

Intelligent Perception Systems for Multi-Modal Data Processing in Industrial Application Contexts

Annaclaudia Bono^{1,2}

¹*Polytechnic University of Bari, Department of Electrical and Information Engineering (DEI), Via E. Orabona 4, Bari, Italy*

²*Institute of Intelligent Industrial Systems and Technologies for Advanced Manufacturing (STIIMA), National Research Council (CNR), Via Amendola 122 D/O, Bari, Italy*

Abstract

Intelligent perception systems represent critical enabling technologies to bring innovation in any physical environment and improve the quality of life tending toward what is known as an intelligent future. This research explores the use of advanced monitoring perception systems with intelligent capabilities in industrial contexts. By combining networks of sensors that generate multimodal data, signal and image processing, and artificial intelligence techniques, the produced know-how can be transferred to industrial application fields for continuously monitoring goods and people and promoting the transition to a sustainable, human-centred, and resilient industry. A significant challenge facing monitoring systems is their inherent complexity, often resulting in usability issues. This study seeks to address this challenge by developing a new methodological framework to make these systems easily accessible and user-friendly for individuals regardless of their level of technical skills. The sector of interest is that of Precision Agriculture (PA), in detail Precision Viticulture (PV), where intelligent systems can make agricultural production more efficient and ensure sustainable practices are adopted to increase food production and meet growing global demand while maintaining high-quality standards. This paper provides an overview of the PhD research proposal considering the main open problems and the main steps that must be considered.


Keywords


Intelligent perception systems, Sustainability, Multimodal data, Image processing, Artificial Intelligence

1. Introduction

In today's rapidly evolving work environment, several challenges need to be addressed to adapt work processes to the changing landscape of technology. These challenges include managing increasingly complex products, facing product deterioration throughout the life cycle, increasing customer demand and intensifying international competition, which requires greater efficiency in responding to the global market's needs. Achieving these goals is reshaping how various tasks and activities are performed, necessitating the acquisition of new skills and the adoption of new working patterns. In this context, it is crucial to explore resilient and technologically advanced workplaces where new technologies can augment the workforce by enabling workers to make the most of their skills and abilities [1]. A key technology that contributes to this vision is machine perception, which focuses on studying and designing systems capable of understanding and interpreting the outside world [2]. The term perception refers to the ability

CAiSE 2024 Doctoral Consortium

 a.bono@phd.poliba.it (A. Bono)

 © 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

of the human brain to organize and explain the information from external stimuli acting on the five senses [3]. The goal of machine perception is to create intelligent systems that could emulate the human perception process [4] to bring significant innovations in a wide range of application sectors, providing smart solutions to improve the quality and efficiency of operations and the welfare of workers [2, 5].

Effective target monitoring, which focuses on detecting, tracking, and analyzing specific objects of interest within a scenario through digital images or videos, is crucial in this evolving work environment. It provides detailed, real-time information on achieving objectives, allowing intelligent perception systems to adapt and respond to changing situations, enhancing the efficiency and reliability of the operations they support. In practice, target monitoring involves the use of sensors, image processing algorithms, and Artificial Intelligence (AI) techniques to identify and monitor objects in real-time.

These systems face various challenges, including variability in environmental conditions, partial occlusion of objects, and changes in lighting [6]. Another issue may be represented by the complexity of the system itself, especially when it comes to making such technologies accessible to a wide audience [7]. These problems can manifest in various ways and can influence the overall effectiveness of the system and its adoption by users. Indeed, addressing the complexity of the system ensures its use by a wide range of users allowing them to fully exploit its benefits.

Intelligent systems can be applied in numerous scenarios. The application field of *quality and production monitoring* is chosen. This type of monitoring refers to the process of supervising and controlling the quality of products and the efficiency of production within a company or production process. Real-time interpretation of multimodal data, such as images and video, can be useful for quality and production assessment. Within these, important is the Precision Agriculture (PA) sector where intelligent systems can support crop management, improving agricultural yields and ensuring that sustainable practices are adopted to increase food production by maintaining high-quality standards [8].

