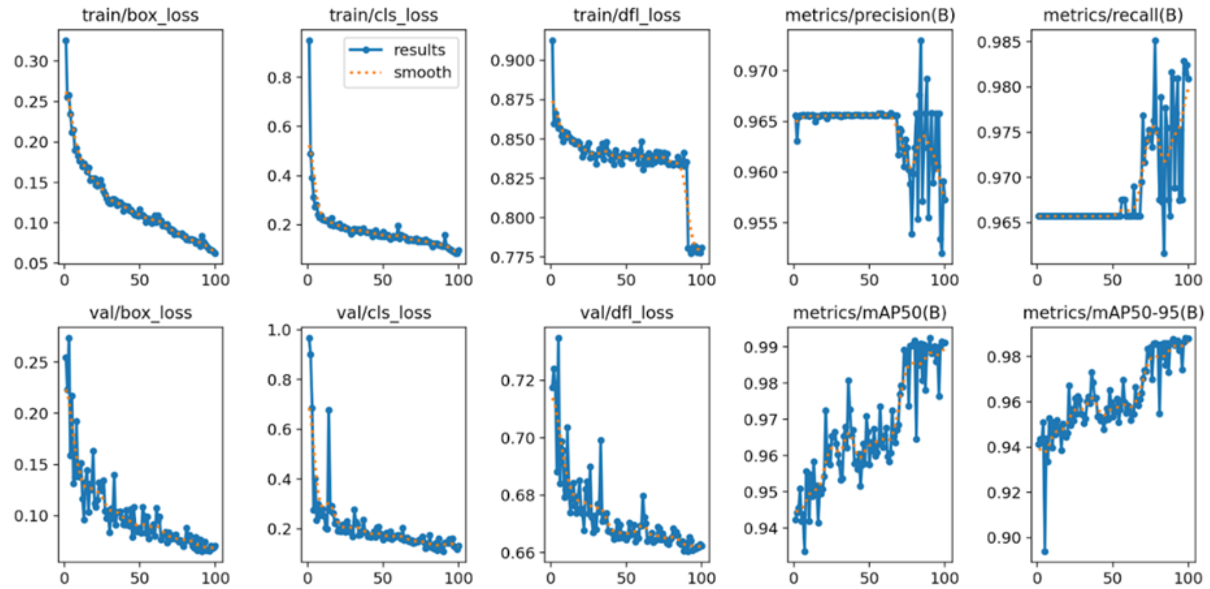


Figure 3.1.17: The training results of this approach are shown. The x-axis represents the epochs of training, and the y-axis represents the metrics referred to at the top of each graph.

Figure 3.1.18. provides a side-by-side comparison of an original image and its interpretation by the best-performing model. It includes the original boundary boxes and those predicted by the model, illustrating the accuracy and effectiveness of our approach.

Labels pred

Figure 3.1.18: Images with Original Labels on the Left and Predicted Images on the Right.

- **Results of isolated vines clusters without augmentation**

This approach involves segmenting the original orthoimage into tiles, each of which is meticulously annotated to identify individual vines.

The details of the deployed approach are given in the table below.

Table 3.1.4: Parameters of training.

Dataset	
Image pre-processing	Segment the orthoimage into tiles
Number of images	780
Learning	
OD model	Yolov8 Nano version
Hyper-parameters	Epochs: 100, Batch size: 16, Image size: 640, Optimizer: auto, Weight decay: 0.0005
Validation	
Data split	70 train set/20 validation set/10 test set
Validation metrics	mAP, precision, recall

Figure 3.1.19. presents a series of these annotated images used in the training phase, showcasing the detailed labelling process.

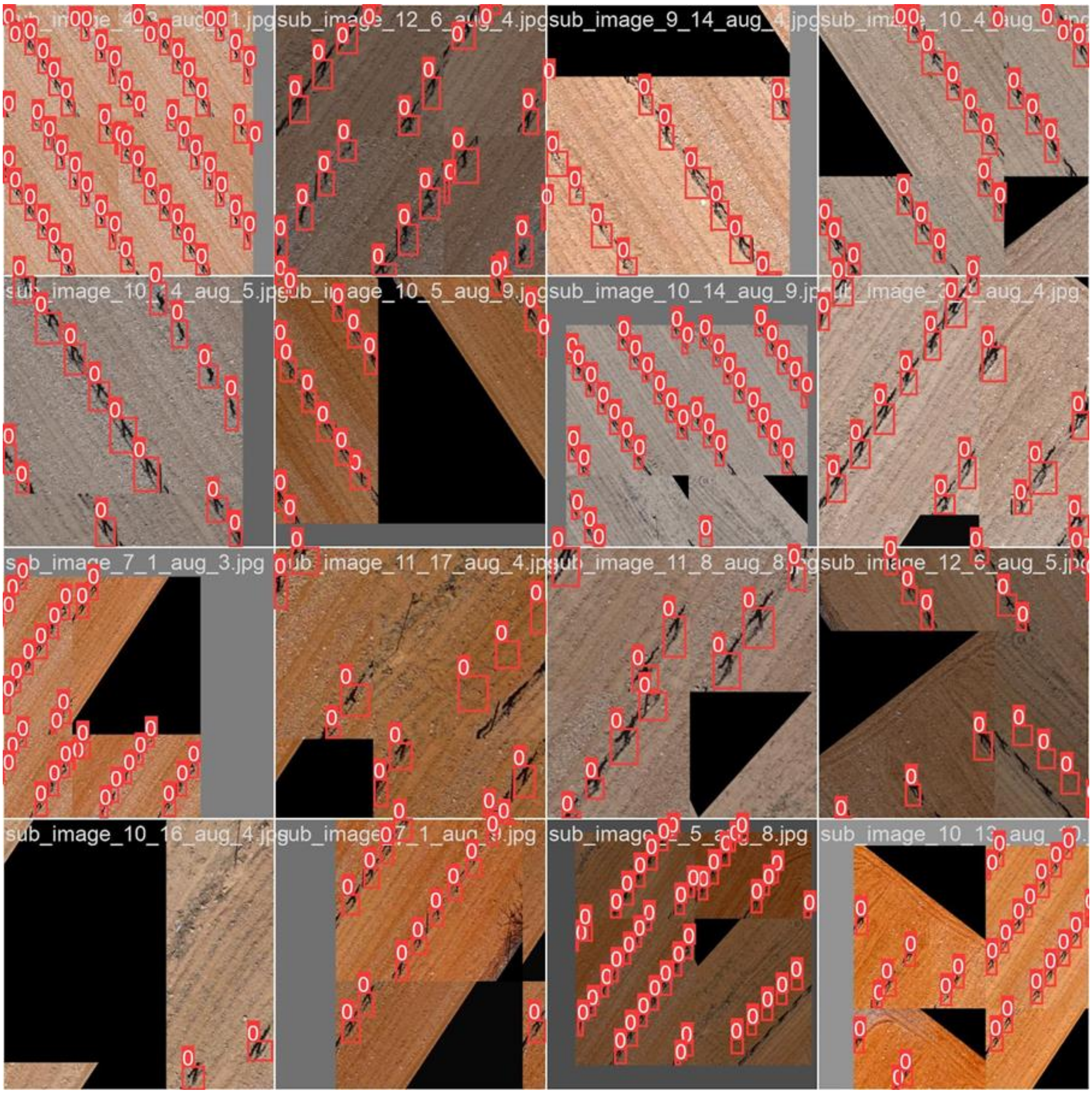


Figure 3.1.19: Batch images of Approach 2 training.

Figure 3.1.20. offers a comprehensive view of the training procedure's performance metrics, evaluated epoch by epoch. The highlight here is the mAP, a key indicator of model accuracy, which peaks at 0.927 in the best-performing epoch out of a total of 100.

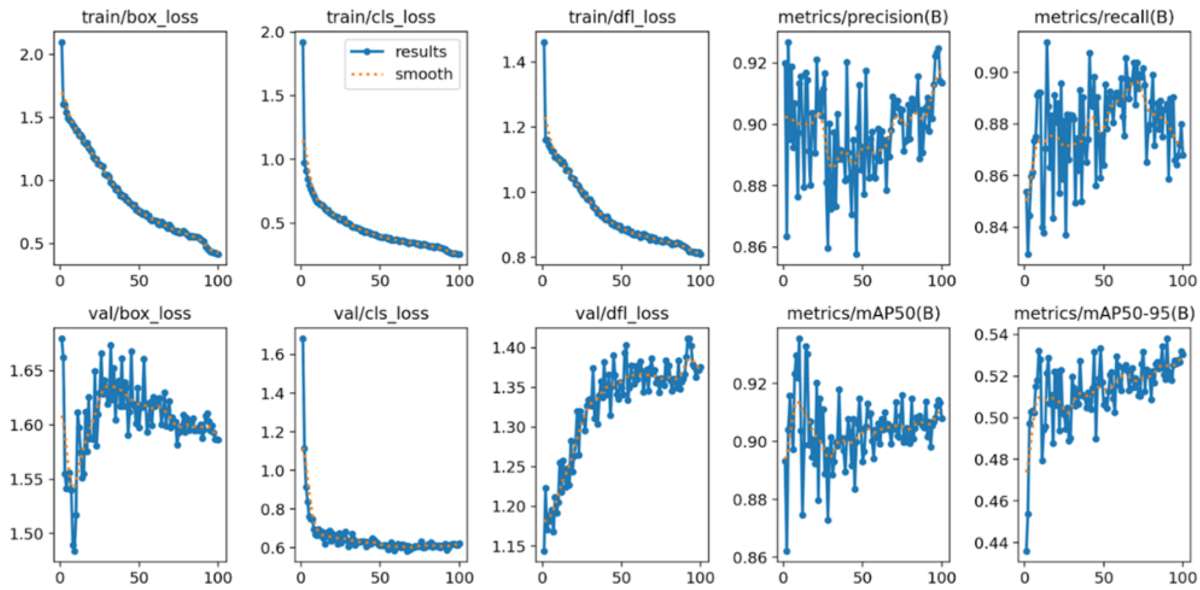


Figure 3.1.20: The training results of this approach are shown. The x-axis represents the epochs of training, and the y-axis represents the metrics referred to at the top of each graph.

Figure 3.1.21. demonstrates the practical application of this approach. It juxtaposes the original image, complete with boundary boxes, against the predictions made by the most effective model, offering a visual comparison between the pre-trained image and the AI's interpretation.

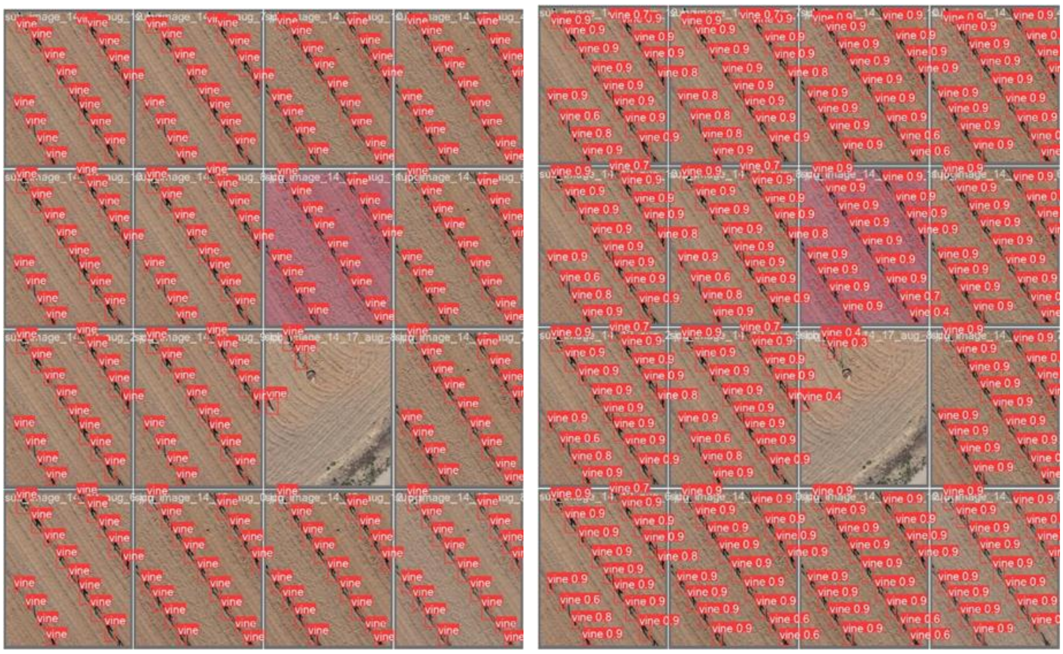


Figure 3.1.21: The training results of this approach are shown. The x-axis represents the epochs of training, and the y-axis represents the metrics referred to at the top of each graph.

- **Results of isolated vines clusters with augmentation**

The approach detailed here involves dividing the original orthoimage into tiles, with each tile being meticulously annotated to identify individual vines. This process is further enhanced by employing augmentation techniques, as described in the methodology section, to train the model. In Figure 3.1.22. we present a collection of these annotated images that were utilized during the training phase. These images showcase the extensive labelling process, now enriched with varied augmentation techniques to improve model robustness.

The details of the deployed approach are given in the table below.

Table 3.1.5: Parameters of training.

Dataset	
Image pre-processing	Segment the orthoimage into tiles and apply augmentation techniques
Number of images	3569
Learning	
OD model	Yolov8 nano version
Hyper-parameters	Epochs: 100, Batch size: 16, Image size: 640, Optimizer: auto, Weight decay: 0.0005
Validation	
Data split	70 train set/20 validation set/10 test set
Validation metrics	mAP, precision, recall

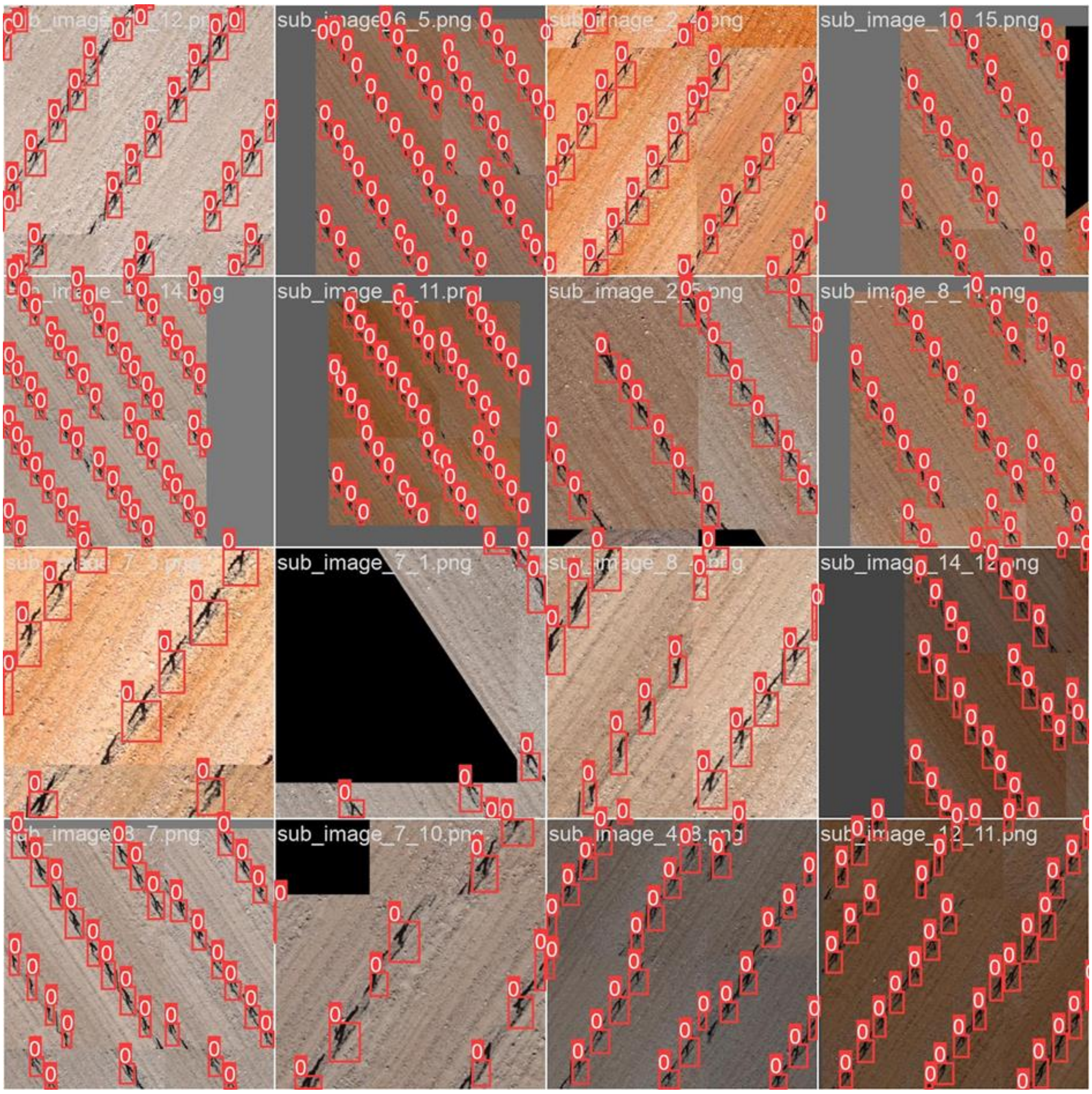


Figure 3.1.22: Batch images of Approach 3 training.

Figure 3.1.23. illustrates the evolution of training performance metrics, tracked across each epoch. A crucial measure of model accuracy mAP, which achieves its peak value of 0.997, marking the best performance among the 100 epochs.

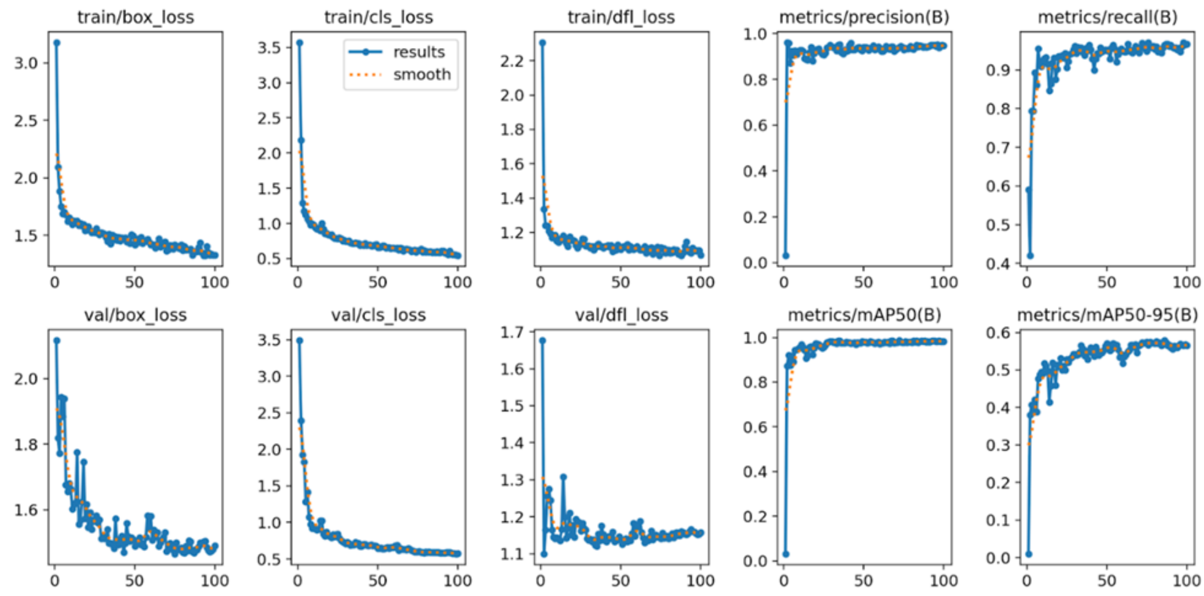


Figure 3.1.23: The training results of this approach are shown. The x-axis represents the epochs of training, and the y-axis represents the metrics referred to at the top of each graph.

To sum upthe results come from the same dataset, which was internally split into training, validation, and testing sets. Data augmentation was used to ensure some level of generalization. While the highest mAP score of 0.995 is very high, it may not fully reflect the model's reliability. Future work will include cross-dataset validation to better assess generalization to other datasets collected under different conditions.

- Conclusions

In conclusion, we have successfully trained and utilized the weights of YOLOv8 in three distinct approaches to detect vineyards from aerial imagery. Each approach exhibited exceptional performance; however, the task presents inherent challenges due to the nature of the data. The complexity arises from factors like the angle and altitude of capture and sun position that cause shadows in different shapes. Specifically, the unique photogrammetric flight planning of UAVs over vineyards and the varying orientation of vines contribute to difficulties in consistent detection. Shadows cast by the vines, not easily discernible from a vertical perspective, emerge as a significant factor influencing model performance.

The key challenges encountered include the variability in vine appearance due to growth stages, lighting, and seasonal changes, and the occlusions caused by foliage and structures. Large-scale datasets demand efficient processing, while the dynamic nature of vine growth necessitates a model capable of adapting to evolving patterns. Environmental factors like lighting and weather variations further complicate the detection process. Addressing these challenges is crucial for the effective monitoring of vineyards using object detection techniques.

Looking ahead, our future work will focus on overcoming these challenges to enhance the model's accuracy and reliability. A promising direction involves the utilization of an external dataset, acknowledging that vines are difficult to recognize but shadows play a crucial role. We plan to leverage image processing techniques to emphasize the contours of these shadows. By processing images to highlight contours with specific shapes and colours (particularly considering the predominantly black nature of shadows), we aim to improve the model's ability to accurately identify vines. This approach, combined with our ongoing research and model refinement, will be pivotal in advancing our object detection capabilities for vineyard monitoring.

3.2. VINEYARD WATER STRESS DUE TO DROUGHT

RELEVANCE AND PROBLEMATIC

The effect of the water deficit, and consequently, the appropriate regulated deficit irrigation (RDI) management approach, depends on the vine phenological growth, plant water stress, and the soil and climatic water conditions.

RDI techniques, like other tools for improving agricultural water use, require the monitoring of crop water status and its progress over time. Midday stem water potential (Y_{stem}) has been proposed as a significant physiological indicator of water status for irrigated and rain-fed vines. Nevertheless, these measurements are performed manually, are time-consuming, and may not be representative of the spatial variability of the water status over the whole vineyard. Moreover, Y_{stem} only represents a snapshot of plant water stress at the time of measurement. Additionally, the changes in irrigation depths with time and the lack of uniformity in water application during the irrigation period emphasize the need for a methodology that would cover the entire season, integrating the short-term variations in vine water status.

Remote sensing (RS) methods based on spectral vegetation indices (VIs) and infrared thermometry are widely used for crop water stress detection because they are non-destructive and have low labour and time requirements. The most affected process by water stress promoted by high temperatures is net photosynthesis. The stomatal conductance, chlorophyll content, and leaf transpiration rate are also influenced by water stress. The accumulated effect of these physiological parameters is evidenced in the leaf spectral response tracked by RS technologies.

At this point, it is important to notice that estimating water stress with thermal imagery or multispectral imagery has a totally different approach and objective. In case of thermal band, this information estimates stomatal conductance, which refers to the water status at the exact time of measurement. Thus, it estimates the instantaneous stress level. However, when multispectral imagery is used, water status is referred to the cumulative stress that the plant has suffered during time along the whole crop season up to the instant of measurement. In any case, punctual field measurements at the same time than the images are acquired should be obtained to correlate the spectral response with the field measurement.

Thus, the objective of this bundle is to implement both methodologies: a) one using thermal imagery that estimates instantaneous water stress and b) the approach using multispectral imagery to estimate cumulative water stress effects.

UAV CHARACTERISTICS

Described in Table 3.2.1. and Figure 3.1.1.

CAMERA CHARACTERISTICS

To obtain the required products the following sensors, included in Chameleon project, can be used:

Multispectral camera, such as Micasense RedEdge MX, for obtaining vegetation indices.

Thermal camera, such as FLIR TAU2.

FLIGHT CHARACTERISTICS

A typical photogrammetric flight planning should be implemented. After selection of the region of interest (RoI), the main parameters to consider for each type of product (multispectral or thermal) are the followings:

- Ground sample distance (GSD)
  - Multispectral products: max. 5 cm.
  - Thermal products: max. 8 cm.
- Overlapping and sidelapping:
  - Multispectral products: 80% (forward overlap) 50% (side overlap)
  - Thermal products: 80% (forward overlap) 50% (side overlap)

In any case, the maximum flight height of 120 m will not be exceeded as well as any other restriction commanded by law.

In case of vineyards, it is preferable to perform the flight perpendicular to the vine rows, to better characterize the whole plant (Figure 3.1.2.).

DATA PROCESSING

The proposed methodology is based on the radiometric characterization of each individual vine in the plot. Geomatic products were obtained using version 2.0.1 of the Agisoft Metashape Professional software (Agisoft LLC, St. Petersburg, Russia). This software allows the generation of dense point clouds, DEMs, and orthomosaics, from the aerial images acquired by UAVs. Fig. 3.2.1. shows a flowchart that summarizes the proposed methodology.

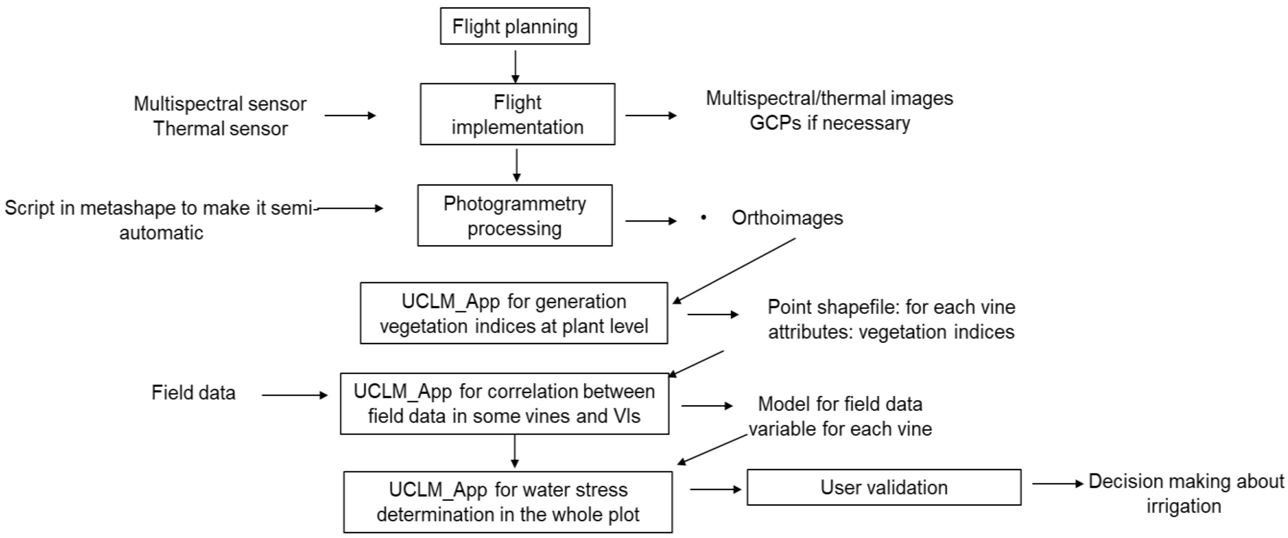


Figure 3.2.1: Flowchart of the proposed methodology

Step 0. Photogrammetric processing

Photogrammetry processing was performed with Agisoft Metashape Professional software (Agisoft LLC, St. Petersburg, Russia) version 2.0.1. A software in Python was implemented to automatize the whole process without the need of human intervention. A two options software was implemented: a) case 1, in which RTK or PPK systems are available in the drone and, therefore, the coordinates of the centre of the images can accurately obtained; and b) case 2, in which GCPs should be placed. In this last case, the coordinates of the images should be generated with a manual step in Metashape by locating the GCPs in the images and performing the firsts steps manually. After aligning images and generating the accurate coordinates of the images, these can be saved in a .csv file and continue the whole process automatically. In this case, only orthoimages (multispectral or thermal) will be generated, because the geometry characterization with multispectral or thermal is poor compared with RGB products. Special care is taken with the generation of thermal products, due to the poor radiometric quality of thermal sensors and low texture of the images, which demands special treatment in the photogrammetry processing.

To evaluate the different vegetation parameters at plant level, it is implemented the approach of frames generation for each vine described in the bundle "Crop growth and development monitoring" and visualized in Fig. 3.1.4.

Step 1. Determining vegetation indices (VIs) based on multispectral or thermal products.

In this case, from the multispectral orthoimages different VIs are calculated. Only radiometric segmentation will be implemented by using the different bands to segmentate vegetation from soil and even shadows (with the calibrated process described in bundle 3.1.). It is important to highlight that the response of the vegetation to accumulated stress should be accounted without the interference of soil effect.

In the case of thermal products, temperature values will be normalized with atmosphere temperature and different VIs based on temperature will be obtained and related with the obtained field measurements.

Step 2. Generation of crop water stress models

Based on the calculated VIs and the field measurements, different models (linear, multilinear, and others to be evaluated) will be generated to correlate this information and, therefore, to estimate water stress from remote sensing data. Thus, for each vine of the plot it can be estimated the instantaneous or accumulated water stress level, with thermal and multispectral products, respectively.

Step 3. Maps generation

After calculating the vegetation parameters for each vine (one value per vine) and applying the model for water stress estimation, a map will be generated that shows the spatial distribution of the plant-based parameter in the whole plot. To do so, a kriging processing of the data is performed. This results in maps for each of the selected parameters that are useful for visual crop growth and monitoring process.

Step 4. Alarm generation

An alarm system is also implemented based on the level of stress, primarily when some areas suffer a higher stress than the average in the plot.

AI DEEP LEARNING PROCESS / ALGHORITHMS

It does not apply.

RESULTS / OUTCOMES

Some of the results obtained after applying the bundle on a vineyard plot are shown in the following figures:

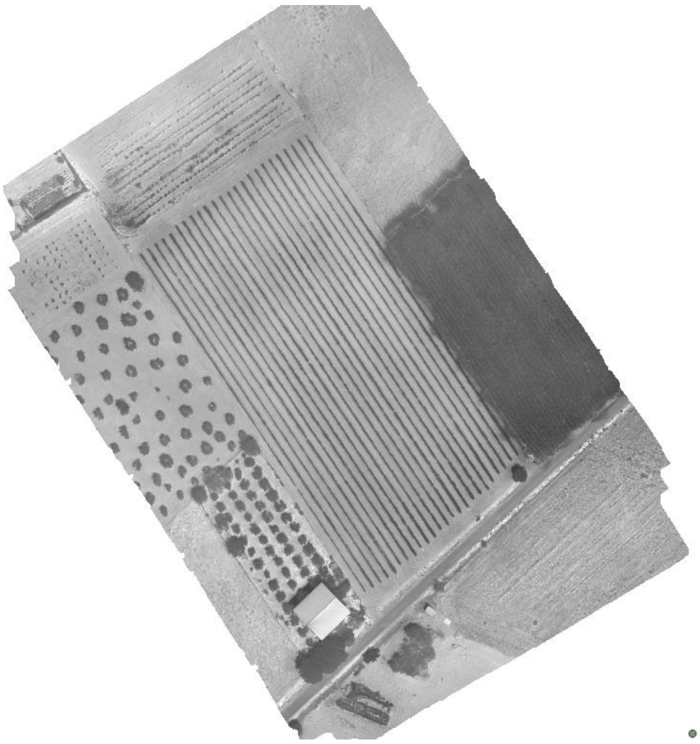


Figure 3.2.2: Orthoimage of temperatures of a vineyard on 18th August 2023

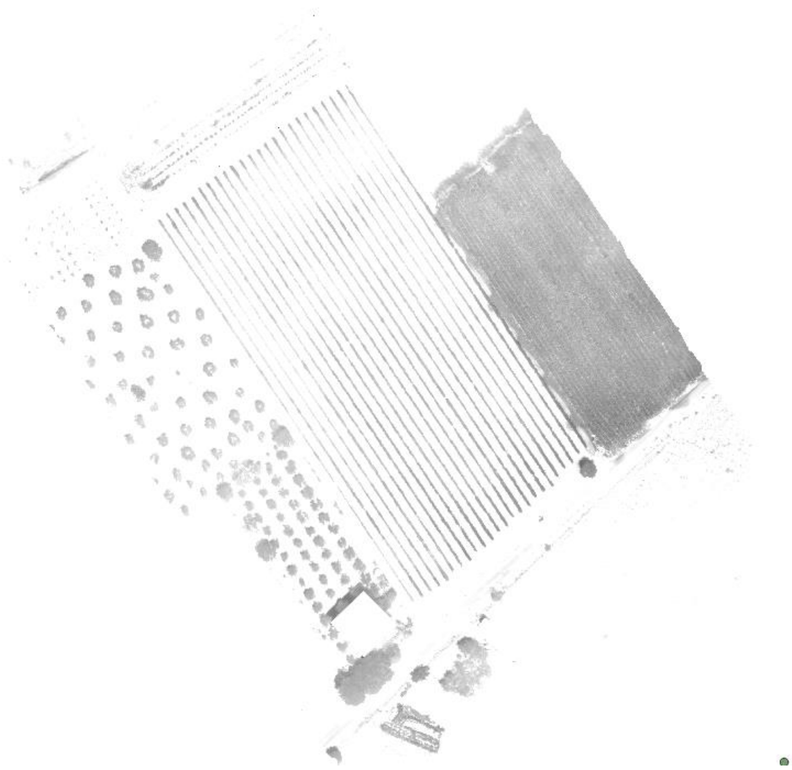


Figure 3.2.3: Orthoimage of temperatures of a vineyard on 18th August 2023, segmenting the vegetation (dark values)

3.3. SOIL ZONIFICATION

RELEVANCE AND PROBLEMATIC

To establish precision agriculture as a management strategy in a farm, it is crucial to perform an accurate plot zoning to apply different inputs depending on the production potential of each of these zones.

There are different strategies to establish the plot zoning, based on the type of measured data:

- Traditional soil sampling, which represent an average soil composition when soil samplings are distributed along the plot, or localized soil characteristics when soil sampling is obtained in a certain point. Spatial distribution of soil characteristics is missing. Different soil samples can be taken in different zones of the plot, defining these zones based on the experience of the farmer.
- Soil reflectivity, which depends on topsoil chemical composition and texture. Remote sensing information has been widely used with this purpose.
- Radar, to determine soil electrical conductivity. Technologies: remote sensing (satellite-based, Sentinel 1), ground penetrating radar, as the most accurate source of information nowadays to perform plot zoning. Soil electrical conductivity is mostly affected by soil moisture. Also, soil moisture is conditioned by soil type. Thus, radar can be used as an indirect measure of soil type.

- Crop as a sensor approach, which uses crop monitoring from previous seasons to determine yield potential of different zones. Technologies: remote sensing (satellite-based, drone based, based both on reflectivity), yield monitoring, with advanced harvesting machinery. In this case, not only soil parameters affect crops development, but also other biotic and abiotic factors. This approach is lately being the preferred by PA practitioners, but special care should be taken with the farmers practices that can affect crop development.

With this information, it is necessary to perform zones grouping, which determines different zones with similar value/values of the above-mentioned parameters. To do so, different clustering algorithms and geostatistical procedures have been developed. Machine learning classification algorithms performs especially well when different variables are considered and when a high amount of data is captured, such as in the case of remote sensing information.

Once performed the different groups or zones in the plot, soil analysis should be performed to analyse the different soil types for each zone and try to relate it with the crop development factors.

The final product of this bundle will be a map that delimits different zones of the plot with differentiated characteristics that lead to different crop growth and development. It will be generated using the crop as a sensor process, measuring GCC and V and generating crop development maps for those parameters at a specific date.

A common process is shared with the bundle "crop growth and development monitoring" bundle. This will result in a higher efficiency of bundles and information usage as well as familiarity of final user with the different products.

UAV CHARACTERISTICS

Described in Table 3.1.1. and Figure 3.1.1.

CAMERA CHARACTERISTICS

RGB camera such as SONY α-6000, for obtaining green canopy cover and volume.

FLIGHT CHARACTERISTICS

Same as in "crop growth and development monitoring"

DATA PROCESSING

The proposed methodology is based on the geometric characterization of each individual vine in the plot. Geomatic products were obtained using version 2.0.1 of the Agisoft Metashape Professional software (Agisoft LLC, St. Petersburg, Russia). Fig. 2.3.1. shows a flowchart that summarizes the proposed methodology.

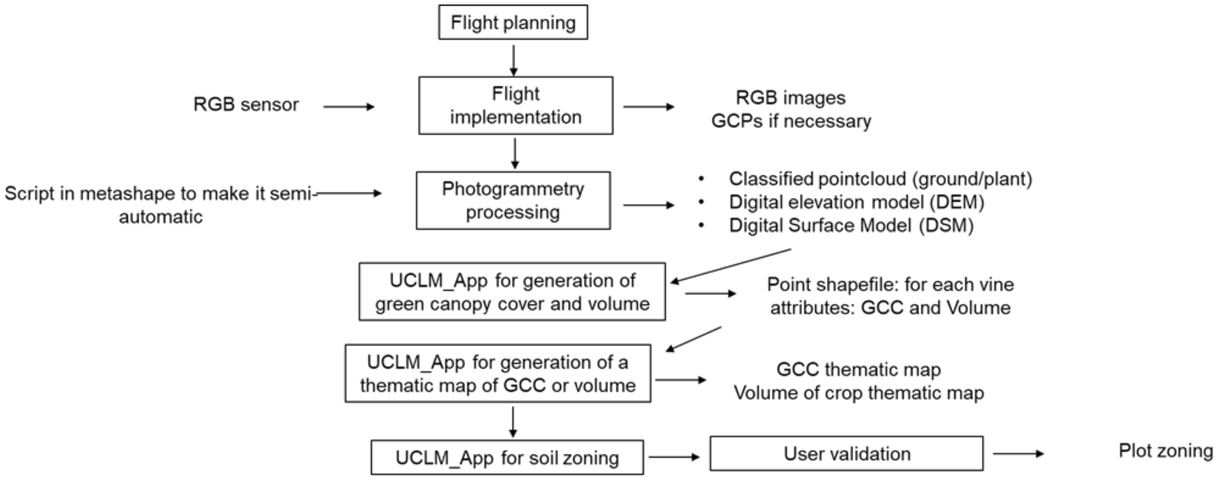


Figure 3.3.1: Flowchart of the proposed methodology

Step 0. Photogrammetric processing

Same as in “crop growth and development monitoring”, only for RGB products.

Step 1. Determining crop parameters based on RGB products.

To determine GCC and V, the difference between DSM and DTM is obtained. The result is the elevation of the vegetation in each pixel. If the elevation of the vegetation in each pixel is multiplied by the area of each pixel, the volume occupied by the vegetation is obtained. It is important to highlight that the volume obtained with this methodology is an orthogonal projection of the vegetation to the soil. This overestimates the volume occupied by the plant because it does not consider that there is a part of the plant without vegetation in the lower part of the plant. If necessary, this can be corrected by measuring in some plants the distance between the soil and the plant. We do not consider it to be necessary because this overestimation is considered for every plant in the plot and could be considered as an offset in any model that relates with parameter with any other variable of interest. Also, it will not affect plot zoning applications.

To calculate the green canopy cover (GCC), it is obtained the existence of vegetation when the difference between DSM and DTM exceeds a limit. This limit has been calibrated and included in the bundle development.

