Research on Grapevine Recognition, Manipulation and Winter Pruning Automation

Coordinator:

Ch.mo Prof. Paolo Ajmone Marsan

Candidate:

Tao TENG

Matriculation n: 4915014

UNIVERSITÀ CATTOLICA DEL SACRO CUORE

Sede di Piacenza

Dottorato di ricerca per il Sistema Agro-alimentare

Ph.D. in Agro-Food System

Cycle XXXV

S.S.D. AGR/09

Research on Grapevine Recognition, Manipulation and Winter Pruning Automation

Coordinator:

Ch.mo Prof. Paolo Ajmone Marsan

Tutor:
Prof. Matteo GATTI
Dr. Fei CHEN

Candidate:

Tao TENG

Matriculation n: 4915014

Academic Year 2021/2022

Table of Contents

List of Tables

List of Figures

Acknowledgments

Seems like yesterday when I started this experience but today is already past three years, and I have acquired a new great experience in agriculture robotics!

First of all, I want to thank my supervisors Prof. Matteo Gatti and Dr. Fei Chen, who gave me the possibility to carry out the doctorate school. Thanks for the support, guidance, and trust and for making the work of this Thesis very smooth.

I am also very grateful to the people who believed in and supported me during these three years of doctorate school and were important to me and so for the realization of this Thesis. Many thanks to the APRIL team from Istituto Italiano di Tecnologia and colleagues for their help and support over these three years. I am grateful to Prof. Stefano Poni, Prof. Darwin Caldwell (Istituto Italiano di Tecnologia), and Dr. Claudio Semini (Istituto Italiano di Tecnologia) for providing support to me.

Finally, I would like to pay special thanks to my parents and friends for their continuous encouragement during periods of self-doubt when things didn't go smoothly. I also want to pay gratitude to myself for the extraordinary gift of hope, strong optimism, and childish enthusiasm. With encouragement, determination, and hope, I was able and will continue to get off the ground and get things done.

Abstract

The main purpose of this thesis is to develop and test autonomous robot systems to realize the perceptual identification and inference of grapevines and then perform the pruning manipulation with the aim to automate grapevine winter pruning.

Grapevine winter pruning is an important practice in viticulture aimed at improving plant growth, productivity, and fruit quality in the next harvest season. Winter pruning requires skilled seasonal workers which are becoming less available in several wine districts, however, the low efficiency of the manual pruning method is no longer suitable for the development of the wine industry. Robotic pruning is a potential solution to deal with the issue of labor shortage and high vineyard management costs. However, numerous challenges are involved in the successful adoption of robotic pruning technologies. As a result, the purpose of this thesis is to perform research on autonomous selective winter spur pruning by using viticulture, pattern recognition, and robotics technologies.

Such a process of automation mainly includes spur detection and recognition, determination and location of pruning points, robotics pruning manipulation, and the development of robotic pruning systems according to the technical analysis of the grapevine pruning robot. Among all the processes, the related research of spur detection and recognition and robotics pruning manipulation is far from the mature stages. Therefore, this thesis mainly carries out in-depth research on these aspects, and the specific contents for research are reported below:

The first topic is to provide hierarchical guidelines for a translation of the human cognition process into algorithms for automated grapevine recognition and robotized pruning.

Multiple combinations of integrated general rules driving the winter pruning process with specifically conceived case studies composed a dedicated database for developing a perception system for the recognition of grapevine architecture and identification of target regions and pruning points.

The second topic is to integrate electric shears into a robot manipulator to autonomously perform the pruning points approach from the output of the first action. We proposed a novel planning framework "Three-phase Approach Planning" (TAP) to guide shears to approach pruning points. The designed motion planner is highly manipulative to approach the target pruning points precisely while avoiding collisions with the grapevine cordon and canes, so that robot doesn't disrupt the condition of the grapevine. Otherwise, the robot can hurt the dormant buds, hence impacting vine growth and productivity.

The third topic is to study the pruning cut control technology of the grapevine pruning robot for reducing mechanical damage to the grapevine and improving the success rate of pruning. To achieve a human-like interactive pruning cut, we designed a learning-based compliant controller to learn impedance profiles from human demonstration, which reduces the possibility of the blade of shears getting stuck in canes.

The main contributions of this thesis are a comprehensive methodology to analyze the grapevine winter pruning operation including perception and manipulation paradigm and hereby design an overall scheme of the grapevine pruning robot system with the combination of a mobile platform and robot manipulator. In addition, it verifies the feasibility and the value in engineering application of the entire scheme of the grapevine pruning robot by means of simulation and experiments, confirming the practical value of related technology in engineering.

Contributors and Funding Sources

Symbols

q_{wb}	state variables of wheeled manipulator *Rolling Panda*
q_m	state variables of mobile base
q_w	wheel joints
q_n	joint vector for the manipulator
μ	distance between the driving wheels and the mobile platform geometric center
ρ	distance from the mobile platform rotation center to the center of mass of the mobile platform
R	the radius of the wheels
J	Jacobian matrices
q_i	ith joint angle
α_i	ith link twist
d_i	ith link length
a_i	ith link offset
$\dot{\theta}_l$	the left wheel velocities
$\dot{\theta}_r$	the right wheel velocities
ξ	joint position vector for all joints of *Rolling Panda*
$\dot{\xi}$	joint velocity vector for all joints of *Rolling Panda*
$\dot{x}$	end effector velocity vector of *Rolling Panda*

M	inertia matrix
C	centripetal and Coriolis matrix
G	gravitational vector
τ	torque vector
q_{hk}	joint position vector of legged manipulator *HyQReal-Kinova*
v_{hk}	Cartesian space velocity vector of legged manipulator *HyQReal-Kinova*
P_x	position of trajectory in Cartesian space
P_v	velocity of trajectory in Cartesian space
P_a	acceleration of trajectory in Cartesian space
x_s	position of the interpolation initial point
v_s	velocity of the interpolation initial point
a_s	acceleration of the interpolation initial point
x_e	position of the interpolation end point
v_e	velocity of the interpolation end point
a_e	acceleration of the interpolation end point
$\boldsymbol{R}_e^r$	3D rotation matrix described by the coordinate frame $\{e\}$ with respect to $\{r\}$
$\boldsymbol{t}_e^r$	translation vector described by the coordinate frame $\{e\}$ with respect to $\{r\}$
$\boldsymbol{H}_{EE}^{ROB}$	homogeneous transformation matrices described by the coordinate frame $\{EE\}$ with respect to frame $\{ROB\}$
φ	rotation angle around the z-axis
W	weighting diagonal matrix
x_{envi}	position of environment
K_d	desired Cartesian stiffness

D_d desired Cartesian damping

Λ Cartesian inertial

ψ Cartesian Coriolis and centrifugal matrix

η demonstration data

1 Introduction

1.1 Background

1.1.1 Motivation

Grapevine is a distinctive crop of the Mediterranean Basin, which is considered to be one of the major fruit crops based on hectares cultivated and economic value [1]. Viticulture has a relevant economic impact in all regions and grape is one of the most valuable products in the Italian tree fruit industry, and the wine industry ranks first in the agro-food chain [2]. Grapes are climbing vines, so they need support to grow over, such as a trellis or pergola. Vineyards in modern wine districts are most frequently based on Vertically Shoot Positioned (VSP, Figure 1.1) canopies: a common and widely used trellis system often requiring high management costs. Winter pruning (WP) is the most fundamental of all vineyard management tasks. Pruning is a primary tool used by grape growers to manipulate vine size and shape which helps to regulate crop load and maintain vine balance [3, 4, 5]. Pruning commences after the harvest, as soon as the leaves begin to fall, and may continue until late March during a relatively long period known as "dormancy" [6, 7]. In a hand-run vineyard, grapevine winter pruning is the second largest labor expense for tree fruit field production after harvesting, accounting for 20% or more of total annual costs (the fixed investment breakdown

for each year, such as land, trellis, irrigation setup, and others, is excluded) [8]. Any grower looking to the future needs to take a close look at labor inputs. Going forward, it is critical to minimize dependence on labor for the long-term sustainability of this industry. Meanwhile, worker safety is another issue in manual vineyard management. Mechanization can limit costs, ameliorate growers' living conditions and improve wine industry sustainability [4].

Figure 1.1: Vertical shoot position (VSP) trellis system.

Grapevine winter pruning is applied when vines are still dormant by removing more than 90% of canes and retaining an appropriate bud-load along the cordon [4]. The primary goal of grapevine winter pruning is to figure out the number and location of the nodes remaining over the winter and involves the removal of plant organs such as spurs and excess one-year-old canes from the previous year's growth. These remaining nodes will grow into new shoots in the next season, and grapes will grow on these new shoots. The structure of a dormant grapevine is illustrated in Figure 1.2. Adopting an appropriate pruning intensity can also improve yield and fruit quality in terms of size, sugar, color, and flavor [6]. It is a
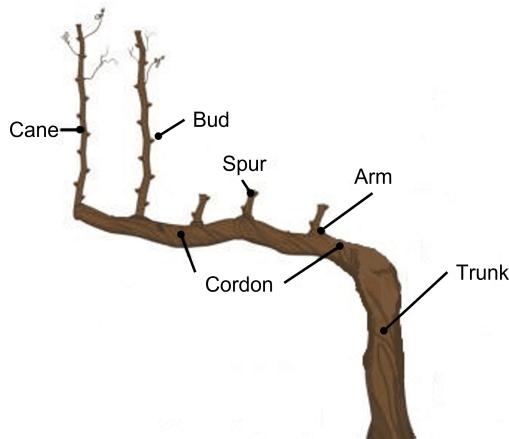
Figure 1.2: General view of the aerial organs of a dormant grapevine.

highly labor-intensive task that requires skilled workers during the winter season, which are becoming less available.

Different pruning strategies have been extensively studied in various grape growing regions and grape varieties to achieve sustainable vine vegetative and reproductive growth, often referred to as vine or vineyard balance [9]. From a physiological perspective, winter pruning affects shoot growth, canopy function, and productivity; bud-load influences leaf-to-fruit ratio at harvest, a reliable indicator of vine balance [10]. There are two basic systems of pruning. How to prune a grapevine will depend on varietal issues especially concerning which part of its canes can be expected to bear fruitful buds. Based on pruning length, the operation is termed cane–pruning (CWP, more than 5 nodes/bearing unit) or spur-pruning (SWP, less than 3 nodes/bearing unit). Technically, cane pruning implies the removal of all previous-year canes except one or two long canes for next season's cropping:

- Select the best-placed canes for renewal and cut them back to 2 nodes each.

- Select the best-quality canes for fruiting.

- Decide how many buds the whole vine should bear.

- Cut back the fruiting canes according to their vigor and size, and that of the vine.

- Remove the rest of the wood.

- Tie the fruiting cane to the setting wire.

Spur pruning shortens several previous-year canes to 2-3 nodes; therefore, depending on several factors, 6 to 14 cuts per meter of the row are required.

- Select the best-placed, most fruitful canes.

- Decide how many buds the whole vine should bear.

- Cut enough canes back to 2-bud spurs.

- Remove the rest of the wood.

However, pruning operations are still primarily dependent on human labor. For each spur, the grower has to decide which cane will keep and then prune down to one, two, or even three buds. For weak spur positions, only one bud will be retained for that season. A healthy cane, usually about 1 cm in diameter, can support two buds. If done properly, the following season the new canes will be pruned to two buds. The operation is labor-intensive and costly, and the pruning decisions may vary from person to person based on the skills and experience of the individual. The availability of farm labor is becoming an issue for growers, especially for time-sensitive and labor-intensive operations [11].

With the adoption of mechanization and automation, the agricultural sector has witnessed a significant increase in production efficiency. To reduce labor costs, mechanical pruning is an effective method that has played an important role in the grape processing industry. Mechanical systems supporting pruning operations in vineyards during the dormant season include (in sequence) mechanical pre-pruning and manual pruning follow-up [12], as shown in Figure 1.3. Mechanical pruning currently consists of a machine-executed non-selective

shortening of all canes followed by a selective manual follow-up allowing a 50-70% reduction of the total work-load associated with spur pruning [4].

(a)   (b)

Figure 1.3: Pruning sequence in a mechanized VSP-trained vineyard: (a) mechanical pre-pruning; (b) manual pruning follow-up.

Research and development of mechanical pruning have been ongoing since the mid-1970s [13] and it alone has further reduced labor costs. However, the lack of specificity in retained nodes may cause the vines to be overcropped (unbalanced) with poor fruit quality [14, 15]. While non-selective vineyard operations can be automated to increase productivity and reduce costs, there are numerous selective operations that are challenging to automate. This limitation of selective pruning capability only provides a partial solution as additional follow operations are often required to complete the task which further increases production cost. Grapevines are perennial plants with indeterminate growth habits leading to canopies that are too complex to analyze even for skilled human eyes, let alone by computer vision algorithms. Therefore, a robotic pruner as a follow-up operation after mechanical pre-pruning could be a practical solution. The development of automated robotic pruning as the mechanical pruning follow-up operation would further lower labor costs and increase specificity in retained node quantity and quality. Robots have the potential to represent a revolution up-

scaling in forthcoming innovative farming systems providing automated solutions combining intelligent robot perception and manipulation.

1.1.2 Challenges

Modern agriculture is confronting a series of challenges related to the adoption of new technologies to further improve the sustainability, profitability, and resilience of the industry. Many of these challenges are related to the labor-intensive nature of particular processes (e.g., harvesting, pruning, and planting) in the production of high-value specialty crops since they rely on a skilled seasonal workforce which sometimes is in limited availability[16, 17]. A large amount of research work has been done to develop automated or semi-automated robotic solutions to systematically analyze the vegetation, thus improving the efficiency and productivity of some of those processes [18]. Enabling robots in viticulture is challenging due to the nature of the environment which is usually more distant and unstructured as compared to a strictly structured factory operating environment [19, 20].

Unlike traditional mechanization, robotic solutions supplement human labor by performing autonomous, highly selective operations such as autonomous precision seeding [21], robotics harvesting [22], pesticide spraying [23]. Nevertheless, due to the particular knowledge required and the unstructured nature of the work environment, there are few robots capable of doing grapevine winter pruning [24]. For a robotic pruning system to work successfully, a number of elements including sensing or perception, planning, and control must be incorporated. These challenges have prevented pruning robots from achieving widespread deployment, there are no commercially available robotic pruning systems for vineyard applications.

Numerous variables, including sensing precision, processing speed, obstacle avoidance, path planning, and pruning sequencing, might influence how effective robotic pruning is [20, 25, 26, 27]. Fast and efficient perception systems need to be developed to increase the

precision of detection and recognition. Normally, the first step of automated pruning is to find the target organ and identify the pruning point location. A proper perception technique is essential to detect the target pruning region in the grapevine, and then select undesirable canes and pruning points based on the desired pruning rules to conduct selective pruning. Once the system determines the pruning points, the system must move a pruning tool to the desired pruning point, an operation that demands high precision, collision avoidance, and error robustness. For robotic pruning problems, approaching the target pruning point when eye-in-hand cameras are used is often divided into two steps [28, 29]. The first step involves moving from the overview waypoint V_0 to an approach waypoint V_1 where a spur is centered in the image. This waypoint V_1 is identified using the approach direction θ, defined as the angle the robot should use to approach the target points from, and distance d. The second step uses an interactive controller [28, 29] to move towards the pruning point until it is reached.

Robot visual perception and object detection in the grapevine winter pruning applications both refer to detecting borders of objects in an image and recognizing its interesting regions. Sensor-based organ recognition in grapevines as well as automated pruning activities is the key element of the automated pruning system [30, 31]. Machine vision is a system combined with sensors and algorithms to collect data on the target objects. Machine vision sensing has been employed in numerous agricultural activities for several decades[32, 33, 34]. Many different sensors have been used in a machine vision system for the detection of agricultural objects, e.g., cameras and Lidar sensor [35, 36, 37]. In specific applications like pruning, thinning, or harvesting, however, where the robot must perceive the plant in order to act on it, there has been only limited success. Focusing on grapevine pruning, it is necessary to have a deeper semantic comprehension of the scenario in order to accurately identify the pruning region. For instance, segmenting canes and spurs from the grapevine structures and accurately measuring important topological metrics like bud distribution and cane lengths. Pruning and perception-based agricultural robots in general have always been challenging for obtaining a comprehensive semantic map of plants in the field[38, 39]. Additionally, variable

illumination conditions and occlusion significantly influence the detection performance [40].

Manipulators currently are widely used in controlled factory environments where the workspace is structured to automate tasks that were once done manually, such as car assembly factories [41]. However, automation of tasks in the field has so far had limited success. This means that robots must be able to operate autonomously in non-artificial surroundings in contrast to robots working in factories where the environment is structured to meet the demands of robots. Unfortunately, outdoor agricultural environments are among the most difficult environment for robots to operate in. The vineyard is unstructured and complicated, with narrow spaces accessible within plant organs for maneuvering, resulting in numerous engineering challenges in developing a system with less spatial requirements. As the pruning cut is made close to the cordon, the pruning manipulator is more likely to collide with the trunk cordon, and cane, which could damage a pruning robot as the cordons are usually stiff and rigid near the trunk. Due to the unstructured and dynamic nature of the agricultural environments, advanced planning algorithms are required to compute a collision-free trajectory navigating the robot manipulator, which should be highly manipulative, to the target pruning points, even when pruning points are close to the cordon and the vine head.

From a multidisciplinary perspective, the main challenge of the thesis is developing and testing innovative robotic technologies for grapevine winter pruning automation. The successful operation of the robotic pruning system depends on grapevine detection, motion planning, approaching sequencing, and pruning manipulation.

1.1.3   State of the Art

In the past decades, research on the development and use of robotic systems for various agricultural tasks has been extensively researched by the scientific community. Bac et al. [42] discussed 50 different robotic harvesting systems developed up to 2014. A recent review by Tinoco et al. [26] focused on robotic manipulators for tree pruning and harvesting,

which mainly highlighted five harvesting and two pruning manipulators that were primarily featured as case studies.

<div align="center">(a)   (b)</div>
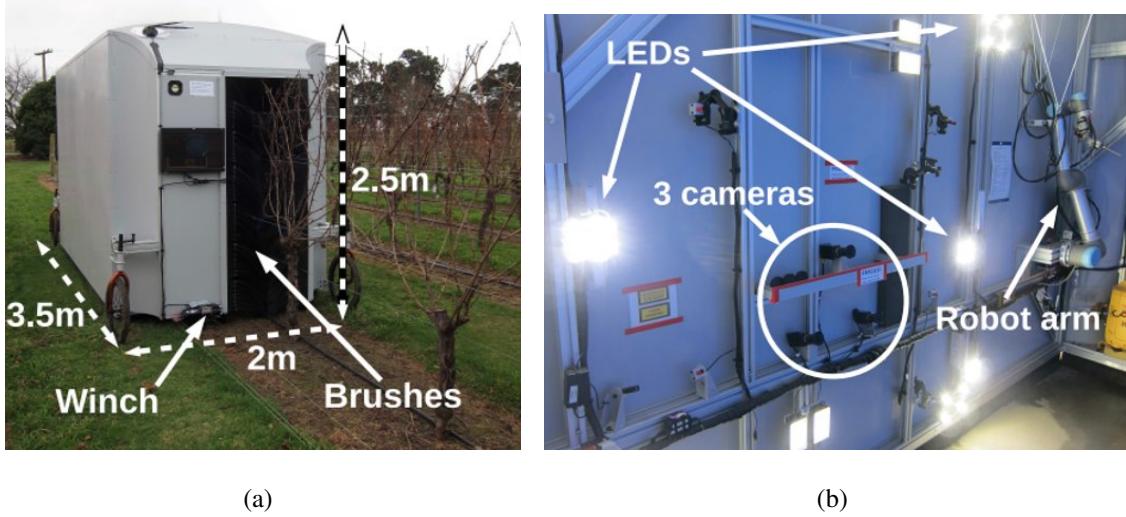
Figure 1.4: Prototype of a grapevine pruning platform, (a) from the outside; (b) from the inside.

However, robotics in viticulture is still at a prototype stage despite the preliminary experience in plant monitoring and operating solutions for foliar applications, mowing, table grape harvest, and grapevine pruning [43]. Despite the obvious advantage of automated pruning and the underlying commercial benefit, automated pruning has not received much attention when compared to harvesting. Automation of the harvesting process is well-documented[44, 45], and only a few attempts of fully automated pruning are known [19, 26, 27, 31]. Some machines are able to carry out spur pruning partially, but manual work is still required to accomplish the remaining tasks [46].

Traditionally, robotic pruning is understood in terms of two related subproblems: perception and manipulation. The perceptual component estimates the position and orientation (pose) of the object to be pruned. The manipulation component reasons where/how to move the manipulator into a pruning configuration and execute the pruning cut. Robotics pruning tasks can be seen as a special case of grasping. The difference to a normal grasping model is

(a)	(b)
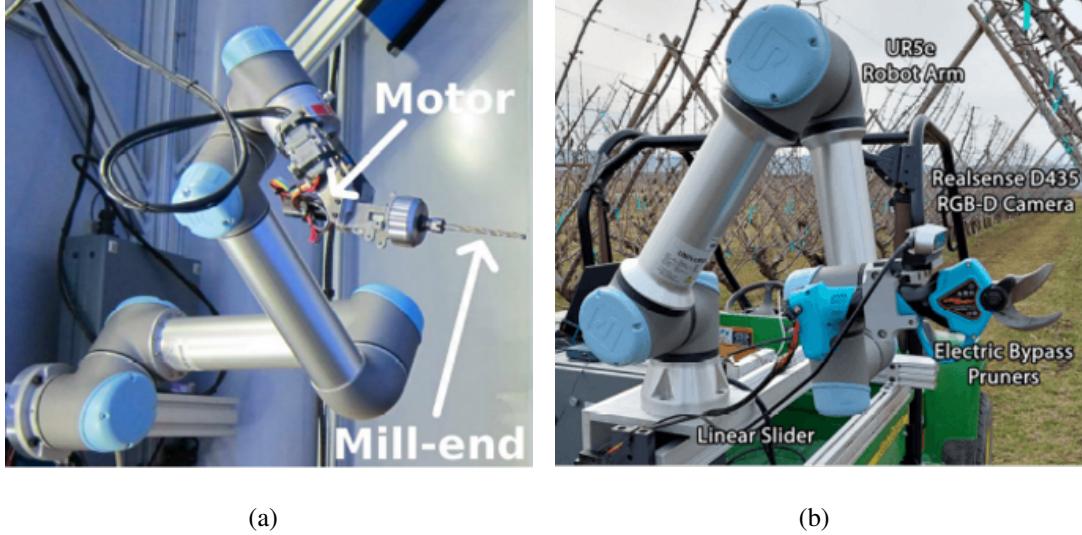
Figure 1.5: Pruning setup: UR5 Manipulator with (a) router mill-end end-effector and (b) shears.

that these methods usually work with big objects with a defined body to grasp, for example, using a cylinder to estimate the geometric shape of the grasped object. Whereas to prune a grapevine, the robot has to be really precise to fit a thin cane into the cutting tool. A close approach to grapevine winter pruning is tree pruning [47, 48, 49].