The agricultural world is facing new challenges, such as the projected increase in the world population to 9.7 billion by 2050, according to a United Nations press release [9]. Estimates show that current agricultural production must increase to secure food supplies for the future, requiring adequate quantities and quality of agricultural products, intensive yet environmentally safe production, and the sustainability of the resources involved. Additionally, the impacts of climate change and the scarcity of resources, such as water and fertilizers, require their efficient use to improve crop yields while reducing environmental impact. Although agricultural yields increased in the second half of the twentieth century, disadvantages emerged related to the request for more fertilizer, pesticides, and water increasing input costs to the farmer and creating environmental problems. As a result, the use of intelligent systems can enable agricultural production to be more efficient and ensure sustainable practices by applying the right treatment in the right place at the right time [10]. In standard agricultural practice interventions are based on the average characteristics of the soil, leading to the risk of applying resources either insufficiently or excessively. Agricultural practices conducted with the use of intelligent systems, on the other hand, aim to adjust timely supplies, taking into account local variability in the physical, chemical and biological characteristics of the field, as well as application times.

An important subset of PA is that of Precision Viticulture (PV) which coincides with the generic purposes of PA: the appropriate management of the inherent variability of crops, the

increase in economic benefits, and the reduction of environmental impact. In detail, PV aims to identify, within a certain degree of stability, the interannual spatial variation of the yields and quality of the grapes, identifying what are the causes that determine this variability and if they are attributable to some specific management practices of the site [11] [12].

The vine is a perennial crop and it is important to promote the development of a framework for the use of intelligent systems to optimize crop management practices, increase economic benefits, minimize environmental impact and provide the farmer with detailed information to allow better soil management and make informed decisions [13]. The control of these crops focuses on plant phenotyping to evaluate and describe the observable characteristics of a grapevine plant related to its physical appearance, structure, and growth characteristics [14]. It is important to perform this characterization directly in the field because vineyards extend over large areas, containing thousands of individual vines, each showing a slightly distinct phenotype. Traditionally, phenotypic evaluations have relied on manual methods which, although providing valuable information, are often labor-intensive, time-consuming, and prone to errors [15]. These crops are difficult to monitor due to the intrinsic difficulties associated with vineyard characteristics: having a discontinuous canopy organized in rows requires higher resolution images to discriminate canopy from the soil and greater computing capacity to manage vineyard spatial information before being used [16]. Therefore, there has been growing interest in developing automated methods for plant characterization through technological advances to provide accurate information on crop structure.

In this context, the main case study of this project is to develop a novel methodological framework for the use of intelligent perception systems to evaluate the phenotypic variations of vine plants over time, identify anomalies in their growth, and establish correlations with the specific conditions that cause them. This will ensure more efficient growth control, allowing farmers to monitor plant conditions better and predict production amounts.

The paper is structured as follows: Section 2 review the related work to provide an introduction to the context of the research topic underlining the current problems to address, Section 3 describe the methodology used in the PhD research, Section 4 presents the research questions, Section 5 shows the planning of the PhD during the years, Section 6 shows some of the methods that will be used and Section 7 presents concluding remarks.

2. Related works

Researchers have explored various methodologies for the use of intelligent perception systems in the agricultural sector. One of the most powerful tools is Remote Sensing which is a technology that can provide a timely assessment of changes in growth by acquiring information about an object or area without directly contacting it [17]. This technology relies on the use of sensors mounted on different types of platforms such as satellites, aerial (aircraft and unmanned aerial vehicles, UAVs), and ground-based (unmanned ground vehicles, UGVs) [18]. The choice of platform is determined by specific application needs, taking into account both advantages and limitations. Satellites, despite their capability for wide-area coverage, are influenced by weather conditions, high costs, and challenges in distinguishing vineyard inter-row paths and vegetation. Low-altitude platforms like manned or unmanned aerial vehicles offer high-resolution imagery,

facilitating the differentiation between vines and weeds. Manned aircraft provide superior spatial resolution and real-time data but are costly and subject to airspace regulations. UAVs, though more cost-effective, cover smaller areas but provide detailed imagery beneficial for canopy analysis. Ground-based platforms, known as proximal sensing, involve sensors placed closer to the target, offering advantages in mobility, adaptability, and control [19]. They can be mobile, mounted on agricultural machinery or fixed using tripods. Proximal sensing meets both small- and large-scale monitoring needs, delivering high-resolution images without flight scheduling or weather constraints.