With the GCC and the volume occupied by each plant, referred to each frame, a mapping of the whole plot for any of these two parameters is generated and utilized for different precision viticulture applications.

With the GCC and the volume occupied by each plant, referred to each frame, a mapping of the whole plot for any of these two parameters is generated and utilized for different precision viticulture applications.

Step 2. Maps generation

After calculating the GCC and V for each vine (one value per vine) a map will be generated that shows the spatial distribution of the plant-based parameter in the whole plot. To do so, a kriging processing of the data is performed. This results in maps for each of the selected parameters that are useful for visual crop growth and monitoring process.

Step 3. Zoning generation

Based on clustering analysis, the user will delimit a specific number of zones in which the plot should be divided. The result will be a shape file of polygons with the limits of each zone.

AI DEEP LEARNING PROCESS / ALGHORITHMS

It does not apply.

RESULTS / OUTCOMES

Some of the results obtained after applying the bundle on a vineyard plot are shown in the following figures:

Figure 3.3.2: Maps of zones of the plot based on the flight of 30th June 2023. The darker, the higher the development of the crop and, thus, the crop requirements

3.4. MONITORING FLORA AT HIGH-ALTITUDE GRAZING AREAS FOR SEASONAL ANIMAL FEEDING

RELEVANCE AND PROBLEMATIC

Grassland, an important vegetation type in terrestrial ecosystems, is the most widely distributed form of land cover with abundant renewable natural resources. Grasslands are the main food sources of livestock products such as beef, lamb, and dairy. In general, accurate measurements of grassland biophysical and biochemical parameters are the basis of grassland monitoring. Traditional measurement methods rely mainly on ground measurements (field surveys), which usually sample the measured area and select numerous plots to present the entire area. These plots can be directly used to generate accurately measured parameters or provide all kinds of precise data related to them. However, these methods are time-consuming and laborious, and they are only precise in small areas.

The application of remote sensing technology in grassland monitoring and management has been ongoing for decades. Compared with traditional ground measurements, remote sensing technology has the overall advantage of convenience, efficiency, and cost effectiveness, especially over large areas.

Leaf area index (LAI) is a dimensionless quantity that characterizes plant canopies. It is defined as the one-sided green leaf area per unit ground surface area (LAI = leaf area / ground area, m^2 / m^2). It is one of the key indices to reflect the growth status of grassland vegetation, as well as one of the most fundamental characteristic parameters in many ecosystem modelling processes.

Fractional Vegetation Cover (FVC), which is the same concept than green canopy cover (GCC) in agricultural applications, is defined as the percentage of the vertical projection of green vegetation over the entire calculated area, which is the basic parameter for describing the characteristics of the grassland ecosystem and for obtaining the condition of grassland vegetation with its changes. Its accurate estimation is of great practical significance for regional grassland environment evaluation, management, and degradation monitoring. Many studies have demonstrated the direct relationship of FVC and LAI as well as the easiness of calculating FVC over LAI with the support of aerial imagery.

Above-ground biomass (AGB). AGB is one of the main parameters of grassland biomass. AGB is defined as the aboveground standing dry mass of live or dead matter from tree or shrub/woody life forms, expressed as a mass per unit area. AGB is one of the significant indices of grassland growth, degradation, easily applied to monitor overgrazing. AGB can be obtained with the support of 3D point clouds derived from photogrammetry processes.

Thus, all these vegetation parameters can be estimated using geomatic products (orthoimages, digital terrain model, digital surface models and/or 3D point clouds) obtained from aerial images. Grassland areas that supply food to several herbs can cover large areas, primarily in the mediterranean areas were the vegetation is not so abundant. Thus, the use of conventional vertical take-off and landing (VTOL) UAVs are limited due to the autonomy of flight. Here, it is crucial the use of mixed technologies based on satellites and drones to, first, delimit the region of interest (RoI) with the use of low-resolution satellite imagery, and second, to accurately estimate vegetation parameters with the use of drones for those delimited RoIs.

The proposed methodology is based on the geometric (FVC, AGB) and radiometric (VIs) characterization of the vegetation. The geometric characterization is mainly related with the quantity of vegetation, while the radiometric characterization is related with vegetation quality. Vegetation quality is of high complexity and requires field testing and validation, so in this bundle the objective of estimating vegetation quality will be limited to generation of the main VIs related with vegetation quality, requiring the validation by final user.

UAV CHARACTERISTICS

No data acquired yet with UAVs for this bundle.

CAMERA CHARACTERISTICS

To obtain the required products the following sensors, included in Chameleon project, can be used:

RGB camera such as SONY α-6000, for obtaining green canopy cover and volume.

Multispectral camera, such as Micasense RedEdge MX, for obtaining vegetation indices.

FLIGHT CHARACTERISTICS

The region of interest to be covered by the UAV flight will be delimited using satellite images information. Those areas of special interest based on NDVI values will be delimited and monitored afterwards with high-resolution products obtained with UAVs.

A typical photogrammetric flight planning should be implemented. After selection of the region of interest (RoI), the main parameters to consider for each type of product (RGB or multispectral) are the followings:

- Ground sample distance (GSD)
  - RGB products: max 2 cm
  - Multispectral products: max. 5 cm.
- Overlapping and sidelapping:
  - RGB products: 80% (forward overlap) 50% (side overlap)
  - Multispectral products: 80% (forward overlap) 50% (side overlap)

In any case, the maximum flight height of 120 m will not be exceeded as well as any other restriction commanded by law.

DATA PROCESSING

Geomatic products were obtained using version 2.0.1 of the Agisoft Metashape Professional software (Agisoft LLC, St. Petersburg, Russia). This software allows the generation of dense point clouds, DTMs, and orthomosaics, from the aerial images acquired by UAVs. Also, it allows a first treatment of these data, such as point cloud classification. Fig. 3.4.1. shows a flowchart that summarizes the proposed methodology.

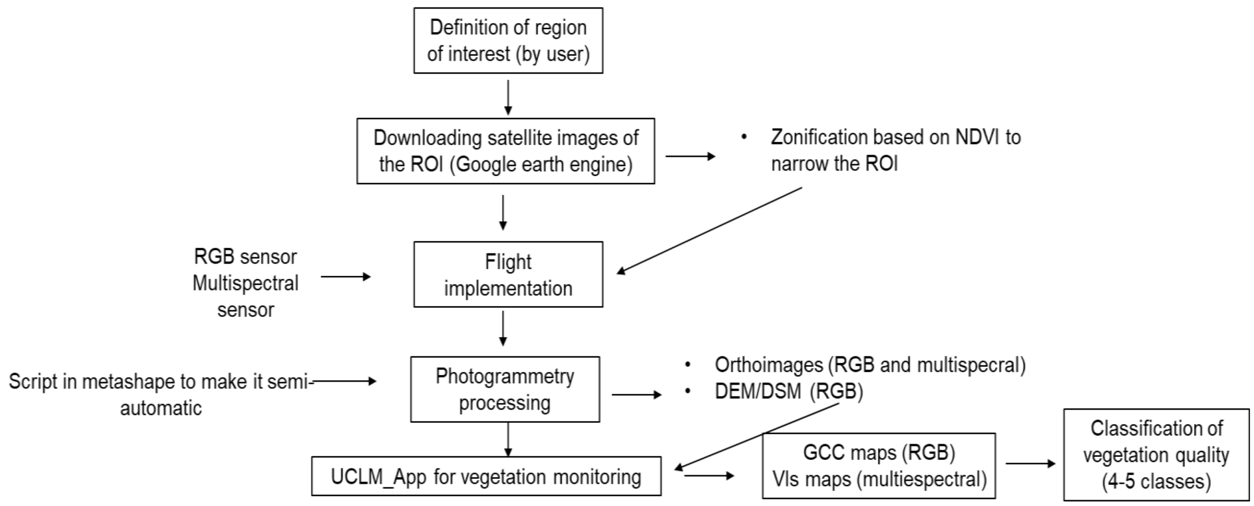


Figure 3.4.1: Flowchart of the proposed methodology

Step 0. Photogrammetric processing

Photogrammetry processing was performed with Agisoft Metashape Professional software (Agisoft LLC, St. Petersburg, Russia) version 2.0.1. A software in Python was implemented to automatize the whole process without the need of human intervention. A two options software was implemented: 1) case 1, in which RTK or PPK systems are available in the drone and, therefore, the coordinates of the centre of the images can accurately obtained; and 2) case 2, in which GCPs should be placed. In this last case, the coordinates of the images should be generated with a manual step in Metashape by locating the GCPs in the images and performing the firsts steps manually. After aligning images and generating the accurate coordinates of the images, these can be saved in a .csv file and continue the whole process automatically.

Summarizing, the following geomatic products are generated for each flight: 1) classified point cloud (ground and vegetation), 2) orthoimage (RGB and/or multispectral), 3) digital terrain model, considering ground points, and 4) digital surface model, considering ground and vegetation points.

A grid to evaluate the different indicators is generated. The size of the grid depends on the type of vegetation to monitor. Also, a first classification between high, medium and low vegetation will be performed to analyse only the type of vegetation required.

Step 1.1. Determining vegetation parameters based on RGB products.

To determine FCV and V, the difference between DSM and DTM is obtained. The result is the elevation of the vegetation in each pixel. If the elevation of the vegetation in each pixel is multiplied by the area of each pixel, the volume occupied by the vegetation is obtained. It is important to highlight that the volume obtained with this methodology is an orthogonal projection of the vegetation to the soil. This overestimates the volume occupied by the plant because it does not consider that there is a part of the plant without vegetation in the lower part of the plant. This is important for trees, but les crucial for bush or low vegetation.

With the GCC and the volume occupied by each plant, referred to the generated grid, a mapping of the whole plot for any of these two parameters is generated. With this map, a clustering technique based on k-means is applied to generate 4-5 classes of vegetation that describes 4-5 classes of amount of vegetation.

At this moment, no information is evaluated yet. The type of soil and vegetation of these abrupt areas can incorporate difficulties on this procedure. In case the bundle testing results, when the information is generated, in inaccuracies, the procedure that uses VIs will be implemented, as below described.

Step 1.2. Determining vegetation indices (VIs) based on multispectral products.

In this case, from the multispectral orthoimages different VIs are calculated. There will be two options:

- Calculate VIs for the generated grid, including soil and plant. It integrates the vegetation response (reflectivity) and the FCV. Thus, it is highly affected by the vigour of the plants. It can establish a mixture between quantity and quality of vegetation, but it is difficult to discriminate between both.
- Calculate VIs for only the vegetation in the grid, that requires a vegetation segmentation. In this case, this segmentation is performed using radiometric

characteristics of the vegetation, establishing a threshold. This threshold will be calibrated whenever the information is available for this type of ecosystem.

Step 2. Maps generation

After calculating the vegetation parameters for the grid, a map will be generated that shows the spatial distribution of the parameters (FCV, AGB, and/or VIs). To do so, a kriging processing of the data is performed. This results in maps for each of the selected parameters that are useful for visual crop growth and monitoring process.

Step 3. Clustering type of vegetation

Generation of FCV, AGB or VIs by themself are useful products to perform vegetation monitoring. However, a clustering process will be implemented to classify: 1) the amount of vegetation, using FCV, AGB or VIs, and 2) quality of vegetation, with VIs of segmented vegetation. Special care should be taken with quality of vegetation since it requires user validation and calibration processes.

AI DEEP LEARNING PROCESS / ALGHORITHMS

It does not apply.

RESULTS / OUTCOMES

The proposed methodology is being implemented in Crete, Greece. Figure 3.4.2. shows the pilot case for the winter time. First, the RoI should be delimited with satellite-based information (Figure 3.4.3).

Figure 3.4.2: Pilot case for the winter time in Crete, Greece.

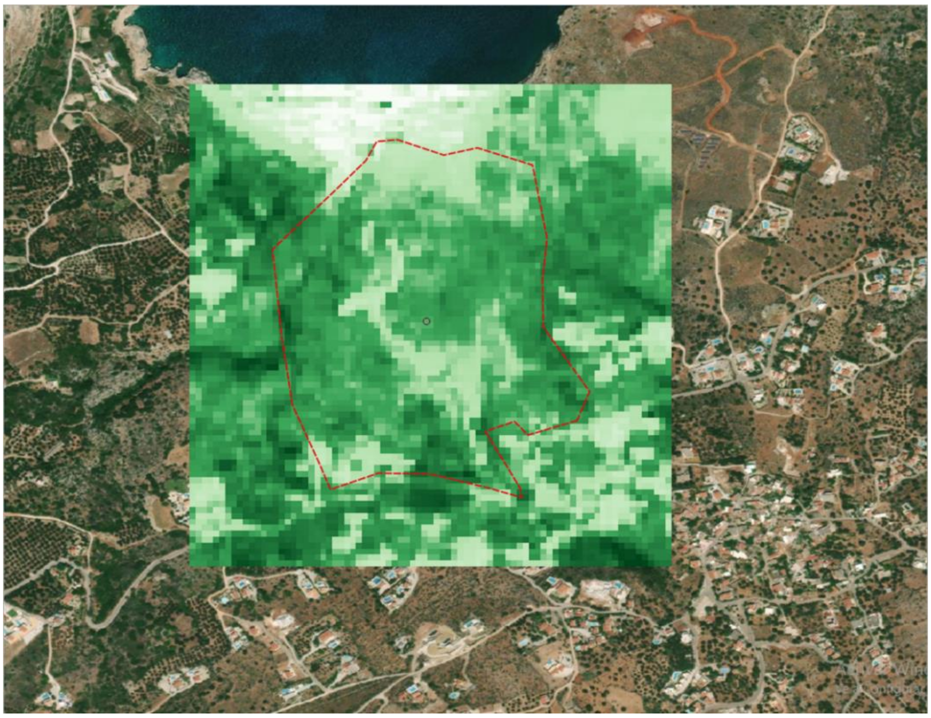


Figure 3.4.3: Map of NDVI of the area to delimit the RoI to be analysed with UAVs

4. BUNDLES, SERVICES FOR LIVESTOCK MONITORING

The Bundles (Business use case) for livestock monitoring of each pilot use case are actions to be conducted on each Pilot with a specific purpose and indicator. The pilot use cases, and the Bundles for livestock monitoring will be developed in CHAMELEON project, according to the conceptualisation and use cases definition of Deliverable 2.1. The CHAMELEON solution for livestock monitoring will be validated and demonstrated under two relevant pilot use cases: i) Spain (Avila) and ii) Greece (Crete) (Table 4).

Table 4: List of bundles for livestock monitoring developed in the CHAMELEON project.

Pilot use case	Bundles (Business use case)
Spain	COLLECTING INFORMATION ABOUT HEALTH STATUS AND STRESS (WILD ANIMALS)
	LAMENESS DETECTION IN COWS
Greece	LIVESTOCK MANAGEMENT (HERD) AND MONITORING (INVIDUAL ANIMAL)
	ANIMALS' HEALTH

Different bundles consist different action steps of the workflow, which was provided in the Deliverable 6.1. Detailed description of development of each bundle for livestock monitoring is provided according to their action level: i) Relevance and problematic; ii) UAV characteristics; iii) Camera characteristics; iv) Flight characteristics; v) Data processing; vi) AI deep learning process/algorithms; vii) Results/Outcomes.

4.1. COLLECTING PARAMETERS RELATED TO THE HEALTH AND STRESS OF LIVESTOCK

This bundle aims to develop new methodologies to enhance the management and performance of livestock farms by monitoring the health and stress levels of the animals. To achieve this, IoT collars placed around the livestock's neck, along with drones, will be used to remotely transmit information, informing the farm managers about the health status of the animals and stress episodes caused by external factors, such as interactions with wild animals.

RELEVANCE AND PROBLEMATIC

Researching and developing new tools to support and ease extensive livestock farming is crucial for making it profitable for the farmers involved.

Despite the clear benefits of extensive livestock farming, coexisting with wildlife can create problematic situations affecting both animals and farmers. Interactions with predators and other stressors can negatively impact livestock health, as well as farm productivity and profitability. The lack of effective tools for monitoring and managing these situations is a gap in livestock management. Therefore, there is a need to develop a control system using IoT collars with various sensors and integrated AI algorithms, allowing proactive intervention in stressful situations and enhancing the resilience of herds against unforeseen events. In addition to IoT devices, drones will be used as a quick and efficient tool for on-site event verification.

Rural depopulation, primarily due to migration to urban areas, reduces available labor for livestock activities, threatening the economic viability of farms and the sustainability of these communities. Additionally, population loss leads to the disappearance of local services, educational institutions, and a dilution of cultural identity. This phenomenon triggers a vicious cycle endangering the survival of rural areas.

Therefore, the goal is to develop an innovative methodology that not only improves efficiency in extensive livestock management but also makes farming more attractive and sustainable in the long term. Integrating IoT collars with drones enhances livestock health and welfare and positions rural communities as leaders in adopting modern, sustainable farming practices. This could boost population retention and attraction, positively impacting the social and economic fabric of these regions.

UAV CHARACTERISTICS

For the performance of this bundle, the use of UAV is optional, with its main utility being to confirm alerts in hard-to-reach areas. Any UAV capable of carrying an RGB camera as a payload is useful.

CAMERA CHARACTERISTICS

To obtain the required products, the following sensors included in the CHAMELEON project are used:

RGB Camera: Sony A6000

FLIGHT CHARACTERISTICS

The plan should include:

GSD: Centimetric.

DATA PROCESSING

Data collection and processing through IoT collars equipped with accelerometers and GPS are a key phase in achieving the goals of this bundle. The data from these devices provide detailed analysis of the behaviour and location of extensive livestock. To optimize data transmission and collar autonomy, the Low Power Wide Area Network (LPWAN) is employed, ensuring efficient connectivity even in remote rural environments. Additionally, AI algorithms are embedded in the device (on the edge), which contributes to reducing energy consumption and enhances the data collection performed by the IoT collars. These algorithms will send alerts for various unusual events by analysing movement patterns (walking, rumination, death, etc.), enabling early detection of potential stress situations and alerting farm managers.

Figure 4.1.1 displays the general flow diagram of the bundle. Figure 3.1.2 presents the input, parameters, output, and display results table of the bundle.

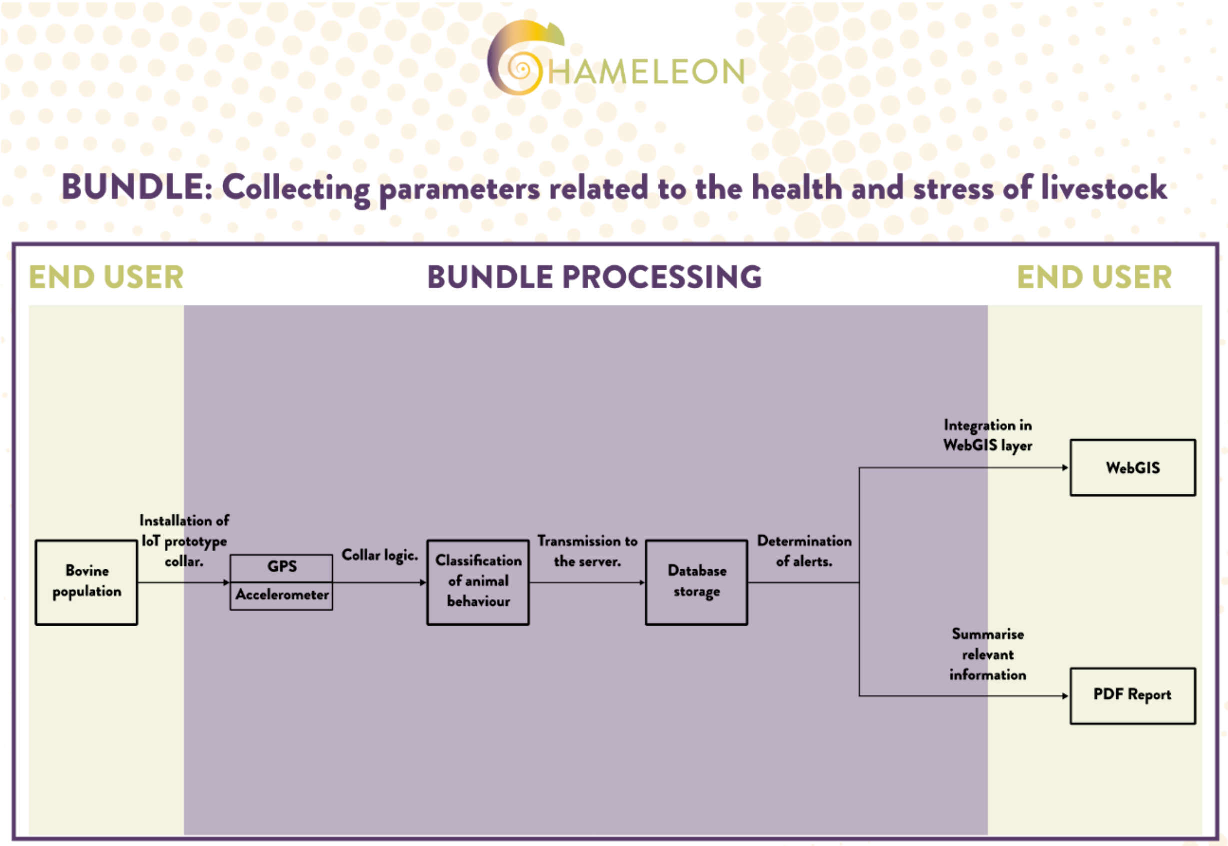
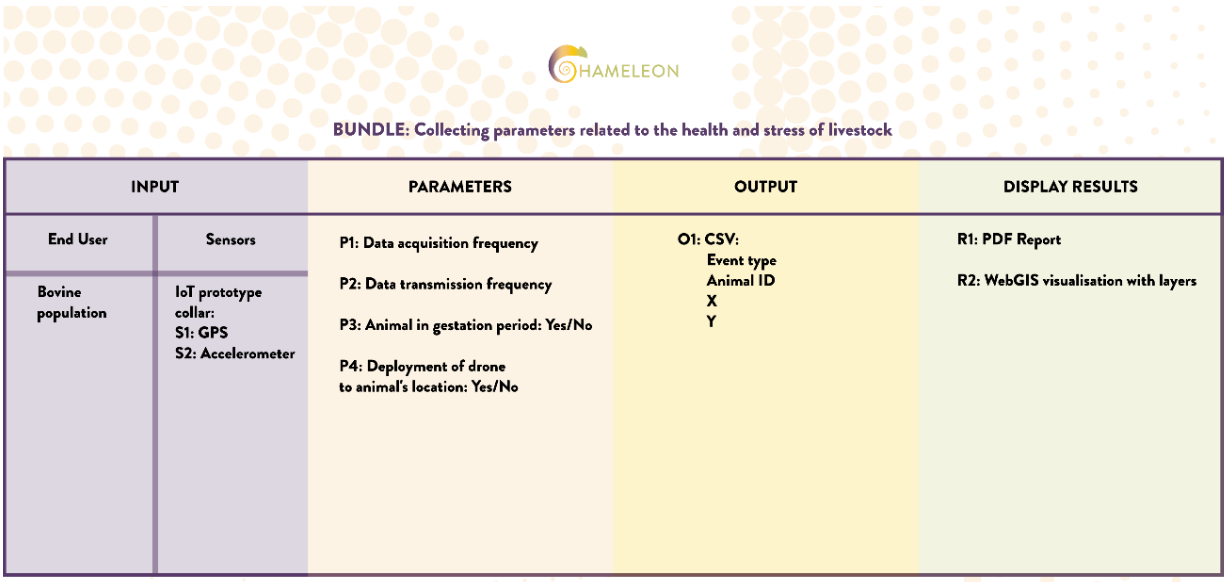


Figure 4.1.1: Workflow of the bundle “Collecting parameters related to the health and stress of livestock”.



INPUT		PARAMETERS	OUTPUT	DISPLAY RESULTS
End User	Sensors			
Bovine population	IoT prototype collar: S1: GPS S2: Accelerometer	P1: Data acquisition frequency P2: Data transmission frequency P3: Animal in gestation period: Yes/No P4: Deployment of drone to animal's location: Yes/No	O1: CSV: Event type Animal ID X Y	R1: PDF Report R2: WebGIS visualisation with layers

Figure 4.1.2: Input, parameters, output and display results of the bundle “Collecting parameters related to the health and stress of livestock”.

AI DEEP LEARNING PROCESS / ALGHORITHMS

Advanced AI algorithms will be applied to analyse the data in real time. These algorithms will focus on detecting movement patterns, classifying livestock activities into specific categories such as grazing, resting, feeding, and potentially stressful behaviours. This analysis will allow a detailed understanding of the herd dynamics, enabling early identification of significant deviations.

To develop the embedded AI algorithms in the collar, the first step is real-time data capture using an accelerometer with Bluetooth connectivity. The raw accelerometer data is then used to train a neural network specialized in classifying various animal states (walking, eating, no movement, etc.).

Figure 4.1.3: Top-left: Livestock collar. Top-right: Cattle grazing with collars. Bottom: Real-time monitoring of accelerometer data.

The Edge Impulse platform will be used for training and developing the AI algorithms, which allows the creation of a C++ library for local and autonomous implementation of these algorithms in the IoT collar (on the edge).

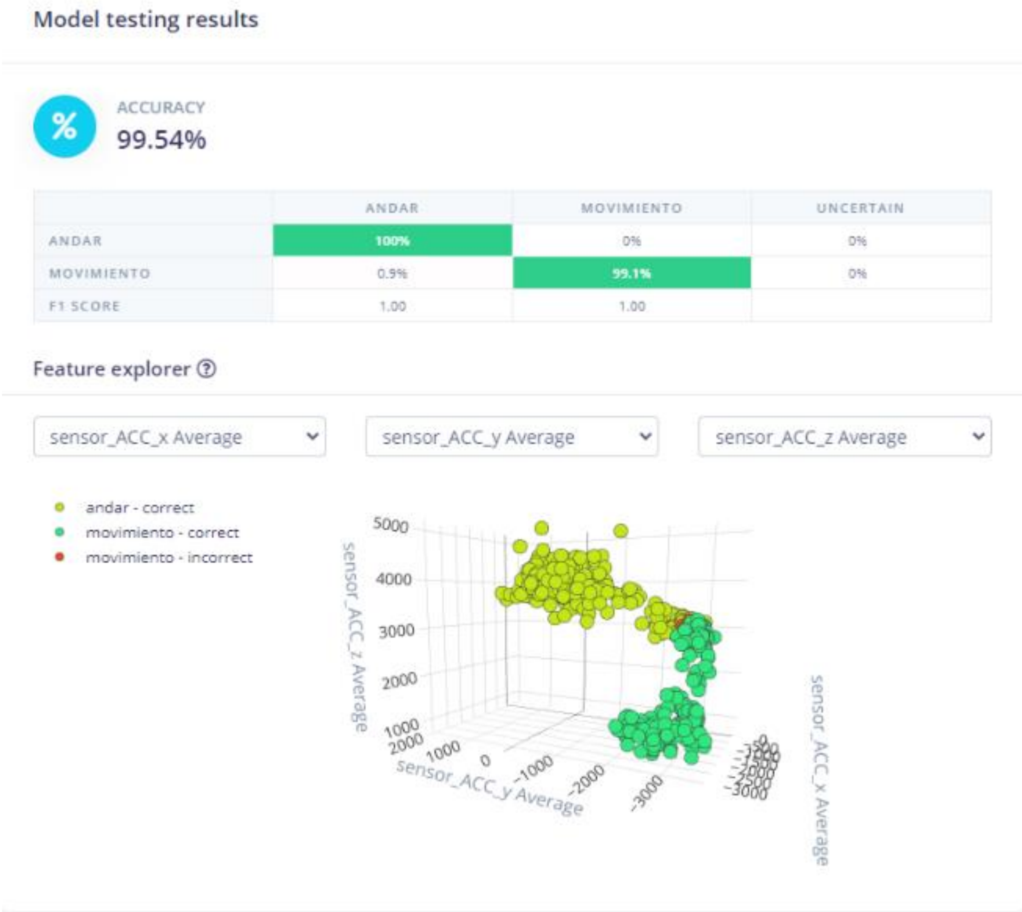


Figure 4.1.4: Classification of cattle movements based on accelerometer data, on the Edge Impulse platform.

This training phase will enable the algorithms to recognize specific patterns associated with different types of livestock movement and behaviour. The detection of abnormal patterns, such as erratic movements or signs of stress, will trigger immediate alerts, sent via the LPWAN network, for timely intervention.

RESULTS / OUTCOMES

The expected outcome of this bundle development is a platform that allows the end user to know the health and stress status of their cattle. To achieve this, the usual behaviour of these healthy animals will be monitored with a CHAMELEON IoT collar prototype. Unusual behaviour patterns in an animal may indicate a pathology that requires evaluation by a veterinary specialist.

From the moment the collars are installed, data is gathered using the CHAMELEON IoT collar prototype. This collar integrates sensors: a) GNSS; and b) accelerometer. The data is filtered and classified into a type of behaviour in real time. In the current phase of the project, the aim is to detect three states: a) Complete absence of movement (inactive cow); b) Head movement (stationary cow); c) Body movement (moving cow). In the future, classifications for additional movements could be integrated, such as: d) Walking; e) Ruminating; f) Jumping; g) Running; h) Eating.

This classification is recorded and processed periodically on the collar, in different time windows. It is also sent periodically via LPWAN to the CHAMELEON server, with a more continuous information flow when the logic of the classification requires it.

Alerts are triggered when certain continuous events occur. For example: a) illness/animal trapped: the animal moves much more than the rest; b) no movement (inactive cow): the animal has moved less than 10 m in 12h. The time window for each alert can also be modified, for instance, in the case of cows close to giving birth, their classification is evaluated more frequently. Technically, this is implemented with different “flags” that control the data sending frequency.

The alerts consist of a CSV document containing the information: a) Type of alert; b) Animal identifier; c) X and Y coordinates of the animal. These alerts are displayed on the WebGIS, and a notification is sent to the end user. There is also an option to send a drone to the animal's location to confirm with an RGB image/video if it is a real alert or a false positive.

The implementation of this advanced data processing technology not only enhances monitoring capabilities but also establishes a proactive system for managing livestock health and welfare. The combination of GPS precision with AI algorithm intelligence marks a significant step towards efficient and sustainable management of extensive livestock farming, contributing to animal health and the overall success of the project.

4.2. LAMENESS DETECTION IN COWS

RELEVANCE AND PROBLEMATIC

The task led by AiDEAS focuses on developing a model for detecting lameness in cows using artificial intelligence (AI) techniques. The initiative aims to significantly enhance animal welfare and increase the efficiency and sustainability of farming operations. By employing advanced AI methodologies, AiDEAS is committed to transforming the approach to early, non-intrusive detection and management of cow lameness, thereby setting new standards in animal healthcare.

This project's scope revolves around utilizing UAV-derived imaging and AI analytics for accurate lameness detection. State-of-the-art pose estimation models such as MMPose for real-time identification and analysis of key animal body parts are crucial in evaluating cow health. Additionally, the project will employ deep learning techniques to perform feature extraction from key points, establishing routines that help in classifying the animal as lame or not based on their movement patterns and gait characteristics. This method ensures a more targeted and accurate assessment of animal health, aligning with the goal of early and precise lameness detection. This approach aims to provide a comprehensive analysis of animal gait through advanced computer vision models, ensuring precise and efficient lameness detection. The integration of these AI-driven methodologies with behavioural analysis will further enhance decision-making in farming practices related to feeding and overall animal management.

PROSPOSED METHODOLOGY – ARCHITECTURAL ASPECTS

This section provides a detailed description of the proposed methodology. Specifically, the identification of end users' needs and conclusions are presented in detail. Then, the investigation of existing commercial cameras, the area of interest for camera placement and the camera points of view for data collection are outlined. Finally, the necessary video-processing procedures are discussed in detail. In particular, video recording, application of pose estimation models, routines for isolating animals in each frame, selection of appropriate frames, motion analysis, quality adjustment of the selected frames, feature extraction of the identified gait characteristics and the next steps for detecting lameness in cows are illustrated.

ASSESSMENT QUESTIONNAIRE

To enhance our understanding of the unique needs and challenges encountered by cow farmers in managing lameness, AiDEAS has meticulously developed a comprehensive needs assessment questionnaire. Following the creation of this questionnaire, our partner USAL took the initiative to translate it into Spanish, ensuring its accessibility and relevance for a wider range of end users. This collaborative effort emphasizes our commitment to inclusivity and effective communication across diverse farming communities.

The questionnaire, designed with precision and executed in collaboration with our partners, is an essential tool for capturing detailed insights directly from those at the forefront of cow farming. It is structured into two main sections for clarity and focus. The first section aims to delve deep into understanding the primary issue of lameness detection, seeking information on farmers' experiences, observations, and current practices. The second section shifts the focus towards deep learning applications, gathering specific insights to align our AI solutions with the practical needs of the farming industry.

The inclusive approach of translating the questionnaire into Spanish by our partner USAL and sharing it with the end user has enriched the feedback process. Below, we present the questionnaire along with the responses and comments from the end user. This valuable data, collected from a diverse set of perspectives, is instrumental in guiding our project towards achieving both technological innovation and practical applicability in enhancing animal welfare.

Questionnaire for Cow End-User

Section 1: Understanding the Problem
#Question:
How many cows do you have in your herd?
#End-user reply:
150 Heifers, but this is an average, it varies depending on the season.
#Question:
How prevalent is lameness in your cow herd (give a probability)?
#End-user reply:
At this moment we have 0. We usually don't have cows affected by lameness.
#Question:
What are the main causes of lameness in your cows?
#End-user reply:
Bumps between them, hoof growth

#Question:
What methods or tools are you currently using to detect and monitor lameness?
#End-user reply:
Visual
#Question:
What are the limitations or challenges of your current lameness detection methods?
#End-user reply:
If it's the weekend, it might be problematic to treat the animal
#Question:
How do you assess the impact of lameness on your cow herd's productivity and well-being?
#End-user reply:
Lack of access to food, weight loss, animals not being retained for replacement
#Question:
Are you using any other sensors, such as accelerometers, to monitor cow movements?
#End-user reply:
GPS, virtual fencing

Section 2: Deep Learning Considerations

#Question:
Do you have any labelled data or a database of cow movements, including instances of lameness? If yes, please provide details on the size and format of this data.
#End-user reply:
No
#Question:
Are you open to collecting new data using stable cameras or UAVs to train the deep learning model? If yes, please describe your ideal data collection setup.
#End-user reply:
Yes, through digitalization project
#Question:
Are there any specific patterns or behaviours that you have observed in lame cows, which could help train the deep learning model?
#End-user reply:
No
#Question:
How accurate do you expect the lameness detection model to be in order to be useful in your daily operations?
#End-user reply:
-
#Question:
How quickly should the model be able to detect lameness after it has occurred?
#End-user reply:
Daily notice
#Question:
What kind of output or alert would you prefer to receive from the model when lameness is detected?

#End-user reply:
Mobile Alerts

Summary of end-user perspective and our conclusions

Section 1: Understanding the Problem

In our herd, there are approximately 150 heifers on average, though this number fluctuates seasonally. Currently, lameness is not prevalent among our cows, with no reported cases at this moment. The main causes of lameness that we have observed include instances where cows bump into each other and issues related to hoof growth. We rely on visual methods for lameness detection, and one challenge we face is the difficulty in treating animals, especially during weekends. When assessing the impact of lameness on our herd's productivity and well-being, we look out for signs such as lack of access to food, weight loss, and an inability to retain animals for replacement purposes. While we do not use specific sensors like accelerometers, we utilize GPS and virtual fencing technologies.

Section 2: Deep Learning Considerations

Currently, we do not possess any labelled data or a database of cow movements, including instances of lameness. However, we are open to collecting new data using stable cameras or UAVs to facilitate the training of the deep learning model. We are particularly interested in integrating this data collection process into our ongoing digitalization project. Regarding specific patterns or behaviours related to lame cows, we haven't identified any distinctive traits yet. As for the accuracy of the lameness detection model, we do not have a specific percentage in mind; however, timely detection is crucial. Hence, one potential approach to address this existing challenge might involve identifying a specific percentage of deviation from the normal cow gait. Ideally, the model should be able to notify us of lameness on a daily basis. We prefer receiving mobile alerts as outputs or alerts from the model when it detects lameness, allowing for swift and efficient responses in our daily operations.

DATA COLLECTION

Investigation of commercial cameras

The necessity of monitoring project cows' movements to detect lameness prompted us to explore options for suitable commercial cameras. After thorough research, USAL with the support of AIDEAS determined that the ideal choice for our project is the Reolink Duo 2 LTE camera. Its technical specifications are as follows:

- Panoramic 4G Battery/Solar Powered Camera
- Dual Lens
- 180° Panorama*
- 2K+ 6MP Super HD
- 4G LTE Network

The images captured by the two lenses are seamlessly merged into a 180° panorama*, providing an ultra-wide-angle view without any gaps and minimal distortion. The Reolink Duo 2 LTE effectively integrates two cameras into one, eliminating blind spots.

Figure 4.2.1: Wi-Fi Camera.

Positioning of the camera

Placing the camera inside a cow pen poses a significant challenge, especially due to the cows' sensitivity and the disruption of their daily routine caused by the presence of the camera. Moreover, the camera must be positioned at a low height and have the capability to record the side view of the cow vertically as it moves. Meeting these requirements adds to the complexity of the task.

The area of interest and the initial placement of the camera is crucial considerations. Below, the area of interest and two points of view of the camera recordings are presented, which comprise our dataset.

Figure 4.2.2: Area of interest and initial camera placement.

In Figure 4.2.3. we can observe the 1st point of camera view. Specifically, from this point, 1173 videos have been recorded with an average duration of 1 minute.

Figure 4.2.3: 1st point of camera view.

Subsequently, in Figure 4.2.4., we observe the second camera positioning point. From this point, we have 925 more recordings with a similar average duration.

Figure 4.2.4: 2nd point of camera view.

DATA PROCESSING

Procedures and Task Breakdown for Lameness Detection in Cows

The procedure to detect lameness in cows using AI involves a systematic process of video analysis and data extraction. This process is broken down into smaller, manageable tasks, ensuring precision and efficiency at each stage. The primary goal is to analyse cow movements through video capture and apply advanced AI techniques to determine if the gait is normal or indicative of lameness.

Step-by-Step Procedure

- **Video Capture**

The initial step involves capturing videos of the cows' using cameras. These videos provide the raw data required for subsequent analysis.

- **Application of Pose Estimation Models**

Once the videos are captured, they are processed through pose estimation models. These models are designed to accurately extract key points of the cows' movements, focusing on specific body parts crucial for gait analysis.
The use of MMPose for the detection and extraction of key points in the context of analysing animal gait patterns is a significant advancement. By leveraging MMPose, a sophisticated human pose estimation framework, researchers can apply it to animals, extracting critical anatomical key points for analysis. In our specific case, 17 key points per animal were extracted, allowing for a detailed characterization of their movements.

These key points likely correspond to specific anatomical landmarks on the animal's body, such as joints, which are essential for understanding the posture and movements of the animals accurately. The extraction of these key points serves as the foundation for in-depth analysis, enabling us to quantify and assess various aspects of the animals' gait patterns.

With 17 key points per animal, we have the ability to track the positions of these anatomical landmarks over time and analyse their spatial relationships during motion. This detailed information can facilitate the calculation of indicators, such as step width, step length, stance time, and asymmetry, with a high degree of accuracy and precision.

Moreover, having access to these key points allows for the application of advanced machine learning algorithms and statistical methods. These techniques can uncover patterns, trends, and anomalies in the animals' gait data, leading to valuable insights into their movement patterns and overall health.