The prior work on robotic pruning remains relatively sparse, though recently there has been an increase in interest (some prototypes are shown in Figure 1.4 and Figure 1.5 [50, 51]), covering aspects such as tree modeling for pruning point detection [52, 53], pruning manipulator design [51, 54, 55, 56], path planning [57], and manipulator control [19, 58]. Some examples of research prototype end-to-end systems for fruit-related pruning include Botterill et al. [50], vision system [46], all of which focus on grapevines. There has also been work on automated pruning for landscapes/gardens, such as the Trimbot2020 system [59] that focuses on the development of intelligent outdoor hedge, rose, and bush trimming capabilities, allowing the robot to navigate over varying garden terrain, approaching hedges to restore them to their ideal tidy state, and approaching topiary-styled bushes to restore them to their ideal shape.

A tractor-mounted robotic grapevine pruning equipment has been developed in New Zealand [50]. The system can reconstruct the three-dimensional model of the vine using multiple stereo cameras and identify various plant features to figure out pruning points on canes. So pruning is made and specific canes are chosen by an operating manipulator mounted on a platform that moves along the row, by using Bidirectional RRT (BiRRT) [60] for motion planning with open-loop control and Euclidean distance between pruning points for optimizing the cut sequence. This process is carried out in a controlled environment inside a box, which covers completely the whole plant from its surroundings, resulting in constant illumination and background. Moreover, cane detection, scalability, and performance of the 3D reconstruction of grapevine, vineyard trials, and computational performance were evaluated, and while it did not work well enough to replace human pruners, it did demonstrate many technologies that will enable pruning to be fully automated. Corbett-Davies et al. [61] describe the AI system for deciding where to prune, the system uses stereo cameras for detecting canes, and it can retain spurs with an adequate length. This commercial prototype system uses stereo cameras to identify canes, which are then cut by shears equipped with robot arms. Spur pruning requires different canes to be shortened at different heights. A task-based kinematic design of a grapevine pruning manipulator was kinematically designed for grapevine pruning of grape vineyards through a systematic seven-stage design procedure[56]. Finally, the manipulator's workspace and kinematic performance were evaluated via simulations. Although dynamics, vibration, control, etc. need to be further studied, this research offers a foundation for the task-based kinematic design of a pruning manipulator. The systems mentioned above are mainly prototypes and a solution to fully automate spur pruning in vineyards is still required. Moreover, if these solutions were tested for the accuracy of pruning and the optimization of engineering processes, seasonal effects on grapevine growth, productivity, canopy efficiency, and fruit composition were never assessed.

1.2 Research Objectives

The general purpose of this thesis is to develop and test innovative robotic mobile manipulation technologies for grapevine winter pruning automation, which is definitely a challenging task. Having defined the above state of the art as well as the general purpose, the thesis targets the following strongly interdisciplinary specific aims under a strong multidisciplinary approach. The key technologies involved are summarized in Figure 1.6.

i) Translating human cognitive processes of grapevine spur winter pruning into deep learning techniques. Recognition of grapevine architecture and identification of target regions (TRs) and cutting points (CPs) will be developed and tested based on a dedicated database of case studies representing the impact of anatomical and cultural factors on grapevine architecture.

ii) Producing solutions for integrating electric shears into a robot manipulator, to autonomously perform the pruning point approach from the output of the perception system. Furthermore, a set of highly manipulative planning strategies is essential, which should approach pruning points precisely while avoiding collisions with the grapevine cordon and canes, so that robot doesn't disrupt the grapevine condition.

iii) Theoretically advance the state of the art of interactive grapevine pruning cut control by studying imitation learning and compliant control technology for reducing mechanical damage to the grapevine and improving the success rate of pruning.

Remark 1 *The research of this thesis will focus on spur winter pruning, which is suitable for many grape varieties; grapevine winter pruning will specifically refer to spur pruning in the present manuscript.*
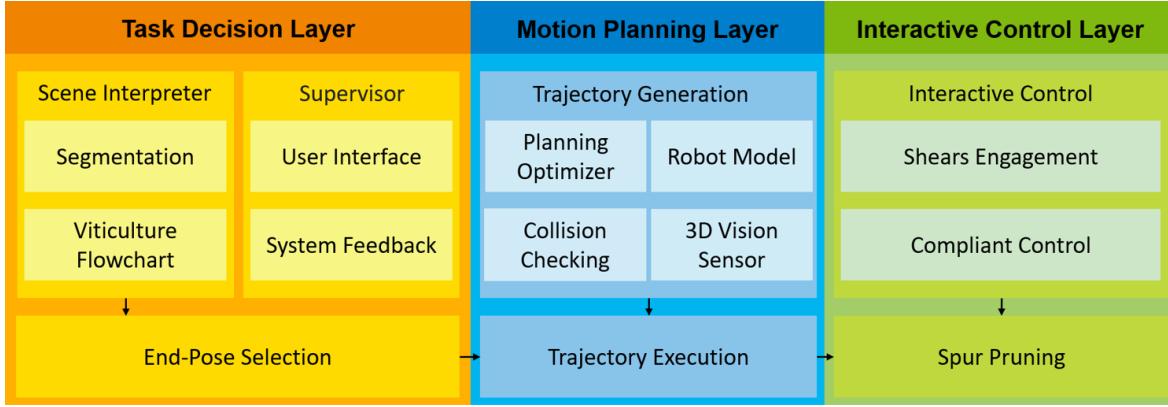
Figure 1.6: Key technology components of grapevine winter pruning pipeline.

1.3   Outline of the Thesis

This dissertation is divided into seven chapters and is organized as follows:

- **Chapter 1** presents the basic background of grapevine winter pruning. The main problems and challenges of robotic winter pruning as well as the corresponding main solutions are clarified.

- **Chapter 2** introduces the design and mathematical model of the pruning robot systems, which include both wheeled and legged mobile manipulators with customized shears.

- **Chapter 3** summarizes the rules of grapevine winter pruning, with particular emphasis on spur pruning, which provides hierarchical guidelines for spur pruning allowing translation of the human cognition process and physiological bases into algorithms for automated plant recognition and robotized pruning.

- **Chapter 4** presents a novel method to create 2D plant models, based on grapevine semantic segmentation, containing the topographical and geometrical information between the different grapevine organs. Combine to generate potential pruning points.

Furthermore, the optimal pruning orientation is generated based on the 3D geometric features of the grapevine, thereby obtaining the target pruning pose.

- **Chapter 5** illustrates a methodology to approach the target pruning point where a three-phase approach strategy is proposed to guide shears to enclose the pruning point decently. Experimental results are presented to demonstrate the effectiveness of the prototype.

- **Chapter 6** introduces a novel compliant control where stiffness is actively controlled by means of learning from demonstration. Experimental tests with the prototype show the pruning cut performance in terms of pruning points engagement.

- **Chapter 7** summarizes the main results of the previous chapters extending them with a more general discussion and giving directions for future ongoing research.

2 Design and Analysis of Robotic Pruning Systems

Robotic pruning is a selective operation that provides precise cuts using a pruning tool attached to a robotic manipulator. A robotic pruner typically consists of a vision sensor, a manipulator, a mobile platform, and an end-effector tool. It must be highly manipulative to avoid collisions so the robot should be a kinematically redundant manipulator. A 7-Degrees of Freedom (DoFs) manipulator, with one DoF of redundancy, could be a good choice since more degrees of redundancy optimize the trajectory planning and design computational costs extremely [62]. Additionally, for this manipulator to reach the desired pruning points, a sizable workspace is required. Both wheeled and legged robots are currently evaluated as mobile platforms, furthermore, legged robots are promising due to their higher mobility on rough terrains [63]. Wheeled Mobile manipulator will be deployed in laboratory and greenhouse conditions. Whereas, the legged mobile manipulator will be deployed in field conditions, due to the dexterous movement characteristics of the legged system.

We present integrated systems that advance the state of the art in perception and manipulation capabilities for pruning. Our system, shown in Figure 2.1, once driven in front of a target grapevine, is expected to autonomously detect pruning targets, move towards them, and then execute precision cuts.

(a)   (b)
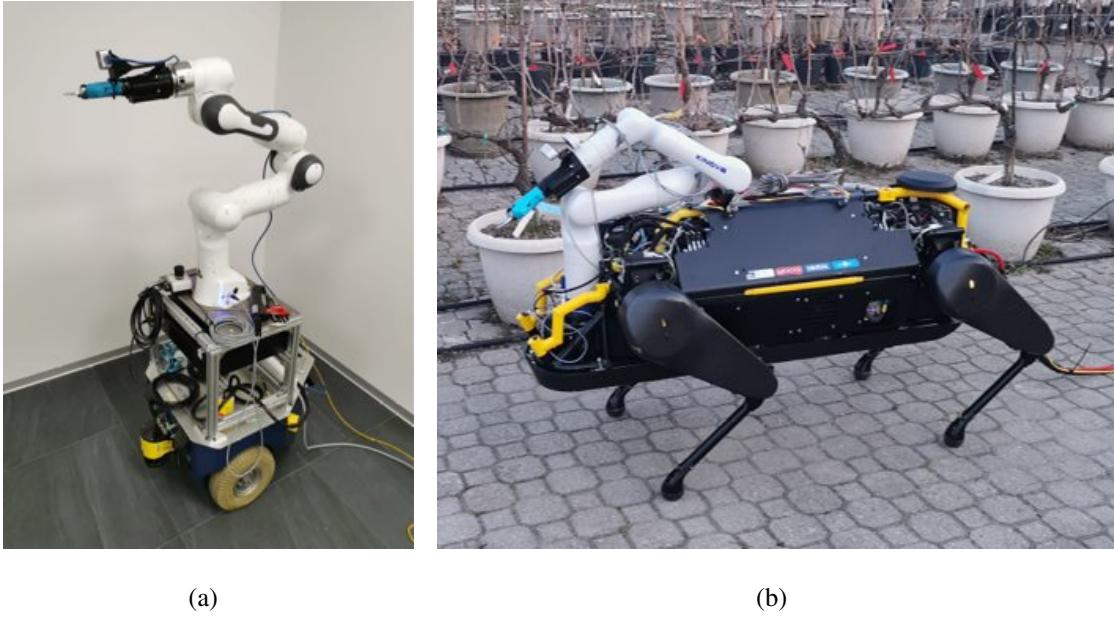
Figure 2.1: Our autonomous pruning system (a) *Rolling Panda* and (b) *HyQReal-Kinova*.

2.1   Robot Platforms

2.1.1   Wheeled Manipulator *Rolling Panda*

As shown in Figure 2.1 (a), *Rolling Panda* consists of two parts: the velocity-controlled two-wheel non-holonomic mobile robot, MP-500 (Neobotix GmbH. Co.) and 7-DoFs robot manipulator (Franka Emika. Co.), equipped with shears (end-effector). Both the mobile robot and manipulator have their own controller interfaces which are simple and user-friendly programming. Regarding MP-500, there is a Robot Operating System (ROS) [64] interface for low-level, real-time velocity controller and localization algorithms[1]. The localization algorithm returns the odometry and twists information of the mobile base's central frame in relation to its global frame. The Franka ROS Interface provides utilities for controlling and managing the Franka Emika Panda [2]. The control frequencies of Franka Emika Panda and

---

[1]https://robots.ros.org/neobotix-mp-500/

[2]https://frankaemika.github.io/docs/

MP-500 are 1 $k\,Hz$ and 50 Hz, respectively. The ROS master laptop, used for the controller, is a core-i7 processor 1.8 GHz with 32 GB RAM.

Kinematic Model
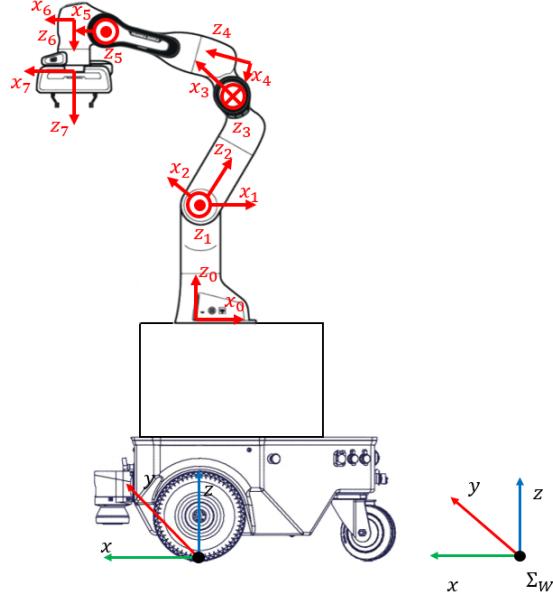
Figure 2.2: Whole-body kinematic model of *Rolling Panda*.

Considering components of the system (non-holonomic move base and 7-DoFs manipulator, as shown in Figure 2.2). The moving base of the *Rolling Panda* is able to realize rigid body planner motion, hence a unified coordinate system is defined as follows:

$$q_{wb} = \left[\begin{array}{ccc} q_m^\top & q_w^\top & q_n^\top \end{array} \right]^\top \tag{2.1}$$

where $q_m = \left[\begin{array}{ccc} x_m & y_m & \phi \end{array} \right]^\top \in \mathbb{R}^3$ is the coordinate of the rotation central frame of the mobile base (as shown in Figure 2.3), $q_w = \left[\begin{array}{cc} \theta_l & \theta_r \end{array} \right]^\top \in \mathbb{R}^2$ is spinning of the wheel joints and $q_n \in \mathbb{R}^n$ is the joint vector for the manipulator. μ is the distance between the driving wheels and the mobile platform geometric center, ρ is the distance from the mobile platform rotation center to the center of mass of the mobile platform, and R is the radius of the wheels.
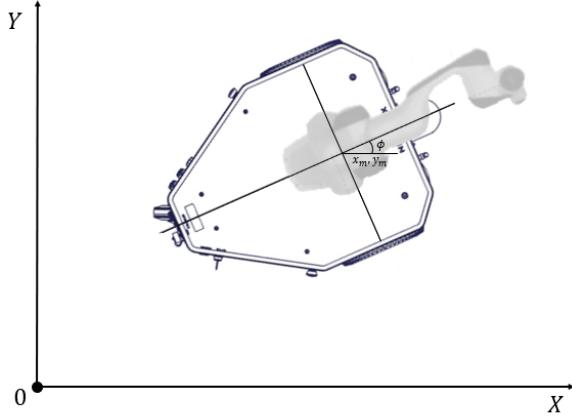
Figure 2.3: Schematic drawing of *Rolling Panda*. x_m, y_m, and ϕ represent the center point coordinate and rotation angle of the mobile base under the global frame.

The Denavit-Hartenberg (D-H) parameters of the prototype in Figure 2.2 are listed in Table 2.1, from which the manipulator's Jacobian matrix is derived using RBDL software [65] and is shown in Equation (2.2). In Table 2.1 q_i is the angle from x_{i-1} to x_i about z_{i-1}, α_i is the angle from z_{i-1} to z_i about x_i, d_i is the distance from x_{i-1} to x_i along z_{i-1}, and finally a_i is the distance from z_i to z_{i-1} along x_i [66].

$$J(q_n) = \left[\begin{array}{c} [J_T]_{3\times7}(q_n) \\ [J_R]_{3\times7}(q_n) \end{array} \right] \tag{2.2}$$

where $J_T(q_n)$ and $J_R(q_n)$ are the translational and rotational Jacobian sub-matrices of the manipulator respectively.

Due to its inherent properties, a differential mobile platform can not move sideways [67]. Hence, the velocity of the mobile base in the lateral directions should be zero.

$$-\dot{x}_m \sin\phi + \dot{y}_m \cos\phi - \rho\dot{\phi} = 0 \tag{2.3}$$

The other two constraints are a pure rolling constraint, relating the base velocities $\dot{x}_m$,

Table 2.1: The Denavit-Hartenberg parameters of Franka Emika Panda.

q_i	α_i	d_i	a_i	Joint
q_1	0	0.333	0	1
q_2	$-\frac{\pi}{2}$	0	0	2
q_3	$\frac{\pi}{2}$	0.316	0	3
q_4	$\frac{\pi}{2}$	0	0.0825	4
q_5	$-\frac{\pi}{2}$	0.384	-0.0825	5
q_6	$\frac{\pi}{2}$	0	0	6
q_7	$\frac{\pi}{2}$	0	0.088	7

$\dot{y}_m$, $\dot{\phi}$ while the wheel velocities $\dot{\theta}_l, \dot{\theta}_r$, ensure the no-slip condition at each rolling wheel in the forward directions.

$$\dot{x}_m \cos \phi + \dot{y}_m \sin \phi - \mu \dot{\phi} = R \dot{\theta}_l$$
$$\dot{x}_m \cos \phi + \dot{y}_m \sin \phi + \mu \dot{\phi} = R \dot{\theta}_r$$

(2.4)

We can set the constraint matrix between the rigid body motion of the mobile base and the generalized coordinates to satisfy the following equation. Combining constraint equation and unified coordinates in a general formulation, we have

$$A(q_{wb})\dot{q}_{wb} = 0$$

(2.5)

where $A(q_{wb}) \in \mathbb{R}^{3 \times (5+n)}$ is the full-ranked constraint matrix.

$$A(q_{wb}) = \begin{bmatrix} -\sin \phi & \cos \phi & -\rho & 0 & 0 & \cdots & 0 \\ -\cos \phi & -\sin \phi & -\mu & R & 0 & \cdots & 0 \\ -\cos \phi & -\sin \phi & \mu & 0 & R & \cdots & 0 \end{bmatrix}$$

(2.6)

The following transform equation can be derived using the null-space technique

$$\dot{q}_{wb} = S(q_{wb})\dot{\xi}$$

(2.7)

where $S(q_{wb})$ satisfies $A(q_{wb})S(q_{wb}) = 0$, and

$$S = \begin{bmatrix} \varrho(\mu\cos\phi - \rho\sin\phi) & \varrho(\mu\cos\phi + \rho\sin\phi) & 0 & \cdots & 0 \\ \varrho(\mu\sin\phi + \rho\cos\phi) & \varrho(\mu\sin\phi - \rho\cos\phi) & 0 & \cdots & 0 \\ \varrho & -\varrho & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.8}$$

where $\varrho = R/(2\mu)$. The set of feasible velocities may be expressed in terms of a suitable vector, $\dot{\xi} = [\; \dot{q}_w^\top \;\; \dot{q}_n^\top \;]^\top \in \mathbb{R}^{2+n}$ denotes the velocity for all actuators of *Rolling Panda* in joint space.

Jacobian matrix maps actual joint velocity space onto Cartesian velocity space, $J_\xi \in \mathbb{R}^{6\times(2+n)}$, can be derived as:

$$J_\xi(q_{wb}) = J_{q_{wb}}(q_{wb})S(q_{wb}) \tag{2.9}$$

where $J_{q_{wb}}(q_{wb})$ is the Jacobian matrix for q_{wb}.

Dynamic Model

The following is the unconstrained equation of motion for a non-holonomic handheld manipulator [68]:

$$M(q_{wb})\ddot{q}_{wb} + C(q_{wb}, \dot{q}_{wb})\dot{q}_{wb} + G(q_{wb}) = B(q_{wb})\tau + \tau_{dis} + \Lambda^\top(q_{wb})\lambda \tag{2.10}$$

where $M(q_{wb})$ denotes an $n \times n$ symmetric positive definite inertia matrix, $C(q_{wb}, \dot{q}_{wb})$ is the centripetal and Coriolis matrix, $G(q_{wb})$ denotes the gravitational vector, τ_{dis} denotes the vector of bounded unknown disturbances including unstructured unmodeled dynamics,

$B(q_{wb})$ denotes the input transfer matrix, τ denotes the torque vector, $\Lambda^\top(q_{wb})$ denotes the kinematic constraints matrix, and λ denotes the Lagrange multipliers vector.

Finally, the dynamic motion of *Rolling Panda* with respect to ξ and $\dot{\xi}$ can be reformulated by removing the generalized constrains, $\Lambda(q_{wb})^\top \lambda$, in Equation (2.10) by combining Equation (2.7) and Equation (2.10), as follow:

$$M_\xi(q_{wb})\ddot{\xi} + C_\xi(q_{wb}, \dot{q}_{wb})\dot{\xi} + G_\xi(q_{wb}) = u + S(q_{wb})^\top \tau_{dis} \tag{2.11}$$

where
$$M_\xi(q_{wb}) = S(q_{wb})^\top M(q_{wb}) S(q_{wb})$$
$$C_\xi(q_{wb}, \dot{q}_{wb}) = S(q_{wb})^\top [M(q_{wb})\dot{S}(q_{wb}) + C(q_{wb}, \dot{q}_{wb})S(q_{wb})]$$
$$G_\xi = S(q_{wb})^\top G(q_{wb})$$
$$u = S(q_{wb})^\top B(q_{wb})\tau.$$

2.1.2   Legged Manipulator *HyQReal-Kinova*

With regard to viticulture, terrain variation could severely impair the stability of traditional platforms and pruning-manipulator performance. Robots must overcome the challenge of moving in rough terrain when performing tasks in complex wild environments. Due to their mobility restrictions over rough terrain, today's remote-controlled vehicles with wheels and tracks remain limited in such tasks [69]. As an alternative, legged robots present the potential of becoming the new generation of rough terrain vehicles that are capable of autonomous, semi-autonomous, or remote-controlled operations in challenging terrains [70]. The next generation of all-terrain vehicles, with legs instead of wheels and tracks, is finally achieving performance levels that show superior mobility on rough terrain.