Each of the platforms mentioned can be equipped with different types of sensors; typical sensors are optoelectronic, such as LiDAR and RGB-D [20]. Both these types of sensors allow 3D color mapping which is crucial to quantify the geometric characteristics of plants like the structure of the canopy, the leaf area index (LAI) and the height of the plants. While LiDAR stands out for its robustness, accuracy, and high resolution in 3D canopy reconstruction, it may come with higher costs, complexity, and longer imaging times. As an alternative, RGB-D cameras have emerged as a solution. Using principles like Stereo Vision or Time-of-Flight (ToF), these sensors offer a balance between precision and affordability. Despite being less precise in generating 3D data compared to LiDAR, RGB-D cameras have gained popularity due to their cost-effectiveness and ease of use. The 3D point clouds provided by both types of sensors are invaluable in the precision agriculture sector. They allow to have a detailed view of the plant to evaluate their health and identify problems in growth or vegetative stress. These 3D images can be processed with computer vision and AI techniques to extract information of interest.

In literature, various works based on using remote sensing tools in precision agriculture are presented. In most of these, UAV platforms are considered due to their advantages compared to the other aerial platforms. For example in [21] the reliability of LAI estimation by processing dense 3D point clouds provided by a UAV-based multispectral imagery was evaluated; in [22] the spatial variability of biomass in a vineyard was studied using images produced by multispectral and RGB cameras on a UAV; in [23] an approach was proposed that exploits point clouds to detect the positions of the trunk of plants and evaluate their characteristics such as the height, the width and the volume of the canopy by exploiting a UAV equipped first with an RGB camera and in a second moment with a LIDAR. Although UAV technology provides detailed information, it doesn't have the necessary resolution to observe details such as leaves and fruits. This problem is faced by using terrestrial platforms. For example, in [24] an RGB-D camera system was used to reconstruct 3D models of vine plants to determine shoot volume; in [25] a methodology was developed for the automated segmentation of bunches of grapes in color images coming from an RGB-D camera mounted on board an agricultural vehicle; in [26] the variation in biomass following the trimming operation within a vineyard was evaluated using RGB-D images acquired by placing the sensor on a tripod; in [27], methods were developed for automated vine phenotyping, to estimate canopy volume, identify and counting bunches using an Intel RealSense RGB-D R200 imaging system. The work in [28] discusses the application of Machine Learning (ML) techniques in viticulture which mainly focus on the detection, counting and prediction of grape yield. Several methods in image analysis use convolutional neural networks (CNN) for image processing, which consist of developing segmentation, shape recognition, and feature extraction algorithms starting from natural images. For example, in [29], an algorithm for grapevine flower counting is developed to forecast crop yield. In [30],

an approach to segmenting UAV images is proposed to map diseased areas and guarantee the healthy state of the plant by continuous monitoring.

However despite the agricultural sector being an active research field, there is a lack of methodological frameworks for intelligent perception systems to improve vineyard management and final production. The current use of these types of systems faces several significant limitations such as the high cost of sensors, the difficult set-up and use of the system, the need for specialized personnel, the variability of crops, the reliability of the measurements provided and the varying environmental conditions, including the impact of climate change. In this context, this study aims to develop a new methodological framework for the use of intelligent systems in the agricultural sector able to overcome these bottlenecks and provide real support for farmers.

3. Research Methodology

The PhD proposal uses as research methodology the Design Science Research (DSR) one [31] which includes five main activities shown in Figure 1. This approach is tailored to address challenges and promote innovative solutions through an iterative process of designing, developing, evaluating, and improving artefacts. In this case, *the artefact is represented by a novel methodological framework for intelligent plant monitoring that can provide robust and reliable decision support to farmers*. The core of the proposed framework (Figure 1) focuses on overcoming inefficiencies and inaccuracies related to traditional vineyard management practices. Among the main objectives, it is important to integrate control over climate change and environmental parameters, while promoting sustainability practices. Fundamentals are also the cost-effectiveness, the ease of installation and use. The DSR ensures a continuous improvement cycle where at each iteration there is further refinement, based on practical tests in real scenarios and constant input from experts, such as agronomists. These experts contribute with baseline measurements and detailed feedback, which are essential for the evolution and validation of the framework.