- **Routine Setting and Standards Implementation**

  - Isolation of Individual Animals: The first routine involves segregating the key points of each animal into separate files. Given that a single video may contain multiple cows, it's essential to isolate the data for each cow to ensure precise analysis.

    Specifically, in the routine processing of our multi-cow video datasets, the first step involves preparing and refining our primary data sources which are a video file, and a corresponding JSON file. The JSON file is crucial as it contains structured data of key points representing specific features or positions of cows within each frame. Given the complexities of these JSON files, which often comprise concatenated JSON objects, our initial focus is on reformatting them. This is meticulously done by identifying and separating these objects. Then we restructure them into a standardized JSON format. Simultaneously, the video file is loaded using OpenCV, a robust library for video and image processing, setting the stage for subsequent annotation steps.

    The core of our process lies in the intricate extraction and separation of key points for each cow. As we process each video frame, we align the frame ID with the corresponding data in the JSON file. This alignment allows us to precisely extract key points associated with individual cows, complete with their coordinates and confidence scores. This step is critical for maintaining accurate tracking of these points across the video frames. Then, we proceed to annotate the video frames with these key points. The annotation involves overlaying points on the video frames at the coordinates which are specified in the JSON data, ensuring that only key points with a confidence score above a certain threshold are used. This selective approach guarantees the accuracy and reliability of our annotations.

    Finally, an equally important aspect of our procedure is the organization and storage of the processed data. We meticulously save the key points data in a new, well-structured JSON file. In this file, the key points are organized separately for each cow, facilitating ease of access and analysis. This methodical organization of data not only streamlines our workflow but also provides a foundation for various analytical and research applications. The annotated video, alongside the structured JSON data, becomes a valuable asset for in-depth studies in animal behaviour, veterinary sciences, and automated monitoring systems, reflecting the effectiveness of our processing routine in handling complex multi-cow video data sets.

  - Selection of Side-View Frames: The process includes retaining frames where the cow is viewed from the side. Side views provide the most relevant perspective for analysing gait and identifying any abnormalities.

In particular to further enhance our analysis of cow behaviours and postures in video data, our routine focuses on identifying frames where the animals are positioned in a side view. This specific orientation is critical for certain types of analysis, such as assessing the physical condition or gait of the animals. The side view provides a comprehensive perspective of the cow's body structure and movement patterns, which is often necessary for detailed veterinary assessments or behaviour studies. In the following Figure 4.2.5., we will showcase an example image that aligns with industry standards and existing literature, demonstrating the ideal side-view orientation of a cow that we aim to capture in our video frames. This example serves as a benchmark for the type of imagery we seek to extract from our video data.

Figure 4.2.5: Side profile of the cow.

Building upon our methodology, the next step involves employing a refined approach to identify frames that capture the cow's profile, which is essential in gait analysis. This is achieved by analysing the length-to-height ratio of each cow within the frame. By focusing on this ratio, we can effectively discern frames where the cow's side view is prominently displayed. This technique ensures a consistent and standardized perspective for analysis, which is crucial for accurate assessments. The length/height ratio serves as a reliable indicator, guiding us to select frames that are most relevant for a detailed study of the cow's gait and overall physical condition. This methodical selection of frames contributes to the precision and effectiveness of our gait analysis, providing data points that are both relevant and consistent.

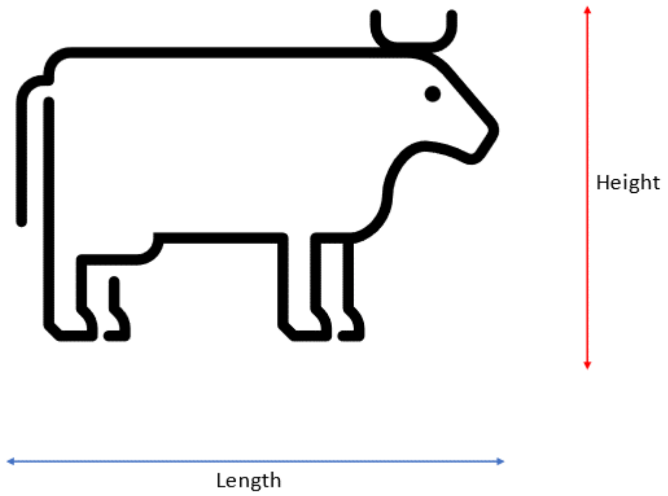


Figure 4.2.6: Sketch of a cow's length and height.

Length/Height Ratio Calculation:

The length/height ratio is calculated by measuring the length and height of the cow in each frame. The length of the cow is typically measured along its spine from the front to the back, while the height is measured vertically from the ground to the cow's back. By dividing the length by the height, you obtain the length/height ratio for each frame.

$$\frac{Length}{Height} = x = \frac{\sqrt{(x4 - x3)^2 + (y4 - y3)^2}}{\sqrt{(x2 - x1)^2 + (y2 - y1)^2}}$$

Identifying Frames with Cow's Profile:

Frames where the length/height ratio falls within a specific range or threshold are identified as frames where the cow's profile is visible. This range is determined based on the characteristics of the cow and the camera setup. Frames falling within this ratio range are selected for further analysis, ensuring that the gait analysis is performed only on frames where the cow's side view is captured.

Vertical Profile Analysis:

Once the frames with the cow's profile are identified, the gait analysis, including the extraction of key points and calculation of gait parameters, is performed specifically on these frames. This focused analysis ensures that the results are consistent and reliable, as they are derived from frames where the cow's gait pattern can be accurately assessed from a vertical profile perspective.

By employing the length/height ratio to select frames for analysis, we can enhance the precision and reliability of our gait analysis. This technique minimizes variability in camera angles and distances, allowing for a standardized evaluation of the cow's gait pattern from a vertical viewpoint.

- Motion Detection: Only frames where the animal is in motion are considered. This ensures that the data used for analysis represents the cow's natural gait.

  The task of quantitatively analysing animal movement has been significantly advanced through the development of a Python-based script. This script is designed to process key point data extracted from video footage, enabling a detailed examination of animal motion on a frame-by-frame basis. The approach is both methodical and precise, offering a high degree of accuracy in distinguishing between states of motion and stillness in animal behaviour.

  At the onset, the script focuses on meticulously extracting key point data. Each data point within this file represents a specific key point in the animal's body, captured in a single frame of the video. The extraction process involves the utilization of regular expressions, a powerful tool for parsing and structuring complex data sets. These expressions are adeptly applied to sift through the file content, isolating and organizing the key point data into a format conducive for detailed analysis. This preprocessing stage is crucial as it sets the foundation for accurate movement analysis.

Analytical Methodology

Central to the script's functionality is the compute_difference function. This function is engineered to assess the movement between each pair of consecutive frames. It accomplishes this by:

I) Transforming the key points from each frame into numerical arrays. This transformation is key to enabling mathematical operations that quantify movement.
II) Computing the Euclidean distance between these arrays from successive frames. This distance effectively measures the degree of movement or change in position of the key points from one frame to the next.
III) Aggregating these distances across all frames and computing their average. This average serves as a quantitative indicator of the overall movement exhibited by the animal throughout the sequence of frames.

Motion Determination and Reporting

The culmination of the script's process is the application of a decision-making logic based on a preset motion threshold. This threshold is a critical parameter, acting as the benchmark to differentiate significant motion from relative stillness. The script evaluates the average movement calculated against this threshold. If the average exceeds the threshold, it is interpreted as an indication of predominant motion; if it falls below, the animal is deemed to be primarily stationary. This threshold, while preset, offers flexibility and can be adjusted to suit the specific nuances and requirements of the study.

This Python script represents a significant advancement in the field of animal movement analysis. By methodically processing frame-by-frame key point data, it provides a nuanced and quantitative understanding of animal motion. This tool is invaluable for professionals in wildlife research, behavioural studies, and other related fields, offering a new level of insight into the dynamics of animal movement.

- Quality Check of Keypoints: An assessment of the key points' quality is conducted to filter out any noisy or irrelevant data. This step is crucial for maintaining the accuracy of the analysis. Below we can find in detail the steps of quality check of KeyPoint's.

  The code begins by addressing the challenge of processing and visualizing key point data from animal movement videos. It efficiently loads and parses JSON data files, which contain vital information about the key points. This step is crucial for accurate data handling, especially given the variability in file structures. Once the data is loaded, the code employs Matplotlib to visualize the key point scores across video frames. This visualization is instrumental in assessing the consistency and accuracy of the key points over time, providing a foundational understanding of the data quality.

Error Handling and KeyPoint Analysis

A significant aspect of the code is its robust error handling capabilities, particularly in managing JSON decoding errors. This ensures that data processing is reliable and accurate. Additionally, the code allows for the selection of specific video frames and the extraction of corresponding key points. This feature is essential for detailed frame-by-frame analysis. Furthermore, using OpenCV, the code plots these key points on the video frames themselves, offering a clear visual representation of their spatial distribution and movement.

Data Quality Enhancement and Output

The code's advanced analysis techniques include the detection of outliers in keypoint deviations using statistical methods, crucial for pinpointing inaccuracies. It also computes the derivatives of keypoint positions to identify inconsistencies or sudden changes in movement. To enhance data quality, the code interpolates missing or inconsistent keypoints and employs threshold-based correction to ensure smooth transitions. Finally, the enhanced data is saved back into a JSON file, and the code demonstrates advanced visualization techniques for comparative analysis, such as drawing keypoints on split frames of the video. This not only finalizes the data correction process but also provides an in-depth comparison between different data sets or treatments.

Overall, the code offers a comprehensive solution for analysing, visualizing, and enhancing the quality of keypoints in video data, particularly in the context of cow movement. Its blend of data processing, error handling, visualization, and advanced analysis techniques makes it a significant tool for further analysis.

- **Feature Extraction**

After applying these routines and obtaining the relevant frames for each cow, the next step is to extract specific features from these frames. At this point, we should mention that motion analysis is a powerful tool that can yield valuable insights into an individual's gait pattern, and its applications extend across diverse fields, including sports science, rehabilitation, and biomechanics. Hence, in the context of our specific task, we identified and propose the extraction of 11 indicators related to cows' gait patterns. This is a significant step in understanding and evaluating their movement dynamics for the development of the decision-making tool.

Up to this point, the procedures for computing the indicators have been developed, with a comprehensive presentation of their detailed calculation scheduled for the upcoming Deliverable. A brief overview of these 11 indicators is presented below:

- Step width, which refers to the lateral distance between the midpoint of the heels of two consecutive footfalls. It is often used to assess gait stability. To calculate step width, measure the horizontal distance between the midpoints of the heels of the left and right footfalls.
- Step length is the linear distance between the point of initial contact of one foot and the point of initial contact of the opposite foot. It provides insight into the overall distance covered during each step. Step length can be calculated by measuring the distance between the initial contact points of the left and right foot.
- Step time represents the duration of one complete step cycle, usually measured in seconds. It includes the time from the initial contact of one foot to the next initial contact of the same foot. To calculate step time, divide the total time taken for a step cycle by the number of steps taken.
- Stance time refers to the duration between initial contact and toe-off of the same foot. It provides information about the time a foot spends on the ground during each step. To calculate stance time, measure the time from initial contact to toe-off of a single foot.
- Stride length is the linear distance between the initial contact point of one foot to the next initial contact point of the same foot. It represents the distance covered in two consecutive steps of the same foot. Stride length can be calculated by doubling the step length.
- Stride time is the duration of one complete gait cycle, including two consecutive steps of the same foot. It is measured from the initial contact of one foot to the next initial contact of the same foot. To calculate stride time, measure the time taken for two consecutive steps of the same foot.
- Tracking-Up refers to the alignment of footsteps concerning a specific path or line. It is often used in sports and rehabilitation settings to assess the accuracy of movement. Calculating tracking-up involves comparing the actual footprints to the desired path and evaluating the deviation.
- Abduction refers to the movement of a limb away from the midline of the body. In gait analysis, abduction can be assessed by observing the angle between the thigh or lower leg and the midline of the body during walking.
- Asymmetry of variables involves comparing various gait parameters (such as step width, step length, step time, stance time, and force) between the left and right sides of the body. Calculate the absolute or relative differences between corresponding parameters on the left and right sides to determine asymmetry.
- Inconsistency of variables involves assessing the variability in gait parameters, such as stance time and force, within a single gait cycle. It can be calculated using

statistical methods like standard deviation or coefficient of variation to measure the dispersion of values around the mean.

– Locomotion score is a composite measure that evaluates overall gait performance based on multiple parameters. It can be calculated by assigning weights to individual parameters (such as step length, step time, and stance time) and summing up the weighted scores to obtain a total locomotion score.

- **Building Classifiers and Anomaly Detection Models**

In the upcoming phase, the final stage will involve utilizing the extracted features for constructing machine learning classifiers or anomaly detection models. These models will be specifically designed and trained to distinguish between normal and abnormal gait patterns in cows. Consequently, the application of these models will enable us to make informed decisions regarding whether a cow's gait is normal or indicative of lameness.

This prospective approach to lameness detection in cows outlines a structured plan for the future, leveraging the capabilities of AI and machine learning techniques to offer a non-intrusive, accurate, and efficient solution. By segmenting the procedure into smaller tasks, we aim to ensure a thorough analysis of each aspect of the cow's gait, providing reliable and actionable insights for enhancing animal welfare.

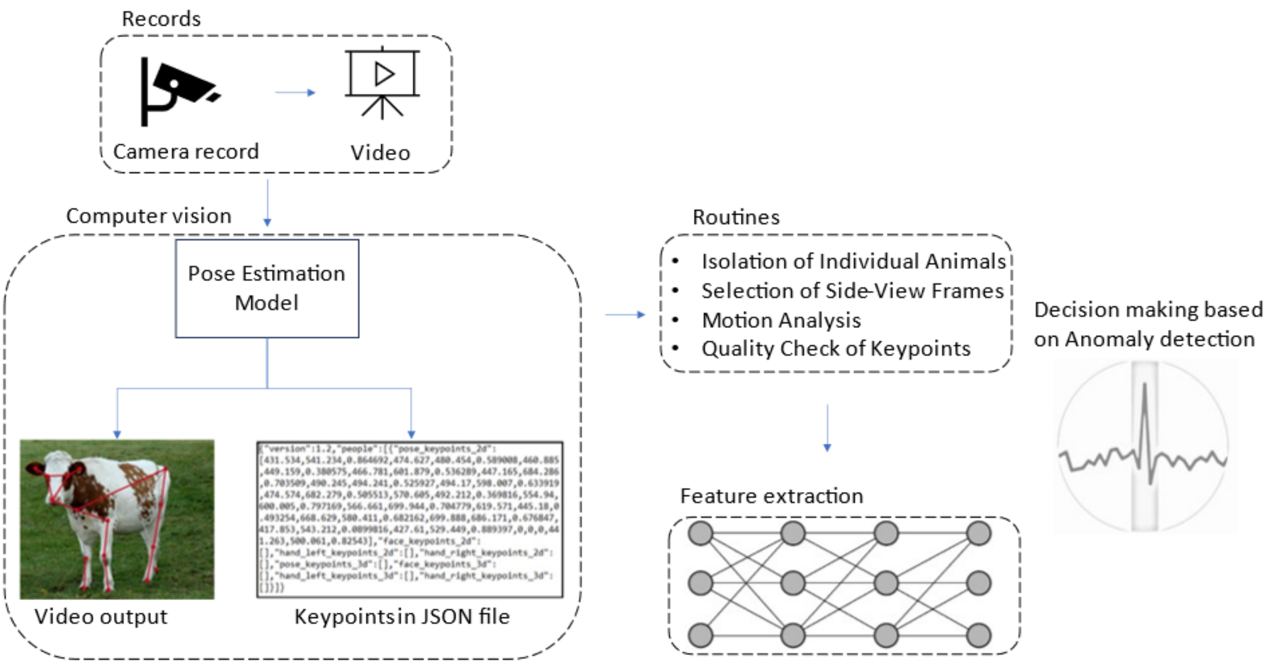


Figure 4.2.7: Pipeline overview

AI DEEP LEARNING PROCESS / ALGHORITHMS

THEORETICAL BACKGROUND OF MODELS

- **Employed models in relevant work.**

In recent years, the advancement of technology has provided a plethora of tools for detecting lameness in animals, particularly in the case of cows. With a focus on precision and efficiency, researchers have introduced various algorithms and models aimed at accurately estimating the poses and actions of cows, contributing to enhanced health monitoring and management on dairy farms. For instance, Gong et al.[9] introduces a YOLOv4-based algorithm for precise cow pose estimation, demonstrating notable performance in detecting multiple cows. Employing 2D pose detection, the algorithm achieves high precision in standing, walking, and lying poses. In the study by Wang et al.[10], the authors present OP-Mask R-CNN, a novel approach for enhancing individual cattle identification. Through optimization strategies and fusion mechanisms, the method improves accuracy and recognition speed compared to the original Mask R-CNN model. Wei et al.[11] introduces an algorithm for precise cow pose estimation in real farm environments, employing a Multi-Scale Temporal Convolutional Network (MS-TCN) with Coord Attention, demonstrating high precision. In the work by Russello et al.[12], the authors extend a static deep-learning model for animal pose estimation, achieving solid performance on occluded data. Fan et al.[13] introduce a concise multi-branch network (CMBN) for efficient cattle pose estimation, outperforming state-of-the-art models while maintaining low parameters and FLOPs. Hua et al.[14] highlight YOLOX-Pose's prowess in extracting cow skeletons with high precision, while PoseC3D, using 3D CNN, achieves accurate recognition of various cow actions.

- **Animal Segmentation with Nano YOLOv8**

Animal segmentation, a fundamental task in computer vision, involves the precise delineation of animals within images or videos. This process is crucial for various applications, including wildlife conservation, agricultural monitoring, and scientific research. The advent of advanced deep learning techniques has revolutionized animal segmentation, enabling the development of highly accurate and efficient models like Nano YOLOv8. Nano YOLOv8, based on the YOLO (You Only Look Once) object detection architecture, is optimized for real-time applications and operates on lightweight networks, making it particularly suitable for resource-constrained environments.

Nano YOLOv8 employs a series of convolutional layers to extract hierarchical features from input images, enabling the model to learn intricate patterns and textures associated with

9. Gong, C., Zhang, Y., Wei, Y., Du, X., Su, L., & Weng, Z. (2022). Multicow pose estimation based on keypoint extraction. PloS one, 17(6), e0269259.

10. Wang, J., Zhang, X., Gao, G., Lv, Y., Li, Q., Li, Z., ... & Chen, G. (2023). Open Pose Mask R-CNN network for Individual Cattle Recognition. IEEE Access.

11. Wei, Y., Zhang, H., Gong, C., Wang, D., Ye, M., & Jia, Y. (2023). Study of Pose Estimation Based on Spatio-Temporal Characteristics of Cow Skeleton. Agriculture, 13(8), 1535.

12. Russello, H., van der Tol, R., & Kootstra, G. (2022). T-LEAP: Occlusion-robust pose estimation of walking cows using temporal information. Computers and Electronics in Agriculture, 192, 106559.

13. Fan, Q., Liu, S., Li, S., & Zhao, C. (2023). Bottom-up cattle pose estimation via concise multi-branch network. Computers and Electronics in Agriculture, 211, 107945.

14. Hua, Z., Wang, Z., Xu, X., Kong, X., & Song, H. (2023). An effective PoseC3D model for typical action recognition of dairy cows based on skeleton features. Computers and Electronics in Agriculture, 212, 108152.

different animals. Its architecture incorporates techniques such as feature pyramid networks (FPN) and anchor box priors, enhancing its ability to handle objects of various sizes and aspect ratios. Additionally, Nano YOLOv8 integrates a highly efficient backbone network, reducing the computational burden while preserving the model's accuracy.

In the realm of animal segmentation, cows present specific challenges due to their diverse appearances, unpredictable movements, and varying environmental conditions. Nano YOLOv8's robustness lies in its adaptability to these challenges. By leveraging advanced object detection algorithms, the model excels in accurately delineating cows from cluttered backgrounds. Its ability to consider both global context and local details ensures precise segmentation results, even when cows exhibit complex poses or occlusions. Moreover, Nano YOLOv8 incorporates techniques from instance segmentation, allowing it to differentiate between individual cows within a herd accurately.

This technology holds immense promise for agricultural applications, enabling comprehensive monitoring of cows. Researchers can leverage Nano YOLOv8's capabilities for tasks such as health assessment, behaviour analysis, and tracking individual animals over time. By providing detailed insights into cow behaviour and well-being, Nano YOLOv8 contributes significantly to advancing precision livestock farming practices and enhancing animal welfare studies. Its efficiency and accuracy in cow segmentation make it a valuable tool in the scientific community, driving innovations in both computer vision and animal science.

- **Advancements in Pose Estimation Using MMpose for Animal Monitoring: A Focus on Cows**

Pose estimation, a critical component in computer vision and animal behaviour analysis, has witnessed significant advancements with the introduction of MMpose, a state-of-the-art framework designed for human pose estimation and beyond. In recent years, researchers have begun exploring its application in animal studies, with a specific emphasis on cows, an essential livestock species. Pose estimation in cows is vital for various reasons, including health monitoring, behaviour analysis, and precision agriculture. MMpose, leveraging advanced deep learning techniques, enables accurate and efficient estimation of cow poses, facilitating real-time data collection and analysis.

The technical steps involved in pose estimation using MMpose for cows are intricate yet innovative. Initially, a dataset comprising annotated cow images is curated, ensuring diverse poses and environmental conditions. This dataset serves as the foundation for training the deep neural networks within MMpose. The framework employs Convolutional Neural Networks (CNNs) for feature extraction, capturing intricate details of cow anatomy. These features are then fed into the Pose Estimation Network (PENet), a key component of MMpose, which refines the features and predicts keypoints corresponding to different body parts. PENet utilizes a carefully designed architecture, often involving stacked hourglass networks, facilitating precise localization of keypoints.

The theory underpinning MMpose revolves around the concept of keypoint detection. Keypoints represent specific anatomical landmarks, such as joints and facial features, essential for understanding the pose. MMpose employs the concept of heatmap regression, where the network predicts heatmaps representing the likelihood of keypoints' presence at various spatial locations. Simultaneously, offset vectors are predicted to refine the keypoints' positions, enhancing accuracy. This combination of heatmap regression and offset vectors enables MMpose to handle complex and varied cow poses effectively. Additionally, the integration of graph convolutional networks (GCNs) aids in capturing the spatial dependencies between keypoints, enhancing the overall robustness of pose estimation, especially in cases of occlusions or challenging poses.

Furthermore, MMpose incorporates data augmentation techniques and regularization strategies to improve generalization and mitigate overfitting. Augmentation methods such as rotation, scaling, and translation diversify the training data, enabling the model to handle a wide array of pose variations. Regularization techniques, including dropout and weight decay, ensure that the model's performance is optimized during both training and inference phases.

In conclusion, the integration of MMpose in animal studies, particularly in cows, represents a paradigm shift in precision agriculture and animal behaviour analysis. By accurately estimating cow poses, researchers can gain valuable insights into their behaviour, health, and well-being. The technical steps, rooted in advanced deep learning methodologies and innovative network architectures, coupled with the robust theoretical foundations, empower MMpose to revolutionize pose estimation in the context of animal monitoring, paving the way for more efficient and sustainable agricultural practices.

Figure 4.2.8: Keypoints based on the mmPOSE model.

RESULTS / OUTCOMES

- **Animal segmentation: A grazing herd of cows**

As previously mentioned, this technology holds immense promise in agricultural research, enabling detailed monitoring of cows for health assessment, behaviour analysis, and overall herd management. Its efficiency in cow segmentation underscores its significance in advancing precision livestock farming and animal welfare studies.

Hence, we utilized publicly available videos to create and enhance our tools using pre-trained networks. During our search, we discovered a dataset featuring a grazing herd of cows, on which we focused our efforts for animal segmentation tasks.

As we observed in Figure 4.2.9., the pre-trained model in this public video can accurately perform object segmentation, specifically identifying cows with a high degree of precision. The results of the segmentation model show not only the recognition of the cow as a class, but also display the class label and the associated probability at the upper side of the bounding box. This further enhances the utility of the model in distinguishing cows from other objects in the scene. The 'mask' in this context refers to the area within the bounding box that the model has identified as the cow, effectively isolating it from the rest of the image to provide a clear and focused classification.

Figure 4.2.9: Animal Segmentation.

- **Pose estimation.**

As outlined in our methodology, the MMPose model was rigorously applied to the dataset provided. This advanced model successfully extracted 17 distinct keypoints from each animal, offering a nuanced and detailed analysis of their movements. Figure 4.2.10. distinctly showcases the results of applying the MMPose model to our specific case study, highlighting its efficacy. Moreover, the model demonstrates remarkable versatility, effectively analysing cattle behaviour in various scenarios, including close proximity interactions, distant observations, and even in cases where the animals are at an incline. This robust application underscores the model's capability in capturing and interpreting a wide range of cow movements, providing valuable insights into their behaviour under different conditions.

Figure 4.2.10: Application of mmPose to cows.

- **Routines**

Below are presented the results following the application of the aforementioned routines in our herd.

- Animal Keypoints isolation. We worked on animal keypoints isolation to give a specific ID to each cow. Specifically, we can individually identify and isolate each cow, which is demonstrated below. In Figure 4.2.11., we present example frames from our processed video, illustrating how each animal is individually isolated and annotated, providing a clear visual representation of our keypoint extraction and annotation process.

Figure 4.2.11: Isolated cows.

- Selection of Side-View Frames. An integral aspect of our proposed methodology involves the meticulous selection of frames. A frame is deemed acceptable for our analysis when the cow is captured from the side. Utilizing the length/height ratio for frame selection empowers us to improve the precision and reliability of our gait analysis. This approach offers the distinct advantage of reducing variability in camera angles and distances, facilitating a standardized assessment of the cow's gait pattern from a vertical perspective. Figure 4.2.12. illustrates both an unacceptable and an acceptable viewing angle for analysing the cow's gait.

a b

Figure 4.2.12: a) Unacceptable frame and b) Acceptable frame.

Figure 4.2.13. presents a detailed time-series analysis of the length-to-height ratio, a critical metric for identifying the most informative frames in videos of cow gait. This ratio is pivotal in our study, as a higher value reliably indicates that the cow is being viewed laterally — the optimal perspective for gait analysis. We meticulously selected and annotated specific frames from high-definition videos, concentrating exclusively on instances where cows were captured

in a side view. The calculation of the length-to-height ratio in these frames is instrumental. A ratio threshold of 1.18 emerges as a key indicator: it marks the transition where the cow's body length, observed as it initiates a stride, consistently exceeds its height. This threshold is essential, as it ensures the cow is positioned laterally, providing a clear, unobstructed view for precise gait analysis. Thus, a ratio of 1.18 or higher is not just indicative but also guarantees that the cow is in the ideal sideways orientation for our analytical objectives.

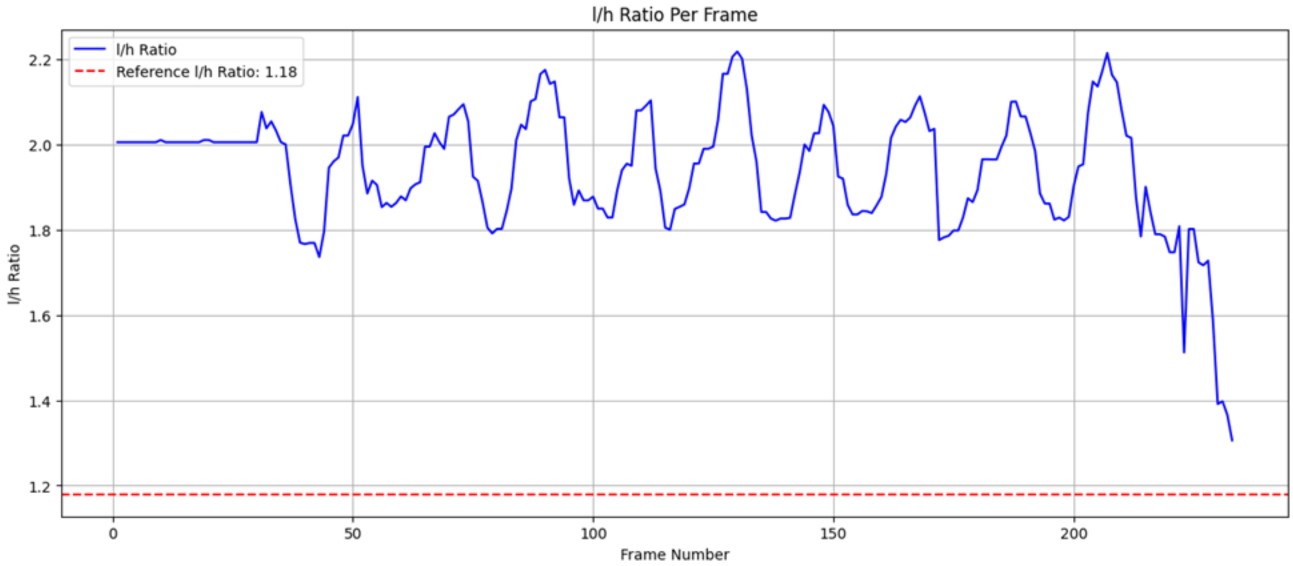


Figure 4.2.13: Time-series of length/height ratio.

- **Motion detection**

The preprocessing stage is essential, setting the stage for precise analysis by enabling the accurate differentiation of motion and stillness in animal behaviour. Figure 4.2.14. illustrates the sequential movement of a cow's gait captured through this process. The series of images clearly depicts the dynamic posture and limb positioning of the cow as it moves across the frame, showcasing the effectiveness of the motion detection technique in highlighting active phases of the cow's natural locomotion. This visual representation underscores the script's ability to distinguish and document the intricacies of animal movement with high accuracy.

Figure 4.2.14: Motion detection.

In Figure 4.2.15. below, a detailed analysis of the speed indicators for all keypoints is presented. This analysis specifically covers a sequence of 120 frames, which corresponds to a duration of 2 seconds. A notable observation from this data is the evident motion in the leg keypoints, as clearly illustrated in the frames above.

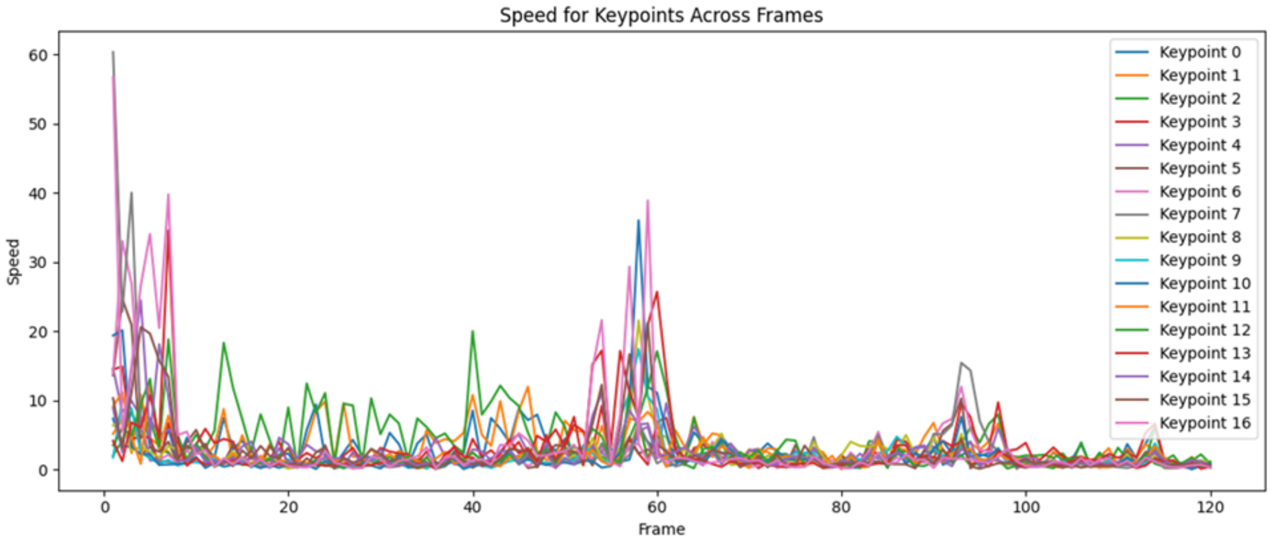


Figure 4.2.15: Speed indicator of all keypoints.

- **Quality check of keypoints**
  - Data Processing and Visualization is pivotal for precise data management, particularly considering the diversity in file structures. After successfully loading the data, the code utilizes Matplotlib to create visualizations of keypoint scores across video frames, as illustrated in Figure 4.2.16. This visualization serves as a valuable tool for evaluating the consistency and accuracy of keypoints over time, establishing a foundational understanding of the data quality.

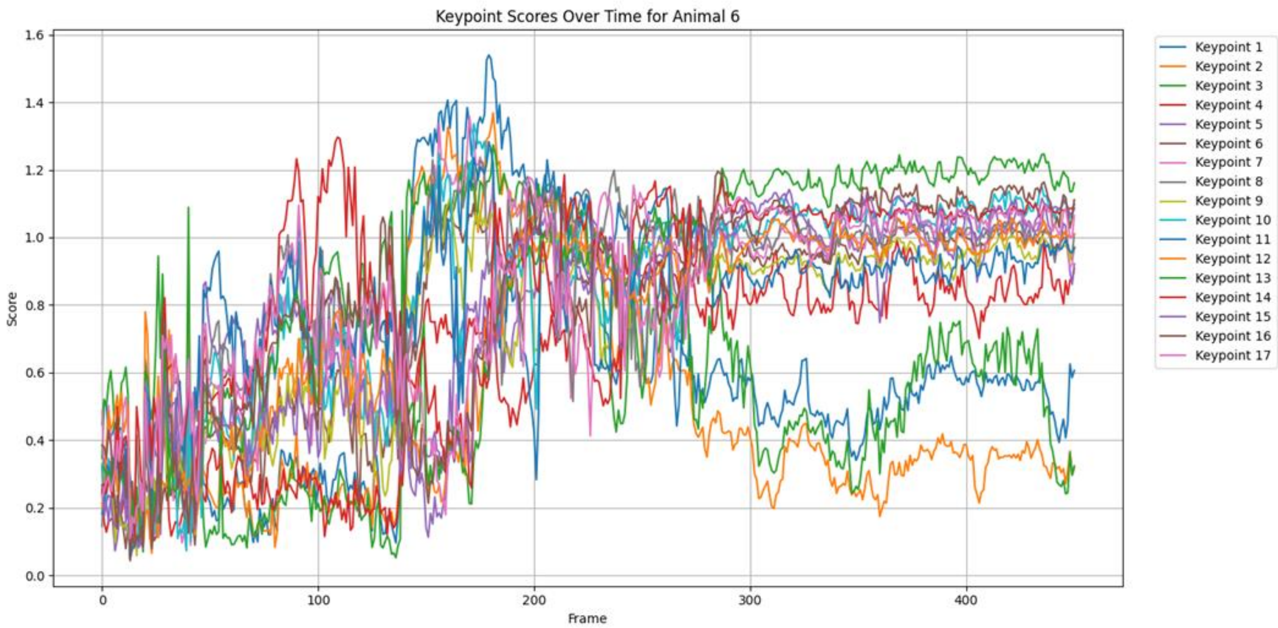


Figure 4.2.16: Keypoints visualization per frame.

- Error Handling and Keypoint Analysis is an integral feature of the code lies in its robust error-handling capabilities, with a specific emphasis on effectively managing JSON decoding errors. This meticulous approach guarantees the reliability and accuracy of data processing. Moreover, the code incorporates the functionality to choose specific video frames and extract corresponding keypoints, a critical feature for conducting detailed frame-by-frame analysis.

  Furthermore, leveraging OpenCV, the code goes a step further by plotting these keypoints directly on the video frames. This application of visual representation provides a clear depiction of the spatial distribution and movement of keypoints, enhancing the overall understanding of their dynamics.

- Data Quality Enhancement and Output. The code incorporates advanced analysis techniques, including the identification of outliers in keypoint deviations through statistical methods, a critical step in pinpointing inaccuracies as depicted in Figure 4.2.17. Additionally, it calculates the derivatives of keypoint positions to detect inconsistencies or abrupt changes in movement. In order to enhance data quality, the code performs interpolation for missing or inconsistent keypoints and applies threshold-based correction to ensure smooth transitions.

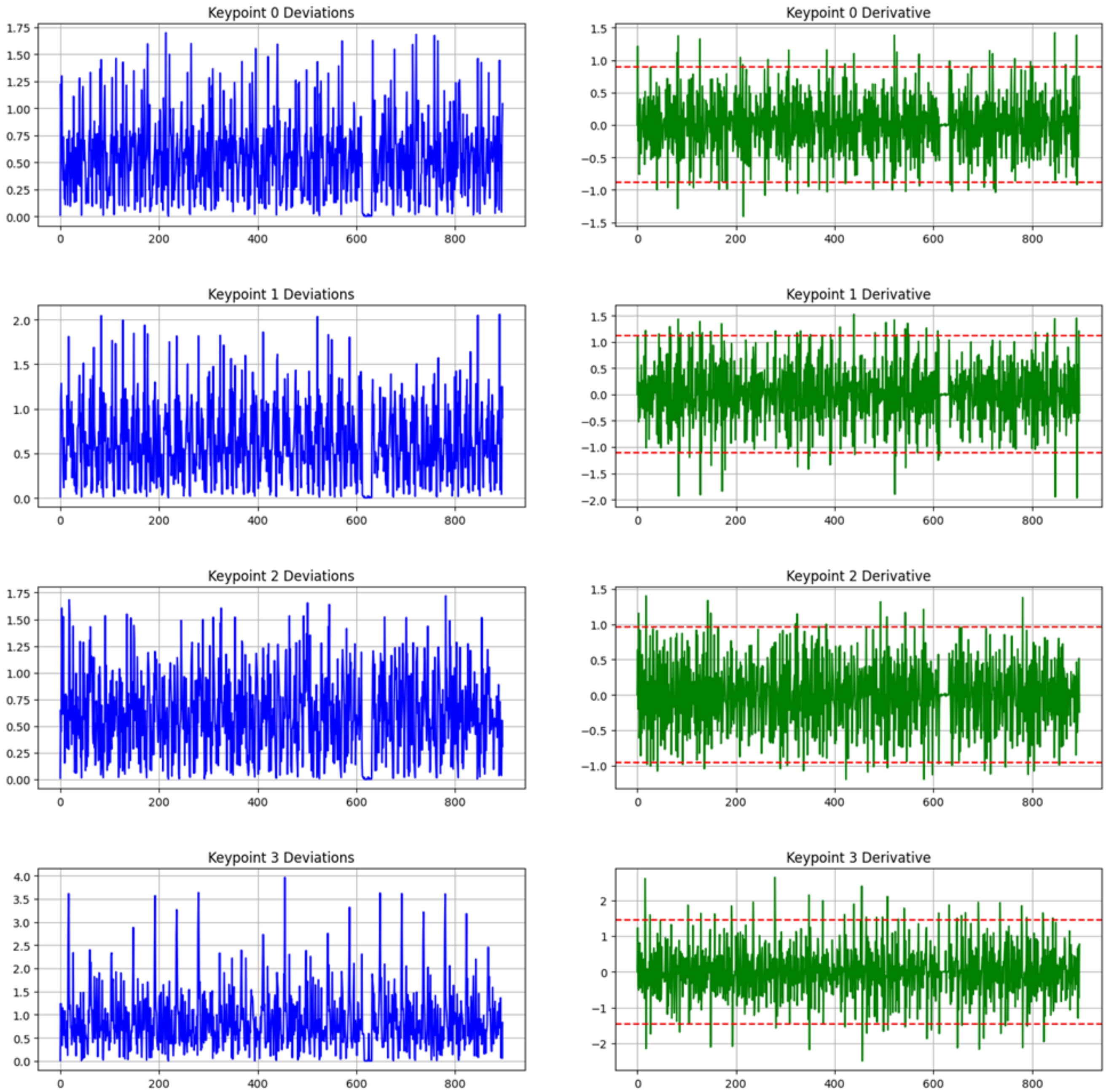


Figure 4.2.17: Example of a Quality Check for Four Keypoints.