As shown in Figure 2.1 (b), *HyQReal-Kinova* consists of two parts: a legged mobile base and a manipulator. HyQReal[3] developed by Istituto Italiano di Tecnologia (IIT) is 1.33 m

---

[3]https://hyq-real.eu/

long and $90\,cm$ tall, and its weight is $130\,kg$. The robot is protected by an aluminum roll cage and a skin made of Kevlar, glass fiber, and plastic [71]. The quadruped has custom-made feet made in special rubber for high traction on the ground and is equipped with a 48 *Volt* battery that powers four electric motors connected to four hydraulic pumps. HyQReal is completely power-autonomous with onboard hydraulics, batteries, and wireless communication. Furthermore, the robot features a higher ruggedness, reliability, and energy efficiency. The manipulator consists of a Kinova Gen3 robot [4], a lightweight manipulator with lower power consumption mounted on a mobile robot; the details of the robot can be found in [72]. The 7 rotational joints provide the dexterity needed for the scanning and pruning actions.

Kinematic Model

The model of a legged mobile manipulator can be formulated as a floating-base B to which the manipulator is attached. The position and orientation of the base frame $\{B\}$ with respect to the world frame $\{W\}$ are expressed as a three-dimensional vector $^{W}p_B \in \mathbb{R}^3$ and a Hamiltonian unit quaternion $^{W}q_B$ respectively. The joint positions of the legged manipulator are collected in the vector $q_{lm} \in \mathbb{R}^{n_{lm}}$, where the number of joints is $n_{lm} =19$. The generalized coordinate vector q_{hk} and the generalized velocity vector v_{hk} are collected as:

$$
q_{hk} = \begin{bmatrix} ^{W}p_B \\ ^{W}q_B \\ q_{lm} \end{bmatrix} \in SE(3) \times \mathbb{R}^{n_{lm}}, \quad v_{hk} = \begin{bmatrix} ^{W}v_B \\ ^{W}\omega_B \\ \dot{q}_{lm} \end{bmatrix} \in \mathbb{R}^{n_v} \tag{2.12}
$$

where the number of velocity vectors is $n_v = 6 + n_{lm}$, and $^{W}v_B \in \mathbb{R}^3$ and $^{W}\omega_B \in \mathbb{R}^3$ are the linear and angular velocity of the floating-base with respect to the world frame expressed in the $\{W\}$ frame. The *HyQReal-Kinova* robot shown in Figure 2.4 has $n_v =25$ DOFs, with six, twelve, and seven DOFs describing the floating base, legs, and the manipulator, respectively.

---

[4]https://www.kinovarobotics.com/product/gen3-robots
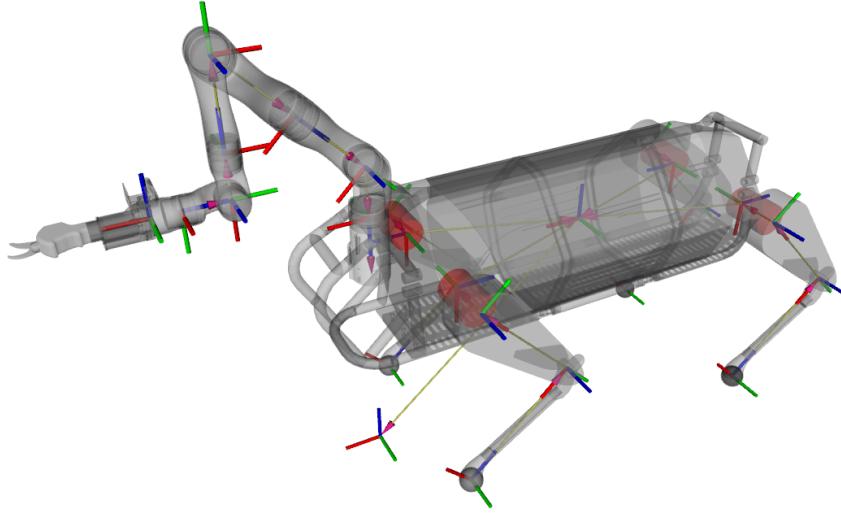
Figure 2.4: Whole-body kinematic model of *HyQReal-Kinova*.

Dynamic Model

The equations of motion of a legged mobile manipulator robot with a floating base can be written as:

$$M(q_{hk})\dot{v} + h(q_{hk}, v_{hk}) = S^\top \tau_{hk} + J_c^\top(q_{hk})\lambda \qquad (2.13)$$

where $M(q_{hk}) \in \mathbb{R}^{n_v \times n_v}$ stands for the inertia matrix, and $h(q_{hk}, v_{hk})$ is a vector of non-linear terms (including Coriolis, centrifugal and gravity forces). The selection matrix $S = \begin{bmatrix} \mathbf{0}_{n_\tau \times (n_v - n_\tau)} & I_{n_\tau \times n_\tau} \end{bmatrix}$ represents the system under-actuation that the floating-base is not directly actuated by joint torques $\tau_{hk} \in \mathbb{R}^{n_\tau}$. If all limb joints are actuated, then the number of actuated joints $n_\tau = n_{lm}$. The vector of contact forces λ is mapped to the joint-space torques through the support Jacobian $J_c \in \mathbb{R}^{3n_c \times n_v}$, which is obtained by stacking the end-effector Jacobians which relate generalized velocities to limb end motion as $J_c = \begin{bmatrix} J_{c_1}^\top & \cdots & J_{c_{n_c}}^\top \end{bmatrix}^\top$, with n_c the number of limbs in contact.

Remark 2 *Legged whole-body manipulation is beyond the scope of this research. Here we only consider decoupled motion policy, mainly focusing on robot arm manipulation, while*

quadruped robots are used as a mobile platform.

2.2 Camera Vision System

Spur recognition is one of the crucial steps toward developing a robotic pruning system for grapevine winter pruning. A key requirement for accurate perception of grapevine modeling and manipulation in complex environments in the outdoors is a robust camera system. It is also the precondition for motion planning to efficiently identify obstacles. Camera-based machine vision system serves as a promising object detection technology and has been widely investigated in the last decade for various tree fruit and grapevines [73, 74, 75, 76, 77, 78]. A major advantage of using machine vision techniques is to provide accurate, reliable, cost-effective, and automatic solutions for orchard management operations in a non-destructive way.

Figure 2.5: Intel RealSense Depth Camera D405.

The Intel RealSense Depth Camera D405[5] (as shown in Figure 2.5) is a short-range stereo camera providing sub-millimeter accuracy for close-range computer vision needs. The D405 operates at an ideal range of $7\,cm$ to $50\,cm$ with minimum object detection down to $0.1\,mm$

---

[5]https://www.intelrealsense.com/depth-camera-d405/

at 7 cm. This is suitable for our grapevine winter pruning application, we need the camera to perform spur detection at close range.

2.3   End-Effector Tool

The end-effector is a critical component of the pruning robot, required to conduct the pruning cut on the selected pruning location. Unlike normal grasping or fruit harvesting, having a proper end-effector is important for grapevine winter pruning because canes are thin, and spurs are usually clustered. Furthermore, the tool must be light enough to be carried by a lightweight robot as an end-effector. Currently, there is no pruning machine on the market that can directly interact with a robot. For this reason, designing a tool capable of cutting grapevine under program control is important for the success of the process.

(a)   (b)
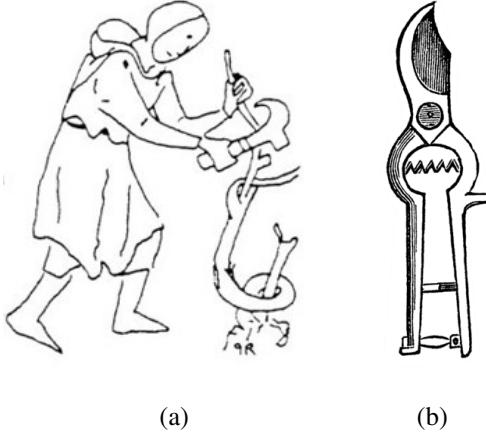
Figure 2.6: Different cutting tools: (a) billhook; (b) pruning shears.

2.3.1   End-Effector Design

Pruning the grapevines necessitates precise cutting at a specific location on the plant. However, before such a cut could be made, it is important to understand the mechanical properties
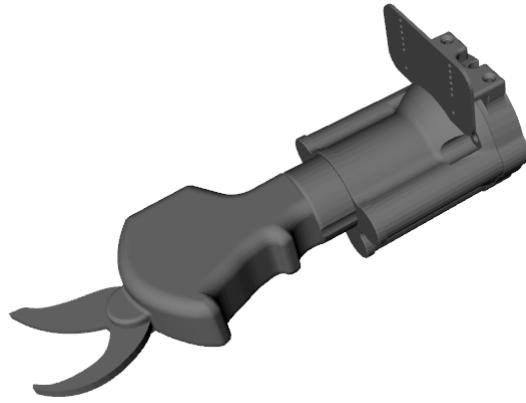
Figure 2.7: 3D model of electric bypass pruners.

of the canes, particularly the force required to cut dormant canes for the proper design of the pruning end-effector.

Winegrowers used to use different cutting tools for each region of Italy. The billhook (Figure 2.6 (a)) was mainly used in the north. In the center and in the south, different shapes of "pennato" were prevalent (in the drawing: it is similar to a billhook, but with the addition of a second blade [79]. This pruning knife shape can be traced back to the Roman era). These tools were difficult to wield: it was easy to injure yourself, to injure the plant too much (with the entry of diseases and pests) and it takes strength. In the nineteenth century, pruning shears (Figure 2.6 (b)) began to be used and the work started to be a little easier. Today we still use them, sometimes electric or pneumatic ones, to reduce fatigue and hand usury.

We modified a commercial electric bypass pruner through a 3D printed support frame so that it can be connected to the robot end effector flange. The 3D model is shown in Figure 2.7.

Finally, the eye-in-hand pruning end-effector consists of a set of electric bypass pruners along with a RealSense D405 RGBD camera shown in Figure 2.8. The camera is located above the blade and the top blade is visible when the shears are open. The shears are rated
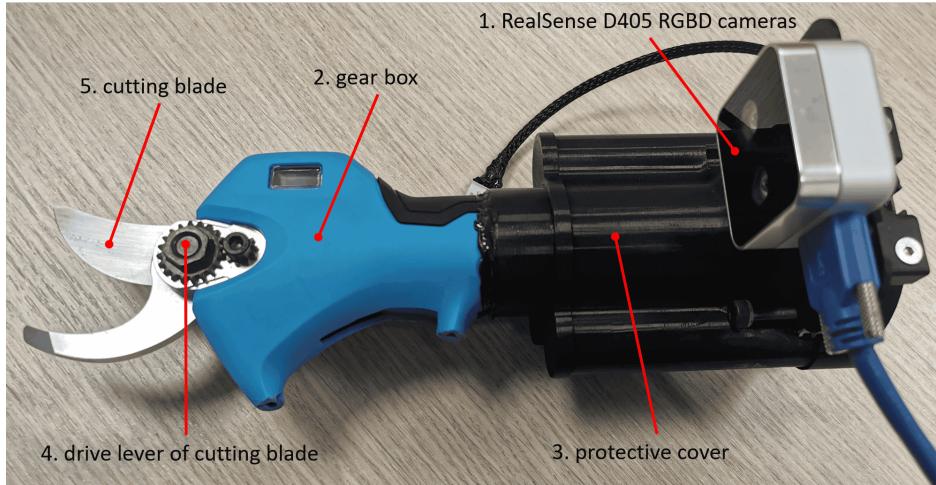
Figure 2.8: Structure of the pruning tool. 1: RealSense D405 RGBD cameras in a 3D printed housing; 2: gear box; 3: protective cover; 4: drive lever of cutting blade; 5: cutting blade.

for cutting branches up to 3.2 cm in diameter.

2.3.2   Hand-Eye Calibration

In robot applications of grapevine pruning, hand-eye calibration is a very basic and critical issue. The purpose of hand-eye calibration is to obtain the relationship between the robot coordinate system and the camera coordinate system and finally transfer the results of visual recognition under the camera frame to the robot coordinates system. The following Figure 2.9 shows a representation of the elements involved in hand-eye calibration.

Eye-in-hand to compute the static transform between the end effector frames of a robot and that of a tracking system, e.g. the optical frame of an RGB camera used to track AR markers. In this case, the camera is mounted on the end-effector, and we place the visual target so that it is fixed relative to the base of the robot.

The following equation describes how to transform a single 3D point from the camera to the robot base coordinate system:
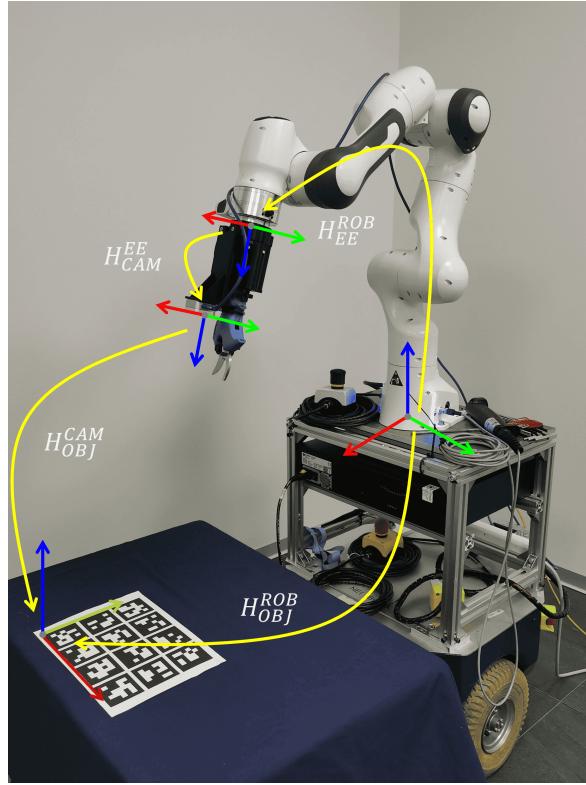
Figure 2.9: Schematic diagram of hand-eye calibration.

$$
\begin{bmatrix} x^r \\ y^r \\ z^r \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_e^r & \boldsymbol{t}_e^r \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{R}_c^e & \boldsymbol{t}_c^e \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x^c \\ y^c \\ z^c \\ 1 \end{bmatrix} \tag{2.14}
$$

$$
\boldsymbol{p}^{ROB} = \boldsymbol{H}_{EE}^{ROB} \cdot \boldsymbol{H}_{CAM}^{EE} \cdot p^{CAM}
$$

To convert the entire point cloud from the camera coordinate system to the robot base coordinate system, apply the equation above to each point in the point cloud. To transform the pose of the object relative to the camera use the following equation:

$$\begin{bmatrix} \boldsymbol{R}_o^r & \boldsymbol{t}_o^r \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_e^r & \boldsymbol{t}_e^r \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{R}_c^e & \boldsymbol{t}_c^e \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{R}_o^c & \boldsymbol{t}_o^c \\ 0 & 1 \end{bmatrix} \tag{2.15}$$

$$\boldsymbol{H}_{OBJ}^{ROB} = \boldsymbol{H}_{EE}^{ROB} \cdot \boldsymbol{H}_{CAM}^{EE} \cdot \boldsymbol{H}_{OBJ}^{CAM}$$

The resulting pose is the one that the robot Tool Center Point (TCP) should attain for pruning. This allows us to express the pose of the camera relative to the end-effector $\boldsymbol{H}_{CAM}^{EE}$ as a function of the robot to end-effector $\boldsymbol{H}_{EE}^{ROB}$, camera to calibration object $\boldsymbol{H}_{OBJ}^{CAM}$ and one constant, unknown pose $\boldsymbol{H}_{OBJ}^{ROB}$. we can solve for $\boldsymbol{H}_{CAM}^{EE}$ using optimization technique [80].

2.4   Conclusion

This chapter defines the grapevine robotics pruning system, analyzes the feasibility of pruning mechanization, designs the pruning end-effector, and establishes the overall scheme of the grapevine pruning robot. The objective is to apply state-of-the-art mobile manipulation platforms and systems, a wheeled mobile platform with a commercial full torque sensing manipulator and multiple sensors, and a quadruped robot mobile platform with a customized robotic manipulator and multiple sensors for various maintenance and automation work in the greenhouse and vineyard. Moreover, the kinematics and dynamic analysis of the two pruning robots, *Rolling Panda* and *HyQReal-Kinova* were carried out. Due to the advantages of fast and smooth movement, and strong adaptability of the wheeled mobile platform, *Rolling Panda* will be used in laboratory and greenhouse conditions. The legged mobile platform, *HyQReal-Kinova*, will perform grapevine pruning in vineyards with more complex terrain.

3 Grapevine Winter Pruning

Grapevine winter pruning is the process of cutting away a portion of the annual vegetative growth of a grapevine to maintain the desired bud load per vine and the plant architecture. Such a process is also known as "winter/dormant pruning" as performed over dormancy [81]. When people prune grapevines, they first select several canes to keep, and then they make cuts to remove the rest. Canes that are long, not too thick or thin, and will keep the head compact in subsequent years are selected. For manual pruning, a certain amount of knowledge and skills are needed to evaluate the grapevine structure and to decide where to prune, without damaging the vine. Pruning rules define a systematic way to remove older canes to keep the vigor and vine balance in control, which provides hierarchical guidelines for spur pruning allowing translation of the human cognition process and physiological bases into algorithms for automated plant recognition and robotized pruning.

3.1 Principles of Grapevine Winter Pruning

Grapevine pruning is a perennial management practice that keeps selected one-year-old grapevine wood for the following season's crop production. Green shoots become woody in the dormant season and will contain buds that produce fruitful shoots in the next season. Woody canes were green shoots in the previous growing season and will be roughly one year

old when the growing season starts again. The retained parts of these canes contain buds that will produce shoots that bear clusters. Once these canes have produced their fruit for the season, they will not produce again. To keep grapes productive, they need to be pruned to renew the young canes which will produce in the following year. One-year-old grapevine wood is light tan or cinnamon colored and can be contrasted with older wood that often has darker brown colored bark.

Grapevine winter pruning is an extremely time-consuming operation requiring up to 120 hours per hectare depending on vine vigor, training system, equipment, and skilled labor availability [8, 82]. In mechanically harvested, vertically shoot-positioned (VSP) trained vineyards, hand pruning can account for up to 75% of the annual labor demand. Such a high labor requirement can be significantly reduced by 50-90% in case of the adoption of mechanical pruning [83]. From an agronomic perspective, winter mechanical pruning was initially applied on spur-pruned cordons and has been demonstrated to be much more suitable for training systems with free-growing canopy such as single, high-wire cordon and Geneva double courtain (GDC) [4]. Even in the case of manual follow-up, mechanical pruning leads to an increased bud load as compared to hand pruning [84] inducing vegetative and yield self-regulation, such as reduced bud-break rates [85, 86], shoot fruitfulness [87], fruit-set [88], cluster and berry weight [89]. The bigger canopy (see increased vine capacity) resulting from mechanical pruning may support similar yields to manual pruning by maintaining similar fruit composition [87, 90, 91].

As already introduced, winter pruning consists of the removal of part of the previous season's growth over dormancy. The operation aims to (i) regulate vegetative growth, (ii) select fruitful nodes, (iii) regulate shoot vigour, vine capacity, yield, and crop load, (iv) produce grapes of target composition, and (v) maintain a desired canopy architecture depending on the training system [81, 84].

Grapevine pruning methods generally fall into two categories spur pruning and cane pruning, which are associated with training methods by virtue of where the fruitful one-
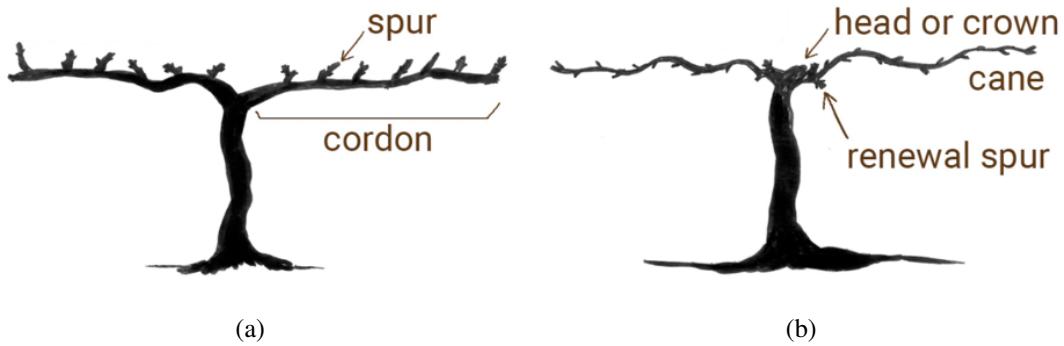
Figure 3.1: Grapevine pruning scheme: (a) spur pruned and (b) cane pruned.

year-old wood originates. The fruitful wood can originate from either the head region or a cordon (illustrated in Figure 3.1). Spur pruning is commonly implemented with cordon training. A cordon is a horizontally trained extension of the trunk that is retained for multiple years, spurs originate along the length of a cordon. Spur pruning is generally associated with pruning and retaining two to four buds on one-year-old wood that originates across the length of the fruiting cordon (as shown in Figure 3.2). Whereas cane pruning is implemented with head training, canes originate from the head region of the vine. Cane pruning does not use cordons. Instead, new one-year-old canes are laid down on the fruiting wire every year.