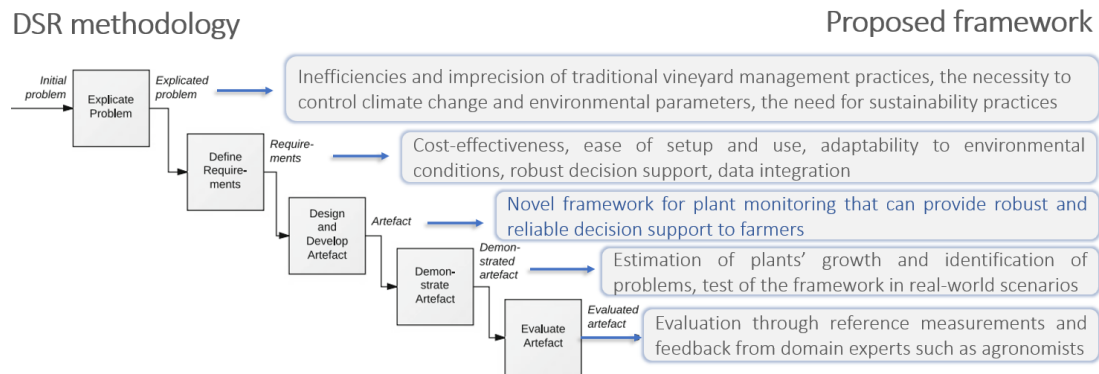


Figure 1: Design Science Research (DSR) methodology proposed in [31] and the proposed framework for its application.

4. Research questions

As stated in Section 3, the main artefact of this research proposal is to develop a new methodological framework for the improvement of plant monitoring by exploiting the multimodality of sensors and emerging artificial intelligence techniques. The methodology wants to enhance the ability to identify problems both in vegetation and fruits early, allowing timely and targeted interventions. This will not only help to quickly identify anomalies but will also try to understand the causes of problems, enhancing the ability to adequately respond to agricultural challenges and improve crop yields. Another important achievement is making the monitoring system more accessible to farmers, regardless of their technological skills, allowing all to use it without the need for specialized personnel and large economic investments. For example, the idea is based on the possibility of integrating low-cost sensors on agricultural machinery, e.g. tractors, which farmers own and use daily within the vineyard. In this way, the system is cost-effective and simple to set up, which is important for agricultural industries facing economic and operational obstacles. In this scenario, the following main research question is: **What framework can be developed to model and support the use of intelligent plant monitoring systems to provide robust and reliable decision-making in the precision viticulture sector?**

The answer to this question, is split into research sub-questions (RQs):

- **RQ.1** What are the key requirements for developing a low-cost and user-friendly framework that supports the heterogeneous nature of vineyards?
- **RQ.2** Which existing platforms and sensors are most suitable for integration into the framework to collect data on the phenotypic characteristics of vine plants?
- **RQ.3** What methods can be employed to effectively integrate and process multimodal data within the framework to support decision-making?
- **RQ.4** How can the framework be designed to address seasonal and environmental variations in vineyards through adaptive data processing techniques?
- **RQ.5** How can the effectiveness of the entire framework be evaluated in accurately describing the phenotypic characteristics of vine plants compared to traditional evaluation methods?

5. Research planning

During the PhD, the study will be divided into different phases. Six Work Packages (WP) have been identified, shown in the Gantt diagram in Figure 2. The next subsections will detail the planning of the three years.

5.1. First year: exploration and definition

The first year, which is ongoing, is dedicated to studying the state of the art of intelligent systems for target monitoring, with attention to the agricultural sector. This phase will provide a comprehensive overview of current technologies and research methods. Based on this analysis, the goal is to define the performance objectives of the monitoring system in alignment with the

desired outcomes. Furthermore, during this initial period, an in-depth analysis is conducted on the sensors available on the market to identify the most suitable ones for specific requirements. Scientific papers are currently being collected to write a review of the literature on intelligent systems developed for the agricultural sector, aiming to better define the context in which the PhD project falls.

5.2. Second year: development and implementation

During the second year, based on the performance objectives defined in the first year, the study will move to the development and implementation phase of the system within the laboratory. The primary goal will be to establish the setup for data acquisition. Following this, the implementation phase will extend to real-world contexts, such as an agricultural field, planning an acquisition campaign to identify key moments for plant growth analysis. The idea is to collect data simultaneously with agronomists, enabling the comparison of the data acquired by the system with that of experts to validate the effectiveness of the system itself. Once data acquisition is completed, the focus will shift to processing in order to define the most appropriate techniques for extracting useful information and highlighting particular issues.