Overall, the code offers a comprehensive solution for analysing, visualizing, and enhancing the quality of keypoints in video data, particularly in the context of cow movement. Figures show acceptable keypoints (Figure 4.2.18.) and noise keypoints (Figure 4.2.19.) of a cow's gait.

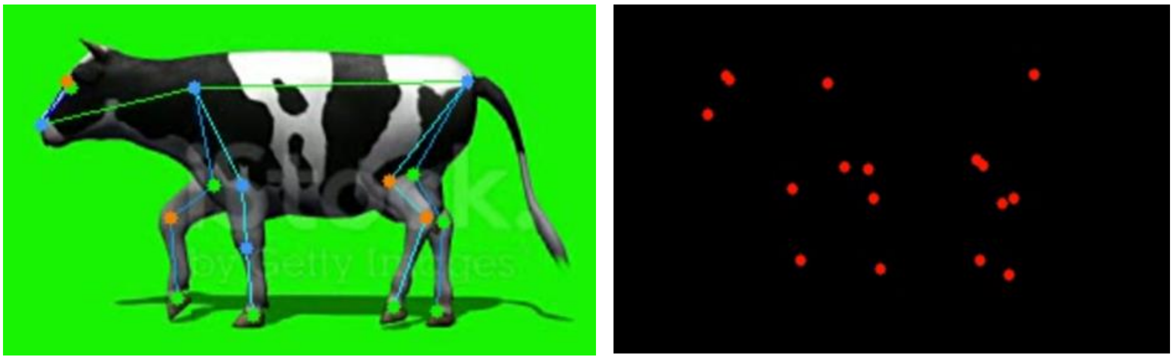


Figure 4.2.18: Acceptable keypoints from cows' gait.

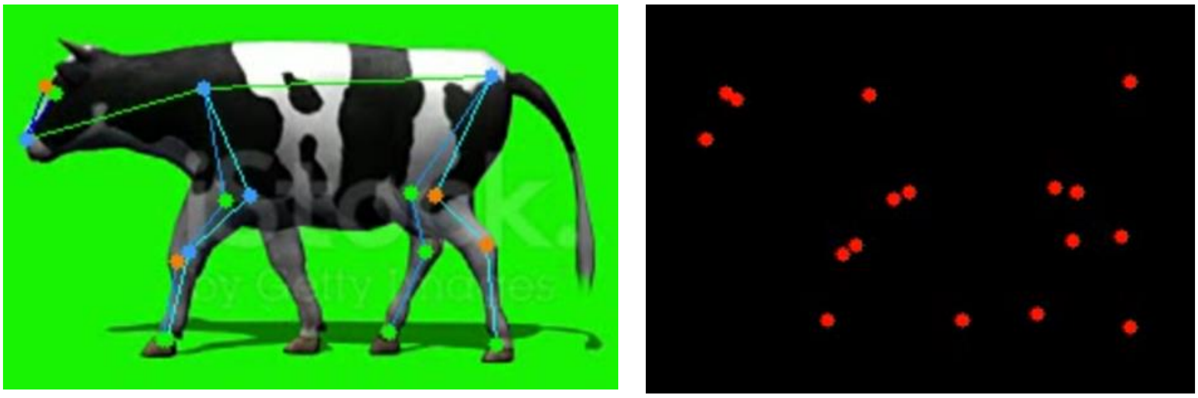


Figure 4.2.19: Noise keypoints from cows' gait.

- **Conclusions**

In conclusion, this deliverable represents a pivotal advancement in precision agriculture and animal behaviour analysis, particularly in the study of cows. By integrating MMPose, a state-of-the-art deep learning tool, we have established a robust architecture that is critical for high-quality, precise pose estimation. This technology is set the way we monitor and understand animal behaviour, offering significant contributions towards more efficient and sustainable agricultural practices.

Our focus has been on the accurate estimation of cow poses, a vital aspect in understanding their behaviour, health, and well-being. Special attention has been given to gait analysis, a key indicator in identifying lameness, a common issue in livestock. We have meticulously selected parameters such as step width, length, stance time, and asymmetry as the most relevant features for our analysis. These parameters are crucial for our future analysis.

In this deliverable, we have laid out a comprehensive pipeline, starting from video capture to the application of pose estimation models. This pipeline includes a series of pre-processing steps, ensuring the isolation of individual animals, selection of appropriate side-view frames, and a rigorous quality check of keypoints. These steps are designed to maintain the integrity and quality of the data, forming the foundation upon which our feature extraction process is built.

Looking ahead, we plan to further enhance the utility of our pose estimation data. In our next deliverable, we will extract the aforementioned indicators and implement various machine learning models for a more nuanced classification of cow gaits. This personalized classification system, categorizing gaits into distinct categories, will greatly aid in informed decision-making concerning animal health and welfare.

Ultimately, the comprehensive approach of this project, from leveraging cutting-edge pose estimation technology to the meticulous planning of data processing steps, demonstrates a holistic and innovative strategy in advancing animal studies within the domain of precision agriculture. The outcomes of this endeavour are twofold: enriching the scientific understanding of animal behaviour and providing practical means to enhance the well-being and productivity of livestock in agricultural settings.

4.3. LIVESTOCK MANAGEMENT (HERD) AND MONITORING (INDIVIDUAL ANIMAL)- ANIMALS' HEALTH

RELEVANCE AND PROBLEMATIC

A primary objective of this project is to leverage advanced technology for the improvement of animal husbandry practices. This initiative is primarily focused on the CHAMELEON platform, a cutting-edge solution designed to revolutionize the way breeders monitor and manage their livestock. The platform aims to provide comprehensive tracking and health assessment capabilities for individual animals, thereby enhancing the overall efficiency and effectiveness of livestock management. These technological advancements are anticipated to be particularly beneficial for breeders, regardless of the size of their animal capital.

To lay the groundwork for this ambitious project, we embarked on an exploration of the latest advancements in the field of computer vision. A meticulous search was conducted to identify existing public datasets that could be instrumental in developing the foundational technology for this initiative. These datasets served as a crucial resource in the creation of an array of computer vision tools. These tools were not only developed but also meticulously adapted and retrained with project-specific data to ensure their efficacy in real-world scenarios. The adaptability and precision of these tools are crucial, as they are intended to support the CHAMELEON platform in various capacities: from monitoring the activities and positions of the herd to assessing the health status of individual members and providing rapid response capabilities to sudden, unforeseen events.

THEORETICAL BACKGROUND

- **State-of-the-art models**

State-of-the-art models in animal detection and tracking often rely on deep learning techniques, particularly convolutional neural networks (CNNs) and their variants. Researchers have developed models with remarkable accuracy and efficiency, enabling scientists to monitor and study animal populations in ways that were previously not possible.

- Animal Detection involves identifying the presence and location of animals within images or video frames. State-of-the-art models utilize advanced CNN architectures, such as Faster R-CNN[15] and YOLO (You Only Look Once)[16], to achieve real-time object detection. These models are trained on large and diverse annotated datasets, enabling them to recognize various animal species in different environmental conditions. Transfer learning techniques, where models pretrained on large datasets are fine-tuned for specific animal detection tasks, have also become popular. This approach allows for accurate detection even with limited labelled data.
- Animal Tracking aims to follow the movement and behaviour of individual animals over time. Deep learning-based tracking models often combine object detection and tracking algorithms. One approach is to employ Siamese networks, which are designed to learn similarity metrics between object representations, making them effective for tracking objects across frames. Recurrent Neural Networks (RNNs)[17] and Long Short-Term Memory Networks (LSTMs)[18] are also integrated to capture temporal dependencies in animal movements, enhancing tracking accuracy. These models can handle complex scenarios, such as occlusions and abrupt changes in motion, making them invaluable for wildlife research and conservation efforts.

- **Related work on sheep monitoring practices**

This brief introduction provides an overview of the impact of computer vision on sheep monitoring practices, setting the stage for a closer examination of specific studies exemplifying these advancements. In a study by Sarwar et al.[19], the exploration involves sheep detection and counting using UAV video. The research employs Region-based Convolutional Neural Networks and compares the results with other techniques. In a paper by Wang et al.[20], a lightweight and high-accuracy deep-learning method is introduced for grazing livestock detection in UAV imagery. Meng et al.[21] propose a two-stage sheep identification method using the YOLO

---

15. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.

16. Li, Y., Fan, Q., Huang, H., Han, Z., & Gu, Q. (2023). A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition. DRONES, 7(5), 304.

17. Mandic, D. P., & Chambers, J. (2001). RECURRENT NEURAL NETWORKS FOR PREDICTION: LEARNING ALGORITHMS, ARCHITECTURES AND STABILITY. John Wiley & Sons, Inc.

18. Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. ARXIV PREPRINT ARXIV:1909.09586.

19. Sarwar, F., Griffin, A., Periasamy, P., Portas, K., & Law, J. (2018, November). Detecting and counting sheep with a convolutional neural network. In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) (pp. 1-6). IEEE.

20. Wang, Y., Ma, L., Wang, Q., Wang, N., Wang, D., Wang, X., ... & Ouyang, G. (2023). A Lightweight and High-Accuracy Deep Learning Method for Grassland Grazing Livestock Detection Using UAV Imagery. Remote Sensing, 15(6), 1593.

21. Meng, X., Tao, P., Han, L., & CaiRang, D. (2022, March). Sheep Identification with Distance Balance in Two Stages Deep Learning. In 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC) (Vol. 6, pp. 1308-1313). IEEE

framework. In a paper by Koklu et al.[22] a CNN-based model using InceptionV3 is proposed to detect sheep breeds from facial images. Simões et al.[23] introduce the paper "DeepWILD: Wildlife Identification, Localisation, and estimation on camera trap videos using Deep learning". The DeepWILD includes a three-step process using MegaDetector for annotation, extending it with Faster R-CNN for detection and classification, and implementing a counting method.

In the literature review, various studies have explored sheep herd monitoring. Our approach stands out from previous initiatives by not solely concentrating on animal detection but also giving precedence to the well-being of animals in challenging terrains within Greece. Furthermore, we adopted the Yolov8 as state-of-the-art object detection model.

- **End-user requirements via questionaries**

We have developed a comprehensive questionnaire tailored for livestock farmers, designed to assess their needs and define new requirements for the project. This effort aims to pinpoint specific areas of improvement within the farming community. The identification of these needs will have a significant impact, shaping the direction of our initiatives and allowing us to better address the challenges faced by livestock farmers, fostering a more sustainable and prosperous livestock sector.

For the purposes of the project, this questionnaire was distributed to one expert farmer. Below, you will find the complete questionnaire, along with the responses provided by the farmer and accompanying technical comments.

Questionnaire

Part A: Problem Definition

A1 Drone-based Monitoring:

#Question

What is the ideal minimum altitude for a drone to monitor the herd without causing stress or fear to the goats?

End-user reply: 100-200m depending on the installed sensors' needs

#Question

Are there any specific terrain features or obstacles that could interfere with drone flight and video capture (e.g., trees, power lines, buildings)?

End-user reply: No, the only terrain features that need to be taken into account are the increasing slope and high altitude (almost 2000m) of the grazing areas that the herds occupy during the summer months.

[22] Koklu, M., Cinar, I., Taspinar, Y. S., & Kursun, R. (2022, June). Identification of sheep breeds by CNN-based pre-trained InceptionV3 model. In 2022 11th Mediterranean Conference on Embedded Computing (MECO) (pp. 01-04). IEEE.

[23] Simões, F., Bouveyron, C., & Precioso, F. (2023). DeepWILD: Wildlife Identification, Localisation and estimation on camera trap videos using Deep learning. Ecological Informatics, 75, 102095.

#Question
What types of activities or behaviours if possible do you want the drone to monitor and detect (e.g., grazing, resting, running)?
End-user reply: Monitor the sheep's posture, moving speed, and their actual geographical location.
#Question
How important is real-time data transmission from the drone to the monitoring station?
End-user reply: On a scale of 1-10, the importance of real-time monitoring is of upmost importance, meaning at least 8 out of 10.
#Question
On average, how fast can a goat move within the herd's area, and what is the typical range of their movement speed?
End-user reply: Sheep typically move at a speed of 7 to 10 km/h, but if they are ill or have any health issues, their speed is significantly reduced.
#Question
If the drone identifies a goat of interest, how long can the system wait before analysing and reporting the goat's behaviour and location (e.g., minutes, hours)? What is the maximum acceptable delay?
End-user reply: Real-time monitoring or near real-time monitoring is preferable (0.5-1 min delay).

A2 Health Status Monitoring:

#Question
What specific health indicators do you want the system to monitor using temperature data (e.g., fever, hypothermia, general health)?
End-user reply: The sheep's posture and moving speed are more important than their temperature.

#Question
How often should the temperature of each goat be measured for accurate health monitoring?
End-user reply: Let's talk about sheep posture and movement speed. It is important to monitor the overall health of the sheep every evening during the summer months, especially when the herds are in high-altitude areas that require a lot of time to reach.
#Question
What is the acceptable margin of error for temperature measurements based on thermal images?
End-user reply: -
#Question
Are there any visible signs of health issues that can be detected using the RGB camera (e.g., changes in posture, limping)?
End-user reply: Yes, the change in posture can be a sign of a health issue for sheep, as the herd owner indicated. For example, if the sheep are constantly moving their tails, it could be a signal that a certain type of fly is bothering them.
#Question
Are there any specific situations or events that should trigger an immediate health status check using the drone's cameras?
End-user reply: Yes, if the sheep's speed is significantly lower than the herd's average speed, this can be an indicator of health issues.

Part B: Data Creation and Deep Learning

B1 Data Collection:

#Question
How many goats are typically in your herd, and what is the average size of the area they occupy?
End-user reply: About 500 sheep need to be monitored at a time.

#Question
How long should the drone be able to operate without recharging or refuelling?
End-user reply: About 40 minutes (5 + 5 to get into place and get back).
#Question
Do you currently have any infrastructure in place to collect temperature data from the goats? If so, please describe it.
End-user reply: -
#Question
What duration of video footage would you consider sufficient for the pilot use case scenario to evaluate the system's effectiveness?
End-user reply: Less than 1 minute, in order to measure the sheep average and individual speed.

B2 Deep Learning:

#Question
What level of accuracy do you expect from the machine learning system for detecting and classifying goat behaviours?
End-user reply: On a scale of 1 to 5, at least 4 is required.
#Question
Would you prefer an automated alert system for specific events (e.g., a goat leaving the designated area, sudden temperature changes)?
End-user reply: Yes.
#Question
How quickly should the system be able to respond to changes in goat behaviour or health indicators?
End-user reply: Immediately.
#Question
Are there any specific data annotation requirements or preferences for the pilot use case scenario (e.g., labelling of goat behaviours, locations, or health indicators)?
End-user reply: Yes, we need the sheep posture, speed, and actual geolocation.

- **Summary of the project requirements**

  - Scanning the Herd from 30-80m: the team extensively explored the strategy of employing drones or a similar scanning system for monitoring the goat herd from a distance of 30-80 meters. The consensus was that this range is optimal as it allows for clear visuals while minimizing stress on the animals. This careful consideration showcases the team's commitment to animal welfare, ensuring that the monitoring process doesn't cause undue distress to the goats.
  - Herd Movement and Modelling: one innovative suggestion put forth was the implementation of a voice-triggered alert system. This system would signal the herd to move backward to a predefined location when necessary. To identify potential health issues within the herd, the team agreed to develop an AI model. This model would focus on recognizing goats that are moving slower than the average speed of the herd. This meticulous approach demonstrates the team's dedication to ensuring the well-being of the goats by promptly addressing any health concerns.
  - Health Monitoring Via Thermal Camera: the team delved into the possibilities offered by thermal cameras for monitoring goat health. Primarily, these cameras would detect body temperature and potentially identify body posture anomalies. This proactive approach to health monitoring showcases the team's commitment to utilizing cutting-edge technology. However, the team also acknowledged the need

for further research to accurately determine the effectiveness of this system in identifying health concerns. This dedication to research highlights the team's emphasis on data-driven decision-making.

- Alert System: the development of a robust alert system was a key focus of the discussions. This system would have dual functionality: firstly, alerting the herd manager in the event of unhealthy goats. The alerts would be generated based on data from both the movement model and thermal camera readings, ensuring a comprehensive assessment of the goats' health. Secondly, the alert system would notify the herd manager if a goat were found to be straying away from the herd. The team is actively defining the specific distance from the herd that would trigger this alert, demonstrating their attention to detail in designing a customized solution.
- Position Monitoring System: in emergency scenarios where swift action is essential, the team discussed the implementation of a two-phase automated mapping system for precise herd location tracking. This system would enable the rapid and accurate determination of the herd's exact position, ensuring a timely response in critical situations. This emphasis on emergency preparedness showcases the team's forward-thinking approach to herd management.

In conclusion, the project discussions reflect a comprehensive and thoughtful approach to goat herd monitoring. By combining innovative technologies, meticulous research, and a focus on animal welfare, the team is poised to create a cutting-edge monitoring system that not only ensures the health and safety of the goats but also facilitates efficient management practices for the herdsman.

PROPOSED METHODOLOGIES

The following section offers an in-depth theoretical exploration of three cutting-edge approaches employed in task 4.4, with a primary emphasis on object detection. This includes an examination of the YOLOv8 model, an advanced system optimized for 3-channel image analysis. Additionally, we delve into a custom-tailored variant of YOLOv8, specifically re-engineered for enhanced performance with 4-channel images. Lastly, we present an overview of the High-Altitude Object Detector, a sophisticated tool built upon the foundations of Waldo's V2 technology. This trio of models represents the forefront of object detection methodologies, each uniquely adapted to their specific application scenarios.

- **Three-channel OD**

In this subsection, we embark on our investigative journey with the YOLOv8 model, marking the initial step in our comprehensive examination of the problem at hand. YOLOv8 stands as an advanced system meticulously optimized for 3-channel image analysis, a domain that predominantly encompasses standard RGB (Red, Green, Blue) images. This model, renowned for its precision and efficiency, serves as a foundational tool in our approach. Our analysis aims to thoroughly understand how our specific problem behaves and responds when processed through the YOLOv8 framework. By dissecting its functionality and performance on 3-channel images, we can gain valuable insights into the model's capabilities and limitations, setting a robust baseline for our further explorations in object detection methodologies.

- Yolov8: the field of object detection remains a dynamic area of research, encompassing numerous scientific and practical applications[24]. The YOLO (You Only Look Once) family is widely employed for robust object detection needs. YOLO adopts a single-stage approach to object detection, offering various backbone sizes that can be selected based on the available processing power. For real-time object detection on resource-constrained devices, YOLO Tiny or YOLO Nano, the smallest models within the YOLO family, can be utilized, albeit with some limitations on performance. Despite its remarkable outcomes, the YOLO approach continues to be an active area of exploration. Ongoing research endeavours are dedicated to achieving optimal scores, specifically in terms of mean average precision (mAP), on existing datasets while minimizing the number of network parameters required. Presently, within the YOLO family, the YOLOv8 network architecture stands out as the most effective, and as such, we will concentrate on this version in our demonstration application.
- Track and count: in our project, we employ a method similar to the one used for tracking and counting objects using YOLOv8. This technique is highly valuable in computer vision, often used in fields like traffic analysis and industrial automation. Our approach begins with object detection using YOLOv8, selected for its effectiveness and ease of integration.
  For tracking, we utilize ByteTrack[25], a tool known for its accuracy in monitoring the movement of individual objects across predefined lines. The combination of YOLOv8 for detection and ByteTrack for tracking ensures precise identification and continuous monitoring of objects, which in our case, are sheep in a herd. This system is also enhanced with counting capabilities, a crucial aspect where we integrate additional tools to accurately count the number of sheep detected and tracked.

- **Four-channel model**

Given the challenges stemming from the unique characteristics of the Greek landscape and high altitudes, it is crucial to explore alternative techniques that offer additional information beyond the conventional RGB 3-channel image. A viable solution to address these challenges involves integrating thermal imagery into the RGB image, thereby creating a 4-channel image. The following is a detailed description of this approach.

- Integration of Thermal and RGB Images into 4-Channel Format for Object Detection: in the realm of multispectral imaging, the fusion of thermal and RGB (Red, Green, Blue) data is pivotal for enhancing the depth and quality of information available to computer vision systems. This fusion process involves consolidating thermal and RGB images into a unified 4-channel format, allowing for a comprehensive representation of the scene. The integration of thermal data, capturing heat signatures, with

---

24. Łysakowski, M., Żywanowski, K., Banaszczyk, A., Nowicki, M. R., Skrzypczyński, P., & Tadeja, S. K. (2023). Real-Time Onboard Object Detection for Augmented Reality: Enhancing Head-Mounted Display with YOLOv8. arXiv preprint arXiv:2306.03537.

25. Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W. and Wang, X., 2022, October. Bytetrack: Multi-object tracking by associating every detection box. In European Conference on Computer Vision (pp. 1-21). Cham: Springer Nature Switzerland.

traditional RGB imagery provides a holistic view that is invaluable for various applications, such as surveillance, environmental monitoring, and industrial inspections. To implement the aforementioned to achieve the integration of thermal and RGB images into a 4-channel format for object detection the following steps will be followed:

A. Thermal and RGB Image Fusion:

The preprocessing pipeline begins with the acquisition of both thermal and RGB images from their respective sensors. To consolidate these disparate data types, a fusion technique is applied. This process involves aligning the spatial information of the RGB and thermal images and then combining them into a single 4-channel image. Various fusion methods, such as pixel-level averaging or weighted blending, can be employed to ensure seamless integration of thermal and RGB data, preserving the unique characteristics of each modality.

B. Channel Reordering and Normalization:

Once the fusion is complete, the resulting 4-channel image typically contains the RGB channels (Red, Green, Blue) and the thermal channel. To ensure uniformity and compatibility with the chosen neural network architecture, the channels might need to be reordered. Common channel orders are RGBT (Red, Green, Blue, Thermal) or TBGR (Thermal, Blue, Green, Red), depending on the specific requirements of the application and the network architecture being used. Additionally, pixel values might be normalized to a common scale to enhance the model's ability to learn from the data effectively.

C. Integration with Object Detection Model:

The processed 4-channel images are then fed into the object detection model, such as the adapted Nano YOLOv8 mentioned earlier. During both training and inference, the model leverages the enhanced information provided by the fusion of thermal and RGB data. The model learns to detect and classify objects based on the combined features from all four channels, enabling it to make more informed and accurate predictions.

D. Advantages of Multispectral Fusion:

The integration of thermal and RGB data into a 4-channel format enhances the capabilities of the object detection system significantly. By leveraging the unique information offered by thermal imaging (such as heat signatures) alongside the visual cues from RGB imagery, the model gains a richer understanding of the environment. This comprehensive view enables the system to excel in tasks that require precise object detection, especially in scenarios where both thermal and visual information are critical, such as search and rescue missions, security applications, and environmental monitoring in challenging conditions.

In summary, the consolidation of thermal and RGB images into a 4-channel format represents a sophisticated preprocessing step that empowers object detection models to operate effectively in diverse and complex environments, providing valuable insights for a wide range of applications. Following the consolidation of thermal and RGB images into a 4-channel format, our next step involved modifying the YOLOv8 model to make it compatible with this enhanced input.

– Adaptation of YOLOv8 for Training and Inference on 4-Channel Images: in the domain of computer vision, the adaptation of advanced neural networks to accommodate specialized image formats is crucial for diverse applications. YOLOv8 is a version of the renowned YOLO (You Only Look Once) object detection algorithm, which has been successfully tailored to handle 4-channel images, a format commonly found in CMYK color models. This adaptation allows YOLOv8 to train and perform inference tasks with precision and efficiency on images containing Cyan, Magenta, Yellow, and Key (black) channels.

Rationale for Incorporating Thermal Data as the Fourth Channel in YOLOv8:

Prior to exploring the integration of YOLOv8 with CMYK-formatted 4-channel images, it's pivotal to understand the significance of including thermal information as the fourth channel. In our case, alongside the conventional RGB (Red, Green, Blue) channels, the fourth channel is dedicated to thermal data. This integration is innovative as it combines visual and thermal imaging, paving the way for a more comprehensive object detection system. While the combination of RGB and thermal imaging has been utilized in other applications such as in security cameras, its application to livestock monitoring represents a novel approach within this field.

Thermal imaging offers crucial insights that are invisible in standard RGB spectra, such as the detection of heat signatures and temperature variances. Additionally, in our case settings, the ability to detect temperature anomalies can aid in health status of animals.

By incorporating thermal data into the YOLOv8 model, we enable the system to detect and analyse objects not just based on their appearance but also on their thermal properties. This dual-data approach enhances the accuracy and functionality of the YOLOv8 algorithm, especially in scenarios where environmental conditions might obscure or alter the visual appearance of objects. The integration of this 4-channel approach, therefore, represents a significant advancement in the field of object detection, offering a more robust and versatile solution for livestock monitoring.

Introduction to CMYK-formatted 4-Channel Images and YOLOv8 Object Detection:

In the realm of digital imaging and computer vision, the CMYK color model stands as a standard representation of colors in the printing process. CMYK refers to the four ink plates used in color printing: Cyan, Magenta, Yellow, and Key (black). Unlike the RGB model commonly used in electronic displays, CMYK images possess four channels, each representing the intensity of one of these ink colors. This format is especially crucial in the printing industry, where precise color reproduction is paramount.

When it comes to object detection, YOLOv8 (You Only Look Once version 8) emerges as a state-of-the-art deep learning algorithm. It excels in real-time object detection tasks, demonstrating remarkable accuracy and speed. In the context of CMYK-formatted 4-channel images, an intriguing challenge arises. While YOLOv8 traditionally operates on 3-channel RGB images, the adaptation to CMYK-formatted images requires a meticulous approach to ensure seamless object detection.

Integration of YOLOv8 with CMYK-Formatted 4-Channel Images:

The integration of YOLOv8 with CMYK-formatted 4-channel images demands a thoughtful preprocessing strategy. Converting CMYK images to the standard 3-channel RGB format is an essential initial step. However, this conversion is not as straightforward as a mere removal of the black channel, as it holds critical information. Instead, a sophisticated method incorporating color space transformations and channel weighting is necessary to retain the vital details during the conversion process.

Once the CMYK image is appropriately transformed into RGB, YOLOv8 can be seamlessly applied for object detection. Leveraging its robust architecture and advanced feature extraction capabilities, YOLOv8 analyses the 3-channel RGB images, accurately identifying and classifying objects within the visual data. This integration enables precise object detection tasks on images initially saved in the CMYK format, bridging the gap between the printing industry's color representation needs and the cutting-edge capabilities of state-of-the-art object detection algorithms.

The following steps will be implemented for the adaptation of YOLOv8 for training and inference on 4-channel images:

A. Data Preprocessing for 4-Channel Images:

To prepare the 4-channel CMYK images for YOLOv8, a careful preprocessing pipeline has been devised. This process involves transforming the CMYK images into a format compatible with the neural network. Techniques such as color space conversion and channel reordering are applied to seamlessly integrate the extra channel information into the model.

B. Modification of Network Architecture:

The architecture of YOLOv8 has been strategically modified to handle 4-channel inputs. This adaptation includes adjusting the input layer to accept 4 channels, ensuring that the network can effectively process the additional color information present in CMYK images. Additionally, the model's internal layers and filters are optimized to make the most out of the enhanced input data.

C. Training Procedure:

During the training phase, the adapted YOLOv8 is trained on a dataset comprising 4-channel images. The training process involves fine-tuning the network's weights and biases, allowing it to learn and recognize patterns specific to CMYK images. Advanced optimization techniques are employed to ensure the model converges efficiently and effectively captures the nuances of the 4-channel data.

D. Inference on 4-Channel Images:

After the training phase, the adapted YOLOv8 excels at making accurate predictions on 4-channel images during the inference stage. The model leverages its learned features to detect and classify objects in CMYK images with a high degree of accuracy. This capability is invaluable in applications where precise object detection is required, such as in the printing industry or any other field utilizing CMYK color representations.

In summary, the adaptation of YOLOv8 to handle 4-channel images exemplifies the intersection of cutting-edge deep learning techniques and specialized image formats. This adaptation empowers the model to excel in tasks demanding intricate color information, opening new avenues for applications that rely on CMYK color models for accurate and detailed object detection.

- **High altitude object detector**

Addressing the challenge of varying the altitude in UAV-derived video capture for multi-object recognition requires specialized treatment. To tackle this, we employed the pre-trained Waldo v2 model, and its detailed description is provided below.

- Introduction to Waldo v2: Waldo v2 represents a cutting-edge localization algorithm designed for precise object detection and tracking in complex environments [[26]]. Building upon its predecessor's capabilities, Waldo v2 employs advanced machine learning techniques and computer vision algorithms to accurately locate and track objects, even in challenging conditions such as high altitudes. This algorithm is particularly valuable in applications where precise object localization is essential, such as ecological research involving wildlife monitoring. In this context, Waldo v2 plays a pivotal role in enhancing the capabilities of drones, enabling them to identify and track specific wildlife, such as beavers, with unprecedented accuracy.
- Integration with YOLOv7 Framework for High-Altitude Beaver Training: The integration of Waldo v2 with the YOLOv7 framework marks a significant milestone in the field of drone technology and computer vision. YOLOv7, renowned for its real-time object detection prowess, provides the computational efficiency and speed required for processing large volumes of visual data in real-world scenarios. By combining the strengths of YOLOv7 with Waldo v2's exceptional localization abilities, drones can now operate at high altitudes and accurately detect, classify, and track beavers in their natural habitat.

In the context of high-altitude beaver training, this integration enables drones to monitor beaver populations with unmatched precision. Traditional monitoring methods often face challenges at elevated altitudes, where visibility and environmental conditions are less than ideal. However, with the amalgamation of Waldo v2 and YOLOv7, drones equipped with this technology can swiftly identify and monitor beavers, providing invaluable data for ecological studies and conservation initiatives. This innovative approach not only revolutionizes wildlife monitoring but also contributes significantly to our understanding of high-altitude ecosystems and the behaviour of beaver populations within them.

[26.] Whereabouts Ascertainment for Low-lying Detectable Objects. The SOTA in FOSS AI for drones. Retrieved in November 2023 from: https://github.com/stephansturges/WALDO

RESULTS

In this section are the results for three approaches: the 3-channel YoloV8 approach, the modified 4-channel YoloV8 approach, and the High-Altitude Object Detector based on Waldo's V2. Each approach utilizes different datasets, which are detailed in the respective subsections below.

- **Results of the 3-channel YoloV8 approach**

Dataset description

The dataset used in this project plays a pivotal role in training and validating the computer vision tools developed for the CHAMELEON platform. This comprehensive dataset, sourced from the Roboflow Universe site, is meticulously curated to suit the specific needs of our project. It includes images of sheep and goats, captured using Unmanned Aerial Vehicles (UAVs) from various altitudes, providing diverse perspectives and conditions for analysis. It encompasses a wide range of images that are critical for the effective training, validation, and testing of our computer vision models.

The dataset is divided into three primary sets:

– Training Set: Comprising 3258 images, which account for 70% of the total dataset.
– Validation Set: Consisting of 958 images, making up 20% of the dataset.
– Test Set: Including 464 images, representing the remaining 10% of the dataset.

Overall, the total number of images in the dataset amounts to 4680. Samples of images from this dataset are illustrated in Figure 3.3.1. Each image in the dataset has been resized to a uniform dimension of 640x640 pixels, using a stretch resizing method to ensure consistency across the dataset. It's important to note that no augmentations were applied to these images, maintaining their original quality and characteristics.

In addition to the standard dataset, we have also incorporated specific videos and images that depict herds of sheep and goats. This inclusion is particularly relevant to the objectives of our project, as it provides a more realistic and practical context for the application of our computer vision tools. These specific visuals have been invaluable in enhancing the robustness of our model, enabling it to accurately track and assess the health of individual animals within these herds. The effectiveness of these additions and their impact on our computer vision tools are thoroughly presented and discussed in the results section of this report.

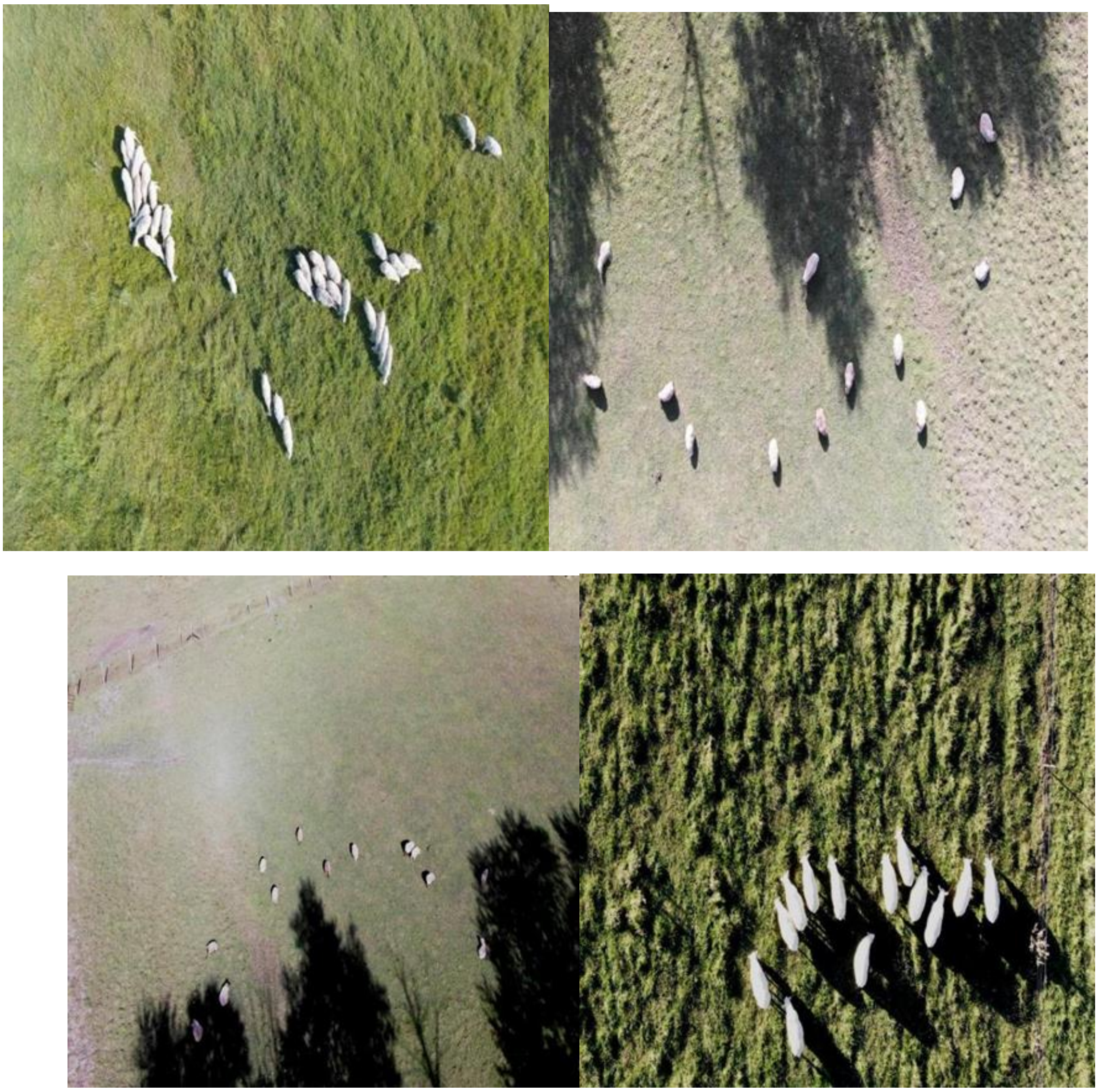


Figure 4.3.1: Sample images from the animal dataset.

External validation dataset from LAMMC

As the work package leader, LAMMC identified a herd for external evaluation of high-altitude animal detection models. The data collection took place in Lithuania and involved the use of both thermal and RGB cameras to capture comprehensive information for evaluation purposes.

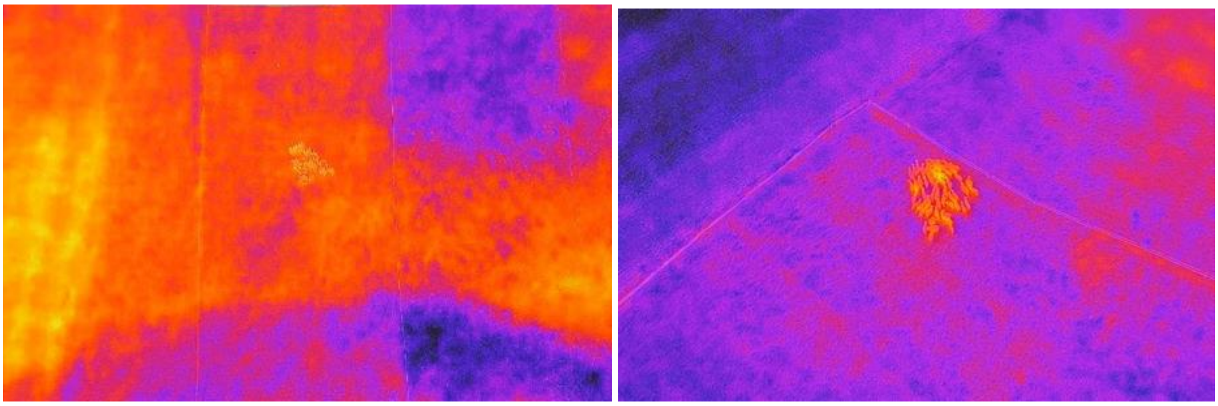


Figure 4.3.2: RGB images and images from the thermal camera

- **Results of Yolov8 trained on the public dataset.**

Our initial approach to solving the task of detecting sheeps and goats from UAVs was to identify relevant databases as previously mentioned and to train object detectors on these images. The model we trained was the YOLOv8 nano version over 100 epochs. On this dataset, the model achieved a mean Average Precision (mAP) of 97.8%. The following Table 4.3.1 illustrates the parameters of this training.

Table 4.3.1: Parameters of training.

Dataset	
Image pre-processing	-
Number of images	4680
Learning	
OD model	Yolov8 Nano version
Hyper-parameters	Epochs: 100, Batch size: 16, Image size: 640, Optimizer: auto, Weight decay: 0.0005
Validation	
Data split	70 train set/20 validation set/10 test set
Validation metrics	mAP, precision, recall

The following Figure 4.3.3. illustrates the results of this training.