Spurs and canes refer to the portions of one-year-old wood that are retained when pruning. Spurs are shorter and contain fewer buds than canes; they are called "spurs" due to their physical appearance after cutting the cane. However, both spurs and canes are bearing units with dormant buds that can give origin to fruitful shoots. Based on the length of the retained bearing units, spur and cane are the two main methods adopted to prune the grapevines, consisting in shortening a long cane to 2-3 or >5 (generally 8 to 15) nodes, respectively. Moreover, bearing units can be arranged in space, with grapevines trained to specific trellis systems depending, among others, on desired quality, harvest method, mechanization needs, and labor requirements [92]. Vertical shoot positioning (VSP) is a trellis system widely adopted in several wine regions and is based on a spur-pruned cordon fixed to a fruiting wire with shoots growing upwards during the season to create a canopy wall varying between 0.8

and 1.5 m high. VSP-trained vines can also be subjected to cane pruning and up to 4 canes per vine kept, depending on vine spacing and optimal bud load [93]. Both the pruning methods are highly selective requiring skilled workers to apply the following decision-making process: (i) assess the previous season's plant status, (ii) identify potential spurs and canes among dormant shoots, (iii) perform the main cuts, (iv) carefully strip-out the dormant shoots that were cut-off, (v) determine the bud load and make the final cuts, (vi) remove older wood and any extra canes, (vii) tie down the cane along the fruiting wire if required by the training system [94]. On the other hand, spur pruning has simpler requirements, and pruning cuts are closely located near the cordon. In addition to bud retention, pruning rules also necessitate qualitative parameters such as cane diameter and the health of canes and buds. In particular, spur pruning automation is the main research topic of this thesis.
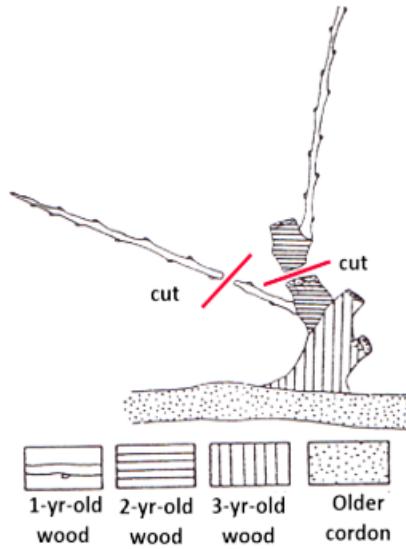
Figure 3.2: Illustration of spur pruning.

Spur pruning is suitable for spurred cordon system, which is widely adopted worldwide in several wine regions. It is suitable for low or medium-fertility soils, even dry, and for those varieties that have good fertility on the proximal buds (basal count nodes) [4]. It is frequent in Tuscany and other Italian central regions where viticulture is primarily based

on Sangiovese cv. At the end of the training of the young vines, the best cane is placed horizontally, becoming a structural element of the plant. In spring, fruiting shoots will form from the retained buds, and will crop grapes and wooden by the end of the season; during winter pruning operations each ventral cane will be removed and upward canes shortened into a spur, forming the cordon. Since the architecture of an adult grapevine is achieved, spurs will be annually renovated through short pruning. This system has been successful and has spread due to the excellent quality management of the vineyard, but also for simple and fast pruning. The action will be driven by the following factors to consider:

- Spur length

  The optimal bud load per vine can be attained by acting on spur number (N) and length (L). So, physiological reactions and different growing pathways might impact vine architecture and next-season pruning operations. In particular, the longer the spur the higher the risk to induce vegetative gradients due to acrotony (preferential growth of distal shoots inhibiting the development of underlying shoots). Moreover, a lower spur distancing along the metr of the row will result in excessive canopy density with negative impacts on canopy architecture, vineyard management, health status, and fruit composition [95].

- Effects of seasonal canopy management on grapevine architecture in winter

  The architecture of a spur-pruned vine is strictly related to canopy management in summer. Excessive or misplaced shoots are preferentially removed early in the season (shoot-thinning) for a resulting canopy with regularly-spaced shoots [96]. Otherwise, unthinned canopies will be composed of a heterogeneous population of shoots according to their insertion on the permanent structures, orientation, and site vigor.

Figure 3.3 is the process of reducing the previous year's two-bud spurs to single shoots, and pruning the remaining new growth to form replacement two-bud spurs. Spur pruning is carried out in winter following this rule:
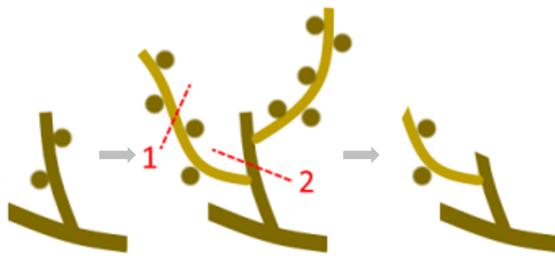
Figure 3.3: Fruiting canes pruned back to 2 buds in winter.

- Select healthy canes, evenly spaced at approximately 15-20 cm apart to form the new spurs. Prune these canes back to two buds from the base, not including the bud at the base (basal cane cut).

- Select upward-facing buds if possible as this is more referable. Make the pruning cuts 1-2 cm above the bud to prevent the buds from drying out. Prune off all other growth from the main laterals. Cuts made with sharp pruning shears will not bruise the vine and will leave a clean, smooth surface.

- Prune the previous year's two-bud spurs in half, removing the top half of the spur (top cane cut).

- Creating a new two-bud spur that will produce the new fruiting canes in the following year. Prune the new growth coming from the spur's remaining lower shoot down to two buds.

3.2   Conclusion

This chapter summarizes the rules of grapevine pruning, with particular emphasis on spur pruning which is commonly implemented with cordon-based training systems. A good spur pruning strategy is to retain 1-year-old spurs that are positioned as close as possible to the cordon, such a condition ensures that clusters are maintained in a confined fruit zone region, which promotes precision spraying, leaf removal, and harvest efficiency. Furthermore, maintaining low-positioned spurs maximizes the amount of exposed canopy leaf area throughout the season. As proved to be a working solution by the human workers, the ideal pruning cut generally should prune these canes back to two buds from the base and prune the previous year's two-bud spurs in half. A hierarchical policy based on the pruning rules could be adopted by the robot vision system to precisely locate these buds and generate the cutting points or their pose.

4   Grapevine Perception System

Grapevine pruning automation is approximated with the help of computer vision and image processing methods. Grapevine image recognition is to use image processing and pattern recognition technology to identify grapevines from complex backgrounds, which can provide the basis for accurate judgment of grapevine segmentation and pruning point positioning. Achieving this through image analysis means that all grapevine organs should be segmented in order to extract the whole canes' bodies for pruning point estimations on them. Semantic segmentation is a very popular method in robotic systems that interact in a wild-free environment. With this approach, grapevine winter pruning automation methods can achieve pruning point estimations or feature extraction mechanisms since the whole area of interest is segmented. It is one of the key technologies for grapevine automatic pruning robots.

4.1   Visual Perception for Grapevine Identification

The overall perception pipeline to perceive the grapevine and generate the pruning pose on pruning location is shown in Figure 4.1. We perform a semantic segmentation of the scene by utilizing optical flow data from RGBD cameras. We feed the grapevine images into a Mask R-CNN network to perform instance segmentation of different grapevine components and

CHAPTER 4. GRAPEVINE PERCEPTION SYSTEM 40

process the resulting instances to obtain pixel estimates for pruning points. Simultaneously, based on the 3D point cloud model of grapevine, the optimal approach orientation is inferred near the pruning point. Finally, the perception system outputs the pruning pose.
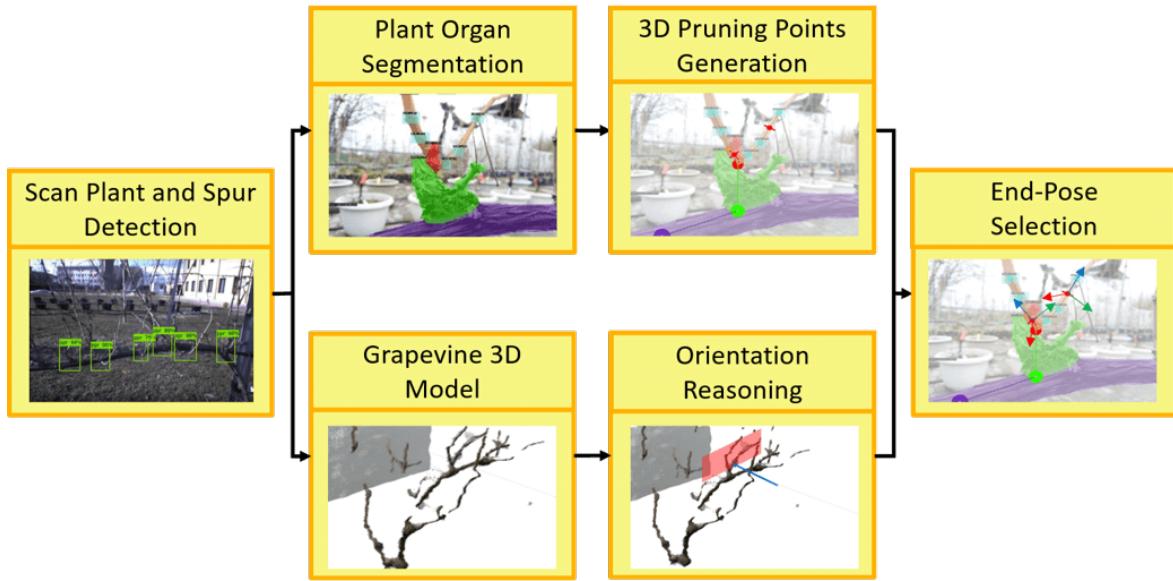
Figure 4.1: A block diagram with all major steps in the perception pipeline. The program flows from left to right.

4.1.1 Ground-Truth Annotation of Image Datasets

A dataset is an important component in the implementation of a machine-learning method. Since no dataset with the plant organ segmentation of spur-pruned grapevines was found, we created a new dataset with the annotation of five classes: cordon, arm, spur, cane, and node. This structure can be seen in Figure 4.2, along with the correct pruning points for the shown example. Each class corresponds to a different visible organ over dormancy.

The data acquisition was performed in the simulated vineyard at UCSC, Piacenza, Italy (Figure 1.1). It is composed of two rows of potted *Vitis vinifera* L. cv. Sangiovese spur-pruned grapevine, running a shoot thinning experiment, where a group of seven plants is
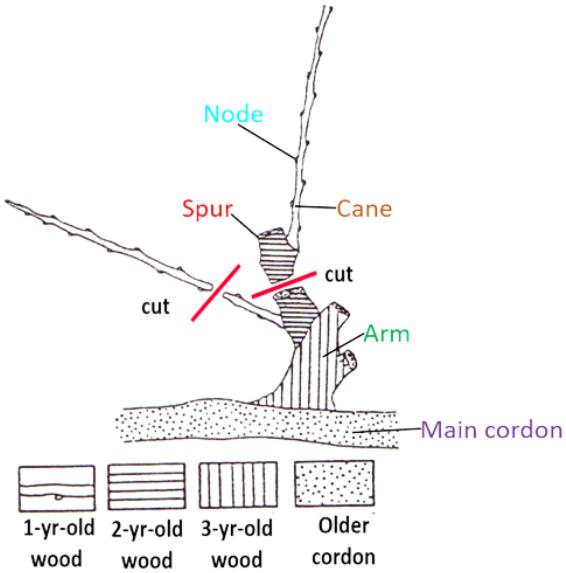
Figure 4.2: Schematic representation of a grapevine pruning region area, illustrating the 5 grapevine organs: Cordon, Arm, Spur, Cane, and Node. The red lines indicate the desired pruning points.

acting as the control group, and the second group of eight grapevines is subjected to shoot thinning. Each plant has around five spurs, and the images are taken from both sides of the grapevines.

Pruning target regions of each image were hand-labeled by viticulture experts and singularly contained in rectangular bounding boxes by using the COCO Annotator tool [97], a web-based tool that is designed to efficiently label images. The dataset is being annotated following the COCO segmentation dataset, with the five mentioned classes, the cordon, the arm, the spur, the cane, and the node.

4.1.2   Deep Learning Model for Grapevine Detection

Object detection systems involve not only recognizing and classifying every object in an image but localizing each one by drawing the appropriate bounding box around it. This makes

object detection a significantly harder task than its traditional computer vision predecessor, image classification. The traditional model-driven method for object detection follows standard feature matching and bundle adjustment pipeline. This pipeline usually leads to high computation and time cost and low tolerance to the scenarios with variable environmental conditions, e.g., brightness, backgrounds, and weather. It performs poorly in scenes where relevant features are not present, or similar [98]. Especially when solving the problem of grapevine identification, factors such as the natural environment in the field and the complex structure of the grapevine all make this problem difficult. With data-driven deep learning technologies, applying deep neural networks for object detection becomes a promising solution. Faster R-CNN [99], R-FCN [100], and SSD [101] are the most widely used object detection models. Other popular models tend to be fairly similar to these three, all relying on deep CNN (DCNN) to do the initial heavy lifting and largely following the same proposal/classification pipeline.

The neural network architecture and backbone used for this work is Detectron2 [102], a state-of-the-art detection algorithm, using Pytorch, Mask R-CNN, and, instead of using several backbones, we settled for just one, which is the ResNext X101-FPN [103]. This ResNext is an improvement on the original ResNet network. The network is trained using the default training procedure of Detectron2. This procedure creates a model, optimizer, scheduler, and data loader with the default configurations provided along with the model. It then loads the pre-trained model weights, initializes logging functions, and starts to follow a standard training workflow with a single-optimizer single-data source iterative optimization. The training hyperparameters are the default ones, with the only changes being the batch size changed to 2, from the original value of 16, and the number of training iterations that were changed to 50000, from the original 270000. These changes were made to adjust to the size of our dataset, which is considered small for the network itself.

4.1.3   Experiments

In order to test each step of the proposed perception system, two experiments were per-formed: **Experiment 1** is about the Pruning Regions (PRs) detection with a first Deep Con-volutional Neural Network (DCNN), and **Experiment 2** is about the grapevine organs seg-mentation with a second DCNN.

Experiment 1: Training and Testing of the DCNN for PR Detection

• **Image Collection**

During winter, a total of 1215 RGB images were acquired on *Vitis vinifera* L. spur-pruned grapevines from 2 different vineyards. In February 2018 and February 2019, 965 and 100 RGB images, were acquired from *Vitis vinifera* L. cv Merlot grapevines planted in 2014 in an experimental vineyard located in Piacenza, Italy. Mature vines presented seven 2-node spurs and were planted along an NS-oriented row with 2.1 $m\times$ 1.2 m spacing (inter- and intra-row). The cordon was set at $0.9\,m$ above the ground. Images were captured with a resolution size of 1280 $\times$ 720 pixels, moving from North to South along the row at a $0.9\,m$ operating distance. In December 2018, 150 RGB images were gathered on eight-year-old *Vitis vinifera* L. cv. Ervi grapevines from a commercial vineyard located at Alseno, Italy. Each vine was pruned to six 2-node spurs for a corresponding bud load of 12 nodes/vine. The east-facing vineyard featured EW-oriented rows and a 2.5 $m\times$0.9 m vine spacing (inter- and intra-row, respectively). Images were acquired from West to East. During each acquisition campaign, all the images were taken at solar noon under a clear sky (Figure 4.3 (a)).

• **Data Annotation**

Pruning target regions of each image were hand-labeled and singularly contained in rect-angular bounding boxes by using the COCO Annotator tool. Every annotation included

(a)   (b)   (c)
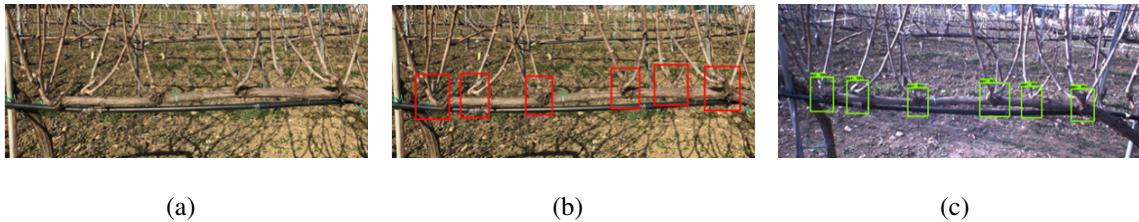
Figure 4.3: Description of the workflow required for fine-tuning a DNN for PR identification: (a) original image; (b) annotated image by experts for training the neural network by using red bounding boxes; (c) example of PR detection through Faster-RCNN 2.0. with green boxes indicating detected pruning regions.

individual spurs avoiding overlapping with adjacent regions, and at least the first 2 basal nodes of each cane (Figure 4.3 (b)). The annotated dataset, with a total of 8361 bounding boxes, as part of a fine-tuning process, was subsequently fed to the neural network Faster R-CNN.

• **Training of the DCNN**

The network was fine-tuned from a pre-trained model of Faster R-CNN [104], trained with the COCO2017 dataset. The fine-tuning hyperparameters were those related to the neural network structure by default adjusted for the following exceptions: the number of training iterations was changed to 50,000 from the original 270,000, the batch size was changed to 1 from an original value of 16, and the decaying learning rate which was set to 0.003 from the start, was changed to 0.0003 at 1000 steps and further decayed to 0.00003 at 2000 steps.

• **Testing of the DCNN**

The fine-tuned algorithm was tested in October 2021 on 2 different datasets referred to mature spur-pruned grapevines of diverse ages, and cultivars and subjected to different growing conditions. Accordingly, a batch of 202 frames was acquired in February 2019 on a subset

of 5 adjacent Merlot vines randomly chosen among those already used for training. The second test dataset, composed of 30 RGB images with a resolution of 4608×3456 pixels, was obtained in December 2020 in Piacenza on a set of 15 *Vitis vinifera* L. cv Sangiovese potted grapevines. The vines were arranged in a single row, trained to a spur-pruned cordon since 2017 with five 2-node spurs and a vine spacing of $0.9\,m$, and a $35°$ NE-SW orientation. The permanent cordon was located $0.9\,m$ from the ground. Each plant was entirely photographed once from both sides at cordon height (Figure 4.3 (a, b)).

Figure 4.4: Description of the pruning regions (PRs) depending on wood type and orientation. PRs defined as "other" are not reported.

For each image, the DNN predicted the potential Pruning Regions (PPRs) through bounding boxes and confidence values (Figure 4.3 (c)); however, only the detections with confidence >70% were considered. Additionally, every PR was progressively numbered and described by: wood type (W), visibility (V), and orientation (O). The wood type included the following categories: cane (cane arising from latent buds on the permanent cordon),

simple spur (spur with ≤ 1 shoot/node), complex spur (spur with >1 shoot/node), and other (PRs not falling in one of the previous categories) (Figure 4.4). For what concerns visibility, PRs were classified as visible or hidden if occluded by other grapevine organs and/or trellis components. Lastly, orientation provided three categories: coplanar (PR lying on the same plane of the row), perpendicular (PR lying on the vertical plane perpendicular to the row), and intermediate (PR lying on a plane in between coplanar and perpendicular conditions).

Experiment 2: Training and Testing of the DCNN for Grapevine Segmentation

• **Image Collection**

In March 2021, 148 RGB images were captured with a resolution size of 4608 × 3456 pixels, on the Sangiovese grapevines already considered for **Experiment 1**. In order to increase the variability among the pruning region complexity, in May 2020 shoot thinning (ST) was performed on 8 out of the fifteen grapevines according to Bernizzoni et al. [96]. The remaining 7 plants acted as an unthinned control (C) (Figure 4.5). The acquisition was performed at solar noon when each PR was individually photographed from a distance of $0.3\ m$ at cordon height; 2 passages per row were performed to consider both the East and West sides. An additional batch of 196 RGB images taken in December 2020 was also considered. Data were randomly captured on the same Sangiovese experimental row with a resolution size of 4608 × 3456 pixels considering different orientations.

• **Data Annotation**

The images were annotated using the COCO Annotator tool and five classes were used to describe the relevant grapevine organs (GO) for pruning purposes: cordon, arm, spur, cane, and node (Figure 4.6 (a)). Each grapevine element belonging to the above-mentioned classes was annotated with a polygon, except for nodes that were considered through bounding boxes (Figure 4.6 (b)). Polygonal annotation was carried out by retracing each element

(a)   (b)   (c)

Figure 4.5: Test set example images from each PR category considered as part of the segmentation network: (a) Control; (b) Shoot Thinning; (c) Light Pruning.

including the outer edge of every organ. To distinguish connected organs within a PR (i.e. arm vs. spur, spur vs. canes) from occlusions and close elements of the background, a few millimeters of overlap between annotated areas were kept for contiguous grapevine organs.

(a)   (b)   (c)
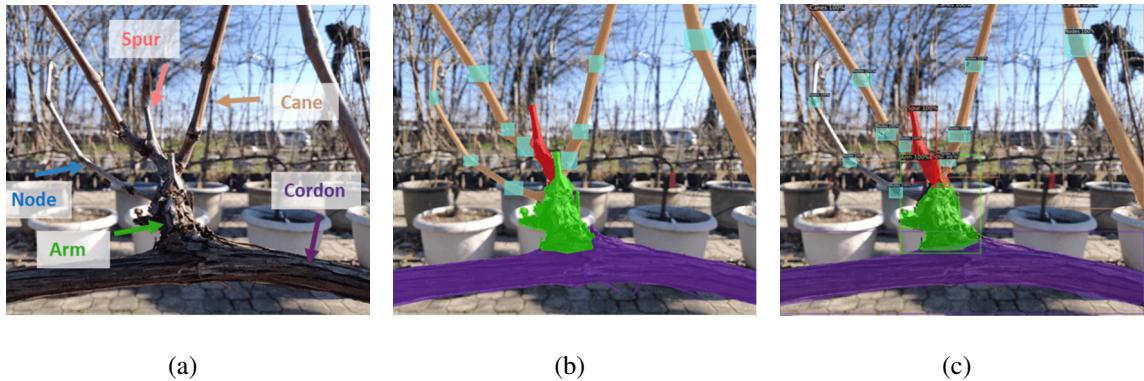
Figure 4.6: Description of the workflow required for fine-tuning a DNN for grapevine organ segmentation: (a) original acquisition with the indication of the 5 relevant classes for winter pruning; (b) annotated image by experts for training the neural network by using 5 different categories: cordon (purple), arm (green), spur (red), cane (brown), node (blue); (c) an example of PR segmentation through Mask R-CNN.