5.3. Third year: validation and optimization

The last year of the PhD program will be dedicated to the validation of the results obtained. This process will involve various activities aimed at confirming the robustness and reliability of the developed system, as well as the consistency of the results with respect to the pre-established objective. In parallel, it will be important to outline any limitations or points of improvement of the systems, to identify areas where changes or upgrades can be made to optimize overall performance. This phase will represent a fundamental moment, as it will provide a fundamental verification of the practical usefulness of the developed system.

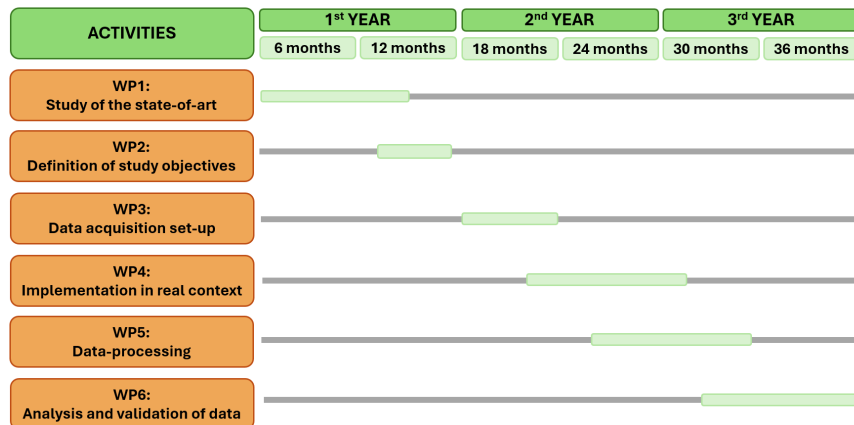


Figure 2: Gantt diagram of the PhD research project.

6. Research methods

In this section, some of the research methods that will be used in the study are explored. An integrated approach combining advanced technologies is expected to be used to improve the precision and effectiveness of vineyard monitoring and management. The following sections will address these methods.

6.1. Data acquisition

In the first six months of the PhD, ground-based platforms were investigated to pursue the outlined specific objectives. The capability of these platforms to provide close views of plants is particularly crucial for monitoring plant changes over time. This modality in data acquisition enables a detailed examination of plant growth dynamics, allowing for a more precise understanding of biomass variations throughout the monitoring period. Consequently, the chosen ground-based approach aligns with the project's goal of closely tracking and analyzing the evolving biomass of vine plants.

RGB-D cameras were investigated because of their advantages compared to LiDAR. They can capture color information like standard RGB cameras and also depth information using infrared sensors to measure the distance between the camera and each pixel in the image. These devices can operate on two principles:

- *Stereoscopic Vision (SV)*: try to mimic human vision by using two cameras facing the scene with some distance between them. The two images produced from these two sensors are compared and, since the distance between the sensors is known, these comparisons give depth information. An example is the Intel Realsense D435 [32] which is a USB-powered camera that has attracted increasing interest for its cost-effectiveness and its wide field of view, which allows to acquire information by analyzing a large section of the space under consideration, with a range of 10 m. This camera can operate in a variety of ambient light conditions and is particularly suited to systems operating at high motion speeds.
- *Time Of Flight (ToF)*: the camera uses the speed of light to calculate depth. A light emitted by the device is swept over the scene and the time required by the light to get back to the camera is then used to estimate the depth. An example is the Microsoft Kinect Azure [33] which is a camera for 3D modeling of environments, supported by advanced artificial intelligence algorithms and a series of additional software (SDK) that allow extending its functionality. It has seven microphones arranged in a circle, as well as the rich image function of two cameras (depth sensor camera and RGB camera) that can measure the colour and depth of the subject.