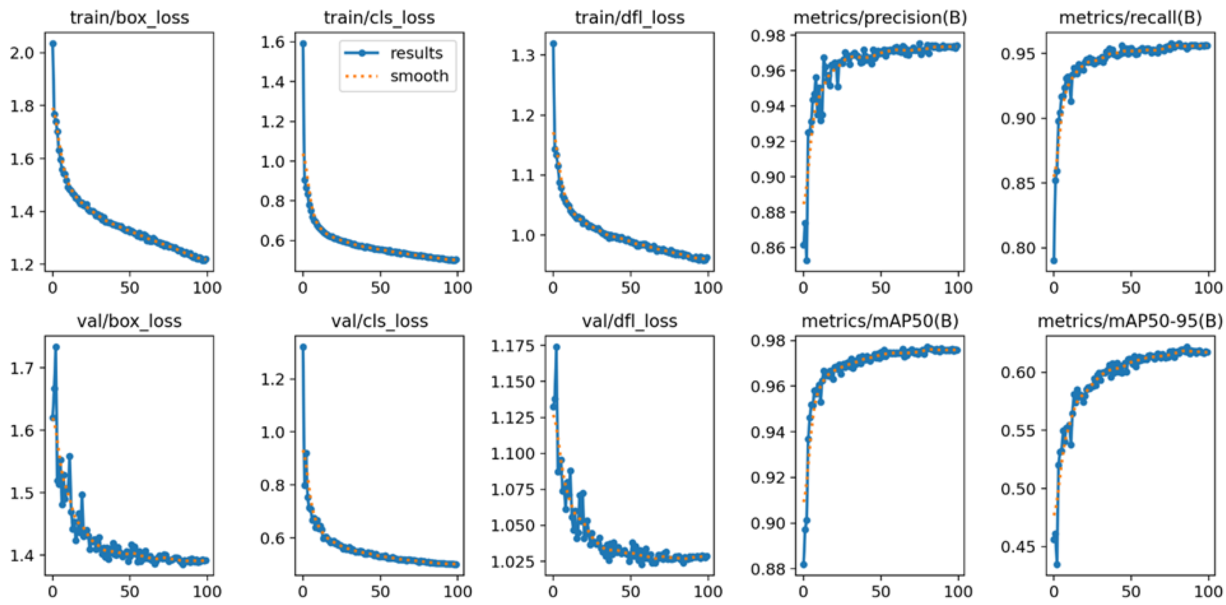


Figure 4.3.3: The training results are shown. The x-axis represents the epochs of training, and the y-axis represents the metrics referred to at the top of each graph.

Below are presented results (4.3.4 to 4.3.7) derived from images and video frames external to the database where the training was conducted. These results demonstrate the model's ability to generalize and accurately detect animals in varied environments and conditions, showcasing its effectiveness beyond the initial training dataset.

Figure 4.3.4: Vertical sheep detection from very high altitude.

Figure 4.3.5: Sheep detection from side point of view.

Figure 4.3.6: Vertical sheep detection from medium altitude

Figure 4.3.7: Vertical sheep detection from high altitude[27].

Track and count

The Figure 4.3.8. below demonstrates the system's effectiveness, emphasizing its precision in detecting, tracking, and counting sheep across various environmental conditions. Notably, the system is designed with a defined boundary line; when a sheep crosses this line, it is recognized and tallied as an object within the system's counting mechanism. This feature ensures highly accurate and reliable counting, even in complex and dynamic outdoor settings.

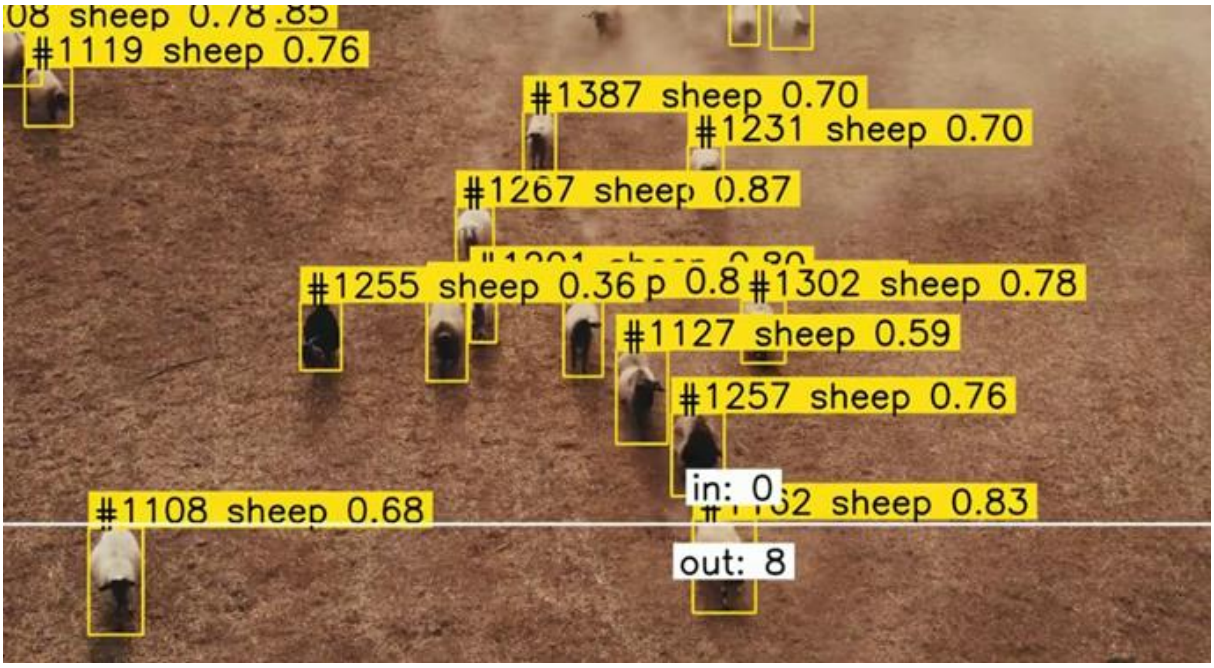


Figure 4.3.8: Animal tracking and counting.

- **Limitations of trained OD model**

During a subsequent pilot test, the algorithms trained on public datasets faced challenges when it came to detecting livestock. This issue arose primarily because of the complex background, which consisted of animal pastures, and the high altitude at which the dataset was captured. The unique combination of these factors proved to be a significant obstacle to accurate detection. Overcoming these issues and addressing the associated challenges are crucial steps in refining the algorithms for successful livestock detection in such environments. In Figure 4.3.9., we observe a compelling example of the terrain's complexity within the pilot use case territory. This illustration showcases four sheep, almost camouflaged against the backdrop of the rugged landscape. Their subtle presence highlights the inherent challenges in identifying animals in such intricate and demanding environments. This figure effectively demonstrates

[27] https://www.pexels.com/el-gr/video/8049412/

the difficulties faced in distinguishing wildlife amidst the natural contours and features of the territory.

Figure 4.3.9: Amidst the challenging terrain of the pilot use case territory, four sheep blend subtly into the background, illustrating the difficulty in discerning animals in such complex environments.

There are more examples explaining potential reasons for model failures, such as training data limitations, occlusion, image quality, and model complexity.

Figure 4.3.10: The Hidden Flock Challenge in Object Detection of Clustered Sheep.

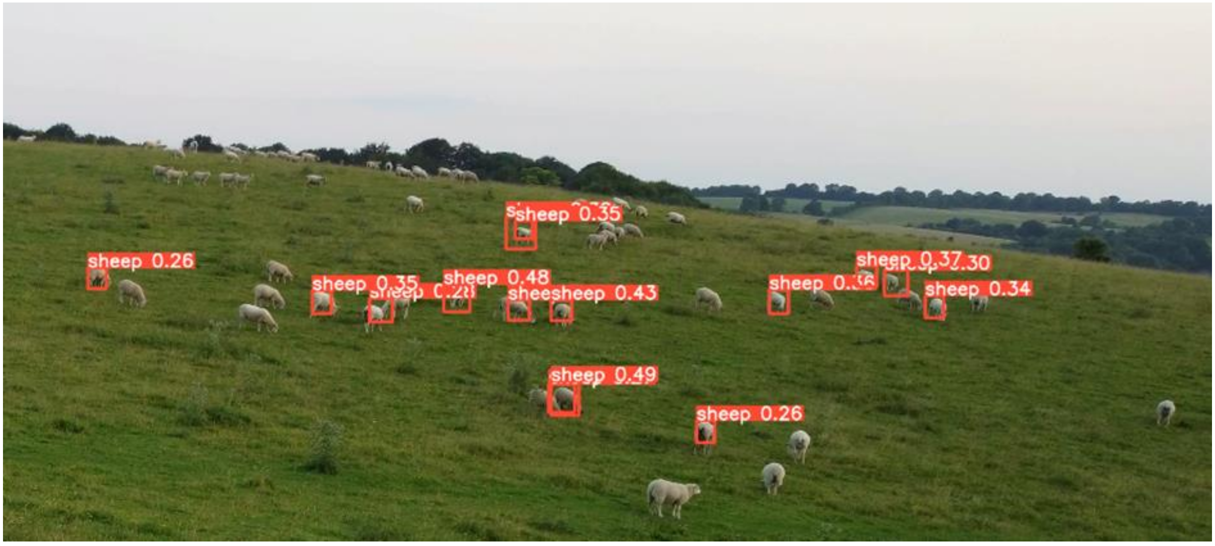


Figure 4.3.11: Partial Herd Recognition- When Object Detection Misses Sheep.

When an object detection model fails to accurately detect sheep, several factors could be at play, excluding the possibility of annotation errors:

- Training Data Limitations: The model's performance is highly dependent on the variety and quality of its training data. If the model wasn't trained on images that capture the wide range of appearances sheep can have, it may not recognize them in new images. For example, if the training data did not include sheep of various breeds, sizes, or sheep in different landscapes and lighting conditions, the model might struggle to generalize what it has learned to new situations.
- Occlusion: In this image, the sheep are clustered together, which can result in occlusion — when one object is partially or fully hidden behind another. Object detection models can have difficulty discerning the boundaries of individual sheep when they are tightly grouped. This is a common issue in object detection tasks where objects of interest tend to group together, like flocks of animals or herds.

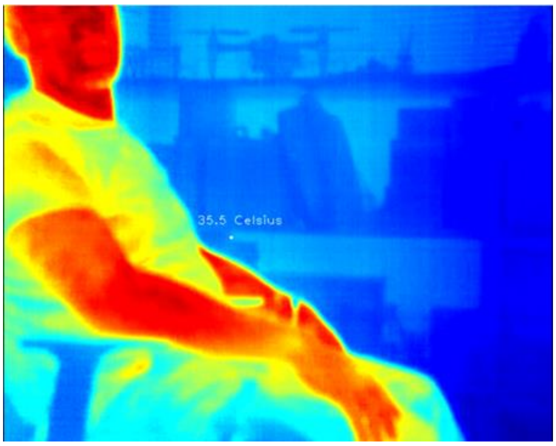
- Resolution and Quality: The effectiveness of an object detection model is also tied to the resolution and overall quality of the image it's analysing. Poor image quality can stem from a variety of factors, including but not limited to low resolution, motion blur, bad lighting, or obstructions. These factors can degrade the model's ability to detect and accurately place bounding boxes around the sheep.
- Model Complexity: The intricacy of the model itself can influence detection success. Basic models might not have the capability to handle the complexity of certain scenes, especially if those scenes contain numerous overlapping objects. Advanced models with deeper architectures and more sophisticated feature detection may be necessary to process such complex visual information effectively.

In terms of risks, failing to detect sheep accurately can have various implications. In agricultural contexts, it might lead to incorrect animal counts, which can affect resource allocation, such as food and space. In conservation efforts, inaccurate detection could misinform researchers about population sizes and the health of ecosystems. In automated surveillance systems, such as those designed to detect predators or track animal movements, misidentification could lead to a lack of timely responses to potential threats. Therefore, ensuring high accuracy in object detection models is crucial for applications where reliable identification is necessary for decision-making.

- **Integration of thermal information**

The aforementioned difficulties are a significant obstacle to accurate detection. To explore the feasibility of the proposed concept that combines RGB and thermal channels, we employed a test dataset to understand how objects could be isolated based on temperature. Initially, we conducted experiments using a human subject, albeit at a close range. The outcomes demonstrated the effectiveness of this approach in those conditions. However, it is crucial to note that these positive results do not guarantee success with our project data.

The primary challenge lies in the fact that the images for our project will be captured from a considerable height of 30 meters. This significant difference in perspective introduces various factors such as distance, angle, and resolution, which could impact the thermal detection process. As a result, further testing and adjustments are necessary to adapt the thermal camera methodology to our specific project requirements and conditions. We specifically applied preprocessing techniques to refine the raw data obtained from the thermal camera. By setting a temperature range between 36.5°C to 39°C, we successfully isolated the human subject. This methodology, initially designed for human detection, is intended to be adapted for the identification of animals as well.



a b

Figure 4.3.12: Testing of thermal cameras: a) raw flir data and b) processed image with a focus on a specific temperature range

In this section, we propose the utilization of thermal imaging technology as a supplementary tool to conventional cameras. Our objective is to evaluate the potential of thermal cameras in providing additional and valuable insights that extend beyond the capabilities of standard optical imaging devices. However, it is imperative to acknowledge existing physical constraints associated with this technology. A notable limitation is the requirement for close proximity to the subjects, particularly wildlife, for optimal thermal data acquisition. This proximity can induce stress or fear in the animals, potentially impacting their natural behaviour. Consequently, our study will rigorously assess the feasibility and effectiveness of integrating thermal imaging in our research methodology, weighing the benefits against the potential disturbances it may cause to the subjects under observation.

- **Four-channel OD**
  - Dataset: The approach described above yielded highly positive results, suggesting the potential for developing an advanced sheep detection model. However, a notable limitation was identified. This limitation stems from the goats being situated in fields, a background that does not align with the topography of our mountainous region. This inconsistency in the background landscape poses a challenge for the model's applicability to our specific terrain.

    Therefore, faced with challenges in data collection, we opted to request a pilot dataset tailored to our specific needs and the requirements of end users with human objects. This dataset was designed to include crucial parameters such as background conditions, altitude information, RGB imagery, thermal camera raw data, and most importantly, human objects. This dataset was instrumental in the development of our model, allowing us to enhance our detection algorithms, despite the challenging background conditions present in the real-world deployment scenario.

    Through the costume-made drones they are developing, the ACCELI partner has collected data sets to detect humans on difficult backgrounds and at varying heights of 30m and above.

– Findings: When testing our models for object detection using only RGB images, we encountered further challenges, primarily stemming from the complicated background and the altitude. The intricacies of the background made it difficult for our models to accurately identify objects, highlighting the complexity of the real-world scenarios our technology is expected to perform in.

Figure 4.3.13: Human detection in high altitude image.

In comparison to the RGB image, the thermal camera image in the current terrain exhibited a clear and unobstructed boundary. It had limited resolution as it relied on temperature variations. This lowest resolution was a result of the high ambient temperature prevalent in the environment.

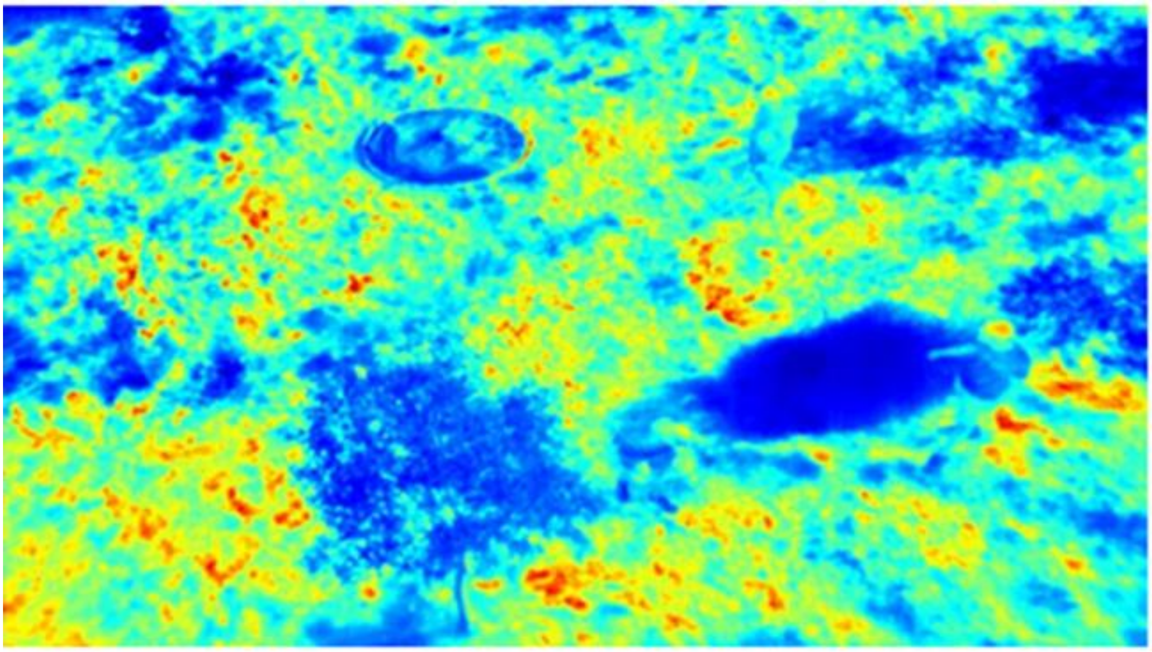
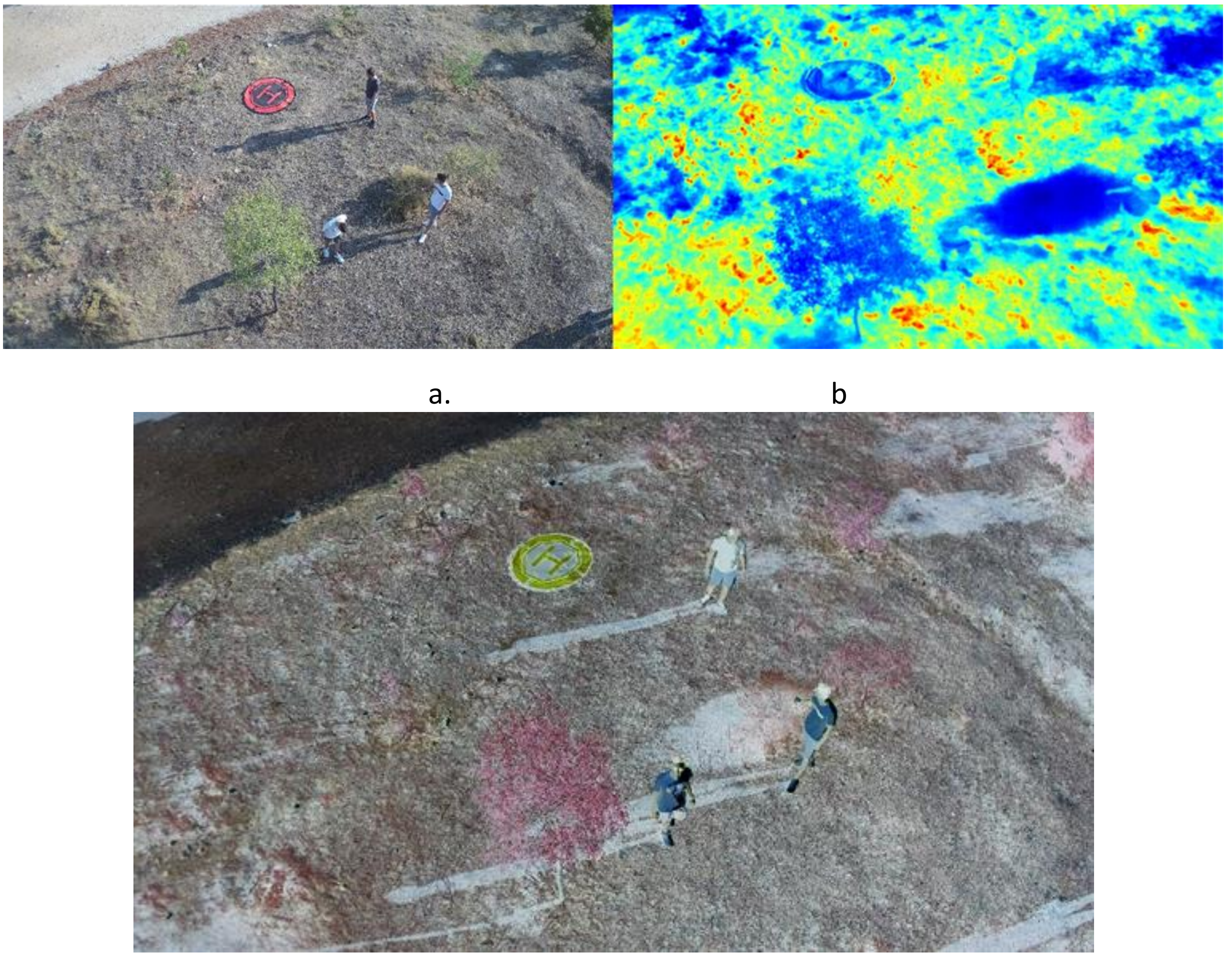


Figure 4.3.14: Frame of thermal camera in specific range of temperature (36-40 degrees Celcius)

Consequently, we decided to merge the data from both the RGB and thermal camera images. By unifying these two sources, we aimed to create a more comprehensive and enhanced input for our models. This approach was undertaken to improve the accuracy and reliability of our object detection algorithms, considering the distinct advantages offered by each type of imagery. Hence, the integration of thermal and RGB images led to a 4-channel image.

This process of consolidating thermal and RGB images into a 4-channel format aimed to provide a holistic input for our models. The resulting images encapsulated both temperature data and color information, enabling our algorithms to leverage the advantages of both modalities for enhanced object detection in challenging environments. In the methodology section, we provided a detailed description of the process used to consolidate the thermal and RGB images into a 4-channel image for the object detection task.



a. b

c

Figure 4.3.15: a) RGB image, b) Image from thermal camera and c) 4-channel image cmyk

In the upcoming phase of our research, we aim to explore the application of Object Detection algorithms on 4-channel data. This direction is motivated by the promising potential this approach has demonstrated in preliminary theoretical assessments. We anticipate that this exploration could lead to substantial enhancements in the accuracy and efficiency of our data analysis methods.

- **Performance of the proposed high altitude object detector**

We have implemented Waldo's V2 training model on the ACCELIGENCE dataset, which is adept at recognizing tiny objects in high-altitude aerial images. This advanced model is capable of detecting even small or difficult-to-see objects, such as a person walking in a forest or animals. The accompanying Figure 16 clearly demonstrates the model's ability to successfully identify humans within a forested area. For our next deliverable, we aim to further develop and apply this technique, tailoring our models to specialize in recognizing herds from high-altitude aerial photographs.

a

b

Figure 4.3.16: Human detection based on Waldo v2 model.

CONCLUSIONS AND FUTURE PLANS

In conclusion, our analysis of three cutting-edge approaches in object detection demonstrates their unique contributions and advancements in the field. For the RGB analysis, we employed the YOLOv8 algorithm, trained specifically on typical RGB data, showcasing its efficiency and effectiveness in standard environmental settings. In contrast, the Waldo V02 model, designed for high-altitude object recognition, brings a distinct advantage in identifying objects from significant heights. Lastly, our exploration with the 4-channel model involved training YOLOv8 on images that combine RGB data with an additional channel, which significantly enhances detection capabilities in complex scenarios. Each of these methodologies represents a significant stride in leveraging advanced machine learning techniques for more accurate and versatile object detection across various environments.

The image illustrates an integration setup where a drone equipped with a camera (RGB and thermal) is responsible for capturing video footage, which is subsequently processed on-board by the Jetson Xavier. The process involves analysing the position of a herd, monitoring health status, and detecting sudden events. The outputs from this processing are system alerts and videos, among other data points. It should also be noted that a dummy Docker version of this system has already been developed and tested, confirming the viability and functionality of the bundle for practical deployment.

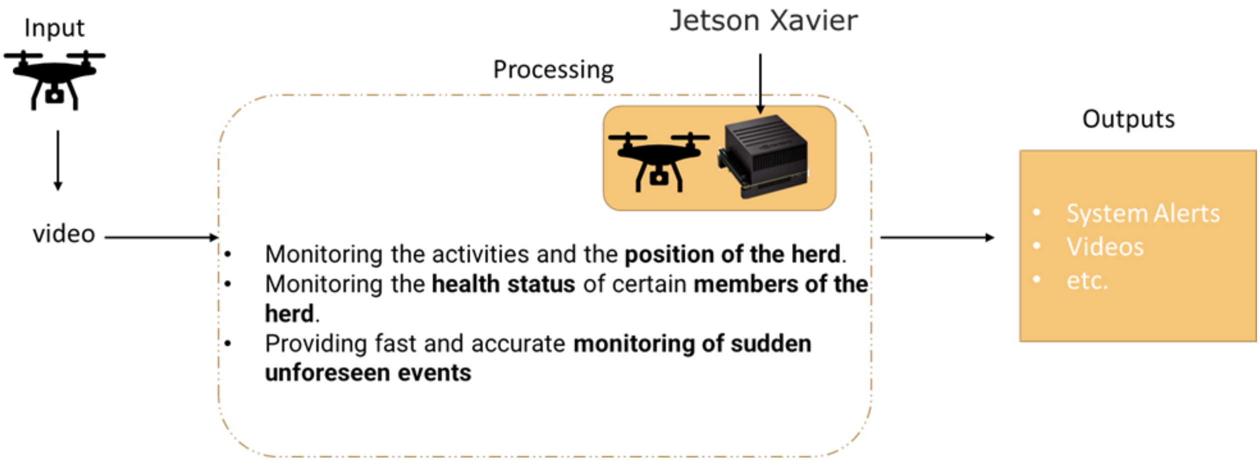


Figure 4.3.17: Architecture design

Our future plan involves collaborating with ACCELI to create a new dataset using the pilot herd. We aim to develop final models capable of recognizing the herd from high-altitude images and videos, as well as assessing their health status. To evaluate the health of the herd, we plan to employ thermal cameras to measure the temperature of each animal. This approach is twofold: firstly, to monitor if any animal deviates from the herd, and secondly, to use metrics from object detectors' boundary boxes to quantify the deviation of each animal's center from the herd's center.

5. BUNDLES, SERVICES FOR FORESTRY

The bundles for forestry are a service created to monitor and solve a specific problem in this sector. The Bundles are conceptualized according to Pilot use case's needs, and finally will be tested in Pilots environment. The CHAMELEON solution for forestry will be validated and demonstrated under two relevant pilot use cases: i) Spain (Avila) and ii) Austria (Vienna) (Table 5).

Table 5: List of bundles for forestry developed in the CHAMELEON project.

Pilot use case	Bundles (Business use case)
Spain	CONTINUITY OF VEGETATION
	CHARACTERIZATION OF WILDLAND-URBAN INTERFACE
	HOT SPOT IDENTIFICATION AT THE BEGINNING OF WILDFIRE
Austria	VEGETATION MONITORING AND CENSUS
	LARGE WOODY DEBRIS ON RIVERS
	HEALTH STATUS OF VEGETATION (MAINLY BARK BEETLE), GAME BROWSING, GROUND COVER, AND FUNGAL GROWTH

Different bundles consist of different action steps of the workflow, which was provided in the Deliverable 6.1. A detailed description of the development of each bundle for forestry is provided according to their action level: Relevance and problematic; UAV characteristics; Camera characteristics; Flight characteristics; Data processing; AI deep learning process/algorithms; Results/Outcomes.

5.1. CONTINUITY OF VEGETATION

RELEVANCE AND PROBLEMATIC

Forests are biodiversity hotspots that store vast amounts of carbon in their above- and below-ground biomass (i.e., carbon sequestration). On the other hand, the escalating frequency and intensity of wildfires are driven by climate change and environmental degradation. As a result, fire behaviour and simulation models are now indispensable for land managers and policymakers. These models predict fire potential, identify high-risk wildfire areas, and guide efficient resource allocation for fuel treatments and forest management.

Specifically, crown fires, which spread through the forest canopy, are especially perilous and challenging to manage. Evaluating crown fire potential is vital in wildfire control. The number of tree-level fuel connections, average fuel connections per tree, and clusters of connected tree fuels are crucial metrics in modifying fire behaviour and curtailing fire spread at a landscape level. Overlapping tree crowns facilitate rapid fire transfer from tree to tree, accelerating crown fire spread. Conversely, tree gaps can hinder fire progression, easing control and suppression efforts. Additionally, crown fire propagation is significantly influenced by weather conditions, affecting the tree spacing needed for fire transmission between adjacent trees.

SATELLITE CHARACTERISTICS

Satellite imagery encompasses a range of sensors, such as RGB, multispectral, hyperspectral, and thermal sensors, supplying vital data that complements images captured by drones for monitoring forestry areas. Satellites provide expansive coverage and regular collection intervals, offering a comprehensive perspective of vast landscapes that enhances the precise, targeted observations made by drones. Multispectral satellite imagery is particularly valuable for evaluating plant health and soil characteristics across large areas. Thermographic sensors not only identify temperature variations pinpointing water stress or irrigation needs but also play a crucial role in early fire detection. The integration of satellite and drone technologies allows for a more comprehensive understanding of extensive environmental landscapes with detailed local specificity.

To effectively manage the extensive geospatial information collected from both satellite and drone technologies, this package leverages the capabilities of Google Earth Engine. This platform excels in processing large datasets, enabling the analysis and interpretation of collected imagery with speed and accuracy. It provides a robust suite of computational tools that empower us to fully utilize the potential of Earth observation data.

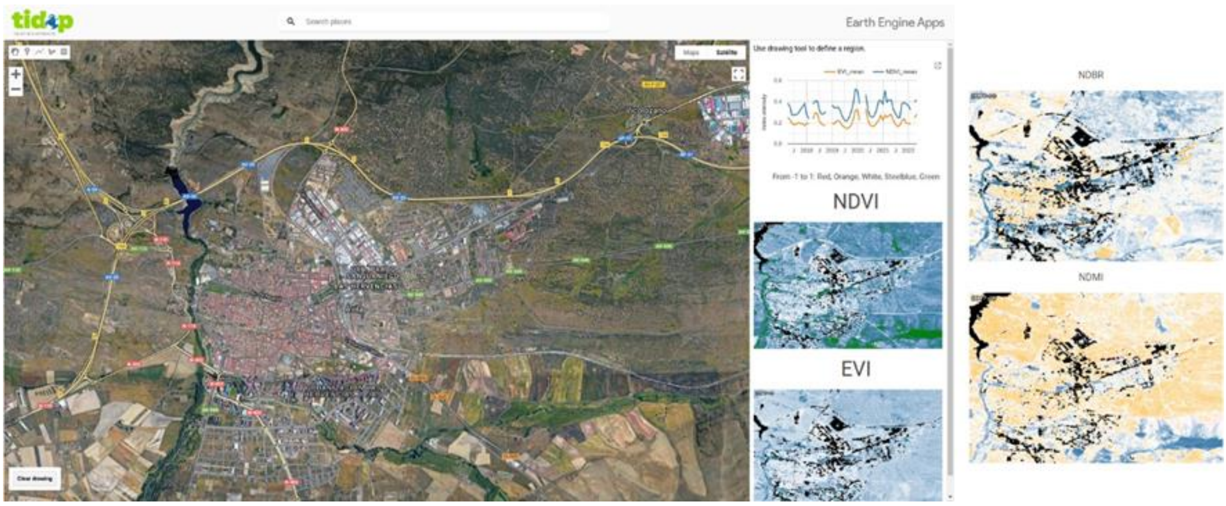


Figure 5.1.1: Viewer developed from images and vegetation indices of the Google Earth Engine platform.

A constellation of Earth observation satellites, including Sentinel-2, Landsat 8 & 9, and Planet satellites, is utilized to derive various indices crucial for monitoring vegetation and terrestrial surfaces. This wealth of information supports the extraction of key vegetation indices such as the Normalized Difference Vegetation Index (NDVI) and the Moisture Stress Index (MSI), vital for assessing plant health and water content in the landscape. A detailed classification of the specific roles and contributions of these satellites is presented in Table 5.1.1.

Table 5.1.1: Satellite mission specifications usable for the bundle “Continuity of vegetation”.

Mission	Constellation	Instruments	Spectral bands	Spatial resolution	Revisit time
Sentinel-2	Sentinel-2A & 2B	MultiSpectral Instrument (MSI)	RGB, NIR, SWIR	10-20 m	5 days
Landsat 8 & 9	Landsat 8 & 9	Operational Land Imager (OLI) & OLI-2	RGB, NIR, SWIR	30 m	8 days
Planet	PlanetScope	Dove	RGB	3-5 m	1 day

Ensuring adherence to quality standards, technical requirements play a fundamental role in shaping satellite data. Careful selection of spectral bands is imperative, aligning them with specific observational needs, such as those required for computing NDVI for plant health assessments. High spatial resolution is essential to distinguish individual landscape features, facilitating accurate cross-referencing with drone data. The revisit time is of critical importance, demanding frequent observations to capture rapid changes in environmental conditions. Cloud cover poses a significant challenge to satellite imagery quality by obscuring the Earth's surface. To counter this, modern satellites are equipped with advanced sensors that exhibit remarkable precision in cloud detection. These sensors can identify the presence of clouds at the pixel level, as exemplified by the Sentinel-2 satellite, which features a capability to estimate cloud probability for individual pixels. This functionality is pivotal for cloud masking, a process that eliminates cloud-covered portions from the image, thereby increasing the proportion of usable data obtained during each satellite pass. These technical specifications are indispensable, ensuring the precision, timeliness, and relevance of satellite data for this bundle, establishing it as a dependable source for environmental analysis and decision-making.

UAV CHARACTERISTICS

For the performance of this bundle, any UAV capable of carrying an RGB camera or a LiDAR as a payload is useful.

CAMERA CHARACTERISTICS

To obtain the required products, either of the following sensors included in the CHAMELEON project may be used:

- RGB Camera: Sony A6000.
- Phoenix RECON-XT

FLIGHT CHARACTERISTICS

Flight requirements vary based on the sensor type used for point cloud acquisition.

For point cloud reconstruction using an RGB camera through photogrammetry, the flight plan should include:

A. Camera orientation: Vertical or oblique flights.

B. GSD: 2 cm/px (minimal requirement). 1 cm/px (optimal requirement).
C. Overlap: 80-85% forward, 40-60% side.

In contrast, when using a LiDAR sensor, the flight plan should include:

A. Point cloud resolution: 80 points/m^2 (minimal requirement). 100 points/m^2 (optimal requirement).

DATA PROCESSING

Figure 5.1.2 displays the general flow diagram of the bundle. Satellite and UAV data are precisely geo-referenced, ready to be fused. Figure 5.1.3 presents the input, parameters, output, and display results table of the bundle.

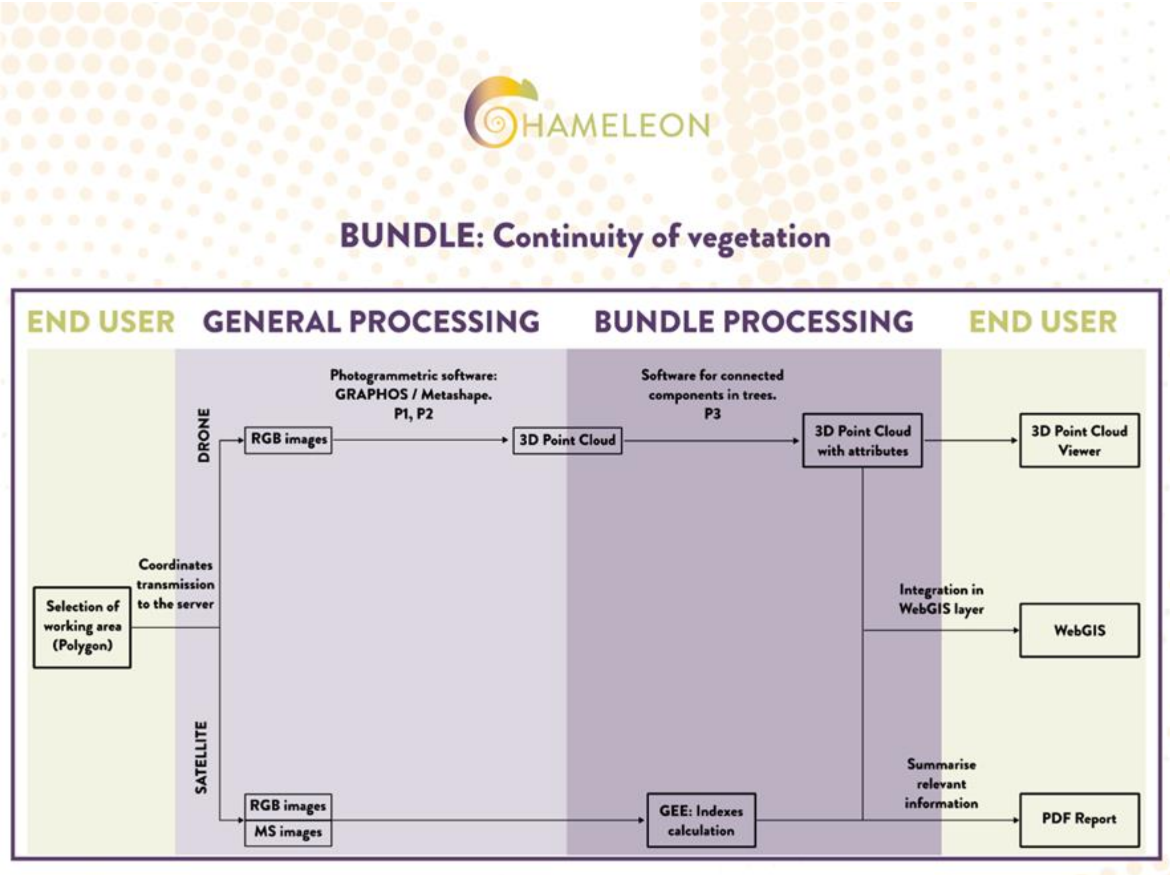


Figure 5.1.2: Workflow of the bundle “Continuity of vegetation”.

CHAMELEON

BUNDLE: Continuity of vegetation

INPUT		PARAMETERS		OUTPUT		DISPLAY RESULTS
End User	Sensors	Graphos / Metashape	Software for connected components in trees	Software for connected components in trees	GEE	R1: PDF Report R2: Representation in custom 3D Point Cloud Viewer R3: WebGIS visualisation with layers
Selection of working area (polygon)	Drone: S1: RGB images -> 3D Point Cloud Satellite: S2: RGB images S3: MS images	P1: Known distances P2: Ground Control Points	P3: Octree	O1: 3D Point Cloud with attributes O2: CSV: Connected Comp. ID Area Volume	O3: Indexes: NDVI SAVI EVI GNDVI NBR NDMI	

Figure 5.1.3: Input, parameters, output and display results of the bundle "Continuity of vegetation".

AI DEEP LEARNING PROCESS / ALGORITHMS

In this bundle, we utilize an artificial intelligence method to evaluate crown fire potential using UAS-Based Structure-from-Motion Photogrammetry, focusing on tree crown connectivity. This method examines the layout and spacing of tree crowns and fuel load distribution to gauge the likelihood of crown fire ignition and spread. We employ the Connected Components algorithm to identify connected point clusters in forested point clouds. The algorithm begins at a vertex, exploring all connected vertices, and repeats this for unvisited vertices until all are covered, using the octree level as a boundary. The octree level, a measure of detail in 3D space division, divides space into octants at each level. Higher octree levels mean larger regions with less detail, while lower levels yield smaller, more detailed regions. Octree levels are vital for efficient organization, storage, and analysis of point cloud data.

Figure 5.1.4: Conceptual workflow of the bundle: canopy connectivity related to the direction of fire propagation and canopy density (the fuel quantity is indicated by the varying darkness of rectangles).