• **Training of the DCNN**

The network was trained on 119 images using COCO2017 pre-trained model weights for the Mask R-CNN. The default training hyperparameters related to the neural network structure were considered. To adapt the model to the relatively small dataset, the number of training iterations was limited to 50000, from the original 270000 and the batch size changed to 2, from the original value of 16.

• **Testing of the DCNN**

The original dataset was randomly split into a training dataset (80%) and a test dataset (20%). Accordingly, 29 images of the test dataset were integrated with 31 images collected in December 2020 as part of a preliminary iteration of the neural network [31]. Such a preliminary batch of images considered the highest morphological variability of grapevine pruning regions encompassing unthinned grapevines (C) spurs subjected to early-season shoot thinning (ST), and light pruning (LP) that is generally undesired because favoring acrotony due to the node-count per spur>2 (Figure 4.6). Therefore, in November 2021 the network was tested on a batch of 60 images representing several canopy management conditions hereafter described as treatment (T). The segmentation output (Figure 4.6 (c)) was composed of inferences provided with an ID, a class label and the corresponding confidence value, and Intersection over Union (IoU) to quantify the overlap between the annotated organ and the model inference.

Evaluation Criteria

For each dataset of **Experiment 1**, the network returned bounding boxes identifying the potential pruning regions (PPRs) of the selected images. Consequently, model predictions (PPRs) were compared with actual PRs, and three possible outcomes were considered: true positive (TP) when the prediction correctly matched with the corresponding PR; false positive (FP) when the prediction did not correspond to a PR; false negative (FN) in case PRs

were not predicted by the DNN. In addition, FPs were divided into the following 6 categories: arm (old wood growing from the cordon), cane (an intermediate portion of a cane), cordon (a portion of the permanent cordon of a target vine), next-row trunk (NRT) (intersection between the cordon in the foreground and a trunk in the background), old cuts (OC) (a portion of the cordon where previous cuts were performed), post (a component of the vineyard trellising). For each FP category, the percentage of false positives (FP%) was calculated as follows:

$$FP(\%) = \frac{FP}{(TP + FP)} \tag{4.1}$$

For each of the 5 classes measured within **Experiment 2**, the output of the grapevine segmentation network was compared to the annotated images. The correctness of a detected grapevine organ was assessed through the IoU overlap with the corresponding ground truth labeling [75, 105]. The IoU overlap was defined according to the following equation:

$$IoU = \frac{(Annotation\ area \cap Inference\ area)}{(Annotation\ area \cup Inference\ area)} \tag{4.2}$$

Within every class, a detected object was assumed as a true positive (TP) when its IoU overlapping with the ground truth annotation was higher than 0.5 [106]. The output was classified as a false negative (FN) when a detected organ did not reach the minimum IoU threshold. The output was classified as false positive (FP) in the case of no overlap with the corresponding ground truth annotation. For FPs, the misclassed grapevine organ or other element was described and considered for further analysis.

In both **Experiments 1** and **2**, the neural network performances were evaluated through recall, precision, and F1 scores, which were calculated for the overall object population of the different datasets [107]:

$$Recall = \frac{TP}{(TP + FN)} \tag{4.3}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{4.4}$$

$$F1 = 2 * \frac{(TP * FP)}{(TP + FP)} \tag{4.5}$$

As part of **Experiment 1**, the same indices were also calculated depending on PR Visibility. In the case of WxV, OxV, and WxOxV interactions, the mean values of the Recall index were calculated and compared by the standard error. In the case of **Experiment 2**, the performance metrics were calculated based on the grapevine organ (GO), treatment (T), and their interaction (GOxT).

Results and Discussion

• **Results of Experiment 1**

The Merlot dataset included 40 pruning regions mostly featured by simple spurs (43%) and coplanar orientation. Simple spurs were also the most common wood type in Sangiovese (73%) where almost half of the PRs were coplanar (51%) with the row's vertical axis. Moreover, most of the PRs were clearly visible in both the Merlot (68%) and Sangiovese (77%) datasets (Figure 4.7).

The PR identification in Merlot was characterized by lower recall (0.66) and higher precision (0.87) rates, with a corresponding F1 score of 0.75 (Table 4.1). In Sangiovese, the DNN performances were represented by the following metrics: 0.59 recall, 0.96 precision, and 0.73 F1 score (Table 4.1). Correct PR's identification was higher in visible spurs with a dramatic recall decrease from 0.72 to 0.53 and from 0.70 to 0.27 when considering the occluded PRs in Merlot and Sangiovese vines, respectively (Table 4.2).

Because of the improvement of DNN performance fostered by visible PRs, the detection model was then assessed as based on the WxV and OxV interactions (Figure 4.8). The
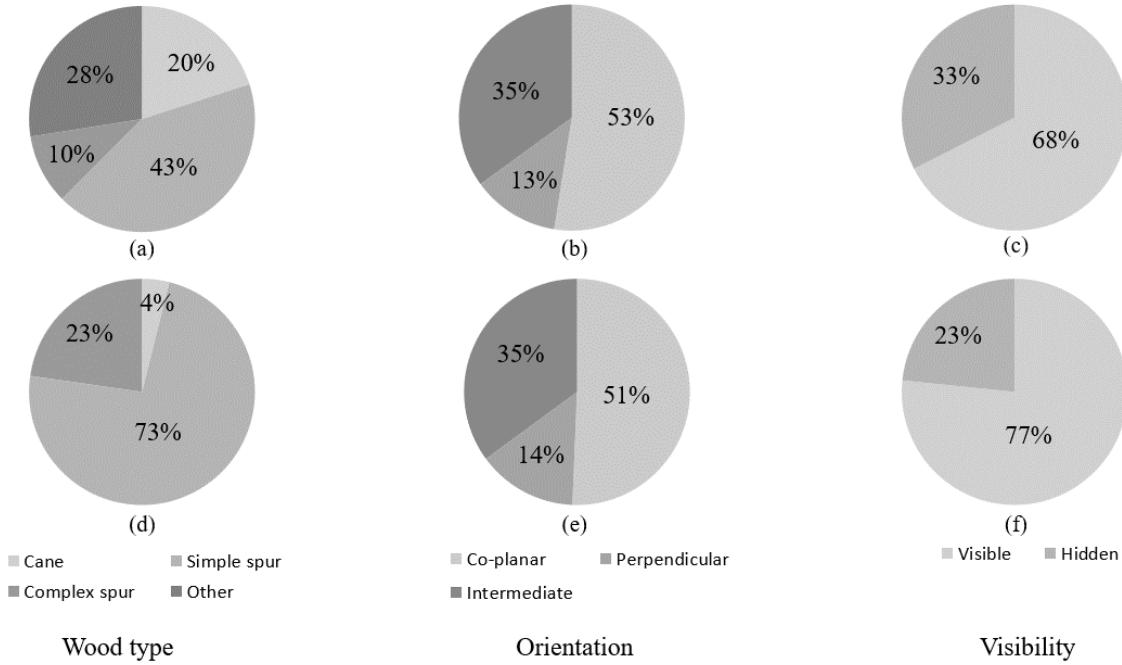
Figure 4.7: Pruning regions (PRs) breakdown according to Wood Type (a, d), Orientation (b, e), and Visibility (c, f) against the Merlot (a–c) and Sangiovese (d–f) datasets. Merlot N = 40, Sangiovese N = 154.

best detection for Merlot grapevines was reported for visible complex spurs with 0.85 recall followed by visible simple spurs and canes. Simple spurs were associated with the lowest standard error (SE=0.07) as compared to the other classes (Figure 4.8 (a)). Visible coplanar spurs showed the highest detection rate (0.75 recall) as compared to perpendicular (0.66) and intermediate PRs (0.65) (Figure 4.8 (b)). Similarly, visible complex spurs of Sangiovese grapevines were associated with the highest recall (0.85) and simple spurs were the second most detected wood type (0.72); moreover, consistency of detection performance was proved by relatively low standard errors (0.06 vs. 0.05). Conversely, the same metrics worsened for visible canes showing the lowest recall (0.25) and inconsistent detection (SE=0.25) (Figure 4.8 (c)). Both Intermediate and coplanar spurs showed the highest detection rate (recall=0.74) as compared to perpendicular PRs (0.67) (Fig. 6d). Notably, the variability of the

Table 4.1: Performance measures of the Faster-RCNN 2.0 vision approach for PR detection against the Merlot and Sangiovese datasets.

	Images	#TP	#FN	#FP	Recall	Precision	F1 Score
Merlot	202	1007	523	149	0.66	0.87	0.75
Sangiovese	30	100	69	4	0.59	0.96	0.73

(a)   (b)

(c)   (d)
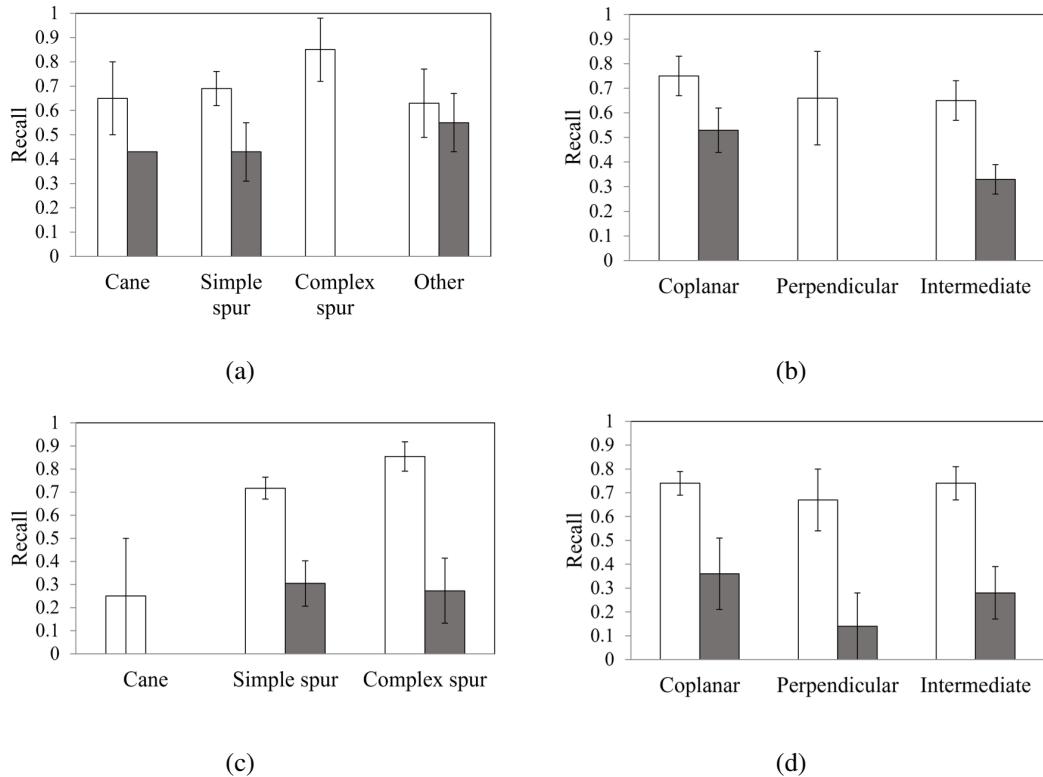
Figure 4.8: Variation over Wood type (a,c) and Orientation (b,d) of the Recall index as a function of PR's Visibility in the Merlot (top) and Sangiovese (bottom) datasets. Visible and hidden PRs are reported in white and grey, respectively. Bars represent the mean value ± SE.

Table 4.2: Performance measures of the PR detection model against the Merlot and Sangiovese datasets depending on PR visibility.

Dataset	Visibility	#TP	#FN	Recall	Confidence
Merlot	Visible	759	299	0.72	0.82 ± 0.01
	Hidden	248	224	0.53	0.81 ± 0.01
Sangiovese	Visible	89	39	0.70	0.82 ± 0.06
	Hidden	11	30	0.27	0.82 ± 0.04

recall index calculated for several wood types and orientations for hidden pruning regions was generally higher as compared to clearly visible PRs (Figure 4.8).

Some categories such as coplanar complex spurs, intermediate and perpendicular canes, and examples belonging to the category other were never observed (Table 4.3). Visible complex spurs with intermediate orientation were detected with 0.97 recall and a SE of 0.03. Similarly, recall values were higher than 0.9 for visible perpendicular simple spurs while the detection performance for the same PR with coplanar orientation did not reach 0.75. Irrespective of their orientation, the percentage of TPs associated with hidden simple spurs ranged between 41% and 43% in intermediate and coplanar spurs, respectively (Table 4.3). Simple and complex spurs were mostly considered in Sangiovese grapevines. When clearly visible, both coplanar and intermediate complex spurs were associated with the highest recall scores (0.85) followed by coplanar simple spurs (0.74). Detection performance for perpendicular and intermediate simple spurs was close to 0.7. Moreover, in both datasets, the recall index was mostly lower than 50% when PRs were hidden.

In Merlot, the most frequent FP category was represented by arms detected in 71 cases out of a total of 149 FPs corresponding to an FP(%) of 6.14% (Table 4.4). Old cuts (OC), canes, and cordons were associated with the following FP(%): 1.73, 0.43, and 0.09%, respectively. Only 4 FPs were categorized in Sangiovese grapevines out of the 154 PRs with a negligible impact on the detection performance.

Table 4.3: Detection rate of the interactions between Wood Type, Orientation, and Visibility in the Merlot (top) and Sangiovese (bottom) datasets.

	Wood Type	Orientation	Visibility	
			Visible	Hidden
Merlot	Simple Spur	Coplanar	0.74 ± 0.11	0.43 ± 0.15
		Perpendicular	0.91	-
		Intermediate	0.51 ± 0.06	0.41
	Complex Spur	Coplanar	-	-
		Perpendicular	0.73	0.27
		Intermediate	0.97	0.03
	Cane	Coplanar	0.76 ± 0.12	0.43
		Perpendicular	-	-
		Intermediate	-	-
	Other	Coplanar	-	0.97 ± 0.13
		Perpendicular	-	-
		Intermediate	0.63 ± 0.14	0.29 ± 0.08
Sangiovese	Simple Spur	Coplanar	0.74 ± 0.06	0.38 ± 0.18
		Perpendicular	0.69 ± 0.13	-
		Intermediate	0.67 ± 0.10	0.40 ± 0.16
	Complex Spur	Coplanar	0.85 ± 0.11	0.50 ± 0.50
		Perpendicular	1	1
		Intermediate	0.85 ± 0.09	0.13 ± 0.13
	Cane	Coplanar	0.33	-
		Perpendicular	-	-
		Intermediate	-	-

Table 4.4: Description of the FPs detected during the DNN testing. NRT: next row trunk; OC: old cuts.

	Class	#FP	Average Confidence	FP (%)
Merlot	Arm	71	0.79 ± 0.16	6.14
	Cane	5	0.77 ± 0.04	0.43
	Cordon	1	0.71	0.09
	NRT	52	0.74 ± 0.23	4.50
	OC	20	0.79 ± 0.04	1.73
Sangiovese	Arm	2	0.74	2.00
	OC	2	0.83 ± 0.04	2.00

• **Discussion of Experiment 1**

The fine-tuned network for PR detection of spur-pruned grapevines was tested on 2 datasets representative of different cordon ages, cultivars, and growing conditions. The overall recall values were relatively similar between the 2 datasets with slightly higher detection rates in Merlot (recall = 0.66) as compared to the younger Sangiovese grapevines (recall = 0.59) (Table 4.1). Indeed, it must be considered that despite being collected in different years, both training and test datasets for the Merlot included grapevines belonging to the same vineyard, suggesting a higher similarity among the PRs. Conversely, even if referring to whole cordon RGB images, taken from a greater distance from the plant with respect to the training setup, the Sangiovese dataset was totally new as part of the life cycle of the model, proving its consistency. Looking at absolute recall values (Table 4.1), the system is less powerful than a branch detection model developed in an apple orchard [75] at 70% confidence threshold, where using pseudo-color images, and pseudo-color and depth images led to 0.84 and 0.89 average recall, respectively. Moreover, PR's visibility affected the detection process in both datasets (Table 4.2). The significant difference between the detection rates of visible and hidden PRs is due to occlusions, a well-known problem in computer vision and in agricultural applications that are frequently performed in unstructured environments [76, 108]. Getting recall scores higher than 0.7 in visible PRs of both the Merlot and Sangiovese testing datasets is an additional confirmation of the detection model consistency. In addition, the occlusion problem mainly depending on PR-to-PR, cordon-to-PR, and trellis elements-to-PR interactions could be tackled by having both sides of the canopy scanned by the vision system, emerging as a relatively easy solution for spur-pruned grapevines where spurs are mainly localized on the upper side of the permanent cordon (Figure 4.3 (a)).

When analyzing the DCNN sensitivity as a function of different factors such as wood type, orientation, and visibility, recall rates were massively improved for some specific categories, with visible intermediate complex spurs showing the highest values in both datasets, followed by visible coplanar simple spurs (Table 4.3). However, complex spurs represented

just a minor part of the actual PRs as well as only 13–14% of the annotations in both the datasets had intermediate orientation, while a larger proportion of actual PRs fell in other categories such as simple spurs and coplanar orientation (Figure 4.7). In addition, regardless of the dataset, the consistency of the detection performances for visible simple spurs is confirmed by the lower standard error associated with the higher count. The poorer cane detection might be due to their scarce representation in the training dataset that was created by including all the PRs belonging to a given number of grapevines irrespective of their different morphology (Figure 4.4). Another interpretation of PR detection results depending on wood type should consider their complexity. In fact, considering individual canes as a major element of a pruning region, the model resulted in better detection of the PRs featured by higher cane numbers suggesting that the DCNN successfully learned how to identify a pruning region based on such a distinctive trait. On the other hand, the same trend would be defined if the model would have just more easily detected bigger pruning regions in terms of encumbrance and area. Similarly, because of the camera orientation considered during the acquisition campaigns, the OxV interactions resulted in the highest recall values for coplanar PRs, and lower values were obtained for intermediate and perpendicular PRs. In fact, due to their cane orientation, coplanar PRs cover a higher image area compared to intermediate and perpendicular PRs with higher overlapping leading to a higher proportion of occluded pixels (Figure 4.4). Merlot WxOxV interactions revealed detection performances of specific PRs (Table 4.3). Although they produce the best detection results (recall = 0.97), visible intermediate complex spurs are not discussed here because they are represented by only 2 elements in the dataset. Considering the most frequent categories with a count higher than 4 (Figure 4.7), with a recall of 0.74, visible coplanar simple spurs were the best-detected pruning regions. In this regard, DCNN consistency was confirmed by similar performances described for the Sangiovese dataset. Indeed, even though visible coplanar and intermediate complex spurs were associated with the highest recall (0.85), the most represented visible coplanar spurs had the second-highest recall (0.74). In fact, there were 55 visible coplanar simple

spurs while only 10 visible perpendicular complex spurs and 13 intermediate complex spurs were considered in the testing dataset. The above-mentioned results suggest that the DCNN performance could be improved by either engineering or agronomic adjustments. First, more training data might result in better performance of the deep learning model[109]; second, improved canopy management in summer can condition the canopy architecture leading to a higher proportion of coplanar simple spurs.

The main wrong detection was represented by arms (Table 4.4), the permanent ramifications growing from the cordon whose number and length might increase over years because of wrong pruning strategies. As part of the overall project pipeline, this misclassification could be considered as a correct identification since the PR detection algorithm is expected to be followed by the segmentation network for analyzing the whole region and recognizing 5 different grapevine organs including arms. However, the arm detection was considered as an FP because the annotation acting as true data required the inclusion of the whole PR (Figure 4.3). Due to the overlap between the permanent cordon in the foreground with the trunks in the background, NRTs were detected by the model as actual pruning regions representing the second most frequent FP category. The incorrect detection of NRT might be decreased by using depth data to filter the image following the study of [110], where a 1.2 m threshold was used to separate apple tree canopies from the background to improve apple detection. Considering the project pipeline, a higher precision might be pursued; however, PR detection will be followed by PR segmentation and the exclusion of wrong detections.

Table 4.5: Overall performance of the neural network for grapevine segmentation with an IoU of 0.5.

Count	#TP	#FN	#FP	Recall	Precision	F1 Score
1359	1069	258	32	0.81	0.97	0.88

- **Results of Experiment 2**

The general performances of the segmentation network were described by a recall of 0.81 and a precision of 0.97 with an F1 score of 0.88 (Table 4.5).

The most recurrent GO in the testing dataset was node followed by cane, spur, arm, and cordon (Table 4.6). FPs related to each class were generally low. The highest recall value was scored by nodes (0.88), followed by cordon and arms (0.81), while spur and cane classes revealed a recall of 0.72 and 0.68, respectively. The precision values ranged from 0.96 (node) to 1 (cordon) with arm and spur segmentations showing intermediate performances.

Table 4.6: Performance measures of the neural network for grapevine segmentation depending on 5 different grapevine organs with an IoU of 0.5.

Organ	Count	#TP	#FN	#FP	Recall	Precision	F1 Score
Cordon	75	61	14	0	0.81	1.00	0.90
Arm	89	71	17	1	0.81	0.99	0.89
Spur	108	77	30	1	0.72	0.99	0.83
Cane	343	229	107	7	0.68	0.97	0.80
Node	744	631	90	23	0.88	0.96	0.92
Total	1359	1069	258	32			

Table 4.7: Performance measures of the neural network for grapevine segmentation depending on canopy management with an IoU of 0.5. C = Control, ST = Shoot Thinning, LP = Light Pruning.