Comparing the Intel RealSense D435 and the Microsoft Azure Kinect, the latter better pursues the research objective. One significant distinction between the two cameras lies in their resolution. The Azure Kinect has a depth sensor resolution of 640×576 pixels, whereas the RealSense D435 features a higher depth resolution of 1280×720 pixels. However, the RealSense D435 captures more detailed but noisy and less accurate depth information compared to the Kinect, highlighting the latter's superior performance despite its lower resolution. Regarding the RGB sensor, the RealSense D435 has a resolution of 1920×1080 pixels, whereas the Kinect

Azure of 2048×1536 pixels. Another distinction is the orientation of their depth sensors: the RealSense D435 captures depth information vertically, while the Kinect Azure does the same horizontally. The field of view (FOV) varies between the RealSense D435 and Kinect Azure cameras, particularly concerning the RGB sensor:

- RealSense D435: Horizontal FOV of 42 degrees and a vertical FOV of 69 degrees.
- Kinect Azure: Horizontal FOV of 74.3 degrees and a vertical FOV of 90 degrees.

It is essential to note that FOV significantly impacts camera selection for specific applications as it influences the camera's vision range. A larger FOV enables more efficient image capture with increased data content, requiring fewer images to cover the entire sample. However, adjustments to FOV may impact other factors like resolution and imaging speed, necessitating a balanced consideration of these specifications. After this preliminary study, the Microsoft Azure Kinect has been selected for its robust depth-sensing technology and expansive field of view which can help to increase the precision and efficiency of image acquisition for grapevine analysis and management. This choice will not preclude the possibility of investigating or integrating further sensory technologies that can complement or enhance the intelligent system, providing additional layers of information and deeper insights.

6.2. Data processing

The acquired data will be processed at a software level to extract the most discriminating characteristics. This is accomplished through the use of signal and image pre-processing and processing algorithms. These algorithms play a crucial role in preparing the data to be suitable for the application of AI techniques. Indeed, machine learning and deep learning algorithms often require well-prepared input data containing relevant features. Depending on the type of sensor taken into consideration, in this case, the RGB-D camera Azure Kinect produced by Microsoft, it is also possible to have a representation of 3D space through point clouds which consist of a set of points in three-dimensional space that represent the surface of an object or an environment. Generally, to take full advantage of the information within them, they require careful pre-processing that includes several operations, depending on the application's specific needs, such as filtering and registration. The latter is a fundamental process for data analysis as it allows for the combination of multiple acquisitions of 3D data from different positions or points of view. This process is crucial for obtaining an accurate and complete representation of the object or environment under investigation. This research aims to combine the depth data captured on field and deep learning techniques to have a complete and detailed study of the vine plant canopy and the estimation of volume changes over time. The processing of the data coming from the sensors will be realized using different tools and programming languages, such as:

- MATLAB: specialized software, that provides several tools particularly useful for applying preprocessing, processing, and machine learning algorithms to data.
- C++: programming language that provides a wide range of features that make it suitable for developing high-performance applications and implementing complex algorithms. Important is the Point Cloud Library (PCL) which is an open-source library for working with point clouds.

- Python: programming language in the field of machine learning. Thanks to its specialized libraries, such as PyTorch, it offers powerful tools for creating, training, and implementing neural networks to process data from sensors.
- Cloud Compare: open-source software for managing 3D meshes, equipped with tools specifically tailored for analyzing point clouds.

In conclusion, the joint use of sensor multimodality, processing algorithm and machine learning opens many possibilities for obtaining detailed and useful information based on the specific application context. The continuous evolution of these technologies offers interesting prospects for the future, allowing for the development of increasingly intelligent systems capable of understanding and interacting with the world around them.

6.3. Decision support

The decision support is fundamental in the development of powerful intelligent systems for monitoring objectives. This method represents the keystone to guaranteeing an accurate and reliable analysis of the results obtained from the previous stages. The comparative analysis of data and results will be done to ensure a complete understanding of how the intelligent system works and its generalization abilities. This knowledge involves different insights such as performance metrics evaluation, the influence of different data sources, different features and models. In addition, the validation of results is a critical step in ensuring the quality of the system. Reference measurements and feedback from domain experts such as agronomists can help validate the system outputs. This validation step can establish the reliability and effectiveness of the framework compared to traditional practices, thereby enhancing its utility and impact in real-world applications.

In conclusion, the decision support method translates the findings into useful and relevant information for agricultural experts. The validity and reliability of the results are fundamental to helping build confidence in the plant monitoring system among stakeholders and end-users in the precision agriculture sector.