Figure 5.1.4 illustrates the conceptual aim of the bundle, focusing on canopy connectivity in relation to the direction of fire propagation and canopy density. It depicts the impact of connected stand mass on fire behaviour, specifically illustrating the ease of crossing vegetation gaps as shown by canopy connectivity. Additionally, it highlights the fuel quantity in each cluster, indicated by the varying darkness of rectangles at the bottom, which represent clusters formed by connected canopies.

RESULTS / OUTCOMES

The Connected Component algorithm is applied at four octree levels, with the outcomes depicted in Figure 5.1.5, showcasing tree connectivity clusters in various random colours.

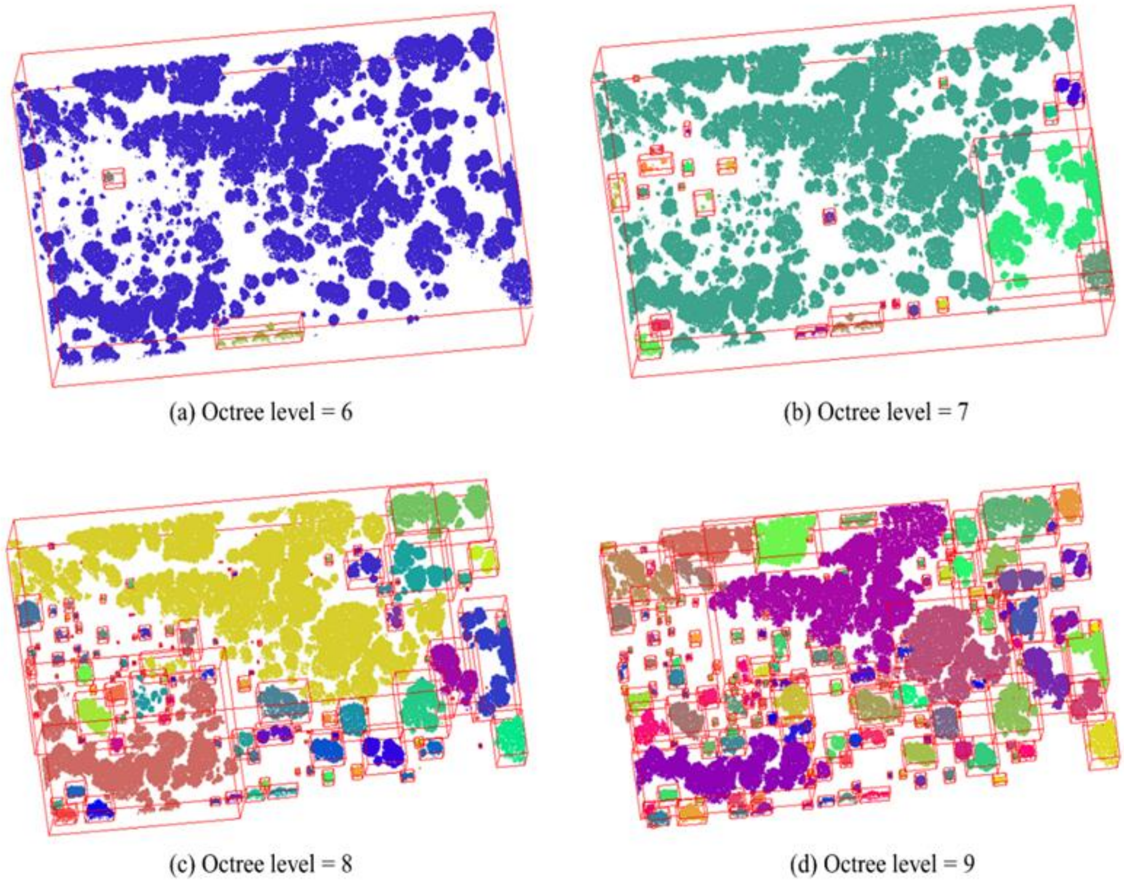


Figure 5.1.5: Forested point cloud generated by UAS-based Structure-from-Motion photogrammetry with the connectivity approach at four octree levels with random colours per cluster (a, b, c, d).

Moreover, crown fire propagation results are significantly impacted by weather conditions, which dictate the tree spacing ranges necessary for fire propagation between adjacent trees.

Integrating advanced physics-based fire behaviour analysis with precise vegetation mapping technologies enables us to investigate individual tree-level fuel characteristics, enhancing our grasp of crown fire risks. We also evaluated trees likely to aid fire spread to canopy fuels. Nonetheless, considering weather parameters and conditions is crucial for accurate analysis. In drier, hotter, and windier areas, these methods risk underestimating the distance fire can spread through adjacent crowns.

5.2. CHARACTERIZATION OF WILDLAND-URBAN INTERFACE

The wildland-urban interface is the perimeter area around urban centers affected by the presence of forest masses. These areas present a high fire risk due to an increase in the density of human activities, which with the presence of abundant vegetation can lead to ignition and/or spread a fire faster. Improving access for firefighters in forested areas involves a multifaceted approach aimed at enhancing their ability to respond swiftly and effectively to wildfires.

RELEVANCE AND PROBLEMATIC

The new climate scenario, together with the changes in land use, has changed the pattern of wildfires, which have become more severe and more extensive. In this context, the wildland-urban interface takes on great importance as it implies a greater risk to human lives as well as to the socioeconomic infrastructure.

The characteristics and spread of a wildfire depend on meteorological, orographic and fuel factors (flammability, quantity and disposition). Fuel management is the main variable that we can modify in this fire triangle.

For all these reasons, prevention and planning in these wildland-urban areas are essential to reduce the intensity and speed of fires, protecting urban areas and creating safe areas for the action of firefighting teams.

Establishing a well-connected network of roads and trails is also crucial. These pathways should be wide enough for fire trucks and emergency vehicles to navigate easily, allowing firefighters quick access to different parts of the forest. Additionally, strategically located access points throughout the forest help reduce travel distances for firefighters, enabling faster response times during emergencies.

This package will be responsible for identifying the wildland-urban interface zones, as well as the accesses available for extinguishing works and the evaluation of the biomass in these zones. The characterization of the vegetation fuel in these areas provides managers with valuable information, allowing them to detect the areas of greatest fire risk and to optimize logistical and economic prevention actions.

SATELLITE CHARACTERISTICS

The images used for vegetation monitoring are of satellite origin and are obtained from the Sentinel-2 constellation, by the European Space Agency (ESA) as part of the elaboration of land use and land cover.

UAV CHARACTERISTICS

For the performance of this bundle, any UAV capable of carrying an RGB camera as a payload is useful.

CAMERA CHARACTERISTICS

To obtain the required products, the following sensors included in the CHAMELEON project are used:

- RGB camera: Sony A6000
- Multispectral camera: Micasense RedEdge MX.

FLIGHT CHARACTERISTICS

To reconstruct the point cloud, a typical photogrammetric flight plan should be implemented, which includes:

- Camera orientation: Vertical or oblique flights
- GSD: 2cm/px (minimal requirement). 1cm/px (optimal requirement)
- Overlap: 80-85% forward, 40-60% side

For the multispectral data, the flight plan should include:

- Camera orientation: Vertical or oblique flights
- GSD: centimetric
- Overlap: 60-70% forward, 20-40% side.

DATA PROCESSING

In this bundle, satellites are used as a coarse approach, while drones provide fine adjustments. The planned analysis of the WUI is 25-50 meters, which translates to only a few satellite pixels. Satellites can offer historical imagery of the WUI's evolution, but drones are more suitable for detailed observation of the WUI, including vegetation type, height, detailed topography, and selecting the most suitable machinery for fuel reduction.

Figure 5.2.1 displays the general flow diagram of the bundle. Figure 5.2.2 presents the input, parameters, output, and display results table of the bundle.

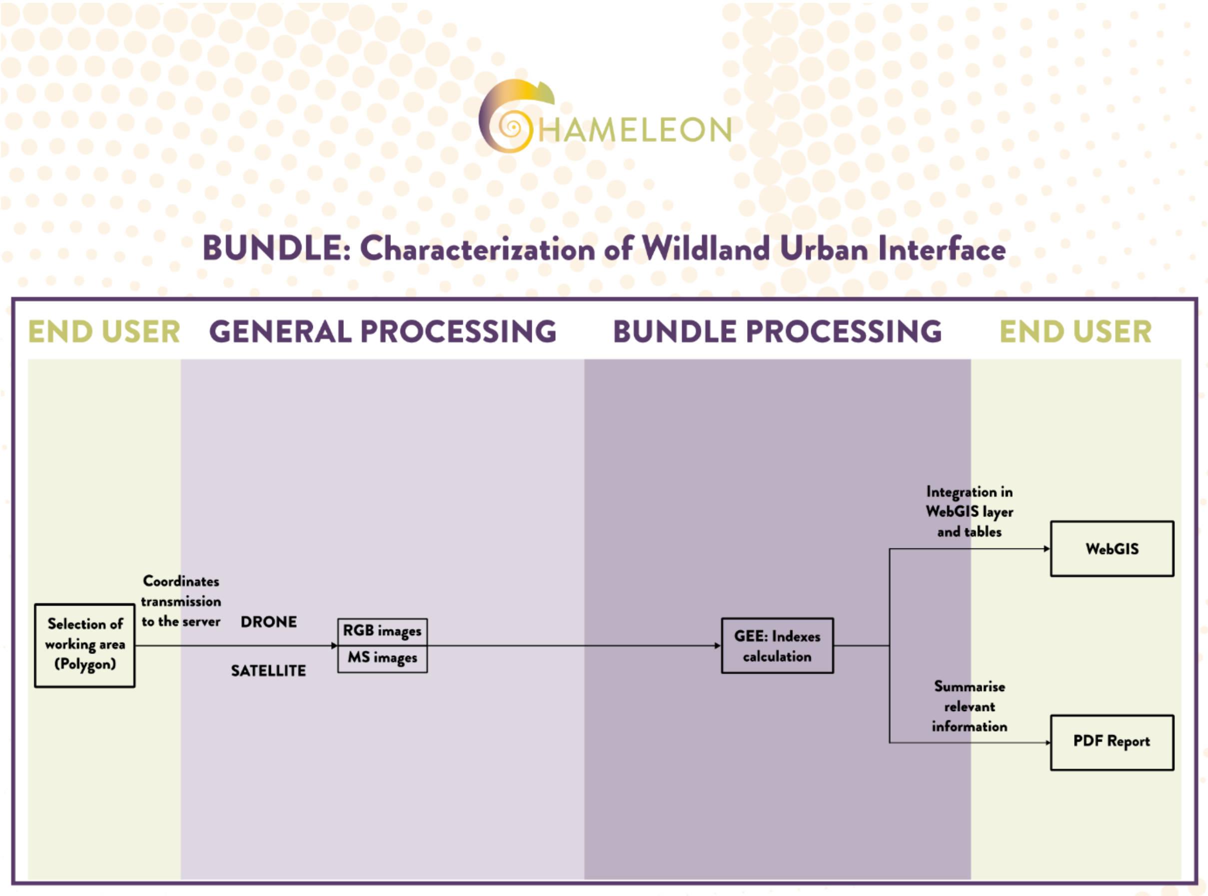


Figure 5.2.1: Workflow of the bundle "Characterization of Wildland Urban Interface".

BUNDLE: Characterization of Wildland Urban Interface

INPUT		PARAMETERS	OUTPUT	DISPLAY RESULTS
End User	Sensors		GEE	
Selection of working area (polygon)	Drone: S1: RGB images S2: MS images Satellite: S3: RGB images S4: MS images	P1: (Optional) Initial and final dates of analysis	O1: Indexes: NDVI SAVI EVI GNDVI NDMI O2: CSV with temporal evolution of indexes	R1: PDF Report R2: WebGIS visualisation with layers

Figure 5.2.2: Input, parameters, output and display results of the bundle "Characterization of Wildland Urban Interface".

AI DEEP LEARNING PROCESS / ALGORITHMS

For the delimitation of the Wildland-Urban Interface, the spatial information on land use developed by ESA's Copernicus program, CORINE Land Cover 2018[28], was used. It is the "V2020_20u1" version, whose spatial resolution is 25ha/100m and comes from the processing of satellite images of the Sentinel-1 and Sentinel-2 constellations taken in 2018.

The delimitation of the WUI involves the construction of perimeter strips around the polygons segmented in the territory as forest and urban. These buffers, whose width is established according to the actions that are intended to be undertaken in these areas, are intersected with each other to obtain the interface.

To identify the areas at greatest risk, a criterion based on the value of the Normalized Difference Vegetation Index (NDVI) has been applied within the WUI. This index indicates the presence of vegetation based on the photosynthetic activity of plants. Thus, areas with a higher

[28.] CORINE Land Cover 2018 (vector), Europe, 6-yearly - version 2020_20u1, May 2020 https://doi.org/10.2909/71c95a07-e296-44fc-b22b-415f42acfdf0

NDVI value will indicate a greater amount of vegetation fuel and therefore will require more immediate or costly management to avoid the fire risk involved.

Also, Normalized Difference Moisture Index (NDMI) is an important metric used in remote sensing to evaluate vegetation vigor by measuring the water content in plants. It is calculated using near-infrared (NIR) and short-wave infrared (SWIR) bands from satellite or aerial imagery. Higher NDMI values typically suggest healthier vegetation due to increased water content, indicating vigorous plant growth. By monitoring changes in NDMI values over time, alterations in vegetation health can be tracked. Sudden declines in NDMI might signal issues such as drought impact, disease outbreaks, or other environmental stressors affecting plant vitality.

The whole delimitation process has been executed using mainly the GDAL library in code programmed in Python language, while access to satellite information has been handled in the background with the Google Earth Engine (GEE) platform.

RESULTS / OUTCOMES

The identification of the wildland-urban interface zones was developed using elaborated land use mapping (Figure 5.2.3).

Figure 5.2.3: Cartography delimiting the wildland-urban interface zones with 25m of urban buffer and 400m of proximity to forest territory.

The presence of vegetation in these areas was identified by analysing satellite images, as shown in Figure 5.2.4.

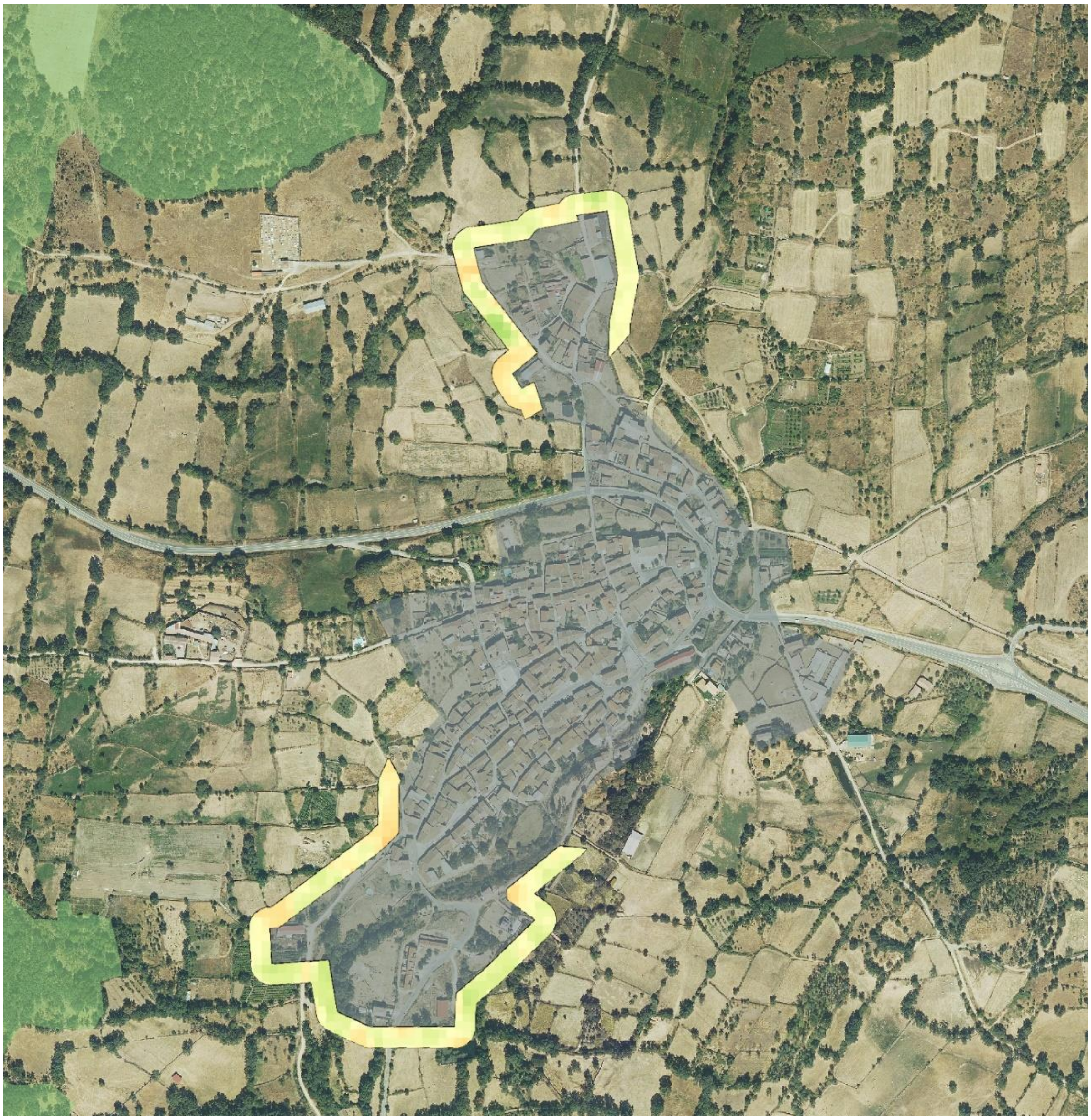


Figure 5.2.4: Distribution of living vegetation by NDVI. Green colors indicate higher photosynthetic activity, while red colors indicate no photosynthetic activity.

5.3. HOT SPOT IDENTIFICATION AT THE BEGINNING OF WILDFIRE

The objective for hotspot identification at the onset of a wildfire is to swiftly detect and pinpoint areas of intense heat or potential ignition within a wildfire-prone region. This early detection is crucial for immediate response and intervention to contain or mitigate the wildfire's spread.

RELEVANCE AND PROBLEMATIC

Early hotspot identification in wildfires is vital for immediate intervention, efficient resource use, environmental protection, public safety, and adaptation to climate change-induced wildfire risks.

Satellites have visualization and hotspot tracking services that are ideal for monitoring fires, burned areas, and detecting fire – risk activities, among others.

Therefore, the capability to detect hotspots has been developed at various scales:

A. Satellite-scale, improving existing services at European level.
B. Drones-scale, allowing precise monitoring of hotspots at night and within extinguished areas.

In terms of resolution, it should be noted that while satellites allow daily monitoring (i.e., images every 24 hours) and with a resolution of several tenths of meters, drones allow real-time monitoring with decimeter accuracy. Therefore, the combination of both data sources and scales is very useful for this objective.

In addition, existing services do not provide information about the burnt area during the occurrence of the fire, in such a way that planning the displacement of firefighters to the area cannot attend the criteria of the different severity levels in the occurring fire. The availability of priori information of the extension of the fire can allow for a more adequate planning of the firefighter operations from the beginning.

SATELLITE CHARACTERISTICS

Satellite products for hotspot detection are differentiated between level 2 products already generated by the space agencies and level 1 products for post-processing and future hotspot generation. These products are listed in the following tables.

Table 5.3.1: Level 2 satellite products. Hotspots of MODIS and VIIRS sensors.

Satellite	Instrument	Temporal resolution	Product	Spatial resolution (m)	Swath width (km)
Terra/ Aqua	MODIS	4 images daily	MCD14DL	1000	2330
SUOMI NPP/ NOAA-20	VIIRS	4 images daily	VNP14IMGTDLNRT	750	3000

Table 5.3.2: Level 1 satellite products. Bands, from Landsat 8/9 and Sentinel 2A/2B /3 satellites, required for hotspot detection.

Satellite	Instrument	Temporal resolution	Spatial resolution (m)	Band	Center Wavelength (µm)
Landsat 8/9	OLI	8 days	30	B5 NIR	0.865
Landsat 8/9	OLI	8 days	30	B7 SWIR	2.2

Satellite	Instrument	Temporal resolution	Spatial resolution (m)	Band	Center Wavelength (µm)
Landsat 8/9	TIRS	8 days	30	B10 Thermal Infrared	10.895
Sentinel 2A/2B	MSI	5 days	20	B8A NIR narrow	0.8647 – 0.864
Sentinel 2A/2B	MSI	5 days	20	B12 SWIR	2.202 – 2.186
Sentinel 3	SLSTR	27 days	1000	S9	10.854

UAV CHARACTERISTICS

For the performance of this bundle, any UAV capable of carrying a thermographic and RGB camera is useful. It would be ideal that the same flight is performed with the thermographic camera together with a multispectral camera, so that the acquisition of all types of data needed at the same moment is ensured.

CAMERA CHARACTERISTICS

To obtain the required products, the following sensors included in the CHAMELEON project are used:

- Thermographic and RGB camera: FLIR Duo Pro R

FLIGHT CHARACTERISTICS

The flight plan should include:

- Camera orientation: Vertical flights.
- GSD: Decimetric.
- Overlap: Less critical, although a minimum is required.

DATA PROCESSING

Figure 5.3.1 displays the general flow diagram of the bundle. Figure 5.3.2 presents the input, parameters, output, and display results table of the bundle.

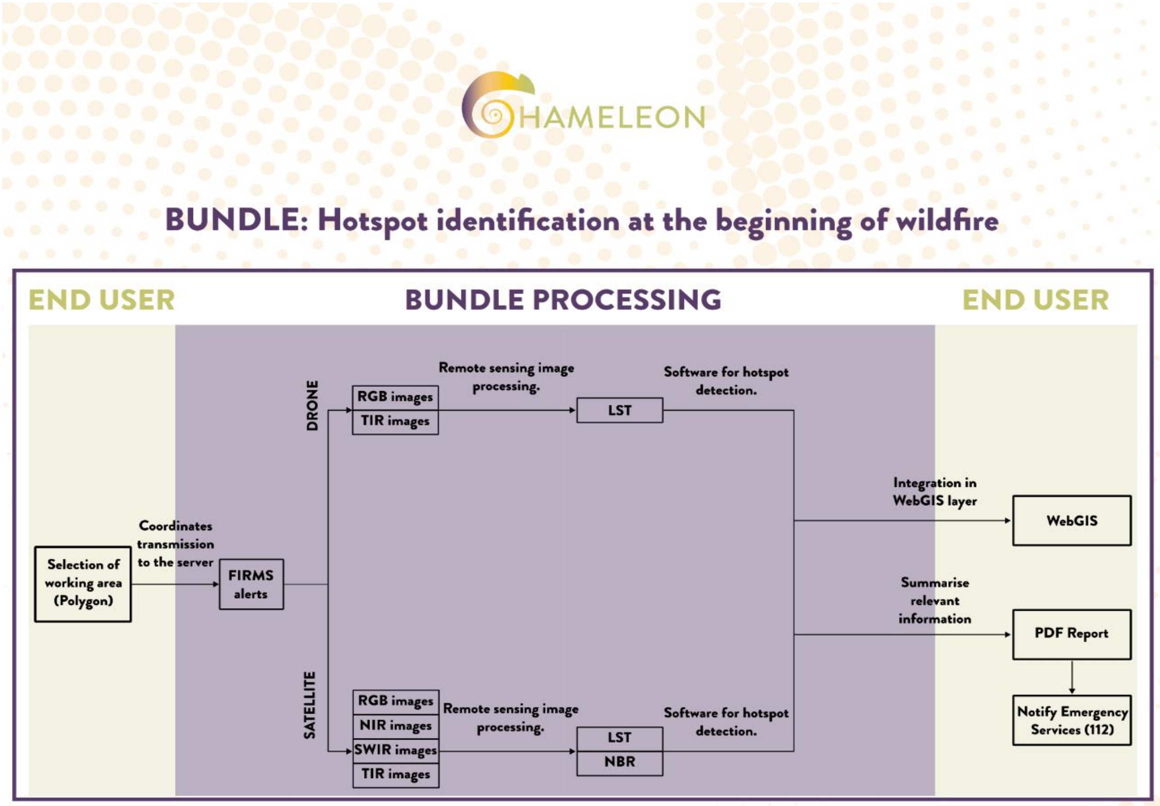


Figure 5.3.1: Workflow of the bundle “Hotspot identification at the beginning of wildfire”.

CHAMELEON

BUNDLE: Hotspot identification at the beginning of wildfire

INPUT		PARAMETERS	OUTPUT	DISPLAY RESULTS
End User	Sensors		Software for hotspot detection	
Selection of working area (polygon)	Drone: S1: RGB images S2: TIR images Satellite: S3: RGB images S4: NIR images S5: SWIR images S6: TIR images	P1: Notify European emergency services (112): Yes/No P2: Load DTM with shading to ease access tasks: Yes/No	O1: SHP: Polygon with burnt area	R1: PDF Report R2: WebGIS visualisation with layers

Figure 5.3.2: Input, parameters, output and display results of the bundle “Hotspot identification at the beginning of wildfire”.

AI DEEP LEARNING PROCESS / ALGORITHMS

The algorithm for hotspot-detection is divided in 3 main processes:

A. Visualisation and download of EFFIS/FIRMS products.
B. Detection of new hotspots with data from higher resolution satellites or drones.

C. Improvement of the spatial resolution of products used by EFFIS/FIRMS for Hotspot detection.

The module developed for the Hotspots detection uses different techniques and programming languages (R, C++, Phyton, GDAL) considering that a coarse-to-fine surface analysis is provided taking advantage of the satellite and drones images, in addition to the interconnection with EFFIS services. The first tests have been carried out in the local environment, aiming at the complete development of the tool prior to its integration in the WebGis environment.

1/ Visualisation and download of EFFIS/FIRMS products:

The first part of this tool focuses on the use of Hotspots from MODIS and VIIRS sensors. Google Earth Engine (GEE) Apps are developed for the download and visualisation of MODIS and VIIRS Hotspots towards their integration in the tool. Figure 3 illustrates the visualisation of the downloaded product in vector format (shapefile) of these Hotspots in a desktop GIS.

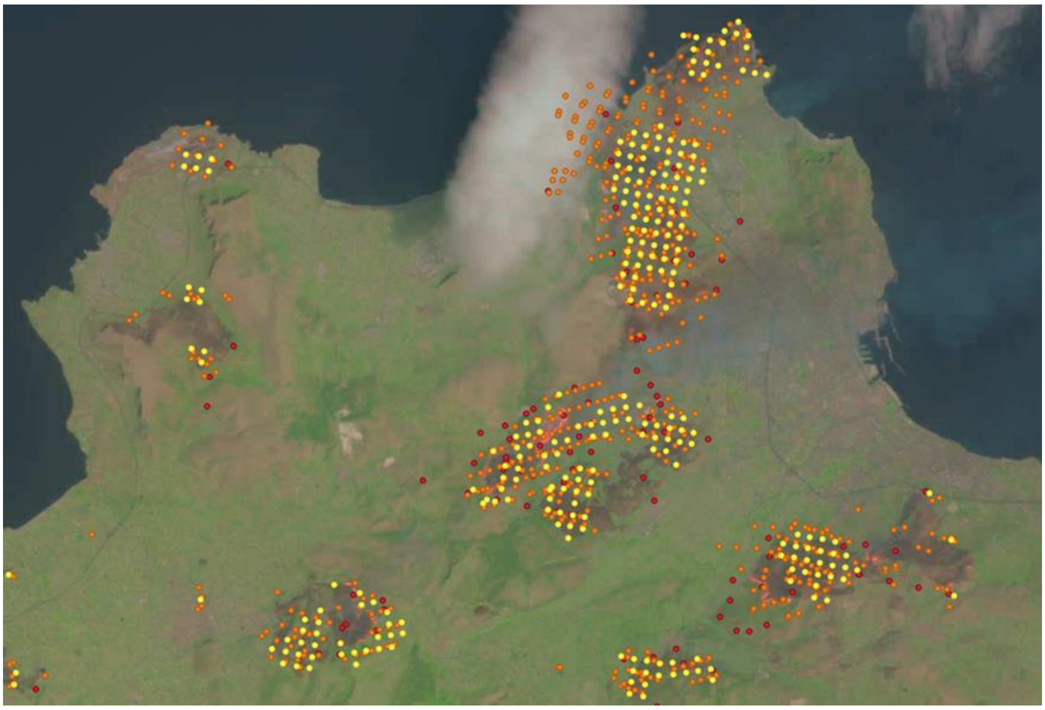


Figure 5.3.3: Shapefile of the Hotspots corresponding to the wildfires in the region of Palermo (Sicily) on 25 July 2023. The Hotspots correspond to the MODIS (in red) and VIIRS sensors (in orange on board the JPSS-1 satellite and in yellow on board the SUOMI NPP satellite).

This application allows the visualization of Hotspots of the last 24 hours as well as to perform a historical search. Finally, it is possible to visualize these points according to their Brightness Temperature (BT) intensity (Figure 5.3.4) and their image acquisition time, thus facilitating the extinction operations.

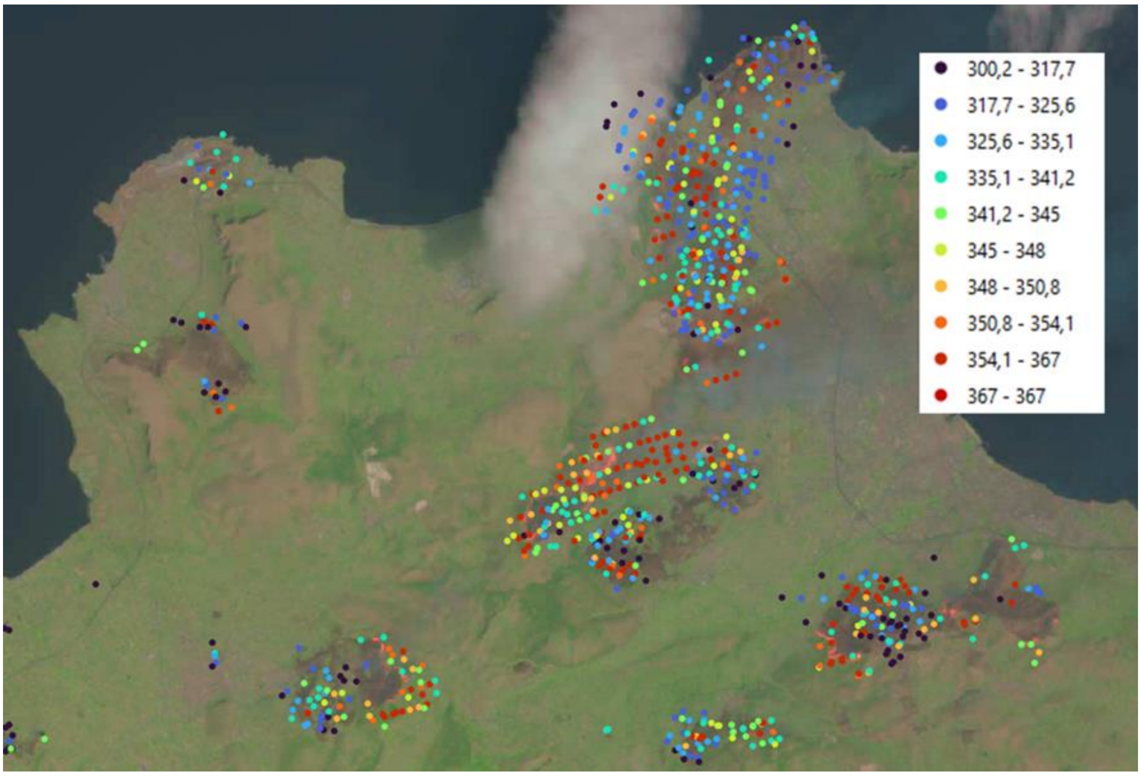


Figure 5.3.4: Shapefile of the Hotspots in the previous figure but categorized by brightness (related to fire intensity) and graded by color.

2/ Detection of new hotspots with data from higher resolution satellites or drones:

To complement existing Hotspot products, a new detection algorithm has been developed in which different levels of layers are considered.

Both the Copernicus Land Monitoring Service (CLMS) and the United States Geological Survey (USGS) provide geographical information on the Earth's surface energy variables. Satellite data from the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA) provide observations from missions, such as Sentinel and Landsat respectively, with higher spatial resolution than the sensors already used for Hotspot detection. The use of these data can provide a new approach because, despite having a lower temporal resolution, they provide valuable information on what is involved in improving spatial resolution. This can also be achieved with the use of drones in situ.

The first active fire detection algorithms have been developed for daytime detection. These algorithms are based on the sensitivity of certain spectral bands in the presence of fire[29]:

- Thermal: The thermal band is used to detect thermal infrared radiation emitted by fires. This band allows Hotspots to be identified on the basis of their temperature, as fires tend to generate extremely high temperatures.

[29] Payra, S., Sharma, A., & Verma, S. (2023). Application of remote sensing to study forest fires. In *Atmospheric Remote Sensing* (pp. 239-260). Elsevier.

- SWIR: This band can partially penetrate through smoke and clouds, which allows imaging of the land surface even under partial coverage conditions, as well as determine the area of ignition within the area affected by the resulting smoke. Another feature of the SWIR band is that its radiometric signature in active fires produces a radiance or reflectance anomaly compared to the background, thus mimicking the concept of thermal anomaly detection using mid-temperature to thermal infrared channels. This is why the SWIR band is used to calculate Normalised Burn Rate (NBR) composites and identify burnt areas on the map.
- NIR: During daytime orbits, the emissive component of the fire mixes with the background, which is dominated by the reflected solar component. To separate them, data from the NIR channel are used, which mostly do not respond to fire-affected pixels, although they are highly correlated with data from the SWIR channel on fire-free surfaces.

The tool developed is based on the algorithm presented in[30] using Landsat 8 and 9 sensors (Figure 5.3.5), with adaptations for its use with different types of sensors and adding thermal information.

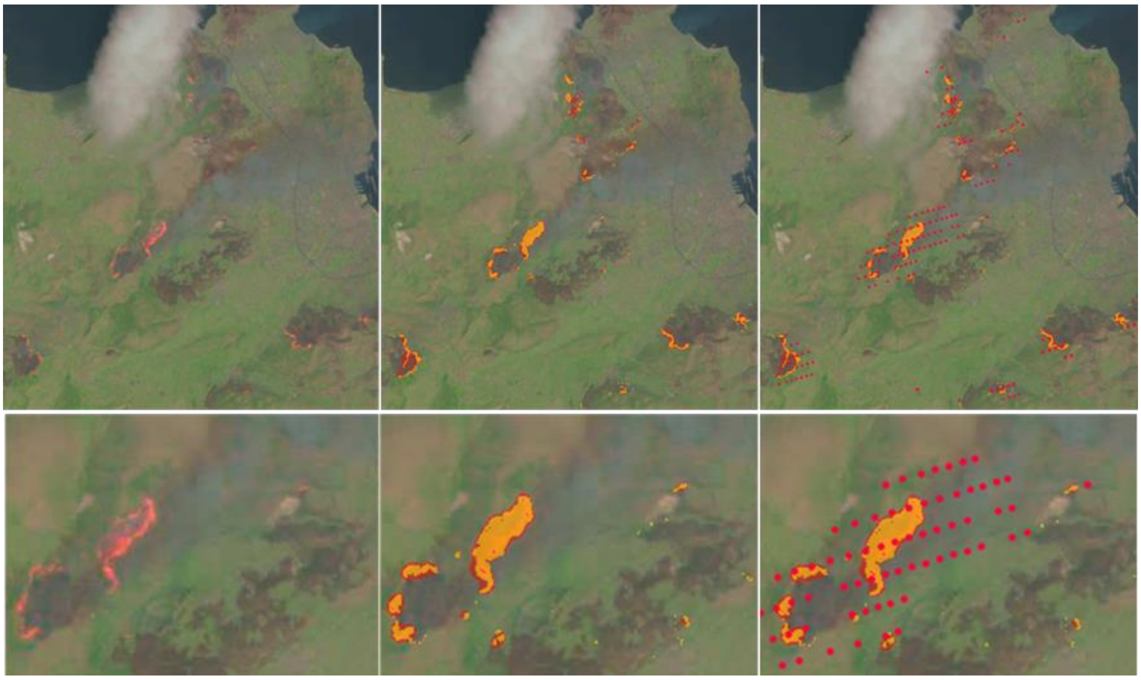


Figure 5.3.5: Landsat 8 RGB image of the Palermo region (Sicily) on 25 July 2023 (left figure) on which Hotspots have been calculated using Landsat 8 satellite data (middle figure). These results have been compared with MODIS and VIIRS Hotspots of the same date (right figure). Top row shows the general area, the bottom row shows a detailed area.

[30] Schroeder, W., Oliva, P., Giglio, L., Quayle, B., Lorenz, E., & Morelli, F. (2016). Active fire detection using Landsat-8/OLI data. *Remote sensing of environment*, *185*, 210-220.

One of the advantages of these new Hotspots lies on the increased accuracy of the geolocation. This has been proven in the detection of Hotspots in Blast Furnaces (Steel Mills) and Rotary Kilns (Cement Factories), where it was demonstrated that the geolocation of these new Hotspots was much more accurate than the Hotspots obtained from FIRMS/EFFIS (Figure 5.3.6).

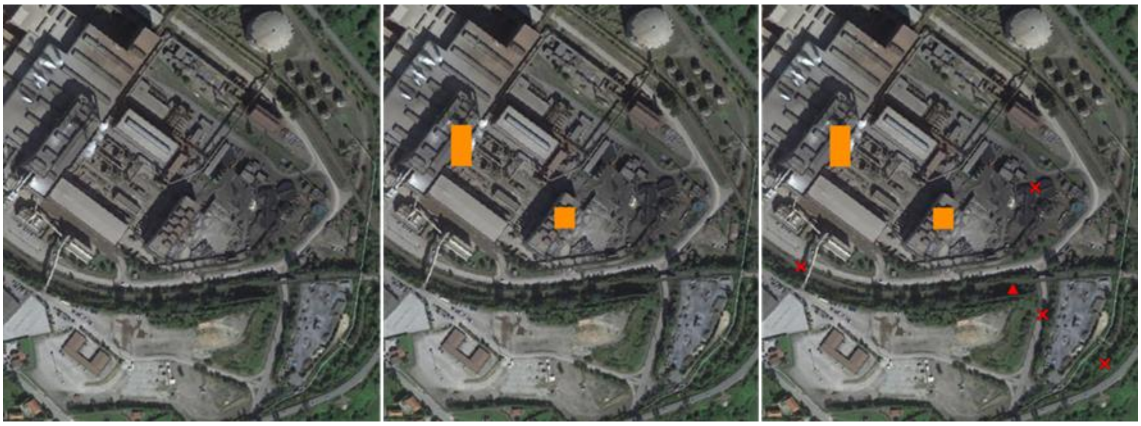


Figure 5.3.6: Satellite image of a steel mill in Avilés-Spain (left figure) in which Hotspots have been calculated using Landsat 8 satellite data (middle figure) and which have been found to be more accurate than MODIS and VIIRS Hotspots (right figure).

Another advantage of these hotspot detection tools developed is that the output file is presented in polygonal rather than point form, so that it is possible to visualize the whole fire front and not just a single point. As in Figure 4.3.4 with the MODIS and VIIRS hotspots, these new hotspot surfaces can be displayed with gradient to show areas of higher or lower fire severity (Figure 5.3.7).