Treatment	Count	#TP	#FN	#FP	Recall	Precision	F1 Score
C	619	493	121	5	0.80	0.99	0.89
ST	590	487	88	15	0.85	0.97	0.90
LP	150	89	49	12	0.64	0.88	0.74
Total	1359	1069	258	32			

For canopy management, the most represented category was control (C) with 619 annotations (Table 4.7) followed by shoot thinning (ST) and light pruning (LP) with 575 and 150 grapevine organs to be segmented. TPs were 493 in C, 487 in ST, and 89 in LP with the high-

est recall values calculated for grapevine organs subjected to ST (0.85), and relatively lower performances described in C (0.80); moreover the segmentation of the grapevines subjected to light pruning led to the lowest recall. With only 5 wrong inferences, precision was highest in C (0.99), with similar responses described for ST organs despite the higher number of false positives (15). Conversely, although the FPs in LP (12) were relatively similar to ST, the precision was much lower (0.88).

To investigate if and how vineyard management influences dormant canopy segmentation, the model was tested against each T × GO combination (Table 4.8). In C cordons, 24 TPs and 6 FNs out of the 30 annotations resulted in a recall of 0.80. Therefore, the F1 score was 0.89. Arm segmentation was described with higher recall (0.87) and F1 score (0.93), while the model resulted in poorer performance to identify spurs (0.72 recall and 0.84 F1 score) and canes (0.64 recall, 0.78 F1 score) characterized by a higher number of annotations (169). No FPs were counted in these grapevine organs, giving a precision of 1. The node class had the highest recall (0.89) having 296 TPs and 37 FNs. The model returned 5 wrong classifications (FPs), lowering precision to 0.98 and limiting the F1 score to 0.93.

As expected, ST showed a lower count than C for annotated canes and nodes, and a similar number of annotated elements for cordons, arms, and spurs. When compared to C, in ST grapevines the recall values increased for cordon (0.91), spur (0.77), and cane (0.76) with no or minor changes for nodes (0.89) and arms (0.85), respectively. Although correct inferences in ST proportionally increased as compared to C, the model errors also increased affecting the precision for most of the classes such as arm, cane, and node, showing the following values: 0.97, 0.94, and 0.97, respectively (Table 4.8).

The light pruning (LP) presented a lower number of annotated GO (Table 4.8). In most cases, FNs were similar to, or higher than TPs. This condition was mirrored by the performance metrics such as recall and F1 score revealing the lowest values within the experiment. Both the recall and F1 scores identified poor segmentation performances for arms and spurs (0.33 and 0.38 recall, respectively), and higher sensitivity for node detection (0.76 recall).

Precision was mostly affected by count varying between 0.75 (spur) and 1 in the case of cordons and arms where no FPs were detected.

Several elements belonging to the grapevine or to the surrounding environment were associated with wrong predictions such as arms, spurs, canes, and nodes (Table 4.9). Nodes were the most wrongly attributed class since 3.6% of the inferences were FPs. The second most frequent incorrect class attribution concerned canes (3.03%), while wrong segmentation of arm and spurs was limited to 1.39% and 1.28%, respectively.

Table 4.8: Performance measures of the neural network for grapevine segmentation depending on canopy management and grapevine organs with an IoU of 0.5. C = Control, ST = Shoot Thinning, LP = Light Pruning.

Treatment	Organ	Count	#TP	#FN	#FP	Recall	Precision	F1 Score
	Cordon	30	24	6	0	0.80	1.00	0.89
	Arm	39	34	5	0	0.87	1.00	0.93
C	Spur	43	31	12	0	0.72	1.00	0.84
	Cane	169	108	61	0	0.64	1.00	0.78
	Node	338	296	37	5	0.89	0.98	0.93
	Cordon	34	31	3	0	0.91	1.00	0.95
	Arm	41	34	6	1	0.85	0.97	0.91
ST	Spur	56	43	13	0	0.77	1.00	0.87
	Cane	141	103	32	6	0.76	0.94	0.84
	Node	318	276	34	8	0.89	0.97	0.93
	Cordon	11	6	5	0	0.55	1.00	0.71
	Arm	9	3	6	0	0.33	1.00	0.50
LP	Spur	9	3	5	1	0.38	0.75	0.50
	Cane	33	18	14	1	0.56	0.95	0.71
	Node	88	59	19	10	0.76	0.86	0.80
Total		1359	1069	258	32			

- **Discussion of Experiment 2**

The current study allowed the fine-tuning and testing of a novel DCNN for grapevine organ identification at 0.5 IoU and 0.7 confidence resulting in the following performance metrics:

Table 4.9: Description of the FPs detected during the testing of the neural network for grapevine segmentation with an IoU of 0.5.

Detected Class	True Class	Count	Confidence (Mean)	FP (%)
Arm	Cane	1	0.91	1.39
Cane	Other object	3	0.89	1.29
Cane	Arm	2	0.95	0.87
Cane	Other grapevine organ	2	0.96	0.87
Node	Other object	11	0.97	1.71
Node	Arm	2	0.99	0.32
Node	Other node	7	0.98	1.10
Node	Other grapevine organ	3	0.99	0.47
Spur	Other grapevine organ	1	0.99	1.28
Total		32		

recall of 0.81 and a precision of 0.97 (Table 4.5). As already mentioned about PR detection, the current results suggest that assuming lower confidence would increase the network sensitivity towards the grapevine organs' identification; as a matter of fact, the general improvement of the detection process would lead to an increased recall at the expense of precision because of the higher number of inferences (TP and FP) regardless of their correctness (Table 4.5). When considering its sensitivity in detecting the 5 organ classes featuring the grapevine canopy over winter, the DCNN resulted in different performances as reported in Table 4.6. With a recall of 0.88 (i.e. 88% of the specific annotations identified), nodes were the best-detected class showing an important improvement on previous results reported by Dìaz et al. [78] that processing RGB images through computer vision and machine learning algorithms identified grapevine buds with a maximum recall of 0.45. Because of the grapevine structure, nodes were the most represented class in the test dataset (Table 4.6). The higher number of nodes in each training image can explain why they are the best-segmented class. Consequently, further improvement of the current DCNN version work will consist in providing more training examples of the under-represented classes such as cordons, arms, and spurs to have a more balanced dataset and consistent results among the 5 classes. In addition to the different abundance of training data, the heterogeneous performances describing our

segmentation process can be explained by the different GO sizes (i.e. thickness and width) characterizing a grapevine canopy over dormancy. Data reported in Table 4.6 describe a higher segmentation rate for bigger organs such as cordons and arms (recall = 0.81) while spurs were less detected (recall = 0.72) because of their thinner structure. The importance of the size of target organs is also confirmed when comparing segmentation performances described for arms and spurs; indeed, because a spur might be considered as the natural continuation of an arm, and the ratio between their count approaches 1 in both training and test datasets, the higher detection described for the arms might depend on the more complex structure characterizing a permanent organ older than 2 years as compared to a 2-year-old spur [84]. Canes, despite being the worst detected organ, by the present algorithm (recall = 0.68) were associated with a higher recall value compared to the results reported by Botterill et al. [50] with the 2D cane detector (0.49). The generally worse segmentation results obtained for spurs and canes can be linked to the higher probability of getting occlusions. Bigger and isolated organs such as cordons and arms are much less subject to occlusion than spurs, relatively thin and short elements surrounded by canes, and canes which are often crossing each other or self-occluding [50]. Precision values in **Experiment 2** are significantly high because of the low number of false positives for each of the five classes.

Canopy management greatly affected the segmentation results showing the best detection performances in ST grapevines where only one shoot per node was kept (Table 4.7). Consequently, an ST canopy has fewer elements to be detected, fewer potential occlusions, and a more standardized canopy that leads to better results when applying computer vision algorithms. However, the three treatments revealed different results in terms of GO segmentation (Table 4.8). Despite slightly improving the overall performances, C followed the same ranking already described in Table 4.6 with recall values decreasing in the following order: node>arm>cordon>spur>cane. Segmentation of ST canopies revealed the highest recall values; specifically, cordons (0.91) were followed by nodes (0.89) and arms (0.85). Node segmentation is described by the same recall value. The reason recall does not decrease in

C treatment is probably due to a lower frequency of occlusion since nodes could only be masked by very thin organs such as canes. In parallel, nodes were successfully segmented also in LP because their morphology did not differ among treatments, while segmentation performances dramatically decreased for the other organs in response to altered growth patterns and PR's morphology induced by highly variable spur length. Shoot thinning is a summer pruning technique reducing disease pressure, and improves canopy microclimate, vine balance, and grape quality to increase the sustainability of viticulture [95]. Moreover, this practice allows more efficient shoot positioning in VSP-trained vines due to the reduced shoot number, making the management of their growth direction and orientation easier and, in turn, facilitating winter pruning operations. ST becomes a quite promising practice in vineyards that will be subjected to automated robotic pruning because of the following reasons: (i) better performances of perception modules such as PR detection and GO segmentation due to the increased proportion of simple spurs and limited frequency of occlusions; (ii) better performances of the manipulation module, by facilitating the motion planning to reach cutting points as well as the end-effector operability; (iii) a significant decrease in cut number per meter of row impacting on robot capacity. On the other hand, such a key role assumed by canopy management supports the idea that, to reach their maximum efficiency, robotic solutions in agriculture need to be coupled with a "robot-ready" orchard [111, 112].

The segmentation network detected few FPs as compared to correct inferences (Table 4.9). The most recurring error consisted of labeling as a node the Other objects such as a variety of small, round, and point-like objects of the image background. The segmentation of "other nodes" as "nodes" mainly included blind buds at the base of longer spurs retained in LP treatment (Figure 4.5 (c)). Due to acrotony, distal shoots of an upward spur show preferential growth during the season, inhibiting bud breaking of the lower nodes that lose the possibility to develop new shoots in the next season even if keeping a relatively similar morphology [113]. The risk associated with this segmentation is that if the old nodes were counted as real, a pruning algorithm could schedule a wrong cut, targeting a spur instead of

a cane.

4.2 Pruning Poses Generation

4.2.1 Potential Pruning Points Localization

Finding the desired potential pruning points requires an additional processing layer after receiving the inference generated by the trained neural network. The approach we take is to build a processing layer that interprets this inference by identifying the connections between the different segments [114]. The data structure used to host these data is a five-class graph,

Table 4.10: Relationships between the different types of organs.

Parent Organ Type	Children Organs Types
Cordon	Arm, Spur, Cane
Arm	Spur, Cane
Spur	Cane
Cane	Node

with the grapevine cordon as the root element, as shown in Figure 4.9. Then, there are arms or spurs connected to the cordon, and canes are connected to spurs or to the cordons. The leaf elements of the graph structure are nodes. The important concept that we want to analyze is the relationship between canes, and nodes on the canes and canes and spurs within a pruning region, in order to obtain accurate pruning points. There are five sets of connections in the graph structure, "cordon to arm" relating the cordon to its connected arm, "cordon to spur" relating the main cordon to its connected spur, "spur to cane" relating the spur to its connected cane, the "arm to spur" set, relating the arm to its connected spurs, and the "cane to node" set relating each cane to its connected nodes. Based on the new five-class

categorization, we defined a new set of more specific and stricter relationships between the different types of organs, to create the new plant model, listed in Table 4.10.
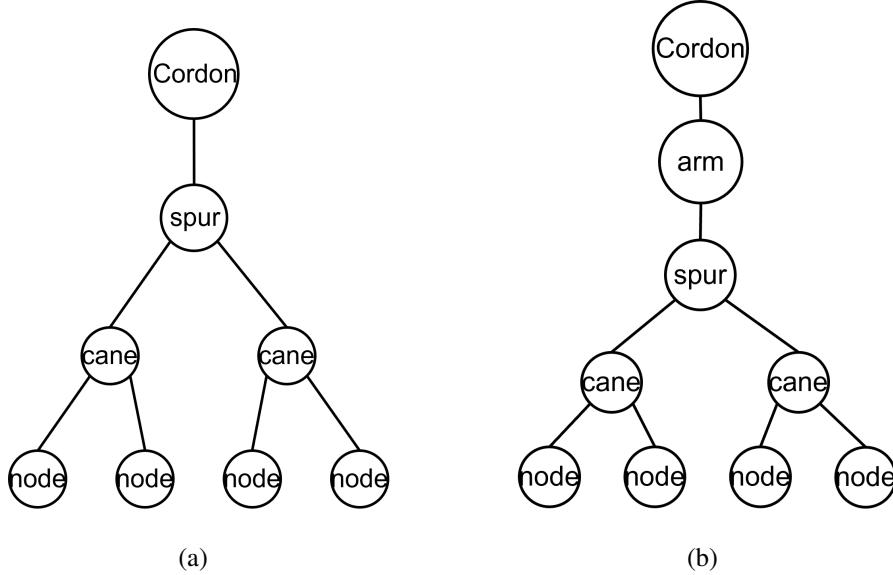
Figure 4.9: The data structure of grapevine graph: (a) grapevine structure with four classes, where the spur is directly connected to the cordon; (b) grapevine structure with five classes.

Figure 4.10 presents an example of a created plant graph Figure 4.10 (c) overlaid on the segmentation output Figure 4.10 (b). This five-shaped graph connects the various grapevine-inferred items, showing the topographical structure of the plant. Each grapevine item consists of a unique identifier number, bounding box coordinates, score, segmentation mask, class identifier, class name, class color, item color, center, thickness, distance from the parent, depth, and parent. This item is based on the common aspects of the neural network classes and extended depending on the class, where the cordon contains a list of arms or spurs, where canes grow on it, and canes are sorted by their distance to the parent, and the canes have a list of nodes, sorted using the same metric. In the end, the graph's root node is the cordon item.

Based on such a structure, the pruning point localization can be carried out after the structure has been established. Each pruning scheme is described by a feature vector of

(a)   (b)   (c)
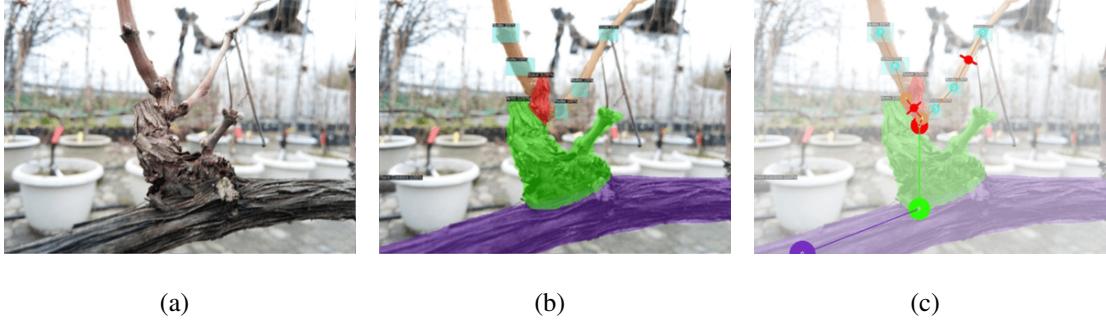
Figure 4.10: An output example of potential pruning points generation: (a) raw image collected in the vineyard; (b) segmentation output, where purple represents the cordon class, green the node arm class, red the spur class, orange the cane class, and light blue the node class; (c) graph connection schematic with potential pruning points marked with red points.

appropriate attributes of the canes that will be kept (length, position, the angle from the head, distance below wires, and whether they grow from the head, the trunk, or another cane). Currently, we have adopted a crude approach for detecting potential pruning points, which are points on canes that locate between two nodes of the same cane, or on the top cane located between the bases of two canes growing from the same spur. An example of this can be seen in Figure 4.10 (c), indicated by the red markers. As described in Chapter 3, by default, a potential pruning point ($\overrightarrow{pp}$) is the midpoint between two of the node points, the second node $\overrightarrow{p_2}$ and the third node $\overrightarrow{p_3}$ on a basal cane, or bases of two canes $\overrightarrow{p_{base1}}$ and $\overrightarrow{p_{base2}}$. For localizing the final pruning point, we decided to select the pruning point located above the second node on the basal cane:

$$\forall \left(\overrightarrow{p_2}, \overrightarrow{p_3} \right)$$
$$\overrightarrow{pp} = \frac{\overrightarrow{p_2} + \overrightarrow{p_3}}{2} \tag{4.6}$$

and the pruning point located on the top cane between the bases of two canes on the same spur:

$$\forall \left(\overrightarrow{p_{base1}}, \overrightarrow{p_{base2}} \right)$$
$$\overrightarrow{pp} = \frac{\overrightarrow{p_{base1}} + \overrightarrow{p_{base2}}}{2} \tag{4.7}$$

Due to the possible curvature of the canes, it may happen that this midpoint is not contained in the cane mask, and if this happens, the point is moved to a point inside the mask.
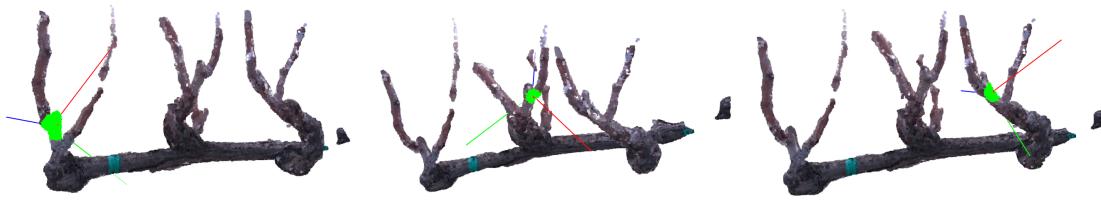
4.2.2 Geometry-Based Pruning Orientation Generation

Robotic pruning must choose the best direction from which to approach the pruning points to minimize occlusion and avoid obstacles that might interfere with the detection along the approach. So, based on the pruning points generated in the previous section, an orientation is needed to form a pruning pose. Given a point cloud and a description of the geometry of a robotic end-effector, the main problem concerning the pruning poses generation is to identify the shears configuration from which a pruning would be formed if the shears were to enclose the target pruning point. Pose generation is an important part of the grapevine robotics pruning pipeline. Estimating the pruning poses in a clustered environment is very challenging and prone to error. For grapevine winter pruning, most of the pruning points are near the cordon body and vine head, causing a narrow space between the canes, which complicated the problem.
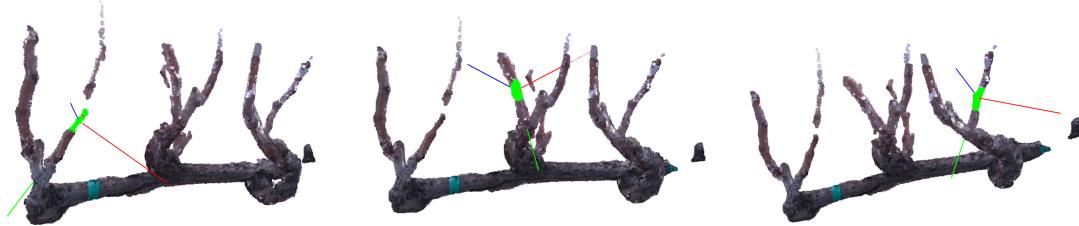
We here adopt a geometric-based approach to analyze the pruning orientation. In order to describe the orientation more conveniently, we use the axis–angle representation of rotation. Based on the geometric features of the local point cloud of the pruning region, firstly, we run the k nearest neighbor (KNN) [115] clustering algorithm to extract the local feature regions of the pruning point and then use the principal component analysis (PCA) [116] algorithm to extract the three main directions of its three-dimensional projection to find its optimal directions toward pruning point. The Rayleigh Quotient is used to solve the maximum and minimum values of the projection variance in PCA to obtain the projection principal vector [116]. The minimum value of the projection variance means that there is less information in this direction indicating fewer obstacles, so we use this eigenvector as the axis target at the pruning region. In order to perform the cut correctly. A rotation angle (φ) of blades is

computed taking into account the slope angle between the straight line connecting the two points and the horizontal direction of the cane. Combined with the rotation angle of the pruning tool generated previously, we can finally get the complete orientation in axis–angle representation, then form the target pruning pose, as shown in Figure 4.11.

$$
\begin{aligned}
\delta_{x,y} &= p_{2_{x,y}} - p_{3_{x,y}} \\[4pt]
\varphi &= \begin{cases}
0, & \text{if } \delta_x = 0 \\[4pt]
\frac{\pi}{2}, & \text{if } \delta_y = 0 \\[4pt]
\arctan \frac{\delta_y}{\delta_x} - \operatorname{sign} \frac{\delta_y}{\delta_x} * \frac{\pi}{2}, & \text{otherwise}
\end{cases}
\end{aligned}
\tag{4.8}
$$

(a) pruning pose for top cane 1   (b) pruning pose for top cane 2   (c) pruning pose for top cane 3

(d) pruning pose for basal cane 1 (e) pruning pose for basal cane 2 (f) pruning pose for basal cane 3

Figure 4.11: Examples of pruning pose generation.

4.3 Conclusion

This chapter describes (i) a methodology for data acquisition and annotation, (ii) a neural network fine-tuning for grapevine segmentation, (iii) an image processing-based method for creating the representative model of grapevines, starting from the inferred segmentation, and (iv) potential pruning pose generation, based on the plant model which is a simplification of the grapevine structure and its geometric features.

The major steps in the perception pipeline involve acquiring static images from the detection position, creating an object segmentation dataset where the primary object is the grapevine in a dormant season scenario, training a neural network for object segmentation, and generating the potential pruning points, which can be done by analyzing the network inference, connecting the different grapevine organs to each other in a graph-based structure that allows the generation of the desired potential pruning points on a spur-pruned grapevine. We presented a novel method to create a 5-class grapevine data structure, based on grapevine semantic segmentation, containing the topographical and geometrical information between the different grapevine organs. Based on the obtained 5-class grapevine structure and pruning rules, we are able to generate potential pruning points. Then, combined with the local geometric structure information of the vine around the pruning point to generate an approaching orientation, we finally obtained the pruning pose.