7. Conclusion

Intelligent perception systems are a key technology to improve significantly production processes in various fields. This PhD proposal aims to develop a novel methodological framework for target monitoring within the precision viticulture sector that could support farmers in vineyard maintenance operations and making predictions about the harvest. The core of this study stands in the in-depth analysis of how data have to be acquired, processed, and modelled to provide decision support and achieve the specific objectives of the considered context. The comprehension of the entire decision-making mechanisms and the relations between the data (in terms of their quality, and quantity) and the chosen architectures for data processing, are fundamental to understanding the applicability of these perception systems in real contexts. The whole approach can be extended beyond the considered application environment, offering monitoring possibilities in other contexts as well.

8. Acknowledgments

I would like to express my gratitude to my PhD tutors, Dr.ssa Tiziana Rita D’Orazio and Prof. Cataldo Guaragnella, for their guidance, support, and inspiration throughout my doctoral journey. I would also like to thank Prof. Jelena Zdravkovic for her fundamental support as tutor in the doctoral consortium during the 36th International Conference on Advanced Information Systems Engineering (CAiSE).

References

- [1] E. Rauch, C. Linder, P. Dallasega, Anthropocentric perspective of production before and within industry 4.0., *Computers & Industrial Engineering* 139 (2020) 105644. doi:<https://doi.org/10.1016/j.cie.2019.01.018>.
- [2] Z. Les, M. Les, *Machine Perception—Machine Perception MU*, Springer, 2020.
- [3] H. Niu, F. Yin, E. Kim, W. Wang, D. Yoon, C. Wang, N. Kim, Advances in flexible sensors for intelligent perception system enhanced by artificial intelligence, *InfoMat* 5 (2023) 5. doi:<https://doi.org/10.1002/inf2.12412>.
- [4] M. Parasher, S. Sharma, A. K. Sharma, J. P. Gupta, Towards human-like machine perception 2. 0., *International Review on Computers and Software* 5 (2010) 476–488.
- [5] M. Molina, What is an intelligent system?, *arXiv* (2020). doi:<https://doi.org/10.48550/arXiv.2009.09083>.
- [6] K. Saleh, S. Szénási, Z. Vámosy, Occlusion handling in generic object detection: A review, in: *2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2021, pp. 000477–000484. doi:10.1109/SAMI50585.2021.9378657.
- [7] R. Akerkar, Intelligent systems: perspectives and research challenges., *CSI Commun*, 2012, pp. 4–9.
- [8] J. V. Stafford, Implementing precision agriculture in the 21st century, *Journal of agricultural engineering research* 76 (2000) 267–275. doi:<https://doi.org/10.1006/jaer.2000.0577>.
- [9] U. Nations, *World population prospects 2022*, 2022. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf [Accessed: (2023-05)].
- [10] R. Gebbers, V. I. Adamchuk, Precision agriculture and food security, *Science* 327 (2010) 828–831.
- [11] J. Arnó Satorra, J. A. Martínez-Casasnovas, M. Ribes-Dasil, J. R. Rosell, Precision viticulture. research topics, challenges and opportunities in site-specific vineyard management, *Spanish Journal of Agricultural Research*, 2009, vol. 7, núm. 4, p. 779-790 (2009). doi:<https://doi.org/10.5424/sjar/2009074-1092>.
- [12] L. Comba, A. Biglia, D. R. Aimonino, P. Gay, Unsupervised detection of vineyards by 3d point-cloud uav photogrammetry for precision agriculture, *Computers and electronics in agriculture* 155 (2018) 84–95. doi:<https://doi.org/10.1016/j.compag.2018.10.005>.
- [13] A. Barriguinha, M. de Castro Neto, A. Gil, Vineyard yield estimation, prediction,