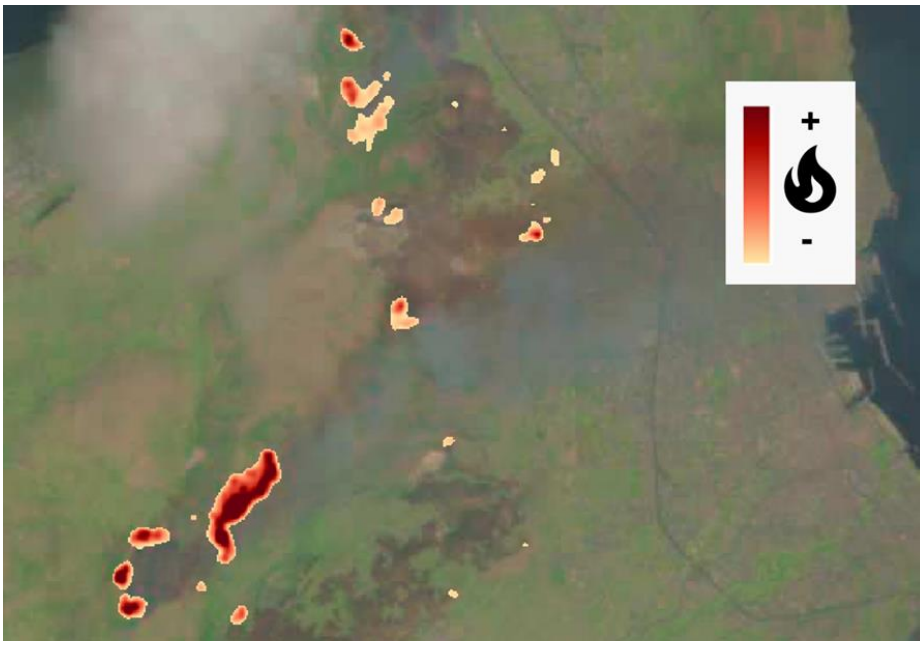


Figure 5.3.7: Gradient display as a function of the brightness of the hotspots calculated with Landsat 8 satellite data.

3/ Improvement of the spatial resolution of products used by EFFIS/FIRMS for Hotspot detection:

Knowing the precise location of Hotspots is essential for firefighting. As mentioned in previous sections, the ideal scenario is to have the temporal resolution of MODIS/VIIRS sensors with the spatial resolution of Landsat/Sentinel sensors or even sensors onboard drones. For this reason, the last section of this tool corresponds to the use of unmixing algorithms with which to increase the spatial resolution of sensors that provide Hotspots with a higher temporal frequency.

The unmixing methodology chosen is the one proposed by[31], but refined to use Landsat, Sentinel, and drone data as inputs. The original methodology is based on improving the spatial resolution of Landsat 8 satellite thermal infrared images, with 30m spatial resolution, by applying unmixing algorithms using the Sentinel 2 RGB combination, with 10m spatial resolution (Figure 5.3.8).

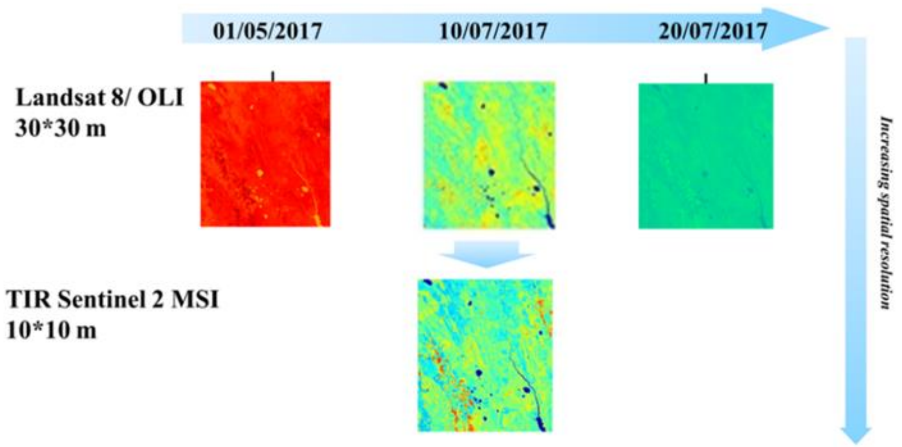


Figure 5.3.8: Conceptual workflow of the unmixing algorithm.

Improvements have been made to this functionality of the tool in order to include the improvement of the spatial resolution of ASTER thermal images using Landsat RGB imagery (Figure 5.3.9), using different strategies. The objective is to improve the original spatial resolution from 90m to the unmixed spatial resolution of 30m. These new unmixing steps correspond to 3x3, 5x5, 7x7 and 9x9 windows (Figure 5.3.10).

[31] Herrero-Huerta, M., Lagüela, S., Alfieri, S. M., & Menenti, M. (2019). Generating high-temporal and spatial resolution TIR image data. *International Journal of Applied Earth Observation and Geoinformation*, *78*, 149-162.

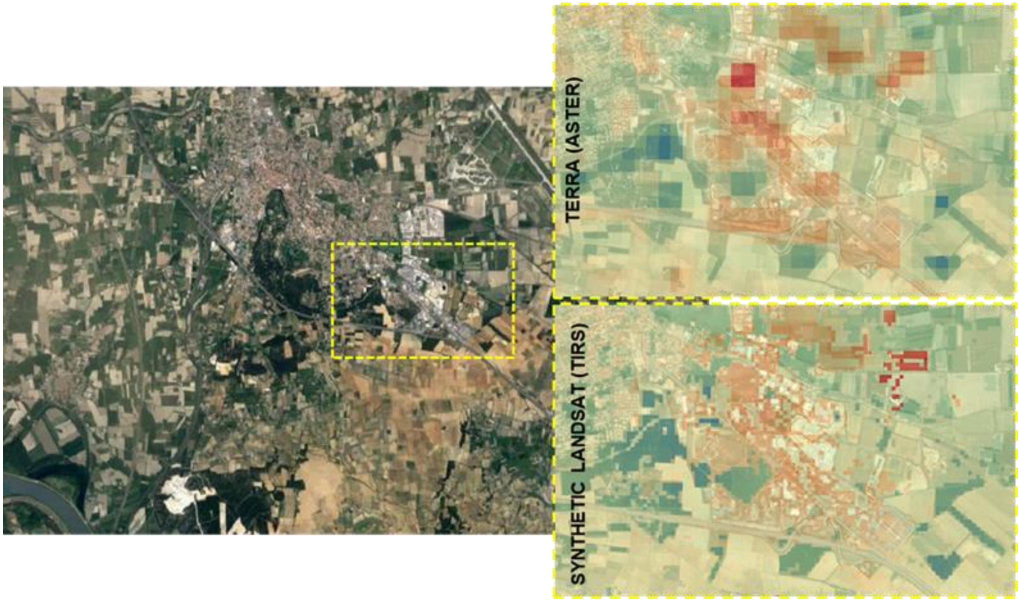


Figure 5.3.9: Test of the unmixing algorithm on an ASTER thermal image (90m) from a Landsat 9 RGB image (30m).

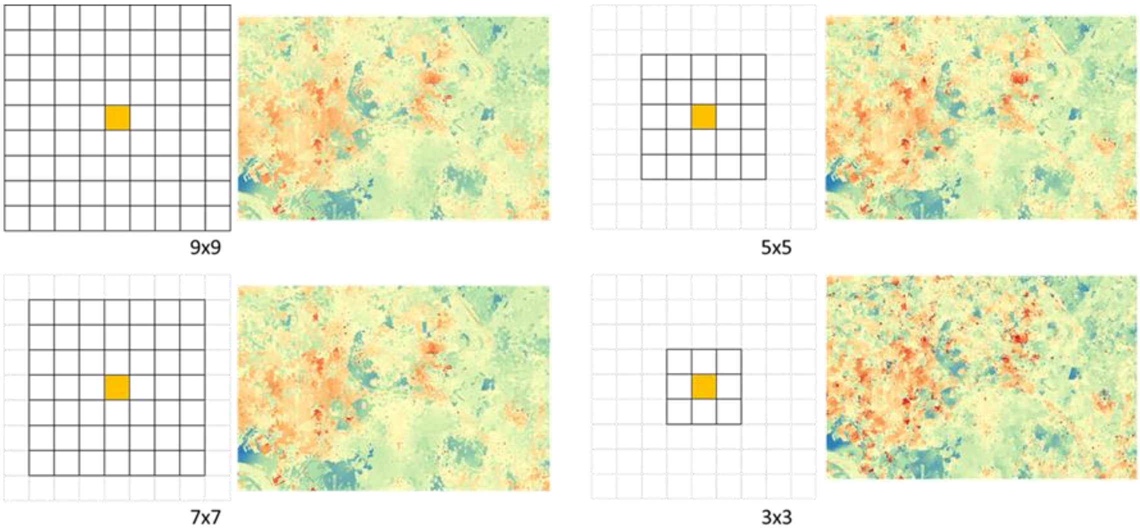


Figure 5.3.10: Example of results using different unmixing steps, corresponding to 3x3, 5x5, 7x7 and 9x9 windows

With respect to route planning and evaluation of access for firefighting, a shadow map is a cartographic representation that displays the distribution of shadows over a terrain. This visual tool is particularly useful for quickly and effectively interpreting the morphology of the terrain. Such maps enable observers to easily grasp the terrain's structure and elevation differences. They facilitate the identification of landforms, the detection of slopes, and the visualization of the terrain's shape, which is valuable for both decision-making and effective communication of geospatial information.

For planning wildfire extinguishing or defense operations, primarily during the response phase when decisions must be made as quickly as possible, visualizing this layer allows for locating the most convenient routes or access points to the affected fire areas. The complexity of the terrain in forested areas, such as steep slopes and prominent geographical features, makes traversing the territory challenging. Therefore, a rapid interpretation of the terrain, as provided by a shadow map, allows users to determine optimal access routes to these areas based on the maneuverability required by the necessary resources for the proposed action.

The Shuttle Radar Topography Mission (SRTM), as described by Farr et al. in 2007[32], represents a collaborative international research project aimed at acquiring digital elevation models spanning nearly the entire globe. The SRTM V3 dataset, also known as SRTM Plus, is made available by NASA JPL, offering a remarkable spatial resolution of 1 arc-second, equivalent to approximately 30 meters. What sets this dataset apart is the meticulous void-filling process it has undergone, relying on open-source data such as ASTER GDEM2, GMTED2010, and NED. In contrast to previous versions that may contain gaps or rely on commercial data for gap-filling, SRTM Plus ensures comprehensive and freely accessible elevation information.

This digital terrain model[33] is the data source chosen as input data for the shadow calculation in GEE (Google Earth Engine). The shadow calculation and the publication of the resulting layer as a WMS service, which queries the CHAMELEON geographic viewer, is performed on the GEE platform.

Thanks to the approach made with these point clouds, the geometric quality of the represented terrain, especially in the wildland-urban interface, is enhanced. This improvement is crucial for simulating wildfires more accurately in these specific areas, as it allows for a precise and dynamic characterization of the wildland-urban interface. With the use of flight platforms operating at lower altitudes, the possibility of further improving this quality in areas or applications requiring this detailed layer is considered, particularly beneficial in wildland-urban interfaces where wildfire simulation is of interest.

RESULTS / OUTCOMES

The visualization of the Hotspots, detected both from satellite and drone data, is performed in a WebGIS environment. The result will always be a map including either the point location of the hotspots or the polygon of the area affected by the fire and colored according to the severity of the fire in the represented area (severity measured as brightness temperature). Thus, hotspots are represented as locations of incidents in the last 24h, with coordinates of the locations and with a typification of the fire level of the forest area under study.

The CHAMELEON WebGIS will contain in its geographic information system the hill shade model map as an additional layer. Users will have this layer at their disposal for a fast interpretation of the morphology of the terrain (Figure 5.3.11). On this visualization, it will be possible to draw the possible accesses to the most inaccessible areas where there is not enough

32. Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., ... & Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of geophysics*, *45*(2).

33. Earth Engine Data Catalog. Retrieved in November 2023 from: https://developers.google.com/earth-engine/datasets/catalog/USGS_SRTMGL1_003

density of communication routes, such as roads, paths, or tracks. The drawing tools palette will allow drawing polylines on the terrain shadow map layer.

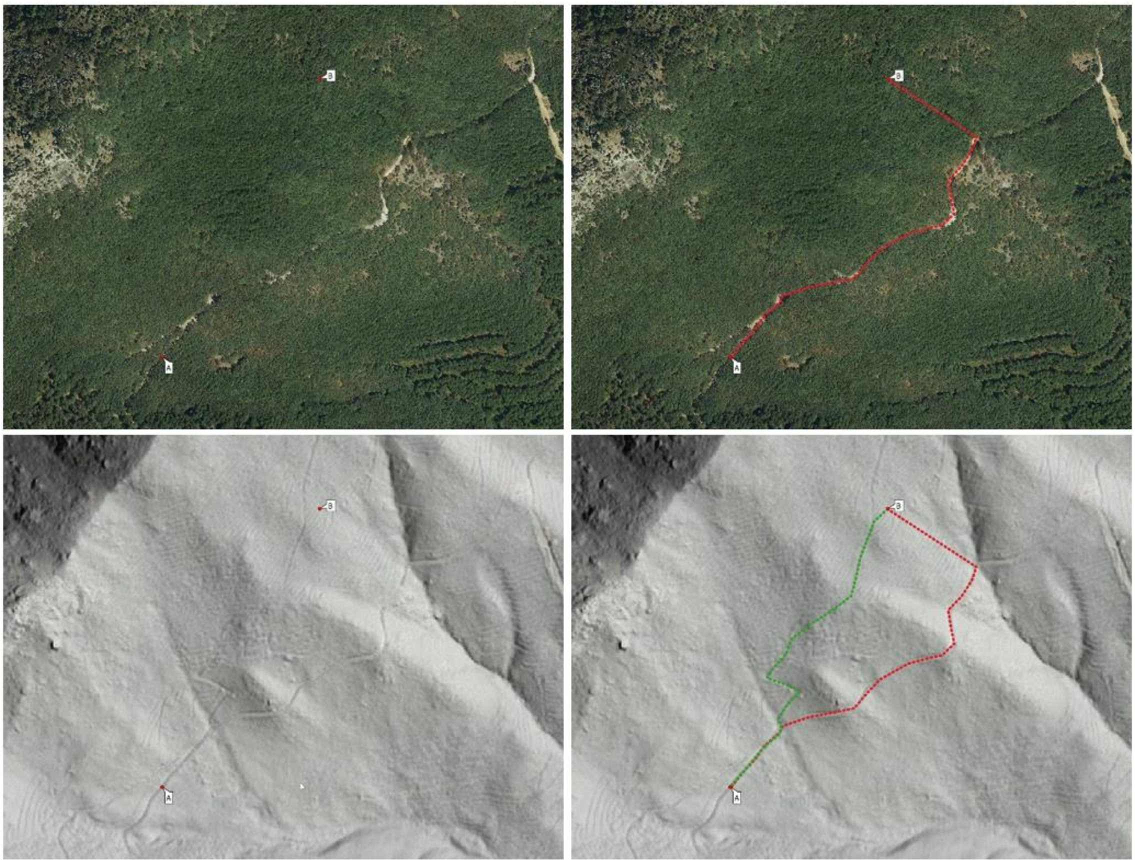


Figure 5.3.11: Orthophoto (top left) and the route planned on it (top right). Hill shade (bottom left) and optimal route planned on it (bottom right).

As can be seen in Figure 5, the interpretation of the terrain, its slopes and old logging roads is not possible on the orthophoto, thus hill shade maps make it possible to plan the optimal access to a hilly area (B).

5.4. VEGETATION MONITORING AND CENSUS

The objective for vegetation monitoring and census bundle is to accurate monitor the vegetation, temporal and spatial changes in vegetation by combining UAV's and satellites obtained data.

RELEVANCE AND PROBLEMATIC

An accurate forestry inventory is a fundamental tool for informed decision-making, sustainable management, and conserving ecosystems and natural resources. It serves as a foundation for balancing human needs with the preservation of the environment. Below are some key aspects highlighting the importance of an accurate forestry inventory:

A. Biodiversity Conservation: Forestry inventories aid in identifying and cataloging the various tree species in a forest. This information is essential for understanding and preserving biodiversity, as different species support diverse ecological niches and contribute to the overall health of the ecosystem.
B. Carbon Sequestration and Climate Change Mitigation: Forests are key in carbon sequestration. Accurate forestry inventories assist in estimating the carbon stored in trees and soil, providing crucial data for climate change mitigation. This information is particularly valuable regarding international agreements like REDD+ (Reducing Emissions from Deforestation and Forest Degradation).
C. Fire Management and Prevention: Understanding the composition and structure of a forest is crucial for effective fire management. Forestry inventories aid in assessing fire risks, developing prevention strategies, and planning controlled burns to maintain a healthy forest ecosystem.
D. Economic Planning and Revenue Generation: Forestry inventories contribute to economic planning by providing data on the potential yield of forest resources. Governments and industries use this information for revenue forecasting, taxation, and sustainable economic development related to forestry.

Moreover, while an accurate forestry inventory is crucial for various environmental and ecological purposes, addressing challenges like deforestation, invasive species, limited resources, and the complexities of climate change is necessary for effective and sustainable vegetation management.

UAV CHARACTERISTICS

For the performance of this bundle, any UAV capable of carrying an RGB camera as a payload is useful.

CAMERA CHARACTERISTICS

To obtain the required products, the following sensors included in the CHAMELEON project are used:

- RGB Camera: Sony A6000
- Multispectral camera: Micasense RedEdge MX

FLIGHT CHARACTERISTICS

To reconstruct the point cloud, a typical photogrammetric flight plan should be implemented, which includes:

- Camera orientation: Vertical or oblique flights.
- GSD: 2 cm/px (minimal requirement). 1 cm/px (optimal requirement).
- Overlap: 80-85% forward, 40-60% side.

For the multispectral data, the flight plan should include:

- Camera orientation: Vertical or oblique flights.
- GSD: Centimetric.
- Overlap: 60-70% forward, 20-40% side.

DATA PROCESSING

In this bundle, we aim to obtain an automatic forest inventory. The data to be processed is a point cloud from forested areas, acquired by Structure from Motion photogrammetry from UAV. The point cloud will be geo-referenced using either GPS-measured GCPs (Ground Control Points) for validation or through RTK/PPK.

Additionally, exploiting the temporal and spectral resolution of satellites in forest management provides a comprehensive understanding of forest dynamics, health, and composition. This information is crucial for making informed decisions related to conservation, sustainable resource management, and mitigating various threats to forest ecosystems. The Vegetation Indexes computed are: NDVI (Normalized Difference Vegetation Index), SAVI (Soil Adjusted Vegetation Index), EVI (Enhanced Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index), NBR (Normalized Burn Ratio), and NDMI (Normalized Difference Moisture Index). The Google Earth Engine platform is used for this purpose.

Figure 5.4.1 displays the general flow diagram of the bundle. Figure 5.4.2 presents the input, parameters, output, and display results table of the bundle.

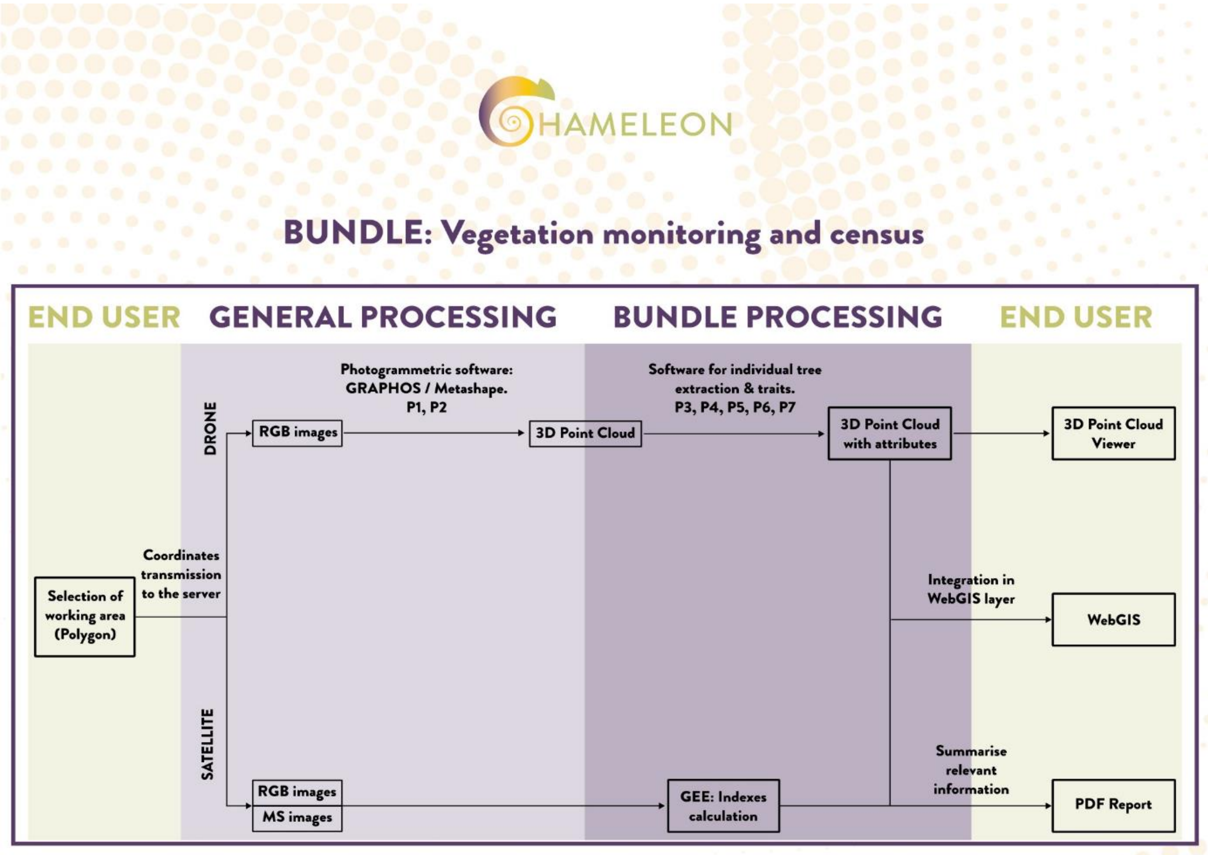


Figure 5.4.1: Workflow of the bundle “Vegetation monitoring and census”.

CHAMELEON

BUNDLE: Vegetation monitoring and census

INPUT		PARAMETERS		OUTPUT		DISPLAY RESULTS
End User	Sensors	Graphos / Metashape	Software for individual tree extraction & traits	Software for individual tree extraction & traits	GEE	
Selection of working area (polygon)	Drone: S1: RGB images -> 3D Point Cloud Satellite: S2: RGB images S3: MS images	P1: Known distances P2: Ground Control Points	P3: Window Size 1 P4: Threshold P5: Local Maximum Filter P6: Window Size 2 P7: Maximum Radius of Crown	O1: 3D Point Cloud with attributes O2: CSV: TreeID Height Vertical variability DBH Crown diameter Crown projection area Volume	O3: Indexes: NDVI SAVI EVI GNDVI NDMI	R1: PDF Report R2: Representation in custom 3D Point Cloud Viewer R3: WebGIS visualisation with layers

Figure 5.4.2: Input, parameters, output and display results of the bundle "Vegetation monitoring and census".

AI DEEP LEARNING PROCESS / ALGORITHMS

Treetop color from satellite imagery

Obtaining tree top color information from satellites can be valuable for various applications in forestry and environmental monitoring:

A. Vegetation Health Monitoring: Assessing the color of tree tops can provide insights into the health and vitality of vegetation. Unusual color patterns may indicate stress, disease, or other environmental factors affecting the trees.
B. Species Identification: Different tree species may exhibit distinct colors, especially in their leaves or canopy. Analyzing tree top color can aid in identifying and mapping tree species across large areas, which is essential for biodiversity studies and forest management.
C. Wildfire Detection and Risk Assessment: Unusual tree top color patterns, especially in terms of increased dryness or dead vegetation, can be indicative of areas at higher risk of wildfires. This information is crucial for wildfire detection and risk assessment.

To stay updated on the latest technologies and applications related to tree top color information from satellites, it's recommended to use data fusion techniques and multisource info to improve accuracy.

Satellites offer imagery with high temporal and spectral resolution, characteristics extremely in demand for forest monitoring. This imagery will provide a downscaling and coarse-to-fine approach to automatically detect geometric and radiometric changes in forests. Data fusion algorithms can improve spatial/temporal resolution (Figure 5.4.3).

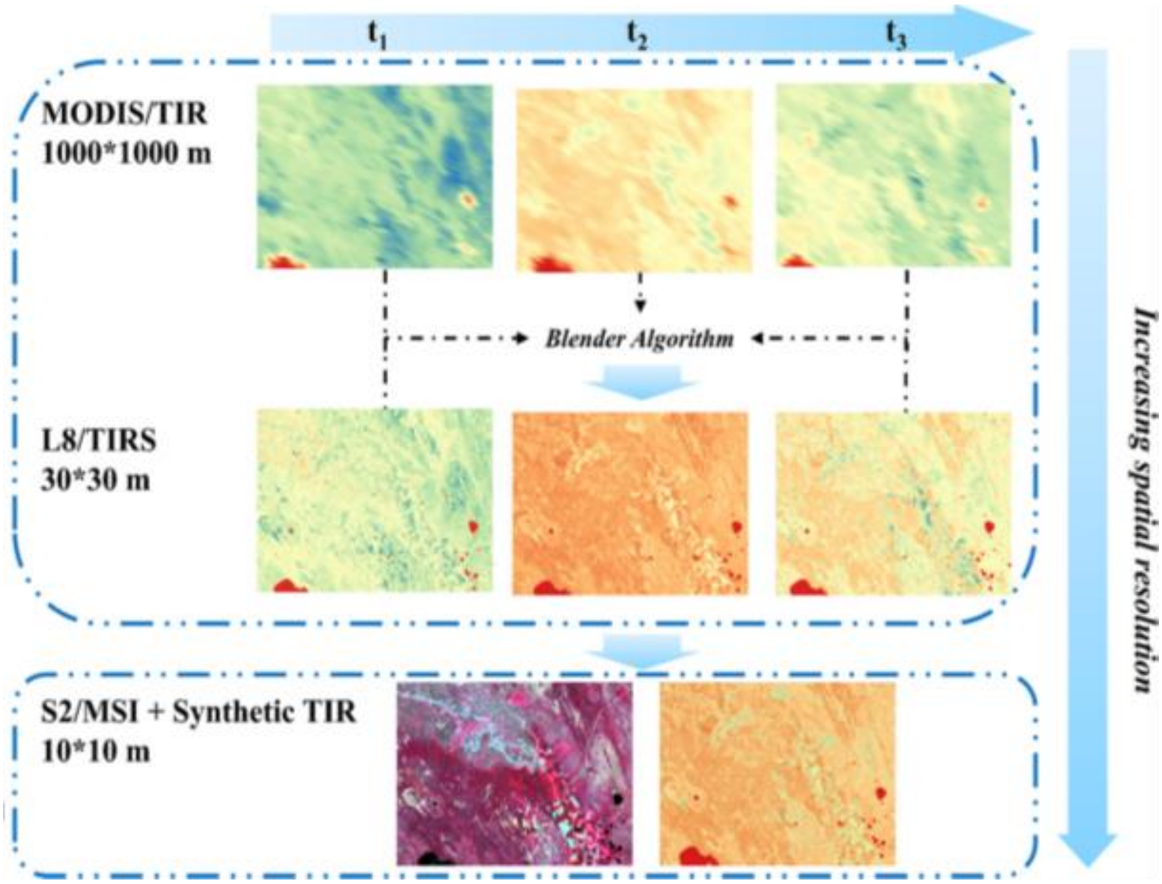


Figure 5.4.3: Conceptual workflow of the temporal and spatial data fusion algorithm already to developed.

Table 5.4.1 collects free satellite observations with the main characteristics. It is worth noting some developments in image matching carried out in TIDOP. In particular, AI algorithms such as R2D2 and D2-Net have proved useful for multimodal image matching.

Table 5.4.1: Overview of remote sensing data usable for the bundle “Vegetation monitoring and census”.

Mission	Constellation	Instruments	Spectral bands	Spatial resolution	Revisit time
Sentinel-2	Sentinel-2A & 2B	MultiSpectral Instrument (MSI)	RGB, NIR, SWIR	10-20 m	5 days
Landsat 8 & 9	Landsat 8 & 9	Operational Land Imager (OLI) & OLI-2	RGB, NIR, SWIR	30 m	8 days
Planet	PlanetScope	Dove	RGB	3-5 m	1 day

Forest inventory

The processing steps of the algorithm to obtain an automatic forest inventory are shown in Figure 5.4.4 and are also described below.

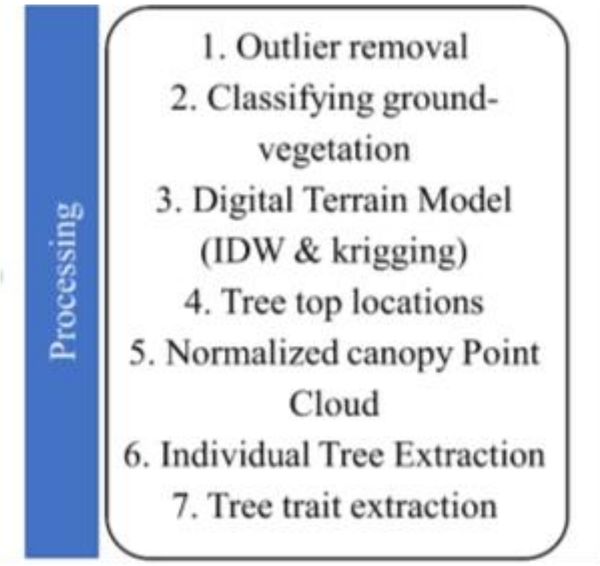


Figure 5.4.4: Processing steps of the algorithm to obtain an automatic forest inventory.

- Statistical outlier removal: The sample point cloud may have outliers and noise, typically due to interference from sources like people and vehicles. These points, not representing actual trees, need to be removed in the first step. This is done by conducting a statistical analysis on each point's neighborhood, using a Gaussian distribution of distances between neighbors to identify and remove outliers.
- Ground/non-ground classification: Here, the Cloth Simulation Filter (CSF) is applied for ground and non-ground classification. This filter requires only a few simple integer and Boolean parameters to configure. The point cloud is inverted, and a rigid cloth is employed to envelop the inverted surface. By analyzing the interaction between the cloth nodes and their corresponding points, the cloth nodes' positions are determined, creating an approximate ground surface representation. Then, ground points are identified by comparing the original points with this surface.
- Digital Terrain Model (DTM) creation: This process begins with the inverse distance weighting algorithm (IDW). It involves a two-step approach: initially determining the spatial covariance structure of sampled points by fitting a variogram with exponential, spherical, or Gaussian curves; then using the weights derived from this structure to interpolate values for unsampled points or blocks across the spatial field.
- Tree top locations: The process of individual tree detection involves locating trees in their spatial positions and extracting their height information using Local Maximum Filter (LMF) algorithm.
- Normalized canopy point cloud: This process mitigates the impact of terrain on above-ground measurements. Normalization aids in comparing vegetation heights and streamlines analyses across different acquisition areas.

- Individual tree extraction: This involves an initial region growing algorithm for point cloud segmentation, followed by detailed steps to enhance the accuracy of individual tree extraction, including 2D border analysis and 3D point cloud recovery within the defined borders. To make the method suitable for different tree forest densities and shapes, specific internal parameters and thresholds need to be tailored accordingly.
- Tree trait extraction: For each tree, we automatically extract traits such as location coordinates, height, vertical variability, crown diameter (based on the maximum distance between points and a circle fitting approach), crown projected area, and canopy volume. These traits are compiled in an Excel file.

RESULTS / OUTCOMES

Figure 5.4.5 illustrates the intermediate results of the steps 1 to 6 of the algorithm, with the seventh step being an Excel file containing tree traits.

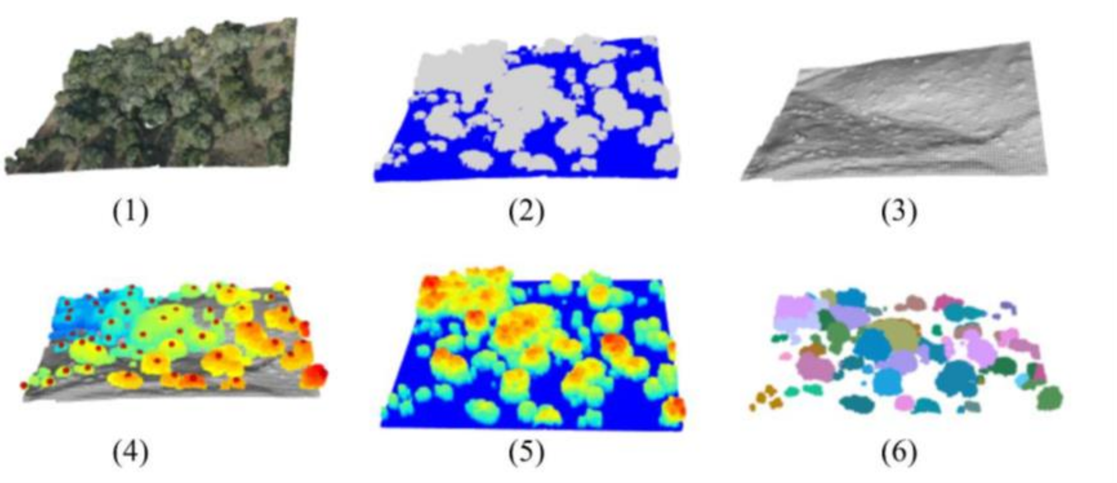


Figure 5.4.5: Intermediate results of steps 1 to 6 of the algorithm to obtain an automatic forest inventory.

The structure of the final results of the algorithm is shown in Figure 5.4.6.

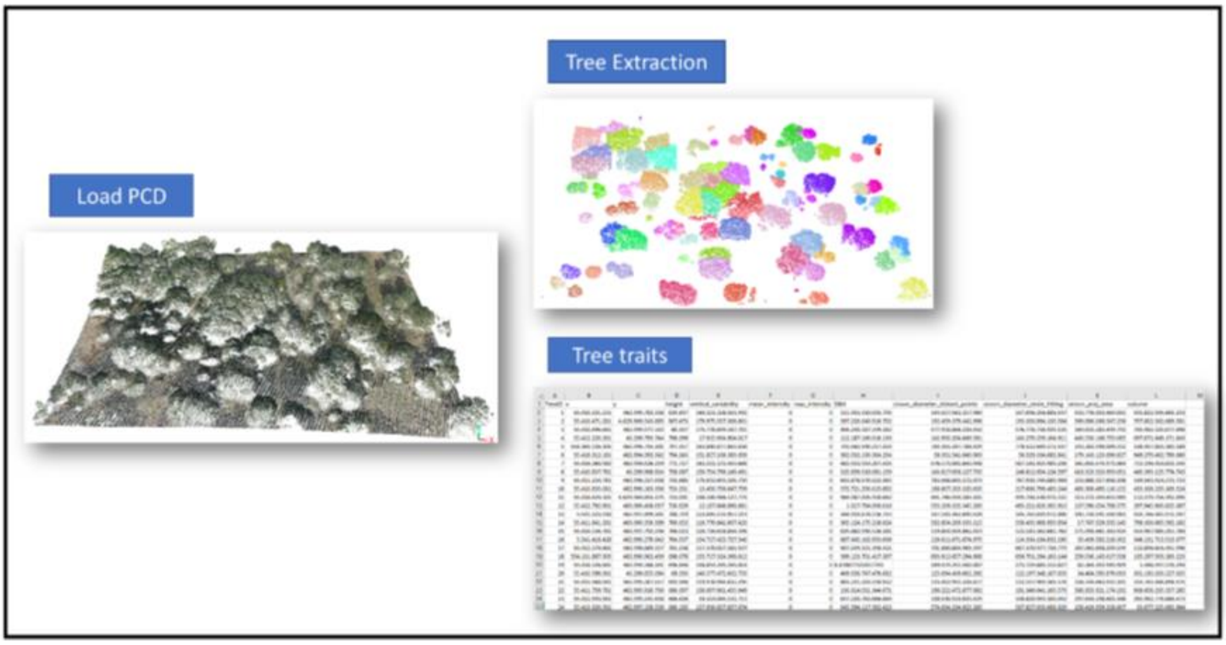


Figure 5.4.6: Structure of the final results of the algorithm to obtain an automatic forest inventory.

The advanced pipeline offers comprehensive insights into forest structure, encompassing precise individual tree extraction. Automatically, crucial tree parameters can be derived, such as:

A. Location coordinates
B. Height
C. Vertical variability
D. Crown diameter
E. Crown projected area
F. Canopy volume.

Additionally, the pipeline extends to include forest parameters, including canopy distribution, canopy volume, and even terrain variations.

5.5. LARGE WOODY DEBRIS ON RIVERS

The large woody debris on rivers is a bundle that will help to identify and locate large fallen trees and woody debris that can alter the normal current of the river, leading to flooding of surrounding areas.

RELEVANCE AND PROBLEMATIC

The objective of this bundle is to automatically detect potentially problematic woody debris in rivers from drone imagery. After extreme climatic events, such as strong storms, it is important to locate fallen trees and woody debris that can alter the normal current of the river, leading to flooding of surrounding areas.

The use of drones for this task offers significant advantages as access to the locations is often difficult and the areas to be covered are extensive. Manually identifying the debris from drone imagery is time-consuming, especially due to the vast areas involved.

The main challenge in automatically detecting woody debris stems from their great variability. These elements can range from logs composed of a large number of fallen trees to individual pieces of varying sizes and types, depending on the ecosystem of the area.

UAV CHARACTERISTICS

For the performance of this bundle, any UAV capable of carrying an RGB camera as a payload is useful.

CAMERA CHARACTERISTICS

To obtain the required products, the following sensors included in the CHAMELEON project are used:

- RGB Camera: Sony A6000

FLIGHT CHARACTERISTICS

Real-time: The plan should include:

- Camera orientation: Vertical flight.
- GSD: 20 cm/px (minimal requirement).

Postprocessing: A typical photogrammetric flight plan should be implemented, which includes:

- Camera orientation: Vertical or oblique flights
- GSD: 2 cm/px (minimal requirement). 1 cm/px (optimal requirement).
- Overlap: 80-85% forward, 40-60% side.

DATA PROCESSING

Methodology overview

The tool presented will take UAV aerial images of the river area as input. There are two different approaches available, depending on whether the process is conducted in real-time or in postprocessing.

Figure 5.5.1 displays the general flow diagram of the bundle. Figure 5.5.2 presents the input, parameters, output, and display results table of the bundle.