5 Motion Planning for Pruning Points Approaching

Once target potential pruning points are localized by the perception module and then converted to 3D pose estimates of the pruning locations, the robot plans to approach the pruning points. The precise pruning of a grapevine is a challenging problem due to the limitations of available electric shears for robotic use and the inherent difficulties of planning pruning operations (i.e. approaching the target object, positioning the shear, setting the correct blade inclination,...). In spur-pruned cordons, The majority of pruning points are located near the cordon body. In addition, many canes grow out of or near the vine head, which makes it a dense area for the pruning manipulator. However, this condition is much more recurrent for long-cane trained grapevines (i.e. Guyor system) while in the case of mature cordons is much less common. The narrow space between adjacent canes may cause manipulator collisions with canes when executing pruning tasks. Pruning robots may be damaged since these areas are typically stiff and inflexible. Therefore, it must be highly manipulative to avoid collisions with the vine cordon, trunk, head, and canes. Otherwise, the collision can hurt the dormant buds, hence impacting vine growing and fruiting negatively. Moreover, robot components may impact vineyard trellis parts such as posts, stakes, wires, and other components. All of them may oppose different resistance to the end-effector depending on size, thickness, and type of materials including wood, metal, plastic, and concrete [94]. Due to the

complexity of this manipulation problem, it is usually formulated as a manipulator motion planning problem and tackled. While manipulator motion planning aims at approaching the target pruning points and avoiding obstacles in the scene.

5.1 Planning Framework for Pruning Points Approaching

5.1.1 Design of Planning Framework for Pruning Points Approaching

The grapevine winter pruning planning framework we proposed is "Three-phase Approach Planning" (TAP), where the path of the end-effector to the end goal pose is divided into three discrete sets of trajectories for three subtasks (as Illustrated in Figure 5.1, moving from home pose to detection pose, then to pre-pruning pose and finally to pruning pose). Firstly, the manipulator moves to the detection pose, where we use the eye-in-hand camera to run the perception pipeline as described in Chapter 4 to generate a target pruning pose. Moving between different poses requires complex joint movements to ensure a collision-free path. Currently, there are many path planners available that can solve the problem of path planning in a complicated configuration space of a robotic manipulator. Commonly, the effective path planners that can fast-search any complicated configuration space are sampling-based, e.g., Probabilistic RoadMap (PRM) [117], Rapidly-Exploring Random Tree (RRT) [118] or their modifications [119], which even consider robot dynamics. These algorithms require the following three functions: (i) a collision detector checks a set of joint angles to determine whether the robot will collide with itself or with objects in its environment; (ii) an inverse kinematics solver, which calculates sets of joint angles, putting the robot's end effector in a particular pose; (iii) a local path planner, which tests whether a path between two points that are nearby in configuration space is collision-free. The local path planner works by interpolating configuration points between the two points, and collision-checking each in turn. We take advantage of the STOMP (Stochastic Trajectory Optimization for Motion

Planning) [120] solver to compute the trajectory corresponding to the desired start and end poses. If these optimizers of STOMP fail, we fall back to the RRT-Connect algorithm [121] to generate a plan from scratch. We use the RRT-Connect implementation from the Open Motion Planning Library [122].
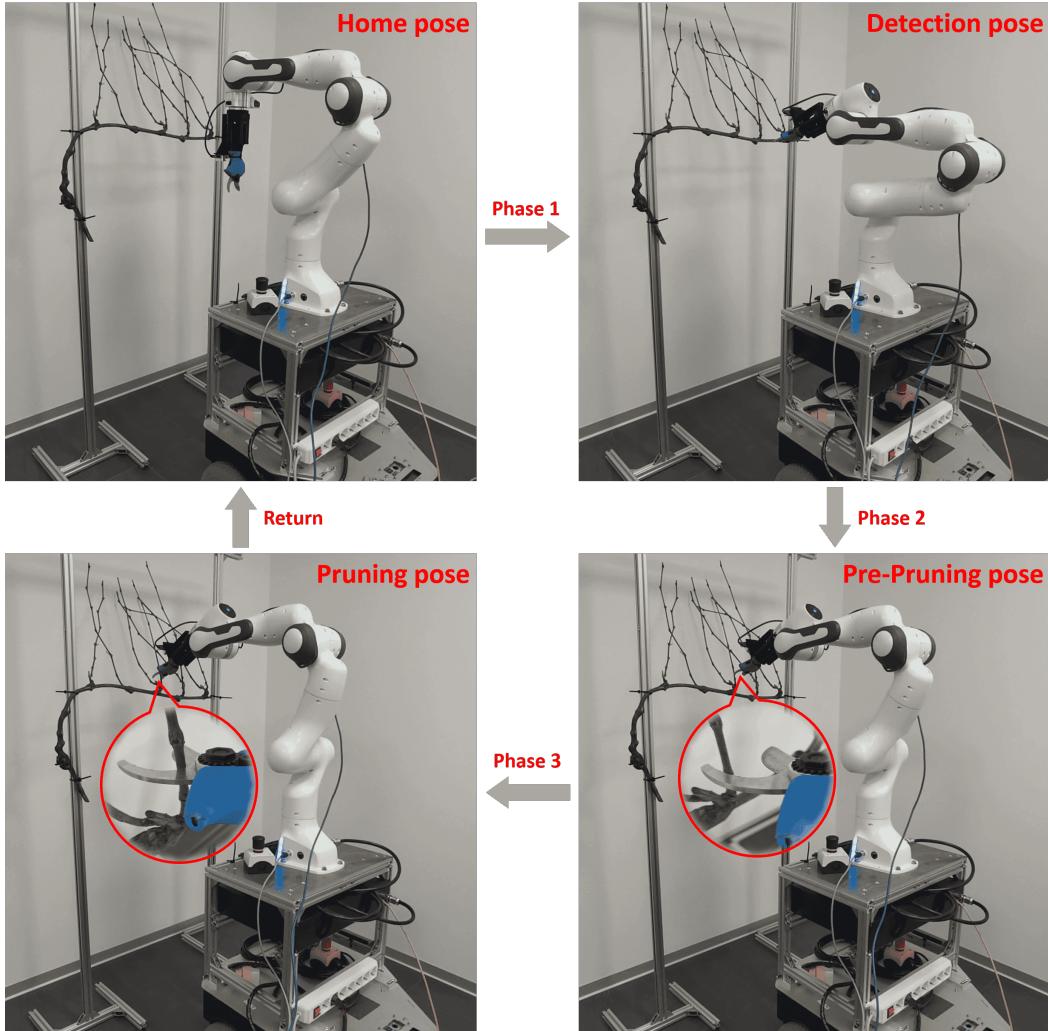
Figure 5.1: Illustration of three-phase approach planning.

Once the poses of the pruning locations were computed, the second phase is to plan the path of the end-effector to the pre-pruning point. The pre-pruning location is a waypoint of the overall approaching path, which is 1.5 cm ahead of the pruning location. The approach
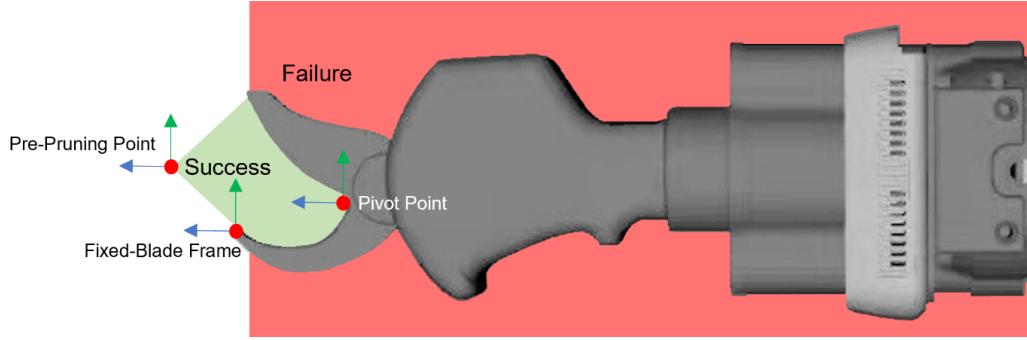
Figure 5.2: The success and failure regions for the electric pruning shear. These are used to query if the cane has entered (or missed) the shear's jaws.

routine moves the shears following the curve generated by quintic polynomial interpolation toward the pre-pruning point. We compute the desired linear and angular velocities using a velocity-based whole-body controller, which will be described in the following subsection, to navigate the end-effector to approach the pre-pruning point. The second phase ends when the robot tool-center-point (TCP) reaches the pre-pruning location while guaranteeing target pruning locate at the "success" region of shears (as shown in Figure 5.2).

Next, in the third phase, the robot shears push forward by calling the learning-based impedance controller, which will be described in Chapter 6, to enclose the target pruning point and then execute the cutting option. Afterward, the robot returns to the detection pose, followed by a stagnation period, and then approaches the next spur. Currently, the order is arbitrary, but in the future, we will optimize the order for speed. Finally, the robot checks whether the whole plant has been pruned. If all pruning is completed, the manipulator will return to the home pose, and then navigate to the next plant. Figure 5.3 shows a flowchart illustrating the process of spur detection and moving the shears to the desired pruning points on the grapevine.

Remark 3 *Because leg whole-body manipulation is outside the scope of this study. We only perform motion planning for the manipulator in the second planning phase while using the*
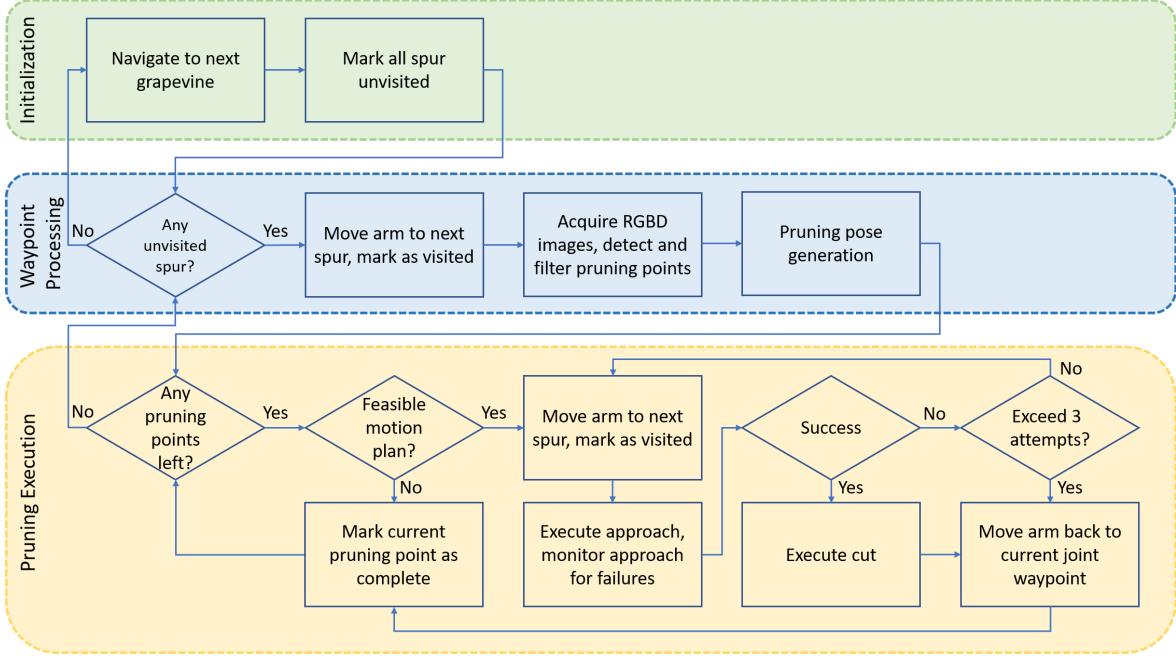
Figure 5.3: Execution flow chart of the pruning system. At each vehicle location, we first initialize the system to its home state. We then alternate between two steps: moving to the next grapevine and detecting pruning points, followed by operating on each of the detected pruning points.

HyQReal-Kinova platform in the vineyard, using the same method as in the first planning phase.

5.1.2   Environment Modeling for Collision Avoidance

Several robotic applications require a 3D model of the environment, which is a key pre-requisite for grapevine winter pruning. The grapevine to be pruned and the surrounding trellis structure, are modeled using an OctoMap [123]. Octomap is an octree-based 3D map creation tool that can display complete 3D graphics including unobstructed and unmapped areas, and sensor data based on occupancy grids can be fused and updated in multiple measurements. This approach uses an efficient data structure based on octrees that enable a

compact memory representation and multi-resolution map queries.

Octree

The octree structure is a data model first proposed by Dr. Hunter in 1978 [124]. The octree structure divides the geometric entities of the three-dimensional space into voxels, and each voxel has the same time and space complexity. Divide to form a direction graph with a root node. In the octree structure, if the divided voxels have the same attributes, the voxel constitutes a leaf node; otherwise, continue to divide the voxel into 8 children. The cube is divided sequentially, and for a spatial object of size ($2n \times 2n \times 2n$), it can be divided up to n times, as shown in the following Figure 5.4:
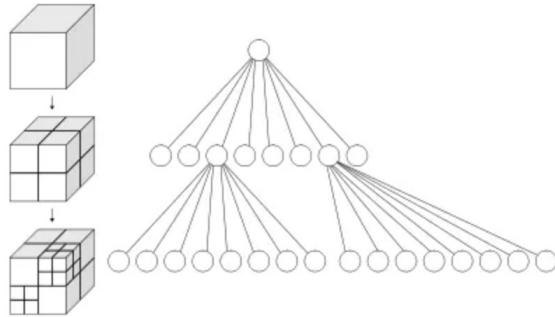
Figure 5.4: Octree structure.

Octomap

The grapevine to be pruned and the surrounding trellis structure are modeled using an OctoMap [123], i.e. a 3D occupancy grid, constructed from point cloud data from the camera, with the following listed features. We currently use data from a single point cloud of grapevine to form OctoMap (illustrated in Figure 5.5). In general, the framework is agnostic to obstacle representation and any convenient representation can be used.
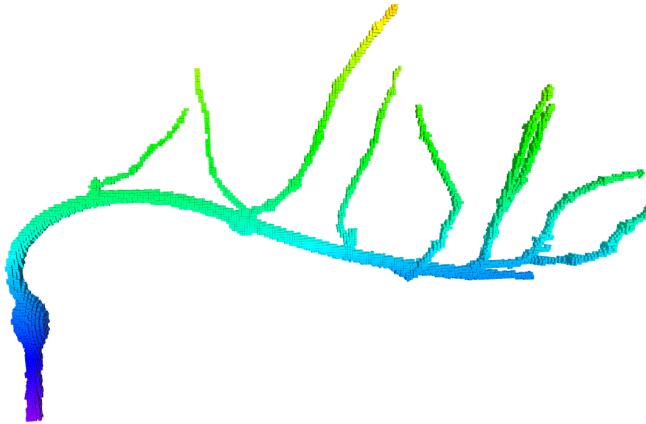
Figure 5.5: Scene representations during the experiments: an image from the simulation showing the OctoMap representation of the tree system.

- **Full 3D model.** The map is able to model arbitrary environments without prior assumptions about them. The representation models occupied areas as well as free space. Unknown areas of the environment are implicitly encoded in the map. While the distinction between free and occupied space is essential for safe robot navigation, information about unknown areas is important, e.g., for autonomous exploration of an environment.

- **Updatable.** It is possible to add new information or sensor readings at any time. Modeling and updating are done in a probabilistic fashion. This accounts for sensor noise or measurements which result from dynamic changes in the environment, e.g., because of dynamic objects. Furthermore, multiple robots are able to contribute to the same map and a previously recorded map is extendable when new areas are explored.

- **Flexible.** The extent of the map does not have to be known in advance. Instead, the map is dynamically expanded as needed. The map is multi-resolution so that, for instance, a high-level planner is able to use a coarse map, while a local planner may operate using a fine resolution. This also allows for efficient visualizations which scale

(a)   (b)
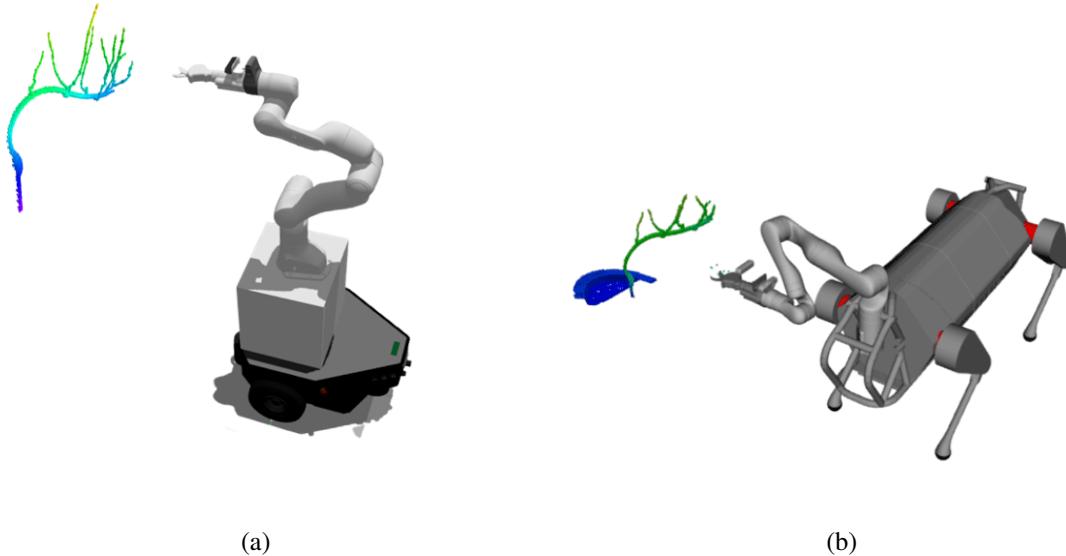
Figure 5.6: Planning scenes with OctoMap: (a) planning scenes of *Rolling Panda*; (b) planning scene of *HyQreal-Kinova*.

from coarse overviews to detailed close-up views.

- **Compact.** The map is stored efficiently, both in memory and on disk. It is possible to generate compressed files for later usage or convenient exchange between robots even under bandwidth constraints.

The preprocessed and adjoined point cloud is fused into a memory-efficient, probabilistic octree-based scene representation that allows for change detection. We compute the forward kinematics for the entire robot along the kinematic chain and reuse intermediate results between different collision checkpoints. Adding more collision checkpoints does therefore not introduce a significant overhead. The full kinematic configuration of the robot is checked against 3D collisions in a dynamically-built OctoMap in an efficient manner. The planning sense is shown in Figure 5.6.

5.1.3   Whole-Body Stack-of-Tasks Control Strategy

Grapevine winter pruning tasks can be accomplished with a mobile manipulator and extends the manipulability and operational space by combining the strengths of the mobile base and manipulator. Two approaches have been investigated. One is a centralized approach, where the loco-manipulation system and the manipulator are seen as unique entities. The motion planner and the controller are then designed for the complete kinematic and dynamic models. The other approach considers the arm separately from the locomotion system. To automate the grapevine winter pruning, a hierarchical control formulation based on multi-task scheduling has been designed for whole-body motion planning.

This combined, more complex structure creates motion planning and control issues due to the high redundancy and versatility of mobile platforms [125]. Previously, the mobile base and manipulator could be considered to be two separate subsystems [126, 127, 128, 129]. Wrock et al. created an automatic switching scheme that used teleoperation methods to achieve decoupled mobile manipulator motion. The decoupled motion refers to a motion pattern in which the mobile base maintains steadfast while the manipulator is operated or vice versa. This pattern is able to provide excellent precision tracking precision, but this takes somewhat longer to complete the task because tracking must be halted during motion. Coordinated movement of the manipulator and mobile base [68, 130] is required to fully utilize the strengths of the mobile base and manipulator and boost efficiency.

De Luca et al. [131] put forward a comprehensive theory to cope with modeling and redundancy resolution for non-holonomic mobile manipulators. A multi-task whole-body regulating strategy based on velocity control was proposed by Li et al [132] for a highly redundant mobile manipulator. The end effector of the mobile manipulator follows the pre-defined trajectory while scheduling low-priority control primitives by using null-space projection. To obtain the inverse kinematics, Roberto et al. [133] provided a systematic mobile manipulator solution that included a selection of redundancy parameters. The solution was

capable of managing collision avoidance, motion restriction, and dexterity enhancement.

In this subsection, a whole-body motion controller enhanced by hierarchical tasks that can regulate a non-holonomic mobile manipulator for grapevine winter pruning is proposed. The desired trajectory in task space is arranged as the top-priority task and is created by using quintic polynomial programming. Contradictions between the primary tasks and the constraint tasks are processed within the stack-of-tasks (SoTs) framework [134, 135, 136] by appropriately establishing a priority planning to all considered tasks and then ensuring that the lower priority tasks are able to project into the null space of the higher-priority tasks.

Joint Limits

While performing the joint-limits avoidance as one of the subtasks, we would like to keep the manipulator as far as possible from its joint limits. Thus, in the case of avoidance of the joint limits [137], the performance criterion can be written as to maintain the manipulator as close as possible to the mid-joint position $\bar{\xi}$:

$$\mathcal{V} = \min_{\xi} \frac{1}{2} \left\| \xi - \bar{\xi} \right\|_W^2 \tag{5.1}$$

where $W = \mathrm{diag}(w)$. The setting of the weighting vector w of Equation (5.1) is very important. If w is too small, the redundant displacement may not be sufficient to avoid the joint limits; if w is too large, the redundant displacement may produce high joint velocities. Therefore, w is usually set based on trial and error.