- and forecasting: A systematic literature review, *Agronomy* 11 (2021). doi:10.3390/agronomy11091789.
- [14] M. Tariq, M. Ahmed, P. Iqbal, Z. Fatima, S. Ahmad, *Crop phenotyping*, Springer, 2020, pp. 45–60.
- [15] J. Campos, F. García-Ruíz, E. Gil, Assessment of vineyard canopy characteristics from vigour maps obtained using uav and satellite imagery, *Sensors* 21 (2021). doi:10.3390/s21072363.
- [16] A. Matese, P. Toscano, S. Di Gennaro, L. Genesio, F. Vaccari, J. Primicerio, C. Belli, A. Zaldei, R. Bianconi, B. Gioli, Intercomparison of uav, aircraft and satellite remote sensing platforms for precision viticulture, *Remote Sensing* 7 (2015) 2971–2990. doi:https://doi.org/10.3390/rs70302971.
- [17] K. Ennouri, A. Kallel, et al., Remote sensing: an advanced technique for crop condition assessment, *Mathematical Problems in Engineering* 2019 (2019).
- [18] H. Jafarbiglu, A. Pourreza, A comprehensive review of remote sensing platforms, sensors, and applications in nut crops, *Computers and Electronics in Agriculture* 197 (2022) 106844. doi:https://doi.org/10.1016/j.compag.2022.106844.
- [19] R. P. Sishodia, R. L. Ray, S. K. Singh, Applications of remote sensing in precision agriculture: A review, *Remote Sensing* 12 (2020). doi:10.3390/rs12193136.
- [20] F. Vulpi, R. Marani, A. Petitti, G. Reina, M. A., An rgb-d multi-view perspective for autonomous agricultural robots, *Computers and Electronics in Agriculture* 202 (2022) 107419. doi:https://doi.org/10.1016/j.compag.2022.107419.
- [21] L. Comba, A. Biglia, D. Ricauda Aimonino, C. Tortia, E. Mania, S. Guidoni, P. Gay, Leaf area index evaluation in vineyards using 3d point clouds from uav imagery, *Precision Agriculture* 21 (2020) 881–896.
- [22] P. Catania, M. V. Ferro, E. Roma, S. Orlando, M. Vallone, Assessment of vine and cover crop vegetation indices using high-resolution images acquired by uav platform, in: *Conference of the Italian Society of Agricultural Engineering*, Springer, 2022, pp. 447–455.
- [23] M. Cantürk, L. Zabawa, D. Pavlic, L. Klingbeil, Uav-based individual plant detection and geometric parameter extraction in vineyards, *Frontiers in Plant Science* 14 (2023) 1244384.
- [24] H. Moreno, J. Bengochea-Guevara, A. Ribeiro, D. Andújar, 3d assessment of vine training systems derived from ground-based rgb-d imagery, *Agriculture* 12 (2022). doi:10.3390/agriculture12060798.
- [25] R. Marani, A. Milella, A. Petitti, G. Reina, Deep neural networks for grape bunch segmentation in natural images from a consumer-grade camera, *Precision Agriculture* 22 (2021) 387–413.
- [26] A. Bono, R. Marani, C. Guaragnella, T. D’Orazio, Biomass characterization with semantic segmentation models and point cloud analysis for precision viticulture, *Computers and Electronics in Agriculture* 218 (2024) 108712.
- [27] A. Milella, R. Marani, A. Petitti, G. Reina, In-field high throughput grapevine phenotyping with a consumer-grade depth camera, *Computers and Electronics in Agriculture* 156 (2019) 293–306. doi:https://doi.org/10.1016/j.compag.2018.11.026.
- [28] L. Mohimont, F. Alin, M. Rondeau, N. Gaveau, L. A. Steffanel, Computer vision and deep learning for precision viticulture, *Agronomy* 12 (2022) 2463. doi:https://doi.org/10.3390/agronomy12102463.

- [29] F. Palacios, G. Bueno, J. Salido, M. P. Diago, I. Hernández, J. Tardaguila, Automated grapevine flower detection and quantification method based on computer vision and deep learning from on-the-go imaging using a mobile sensing platform under field conditions, *Computers and Electronics in Agriculture* 178 (2020) 105796. doi:<https://doi.org/10.1016/j.compag.2020.105796>.
- [30] M. Kerkech, A. Hafiane, R. Canals, Vine disease detection in uav multispectral images using optimized image registration and deep learning segmentation approach, *Computers and Electronics in Agriculture* 174 (2020) 105446. doi:<https://doi.org/10.1016/j.compag.2020.105446>.
- [31] P. Johannesson, E. Perjons, A Method Framework for Design Science Research. In: *An Introduction to Design Science*, Springer, 2014. doi:https://doi.org/10.1007/978-3-319-10632-8_4.
- [32] Intel, Intel realsense depth camera d435, 2022. <https://www.intelrealsense.com/depth-camera-d435>[Accessed: (2023-05)].
- [33] Microsoft, Azure kinect sensor sdk, 2022. <https://learn.microsoft.com/en-us/azure/kinect-dk/sensor-sdk-download> [Accessed: (2023-05)].