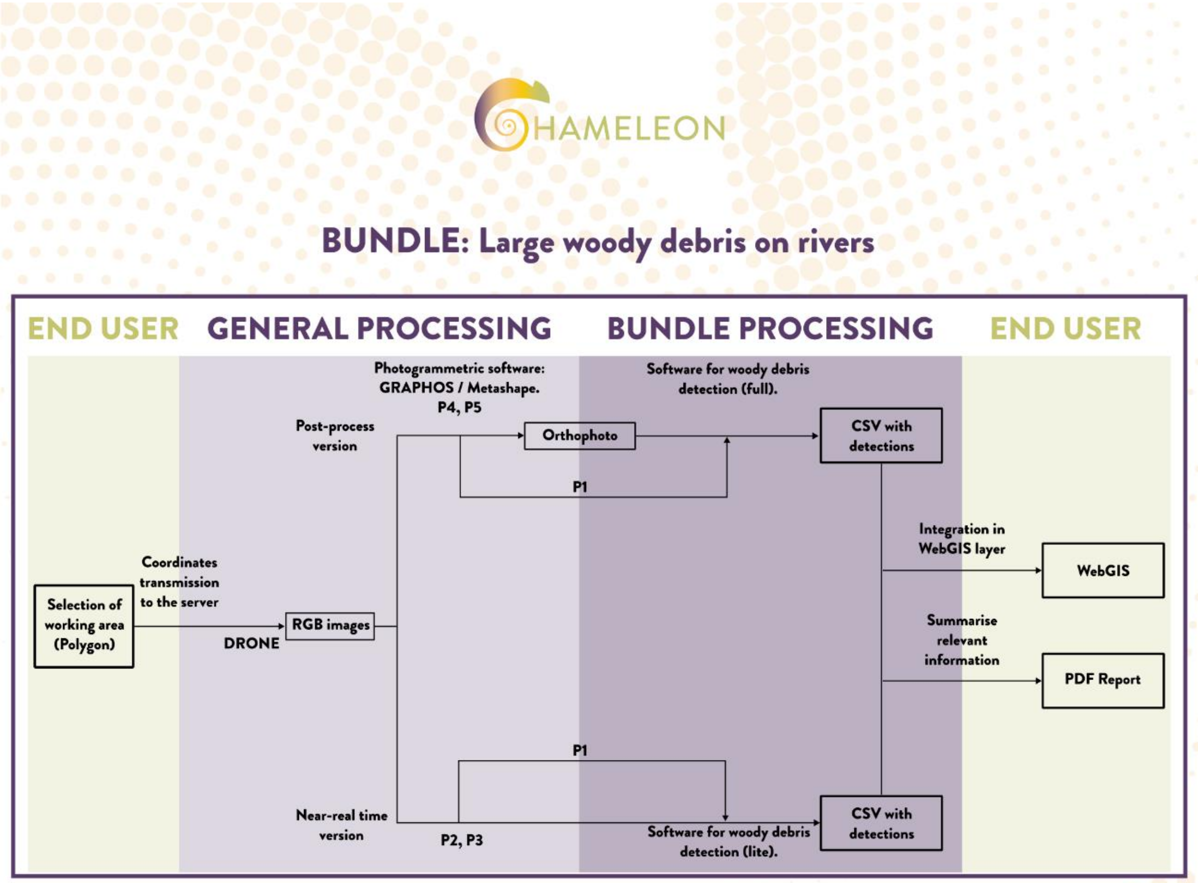


Figure 5.5.1: Workflow of the bundle “Large woody debris on rivers”.

BUNDLE: Large woody debris on rivers

INPUT		PARAMETERS		OUTPUT	DISPLAY RESULTS
End User	Sensors	Near-real time version	Post-process version	O1: CSV: Detection ID X (global coordinate system) Y (global coordinate system) Area Ratio: Occupied / Free Severity	R1: PDF Report R2: WebGIS visualisation with layers
Selection of working area (polygon)	Drone: S1: RGB images	P1: DEM P2: GPS P3: IMU	P1: DEM Graphos / Metashape: P4: Known distances P5: Ground Control Points		

Figure 5.5.2: Input, parameters, output and display results of the bundle “Large woody debris on rivers”.

In the first scenario, the docker script installed on the drone's GPU will process images, GPS and INS data, as well as a DTM of the area. For post-processing, the docker script on the CHAMELEON server will create an orthoimage for the detection process. In every case, the detection will follow these steps:

- Water detection
- Debris detection
- Risk evaluation
- Extraction of coordinates

Water detection

For water detection, the image will be divided into tiles of 128x128 pixels with an approximate GSD of 20 cm. A water detection algorithm, based on UNET, will be used to identify water in each tile. Finally, the tiles will be reassembled, and a morphological cleaning process will be conducted to determine the river's boundary. This boundary will then be used to create an image where areas not containing water are masked.

Debris detection

Debris detection will be performed using the water-only image obtained in the previous step. Eliminating non-water areas will significantly reduce the number of false positives. For this task, we trained a custom YOLOv8x-seg model with our own dataset of woody debris images. This model will be utilized for the detection and segmentation of debris. To process large images, SAHI[34] will be employed to slice the image into smaller tiles.

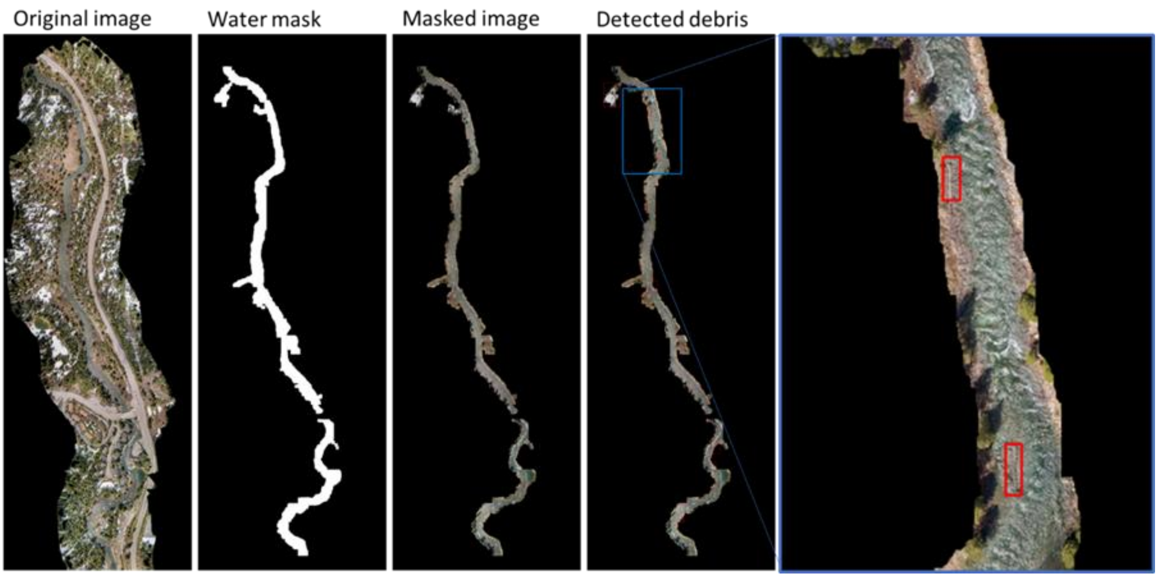


Figure 5.5.3: Debris detection process and results.

Risk evaluation

For each detected element, a risk evaluation will be computed. This score will be determined based on the size of the debris and its impact on the current. Debris will be categorized into

[34] Framework agnostic sliced/tiled inference + interactive ui + error analysis plots. Retrieved in November 2023 from: https://github.com/obss/sahi

three risk levels: low, medium, and high. Low risk corresponds to debris on the river's margin, minimally affecting the current, while high risk is assigned to logs that block the current.

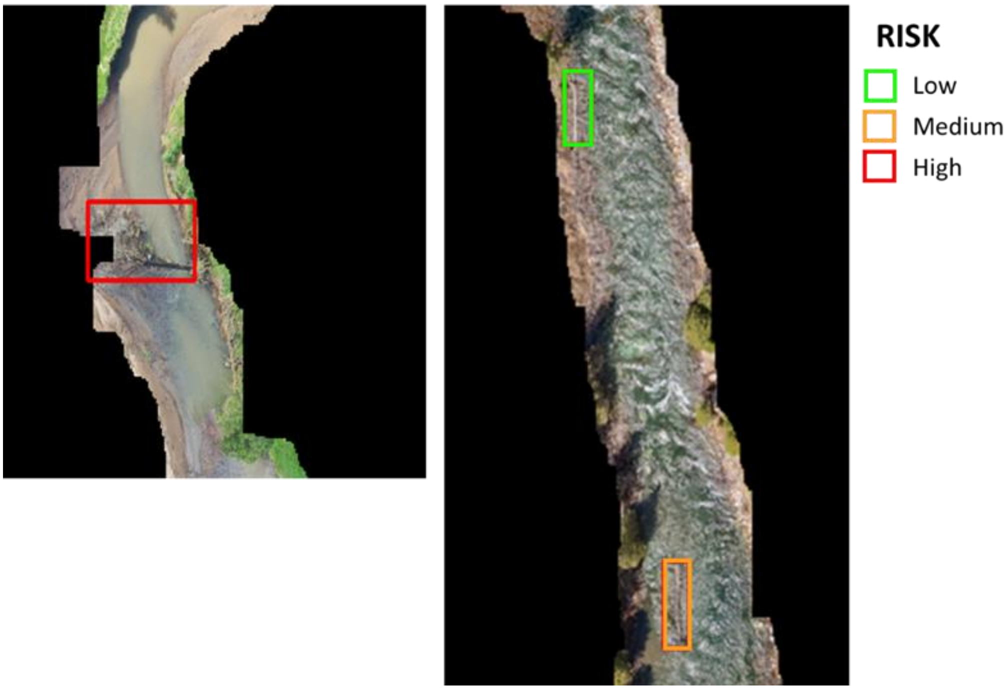


Figure 5.5.4: Risk evaluation examples.

Geolocation of woody debris

Two different approaches are possible, depending on the input data. For real-time processing, the input includes drone images along with GPS and INS data, which facilitate the registration of images to a known coordinate system. Otherwise, the input is an orthoimage, correctly referenced in a coordinate system.

– Real-time processing:

In this scenario, the tool requires a Digital Terrain Model (DTM). This information, combined with the drone's GPS and INS data, allows for the calculation of the drone image footprint. This data is used to estimate the approximate terrain coordinates of detected debris. If INS data is unavailable, image acquisition must be nadiral.

– Postprocessing:

Here, the input is an orthoimage, where the coordinates of each pixel are already known. The coordinates of each detection are obtained automatically with the precision of the orthoimage.

AI DEEP LEARNING PROCESS / ALGORITHMS

The tool includes two deep learning algorithms: one for detecting water and another for detecting woody debris.

UNET for water detection

Watercourse detection is conducted using UNET, with the model running on TensorFlow and the segmentation_models library[35].

The model was trained on a dataset of approximately 2000 images, featuring manually delineated water areas, labeled using Roboflow[36].

Various backbones and image resolutions were tested. The current tool employs Efficientnetb0 as the backbone with a low resolution of 128 pixels.

The area under the curve (AUC) for the test data was 0.96.

YOLOv8 for debris detection

A dataset of 2312 images was used for model training. Some images were fragments of orthoimages obtained by drones, while others were sourced from Google Earth using historical images or from other sources like OpenAerialMap.

Images were cropped to a small size, maintaining a similar scale for the debris. For all images, water was previously masked using the model developed in earlier steps, and then they were manually labeled.

Various object detection algorithms were tested, including YOLO-NAS (using the super-gradients library), YOLOv8 (ultralytics library), Faster-RCNN (detectron2 library), and SSD (super-gradients library).

The final tool employs the YOLOv8x-seg model, as it produced the best results, with an mAP50 ranging from 0.65 to 0.8 across different datasets.

RESULTS / OUTCOMES

For each processed image, the algorithm will generate a CSV file containing the following data for each detected element:

- Polygon of the detected debris in image coordinates.
- Coordinates of the detection in an absolute coordinate system.
- Risk score (low, medium, or high).

This data will facilitate the visualization of results on the WebGIS platform. Users will be able to view the images for each detection by clicking on the corresponding points/icons, allowing them to assess the actual situation of each detection.

In addition to WebGIS visualization, a PDF report will be provided, detailing all detections, their coordinates, and the affected river area.

5.6. HEALTH STATUS OF VEGETATION (MAINLY BARK BEETLE), GAME BROWSING, GROUND COVER, AND FUNGAL GROWTH

35. Segmentation models with pretrained backbones. Keras and TensorFlow Keras. Retrieved in November 2023 from: https://github.com/qubvel/segmentation_models

36. Roboflow. Retrieved in November 2023 from: https://roboflow.com/

This bundle will be responsible for describing the health status of forest stands, specifically coniferous stands, following bark beetle damages. The main objective is timely identifying changes in spruce forest health status affected by bark beetle of *Ips typographus* and classify affected and non-affected trees in near-real time processing; identify new outbreak emerging in bark beetle damaged areas.

RELEVANCE AND PROBLEMATIC

Forest health decline is a phenomenon of global concern involving variety abiotic, biotic and human caused factors. Tree vitality is consistently under influence of a wide variety of stress factors, including climate, soil conditions, air pollution, pathogens, drought, etc.

Insects are important disturbance driver that play roles in the long-term dynamics of forest ecosystems. Bark beetles are widely distributed in European forests, act mostly as secondary biotic agents, affecting stands that are already weakened, however, large *Ips typographus* populations could affect relatively health spruce forests.

In forest health monitoring routine, the traditional visual tree health indicators assessment is often compounded do to stand density, stand high, difficult or restricted access to the forest. The efficient forest health monitoring, pests and bark beetle outbreaks control could be precisely determined with the help of UAV. UAV remote sensing system could provide accurate and valuable information with low operational costs. Using high-quality photos taken with the UAV, enable to identify early emergence of bark beetle outbreaks, location of pest concentrations and the status of their population, as well as trends. Near real-time results helps to make timely decisions and take forest management actions for stakeholders and final users of this bundle.

UAV CHARACTERISTICS

The data for this bundle development were collected by using:

A. Low altitude drone used for small area RGB imaging.
B. Medium-high altitude, fixed wing UAVs for larger area multispectral imaging.

SenseFly eBee plus was used for data collecting (Table 5.6.1).

Table 5.6.1: SenseFly eBee plus Fixed wing UAV characteristics

Technical Specifications	Requirements
Weight	>1kg
Wing type	Fixed wing
Battery (2150 mAh)	Two Lithium-Polymer battery packs &
Camera (including 16 GB SD card, battery, USB cable & charger)	18.2 MP
Multispectral camera	
Operational requirements	
Flight time (Max.)	50 -60 minutes

Cruising speed	50-100 km/h
Radio link range	Up to 3.5 km
Maximum coverage (single flight)	$15km^2$
Wind resistance	45 km/h
Ground sampling Distance (GSD)	1.5cm per pixel
Relative orthomosaic/3D model accuracy	1-3x GSD
Absolute horizontal/vertical accuracy (w/GCPs)	Down to 3 cm / 5 cm
Absolute horizontal/vertical accuracy (no GCPs)	1-5 m
Multi-drone operation	Yes (including mid-air collision avoidance
Automatic 3D flight planning	Yes
Linear landing accuracy	Approx. 5 m
flight planning & control & Terra 3D professional photogrammetry software	

CAMERA CHARACTERISTICS

For this bundle development and data collecting two types of cameras were used: multispectral and RGB. Minimal spectral bands requirements: RGB+NIR.

FLIGHT CHARACTERISTICS

The optimal flight characteristics for successful flight consists of these main steps and modes:

- Flight planning – the area to be captured should be larger than the actual field of interest so that there is sufficient data all the way to the edges of the field.
- Minimal ground resolution requirements GSD, cm/px: 3-10 cm/px.
- Minimal image overlap requirements, %: 85%.

The example of the actual flight route with images capture locations are presented in Figure 5.6.1.

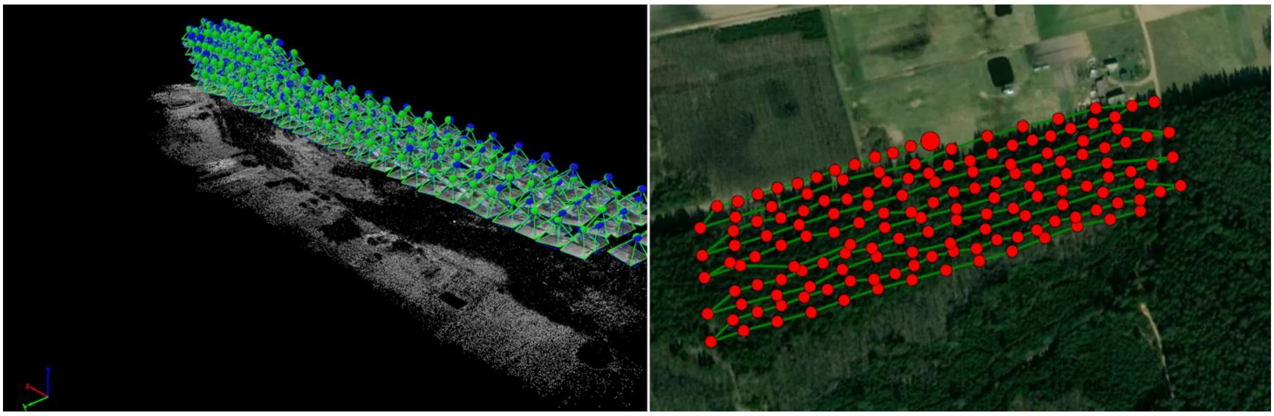


Figure 5.6.1: The example of the actual flight route (lines) with images capture location (dots).

DATA PROCESSING (METHODOLOGY)

Methodology overview. The purpose of this forest sanitary condition assessment tool is recognition of pest outbreaks in the forested area based on UAV-based aerial imagery. To this end, recognition of tree crowns is performed in an orthophoto map employing a custom-trained crown segmentation model. Then, different vegetation indexes are calculated for each segmented tree crown, based on which tree health condition is determined and indicated by its class: healthy, stressed, dead. The location and status of tree crowns are saved in a GIS vector file. All findings are aggregated and summarized, including total inspected area, number of tree crowns, percentage of unhealthy trees based on different vegetation indexes.

Analytical steps. The forest health assessment tool based on RGB and/or multispectral (RGB+NIR) UAV aerial imagery with GSD resolution 3-10 cm/px. The schematic view of key analytics steps and components are presented in Figure 5.6.2.

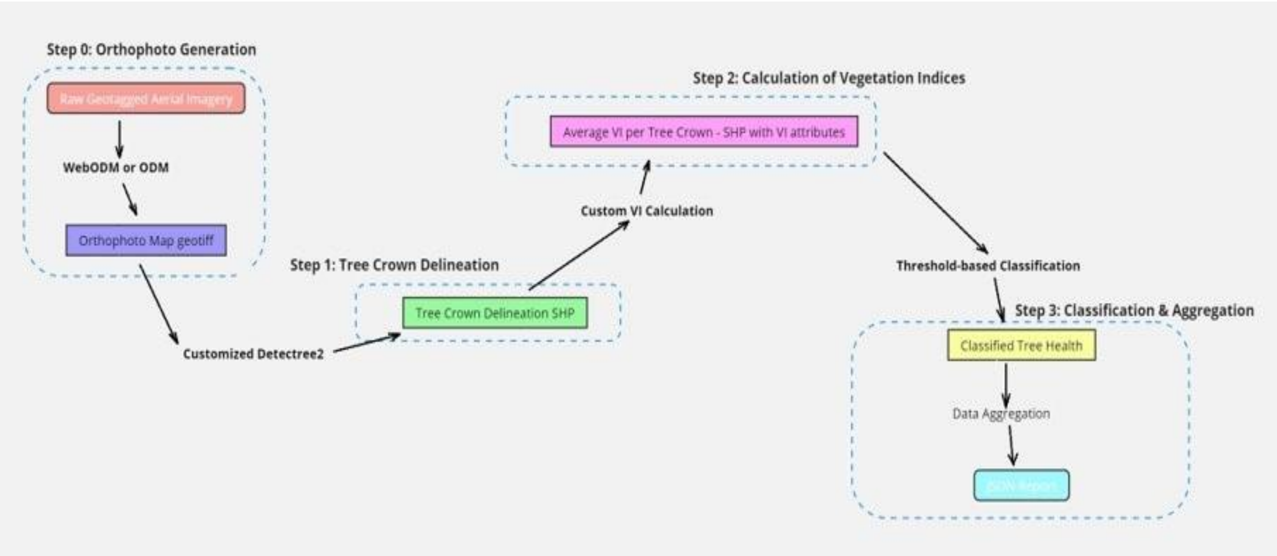


Figure 5.6.2: The schematic view of key analytics steps and components for this bundle development

A. **Orthophoto generation.** Generation of orthophoto map from geotagged aerial imagery (performed externally).
B. **Tree crown delineation.** Segmentation and delineation of tree crowns in aerial orthophoto map.
C. **Calculation of average VI.** Calculation of average vegetation index value for each delineated tree crown based on spectral information:
   - NGRDI – normalized green-red difference index [Green-Red]/[Green+Red];
   - VARI – visible atmospherically resistant index [Green-Red]/[Green+Red-Blue];
   - NDVI – normalized difference vegetation index [NIR-Red]/[NIR+Red].
D. **Classification by tree health.** Classification of delineated tree crowns based on the average vegetation index, where the health class (healthy, stressed, dead) is indicated in the attribute table of the ESRI shapefile vector file:

- Healthy/Stressed/Dead_NGRDI;
- Healthy/Stressed/Dead_VARI;
- Healthy/Stressed/Dead_NDVI.

E. **Reporting.** Data aggregation and calculation of vegetation health statistics based on each vegetation index, reported in json format:

- Total number of delineated tree crowns.
- Percentage of stressed trees based on each vegetation index (NGRDI, VARI, NDVI), calculated as a ratio between stressed and healthy trees.
- Percentage of dead trees based on each vegetation index (NGRDI, VARI, NDVI), calculated as a ratio between dead and healthy trees.
- Total inspected area (calculated from orthophoto map geotiff raster file).
- Percentage of area covered by tree crowns (calculated as a ratio between total area of delineated tree crowns and total geotiff area).

The detailed information about each analytics step and components of the forest health assessment tool are presented in Table 5.6.2.

Table 5.6.2: Analytics steps with description for the forest health assessment tool development

Analytics steps with descriptions						
Step	**Title**	**Tool, origin**	**Input**		**Output**	
			Description	**Format**	**Description**	**Format**
0	Generation of orthophoto maps (performed externally)	ODM, WebODM	Raw geotagged aerial imagery (red, green, blue, NIR, etc.)	jpeg, tiff: 1-4 bands per image (red, green, blue, NIR) or other supported by the photogrammetry software	Orthophoto map	geotiff, 4 bands: red, green, blue, NIR (0-255)
1	Delineation of tree crowns	Detectree2, customized	Orthophoto map	geotiff, 4 bands: red, green, blue, NIR (0-255)	Delineated tree crowns	Polygon, ESRI shapefile SHP
2	Calculation of average vegetation indices for each crown (NGRDI, VARI, NDVI)	Custom	1) Delineated tree crowns 2) Orthophoto map	1) polygon, ESRI shapefile SHP 2) geotiff, 4 bands: red, green, blue, NIR (0-255)	Delineated tree crowns with average NGRDI, VARI, NDVI attributes	Polygon, ESRI shapefile SHP
3	Classification and aggregation of health assessment results based on VIs	Custom	Delineated tree crowns with average VI attributes	Polygon, ESRI shapefile SHP	Calculated tree health condition statistics for the inspected area	json

Step 0 – Orthophoto generation

An open-source tool for orthophoto generation WebODM or ODM were used. If required, this tool can be incorporated into the analytics pipeline. Orthophoto maps generated using any other software are also compatible, if they consist of 3 (RGB) or 4 spectral bands (RGB+NIR) with amplitude values normalized to 0-255. A typical ground resolution for RGB imagery is ~3 cm/px and for multispectral imagery ~10 cm/px. An example of RGB orthophoto map is presented in Figure 4.6.3. panel (A).

Step 1 – Tree crown delineation

Tree crown delineation is performed using an open-source tool Detectree2 as a basis for our customized solution. This tool employs a tree crown segmentation model that we have trained on extensive aerial data of forested areas. Model training was carried out using ground truth data created by manual labelling of individual tree crowns using polygons on orthophoto maps of 3-10 cm/px GSD, totalling ~50 000 tree crowns in the training dataset. The trained model is placed in “lammc\model.pth”. Each segmented tree crown is then delineated, and its contour, or ground projection, is stored in a SHP (ESRI shapefile format) vector file “result\crown_map.shp”, where attribute “confidence” indicates a probability that the model's prediction is correct (where 1 corresponds to 100%). The tool best recognizes tree crowns in orthophoto maps generated from geotagged UAV-based aerial imagery of 3-10 cm/px GSD. This step is performed using lammc\predict_shp.py.

Step 2 – Calculation of average vegetation indices

Based on the tree crown polygon and a multispectral orthophoto map, an average vegetation index value is calculated for each segmented tree. Three vegetation indices are used in this stage, using intensity from red, green, blue, and near-infrared spectral bands:

a. NGRDI – normalized green-red difference index [Green-Red]/[Green+Red]
b. VARI – visible atmospherically resistant index [Green-Red]/[Green+Red-Blue]
c. NDVI – normalized difference vegetation index [NIR-Red]/[NIR+Red]

Different vegetation indices are included in calculations to better cover different environment conditions, as one VI might demonstrate higher sensitivity to forest health state than the others. Also, NGRDI and VARI are based on RGB-only data input, while NDVI requires additional NIR spectral band and thus multispectral camera. As RGB cameras are typically of higher resolution and lower cost than multispectral cameras, the developed system’s functionality to be able to analyse RGB-only datasets have corresponding advantages. As a result of average vegetation index calculations, a new layer with attributes NGRDI, VARI and NDVI is created “crown_map_ndvi.shp”. This step is executed using lammc\predict_vi.py, see an example of output data with attributed in Figure 5.6.3.

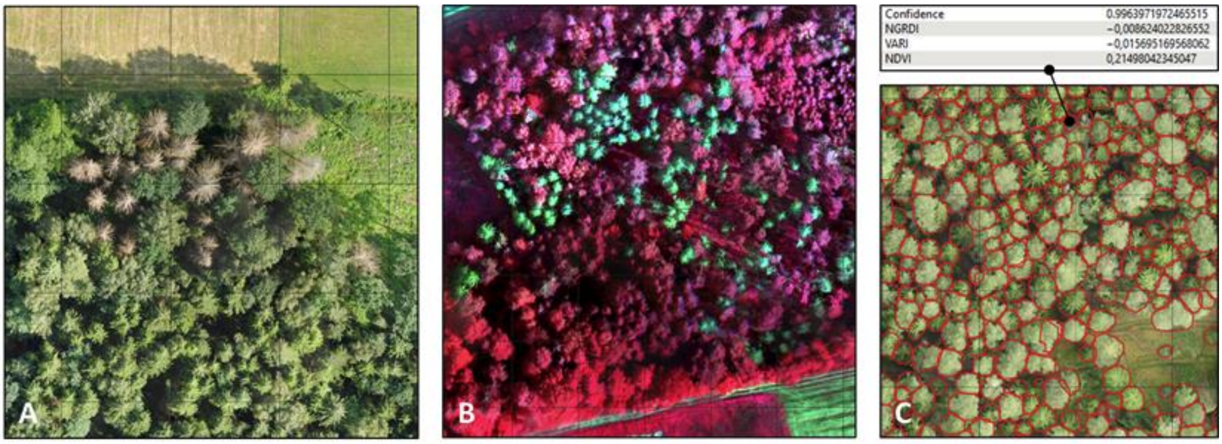


Figure 5.6.3: (A) Segment of RGB and (B) multispectral (CIR render) orthophoto map with stressed and dead trees (indicated in blue-green) and (C) example of results generated during the analytics Step 2. Grid spacing 25m.

Step 3A – Classification by tree health

Based on the average vegetation index values a classification of tree health condition is performed. To this end, threshold values should be defined for each vegetation index and each class: healthy, stressed, dead. Default threshold values are included based on scientific reports, however each user can finetune these threshold values based on their best knowledge and needs. These threshold values are used for report generation at step 3B. No additional output is generated at this step.

Step 3B – Aggregation of health assessment results and reporting

- A report in json format is generated that includes aggregated tree health assessment results of the inspected area. This step is executed using lammc\summarize.py. total_area: total inspected territory equal to the area of analysed geotiff orthophoto map;
- tree_area_p: percentage of inspected area covered in trees (ratio between total tree crown area and geotiff area);
- trees_detected: number of segmented and delineated tree crowns;
- NGRDI/VARI/NDVI statistics for different vegetation indices:
  - healthy_p: percentage of healthy trees;
  - stressed_p: percentage of stressed trees;
- dead_p: percentage of dead trees.

RESULTS / OUTCOMES

Trees health assessment Model training was carried out using ground truth data created by manual labelling of individual tree crowns using polygons on orthophoto maps of 3-10 cm/px GSD. The output values of the tool are the following:

- total_area: total inspected territory equal to the area of analysed geotiff orthophoto map;

- tree_area_p: percentage of inspected area covered in trees (ratio between total tree crown area and geotiff area);
- trees_detected: number of segmented and delineated tree crowns;
- NGRDI/VARI/NDVI statistics for different vegetation indices:
  - healthy_p: percentage of healthy trees;
- stressed_p: percentage of stressed trees;
- dead_p: percentage of dead trees.

At the current model development, about 50 000 tree crowns from more than 20 forest stands were used. The further model training will be performed to increase accuracy of the model.

6. CONCLUSIONS AND IMPLICATIONS

This public document is the first version (v1) of the Deliverable D4.3 CHAMELEON, bundles, services, and stands as one of the outcomes of WP4 implementation. The type of deliverable – document, report. The agriculture, livestock and forestry bundles are the three pillars of the CHAMELEON solution development. This Deliverable 4.3 illustrates the required system analysis and workflow of the agriculture, livestock and forestry related bundles and services, as well as functional and technical specifications of each service and bundles development stage are described. This document does not content the demonstration, pilot, prototype attitude – this will be provided in the second version (v2) of the Deliverable 4.3 [M29].

The document outlines the needs and peculiarity of CHAMELEON pilot cases, addressing local stakeholders' interests: i) Avila (Spain), ii) Crete (Greece), and iii) Vienna (Austria). The bundles and services were planed and adapted according the specific of each pilot area.

The final list of the developed bundle is provided in Table 6 with the current status of each bundle till M18. At this date, two out of all bundles are fully prepared for testing in CHAMELEON architecture: i) soil zonification (in a vineyard), ii) Health status of vegetation (bark beetles damages detection). The rest ones are significantly advanced in Beta version preparation.

Table 6: List of bundles selected to developed in the CHAMELEON project with current status of development.

Pilot Case	Group	Developer	Use Case	Bundle	Progress on Beta version preparation, % (Expected date of final Beta version)
Austria	**Forest**	**USAL**	Crop and vegetation monitoring	Vegetation monitoring and census	50
		USAL	Extreme weather event and drought	Large woody debris on rivers	70
		LAMMC	Health and pests	Health status of vegetation (mainly bark beetle), game browsing, ground cover, and fungal growth	100
	Vineyard	**UCLM**	Crop and vegetation monitoring	Crop growth and development monitoring	80 (December 2023)
		UCLM	Extreme weather event and drought	Vineyard water stress due to drought.	50 (December 2023)
Greece	**Livestock**	**AiDEAS**	Monitoring livestock	Livestock management (herd) and monitoring (individual animal)	70
		AiDEAS	Monitoring livestock	Animals' health	40

	Pasture	**UCLM**	Crop and vegetation monitoring	Monitoring flora at high-altitude grazing areas for seasonal animal feeding	25
Spain	**Forest**	**USAL**	Crop and vegetation monitoring	Continuity of vegetation	60
		USAL	Wildfire	Characterization of Wildland-urban interface.	60
		USAL	Wildfire	Hot spot identification at the beginning of wildfire	50
	Livestock	**USAL**	Monitoring livestock	Collecting information about health status and stress (wild animals)	60
		AiDEAS	Monitoring livestock	Cow lameness detection	60
	Vineyard	**UCLM**	Crop and vegetation monitoring	Crop growth and development monitoring.	80
		UCLM	Extreme weather event and drought	Vineyard water stress due to drought.	50
		UCLM	Soil	Soil zonification	100

7. REFERENCES

1. Abbas, A., Zhang, Z., Zheng, H., Alami, M.M., Alrefaei, A.F., Abbas, Q., Naqvi, S.A.H., Rao, M.J., Mosa, W.F.A., Abbas, Q., et al. (2023). Drones in Plant Disease Assessment, Efficient Monitoring, and Detection: A Way Forward to Smart Agriculture. Agronomy, 13 (1524).
2. Shah, S.A., Lakho, G.M., Keerio, H.A., Sattar, M.N., Hussain, G., Mehdi, M., Vistro, R.B., Mahmoud, E.A., Elansary, H.O. (2023). Application of Drone Surveillance for Advance Agriculture Monitoring by Android Application Using Convolution Neural Network. Agronomy, 13 (1764).
3. Kumar, S.P., Subeesh, A., Jyoti, B., Mehta, C.R. (2023). Applications of Drones in Smart Agriculture (In: Smart Agriculture for Developing Nations). Advanced Technologies and Societal Change. Springer, Singapore.
4. Equinox's Drones. How Drone Technology is Becoming Essential for Forestry. Retrieved in November 2023 from https://www.equinoxsdrones.com/how-drone-technology-is-becoming-essential-for-forestry/
5. PIX4D. Image acquisition - PIX4Dmapper. Retrieved in November 2023 from https://support.pix4d.com/hc/en-us/articles/115002471546-Image-acquisition-PIX4Dmapper
6. Di Gennaro, S.F., Vannini, G.L., Berton, A., Dainelli, R., Toscano, P., Matese, A. (2023). Missing Plant Detection in Vineyards Using UAV Angled RGB Imagery Acquired in Dormant Period. Drones, 7(6), (pp. 349).
7. Di Gennaro, S.F. and Matese, A. (2020). Evaluation of novel precision viticulture tool for canopy biomass estimation and missing plant detection based on 2.5 D and 3D approaches using RGB images acquired by UAV platform. Plant Methods, 16, (pp.1-12).
8. Jurado, J.M., Pádua, L., Feito, F.R. and Sousa, J.J. (2020). Automatic grapevine trunk detection on UAV-based point cloud. Remote Sensing, 12(18), (pp. 3043).
9. Gong, C., Zhang, Y., Wei, Y., Du, X., Su, L., & Weng, Z. (2022). Multicow pose estimation based on keypoint extraction. PloS one, 17(6), (e0269259).
10. Wang, J., Zhang, X., Gao, G., Lv, Y., Li, Q., Li, Z., ... & Chen, G. (2023). Open Pose Mask R-CNN network for Individual Cattle Recognition. IEEE Access.
11. Wei, Y., Zhang, H., Gong, C., Wang, D., Ye, M., & Jia, Y. (2023). Study of Pose Estimation Based on Spatio-Temporal Characteristics of Cow Skeleton. Agriculture, 13(8), (1535).
12. Russello, H., van der Tol, R., Kootstra, G. (2022). T-LEAP: Occlusion-robust pose estimation of walking cows using temporal information. Computers and Electronics in Agriculture, 192, (106559).
13. Fan, Q., Liu, S., Li, S., Zhao, C. (2023). Bottom-up cattle pose estimation via concise multi-branch network. Computers and Electronics in Agriculture, 211, (107945).
14. Hua, Z., Wang, Z., Xu, X., Kong, X., Song, H. (2023). An effective PoseC3D model for typical action recognition of dairy cows based on skeleton features. Computers and Electronics in Agriculture, 212, (108152).
15. Ren, S., He, K., Girshick, R., Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.
16. Li, Y., Fan, Q., Huang, H., Han, Z., & Gu, Q. (2023). A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition. Drones, 7(5), (304).
17. Mandic, D. P., and Chambers, J. (2001). Recurrent neural networks for prediction: learning algorithms, architectures and stability. John Wiley & Sons, Inc.
18. Staudemeyer, R. C., Morris, E. R. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. ARXIV PREPRINT ARXIV:1909.09586.

19. Sarwar, F., Griffin, A., Periasamy, P., Portas, K., Law, J. (2018). Detecting and counting sheep with a convolutional neural network. In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) (pp. 1-6). IEEE.
20. Wang, Y., Ma, L., Wang, Q., Wang, N., Wang, D., Wang, X., ... Ouyang, G. (2023). A Lightweight and High-Accuracy Deep Learning Method for Grassland Grazing Livestock Detection Using UAV Imagery. Remote Sensing, 15(6), (1593).
21. Meng, X., Tao, P., Han, L., CaiRang, D. (2022). Sheep Identification with Distance Balance in Two Stages Deep Learning. In 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC) (Vol. 6, pp. 1308-1313). IEEE.
22. Koklu, M., Cinar, I., Taspinar, Y. S., & Kursun, R. (2022). Identification of sheep breeds by CNN-based pre-trained InceptionV3 model. In 2022 11th Mediterranean Conference on Embedded Computing (MECO) (pp. 01-04). IEEE.
23. Simões, F., Bouveyron, C., & Precioso, F. (2023). DeepWILD: Wildlife Identification, Localisation and estimation on camera trap videos using Deep learning. Ecological Informatics, 75, (102095).
24. Łysakowski, M., Żywanowski, K., Banaszczyk, A., Nowicki, M. R., Skrzypczyński, P., Tadeja, S. K. (2023). Real-Time Onboard Object Detection for Augmented Reality: Enhancing Head-Mounted Display with YOLOv8. arXiv preprint arXiv:2306.03537.
25. Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W. and Wang, X. (2022). Bytetrack: Multi-object tracking by associating every detection box. In European Conference on Computer Vision (pp. 1-21). Cham: Springer Nature Switzerland.
26. Whereabouts Ascertainment for Low-lying Detectable Objects. The SOTA in FOSS AI for drones. Retrieved in November 2023 from: https://github.com/stephansturges/WALDO
27. CORINE Land Cover 2018 (vector), Europe, 6-yearly - version 2020_20u1, May 2020 https://doi.org/10.2909/71c95a07-e296-44fc-b22b-415f42acfdf0
28. Payra, S., Sharma, A., & Verma, S. (2023). Application of remote sensing to study forest fires. In Atmospheric Remote Sensing (pp. 239-260). Elsevier.
29. Schroeder, W., Oliva, P., Giglio, L., Quayle, B., Lorenz, E., & Morelli, F. (2016). Active fire detection using Landsat-8/OLI data. Remote sensing of environment, 185, 210-220.
30. Herrero-Huerta, M., Lagüela, S., Alfieri, S. M., & Menenti, M. (2019). Generating high-temporal and spatial resolution TIR image data. International Journal of Applied Earth Observation and Geoinformation, 78, 149-162.
31. Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., ... & Alsdorf, D. (2007). The shuttle radar topography mission. Reviews of geophysics, 45(2).
32. Earth Engine Data Catalog. Retrieved in November 2023 from: https://developers.google.com/earth-engine/datasets/catalog/USGS_SRTMGL1_003
33. Framework agnostic sliced/tiled inference + interactive ui + error analysis plots. Retrieved in November 2023 from: https://github.com/obss/sahi
34. Segmentation models with pretrained backbones. Keras and TensorFlow Keras. Retrieved in November 2023 from: https://github.com/qubvel/segmentation_models
35. Roboflow. Retrieved in November 2023 from: https://roboflow.com/

A Holistic Approach to Sustainable, Digital EU Agriculture, Forestry, Livestock and Rural Development based on Reconfigurable Aerial Enablers and Edge Artificial Intelligence-on-Demand Systems

The Members of the CHAMELEON Consortium:

Contact:

Project Coordinator: **Pantelis Velanas** Acceligence Ltd.	**pvelanas@accelligence.tech**

Disclaimer