Manipulability

Robot manipulability refers to the manipulating ability of a robot within the task space, and it is used as an indicator to avoid robot singularity. As it is well-known the manipulability measure gives a scalar representation of the gain between joint velocities $\dot{\xi}$ and task velocities $\dot{x}$, and, consequently, measures the ability of the robot to move its end-effector, or,

equivalently, the distance from a singular configuration, i.e., a configuration $\dot{\xi}$ for which the Jacobian is ranked deficient. Singularities occur when a robot loses control in one or more configurations, and they ought to be prevented when planning and controlling robot motion. The further the Jacobian determinant is from zero, the further the robot is from the singularity. Hence, the Jacobian determinant could be a good measure of the robot's manipulability. Yoshikawa has defined the measure of manipulability as in Equation (5.2) [138], The manipulability calculation has been widely used and proven to be an effective way to keep robots away from singular configurations among several methods [139]. The manipulability measurement is defined as:

$$\omega_m\left(\xi\right) = \sqrt{\det\left(J_\xi\left(\xi\right) J_\xi^\top\left(\xi\right)\right)} \tag{5.2}$$

Null-Space Projections for Multi-Tasks Regulation

Robots with redundancy (particularly high redundancy) can deal very effectively with constrained tasks in the Cartesian space. The redundant self-motion can simultaneously fulfill both, the higher-priority task and the additional low-priority constrained tasks.

The task-space augmentation principle incorporates a constraint task that must be performed simultaneously with the end effector task. In this instance, an incremental Jacobian matrix is constructed, the inverse of which yields the targeted joint velocity solution [140]. The relation between the i-th configuration coordinate vector ξ_i and the i-th Cartesian space task vector x_i is as follow:

$$\dot{x} = J_\xi \dot{\xi} \tag{5.3}$$

The generalized Moore-Penrose pseudo inverse $J^+(q)$ is used since the inverse of the nonsquare (analytical) Jacobian J_ξ does not exist in the redundant case. Null-space projec-

tion, for example, can be used to complement optimization criteria for redundant self-motion, resulting in the relation:

$$\dot{\xi} = J_\xi^+ \dot{x}_\xi + \left(I - J_\xi^+ J_\xi\right) \dot{\xi}_0 \tag{5.4}$$

The expression $\left(I - J_\xi^+ J_\xi\right)$ denotes the orthogonal projection into the null space of J_ξ, and $\dot{\xi}_0$ is an arbitrary joint-space velocity satisfying augmented constraint tasks; hence, the second element of the solution is, therefore, the null-space velocity.

Whole-Body Planning Experiment

The following subsections illustrate the detailed experimental setup and robotic system specification. The overall paradigm of the experiment is shown in Figure 5.7.

- **Grapevine Winter Pruning Points Approaching**

Figure 5.8 gives an intuitive and graphical illustration of the correct pruning points. The whole-body controller generates a trajectory to approach the target pruning point using quintic polynomial interpolation programming [141]. The planned trajectory is defined as follows:

$$
\begin{aligned}
P_x(t) &= c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 + c_5 t^5 \\
P_v(t) &= c_1 t + 2c_2 t + 3c_3 t^2 + 4c_4 t^3 + 5c_5 t^4 \\
P_a(t) &= 2c_2 + 6c_3 t + 12c_4 t^2 + 20c_5 t^3
\end{aligned}
\tag{5.5}
$$

where P_x, P_v, and P_a respectively correspond to the position, velocity, and acceleration in Cartesian space. Therefore, Equation (5.5) can be rewritten as:
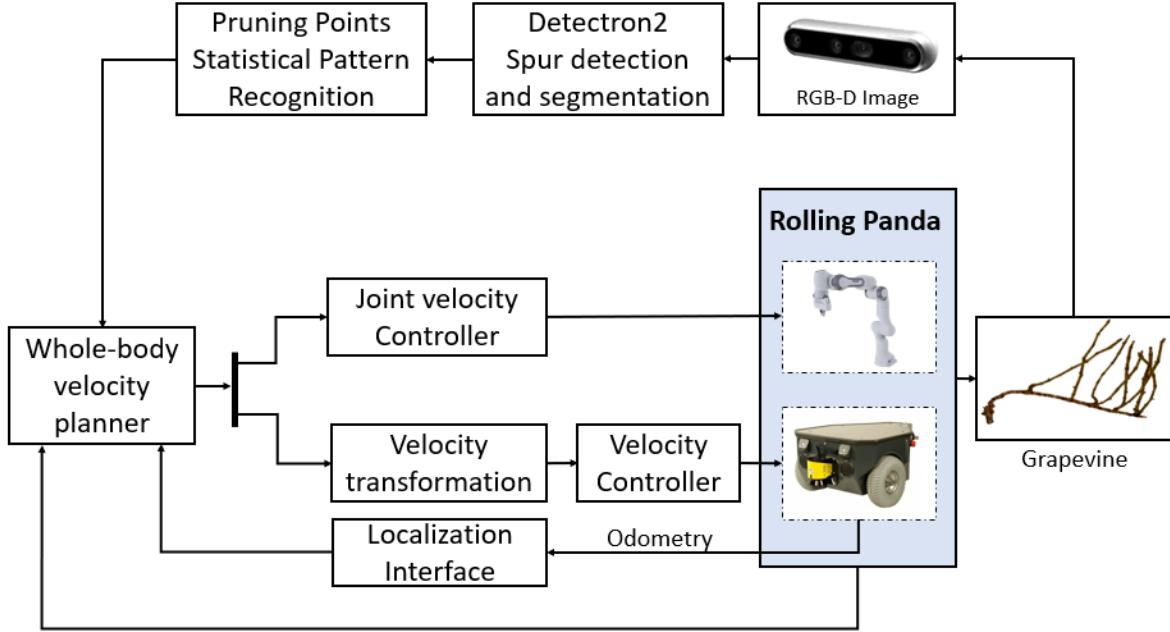
Figure 5.7: Overall paradigm of the whole-body control framework for grapevine winter pruning experiment.

$$
\begin{bmatrix}
1 & t_s & t_s^2 & t_s^3 & t_s^4 & t_s^5 \\
1 & t_e & t_s^2 & t_s^3 & t_s^4 & t_s^5 \\
0 & 1 & 2ts & 3t_s^2 & 4t_s^3 & 5t_s^4 \\
0 & 1 & 2t_e & 3t_e^2 & 4t_e^3 & 5t_e^4 \\
0 & 0 & 2 & 6t_s & 12t_s^2 & 20t_s^3 \\
0 & 0 & 2 & 6t_e & 12t_e^2 & 20t_e^3
\end{bmatrix}
\begin{bmatrix}
c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5
\end{bmatrix}
=
\begin{bmatrix}
x_s \\ x_e \\ v_s \\ v_e \\ a_s \\ a_e
\end{bmatrix}
\tag{5.6}
$$

where x_s, v_s, and a_s correspond to the position, velocity, and acceleration of the interpolation initial point respectively, and x_e, v_e and a_e correspond to the position, velocity, and acceleration of the interpolation end point respectively.

Constraint task 1: Singularities occur when a robot loses control in one or more con-
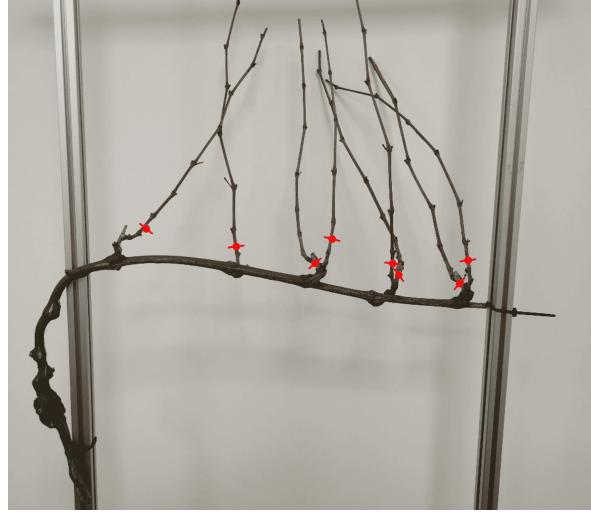
Figure 5.8: Correct grapevine winter pruning points are marked with red dots.

figurations, and they ought to be prevented when planning and controlling robot motion. Since singularities have no effect on the location of the mobile base, only upper manipulator configurations will be considered in this hierarchical control formulation. The term "manipulation" refers to the ability to manipulate something and the measurement is defined as

$$\omega_m\left(\xi_{c1}\right) = \sqrt{\det\left(J_\xi\left(\xi_{c1}\right) J_\xi^\top\left(\xi_{c1}\right)\right)} \tag{5.7}$$

When the manipulability measurement is increased, the manipulator will move away from singularities. The following equation can be used to determine the corresponding joint inputs:

$$\dot{\xi}_{c1} = k_0 \left(\frac{\partial \omega_m\left(\xi_{c1}\right)}{\partial \xi_{c1}}\right)^\top \tag{5.8}$$

where k_0 is a positive gain.

Constraint task 2: Joint limits are physical constraints on robots that must be carefully considered in order to avoid damaging the robotic system. Certainly, only the upper manipulator joints have physical limitations. All joint angles are restricted to a certain range, which

Table 5.1: Joint limits of Franka Emika.

	joint 1	joint 2	joint 3	joint 4	joint 5	joint 6	joint 7
q_{iM}/rad	2.8973	1.7628	2.8973	-0.0698	2.8973	3.7525	2.8973
q_{im}/rad	-2.8973	-1.7628	-2.8973	-3.0718	-2.8973	-0.0175	-2.8973

is listed in Table 5.1. To avoid joint limits, the artificial potential field technique [142] is used to compute the distance between the ith joint and its limits.

We also detect if the configuration approaches a joint limit or falls below a given manipulability threshold, in which case we terminate the approach;

$$d_i = \min\left(\|\gamma_i - \gamma_{li}\|, \|\gamma_i - \gamma_{ui}\|\right), \tag{5.9}$$

where γ_i is joint angle of the i-th joint, γ_{11} and γ_{ul} are the lower and upper joint boundary of the i-th joint, respectively.

The i-th joint's "repulsive velocity" is defined as follows:

$$\dot{\xi}_{i,c2} = \begin{cases} k_i d_i^2, & d_i \leqslant \gamma_{\text{start}} \\ 0, & d_i > \gamma_{\text{start}} \end{cases} \tag{5.10}$$

where k_i denotes a positive gain, and γ_{start} denotes the minimum distance to be free of repulsive force.

So far, we have modeled the main task (*Rolling Panda*'s end effector approach grapevine winter pruning point) and low-level tasks (constraint task 1 and constraint task 2) at the velocity level. The application of mull-space projection technology is able to project constrained tasks to the null space of the main task so that the robot can perform these tasks simultaneously and ensure the priority of the tasks.
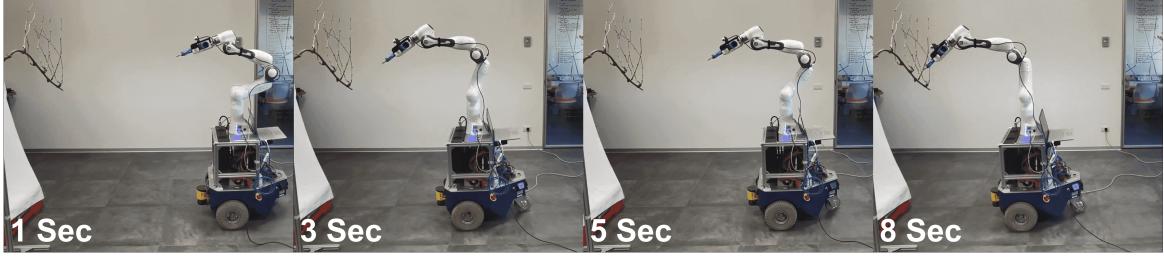
Figure 5.9: Experimental snapshots for grapevine winter pruning point approach using the whole-body controller.

• **Results and Discussion**

The proposed whole-body controller for the *Rolling Panda* can deal with grapevine pruning tasks while satisfying singularity avoidance and joint limitation avoidance, as shown in Figure 5.9. The main task is to move the end-effector along the direction of the z-axis to approach the pruning point and then rotate around the $Roll$-axis to match the orientation of the pruning point. When the target pruning point is given, the whole-body controller can control the overall coordinated movement of the robot to reach the pruning point smoothly.

5.2 Experiment

5.2.1 Experimental Setup

We set up a mock trellis system in a laboratory environment using metal supports to support the vine we got from the simulated vineyard at UCSC, Piacenza, Italy (see Figure 5.10) to compare our planning policy (Three-phase Approach Planning) with a "baseline planner" in a physical environment. For comparison, we used a "baseline planner" that sequences cut points based on Euclidean distance in the workspace and plans trajectories using the Open Motion Planning Library [122], which is a single-phase approach planning. Goal configurations for each approach pose are determined greedily at plan time by choosing the

closest IK solution to the current robot configuration. This setup is based on the structure found in modern vineyards. We then affixed two real unpruned grapevines to the base of the frame. This gives the grapevines compliance properties similar to those found in an actual vineyard. We identified 15 viable pruning target points, including 10 "basal cane cut" and 5 "top cane cut". In total, for each Planner, we ran 120 trials, corresponding to 8 trials each for 15 target points.

The planning algorithm ran on a laptop with a core-i7 processor 1.8 GHz with 32 GB RAM. The Franka robot was limited to 0.1 m/s end-effector velocity for safety. The obstacle model used as input to the planner was given by a 3D model from an RGB-D camera using Octomap. The minimum voxel size for the OctoMap was 0.5 cm.
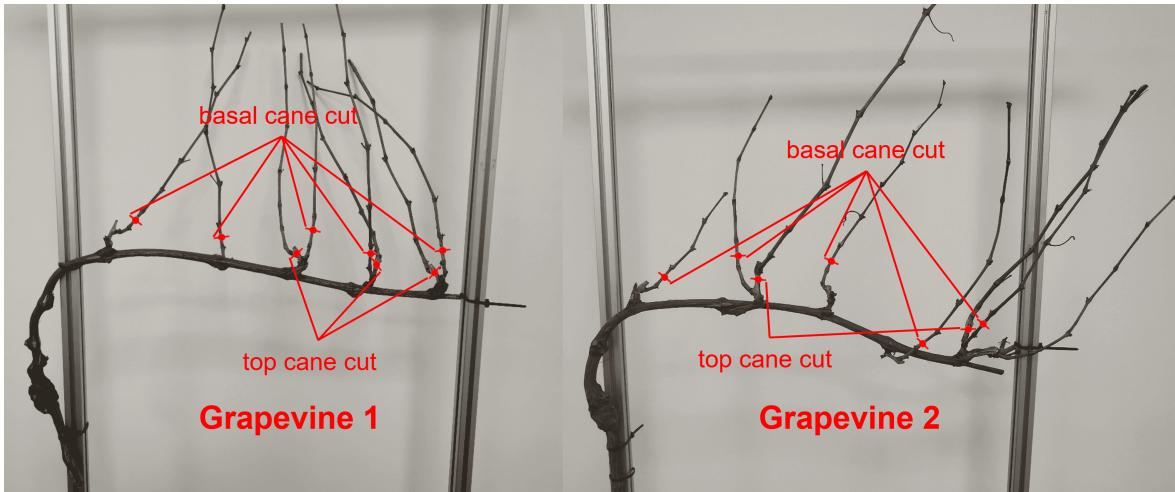
Figure 5.10: Our experimental testbed consists of two grapevines affixed to a trellis system.

5.2.2   Evaluation Criteria

At the end of each approach attempt, the following measures were registered: the result of the attempt (success/failure, by checking if target pruning points locate at the success region of pruning shear) and the reason for any attempt failure (collision, planning failure, or lost pruning point from sight during motion). For each pruning point, additional measures

were registered: the approach cycle time. The approach cycle time is the time it takes from when the robot starts moving to the first waypoint until it has been marked as reachable or unreachable.

- **Cycle time.** Defined as the average pruning points approach cycle time for the multiple approach strategy (detection time is excluded) [47].

- **Ratio of reachable canes** in each approach strategy. Allows a comparison between the success ratio of each approach strategy.

- **Approach attempt failure ratio.** Provides insights into the reasons an approaching attempt fails and the frequency at which failures occur.

5.2.3   Results and Discussion

This section presents the design, evaluation, and practical application of a three-phase planning framework for the task of approaching target pruning points and avoiding colliding with their branches.

We evaluated the performance of an end-to-end pruning system in terms of accuracy (if the potential pruning point falls in the "success" region of shears), and execution time. For comparison, we used a "baseline planner" that sequences cut points based on Euclidean distance in the workspace and plans trajectories as our method does, but with no prior trajectory information. Goal configurations for each approach pose was determined greedily at plan time by choosing the closest IK solution to the current robot configuration. Each experiment involved 15 selected pruning points, including 10 "basal cane cut" and 5 "top cane cut". In total, for each Planner, we ran 120 trials, corresponding to 8 trials each for 15 target points.

Approach Strategies Cycle Time and Success Rate Comparison

Table 5.2 shows results averaged within each of the ten experiments of basal cane cut. Three-phase Approach Planning (TAP) executed higher approach accuracy than the baseline planner (83.75% versus 67.50% success). Although TAP took longer to execute (12.14 s versus 10.06 s), the TAP framework can decompose the planning task which ensures the pruning point is always in the center region of the robot's workspace and improve the success rate of the final pruning point approach by reducing the impact of kinematic miscalibrations. Because keeping the target point in the center of the workspace as much as possible will reduce the error caused by parameter linearization in kinematics calibrations. Besides, the obstacle avoidance task could be completed at the early stage of the whole process, due to the phased planning strategy. After that, whether the shears can enclose the target pruning point mainly depends on whether the pruning point can be successfully landed in the "success" region. Table 5.3 showed results averaged within each of the five experiments of pruning points on spurs. Three-phase Approach Planning (TAP) executed more successful approaches than the baseline planner (55% versus 27.5% success). Similarly, due to phased planning, the Three-phase Approach Planning (TAP) decomposes the planning task and improves the success rate of the final pruning point approach, but at the cost of more time spent executing trajectories (14.75 s versus 24.39 s). Overall, the pruning approach of "top cane cut" is more difficult than the pruning approach of "basal cane cut". Because the shape of the spur is more complex and requires higher manipulability to approach, the target pose generated by the perception system is not always feasible for "top cane cut". In addition, to realize the "top cane cut", it is necessary to engage the spur into the pivot point of the shears in a narrower space. This can easily generate the risk of not reaching the pruning point. Because of the larger diameter of the spur, it is easy to fall into the "failure" region. It is worth noting that when reaching the target pruning pose requires a large-radius rotation of shears, the reachable rate will decrease, which is mainly caused by two reasons: (i) the calibration of the eye-in-hand camera is approximated by least squares estimation of non-linear function,

the rotation will make the error larger; (ii) due to the real-time limitations of the perception pipeline, visual servoing has not been used yet, which affects approaching accuracy.

Table 5.2: Results of 10 basal cane cut experiments.

		1	2	3	4	5	6	7	8	9	10	Avg
Success rate	Ours	75%	100%	100%	75%	62.5%	75%	100%	100%	87.5%	62.5%	83.75%
	Baseline	62.5%	75%	100%	50%	50%	50%	75%	87.5%	62.5%	62.5%	67.50%
Trajectory time (s)	Ours	16.79	9.02	8.56	12.61	9.44	12.14	12.01	12.50	14.08	14.26	12.14
	Baseline	9.68	8.34	6.45	10.12	9.95	10.95	10.74	11.87	12.51	10.02	10.06

Table 5.3: Results of 5 top cane cut experiments.

		1	2	3	4	5	Avg
Success rate	Ours	75%	50%	37.5%	87.5%	25%	55%
	Baseline	50%	25%	0%	50%	12.5%	27.5%
Trajectory time (s)	Ours	19.76	25.33	28.54	18.32	30.01	24.39
	Baseline	11.21	15.26	17.45	10.25	19.58	14.75

Failure Analysis

To be able to investigate the reasons leading to unsuccessful approaches, the following failure analysis methodology is presented. The failure analysis looks into the reasons why an approach strategy fails, resulting in a target pruning point being marked as unreachable, as well as the reasons for the failure of individual approach attempts in the multiple approach strategy. The main failure reason this experiment address is related to unstructured properties of grapevines and occlusion, but one should consider the tight space between canes in which the robot has to operate in a safe manner, minimizing the harm caused to the plants around. Moreover, other failures, caused by the planner, are considered and measured. The failure reasons are then separated into three categories: (i) planning the approach pose (e.g., could

(a)   (b)   (c)
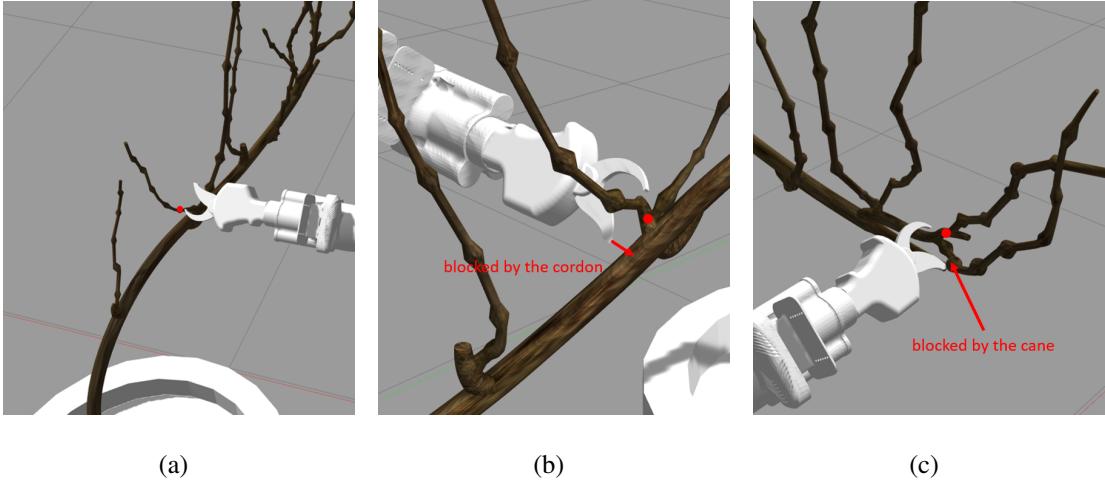
Figure 5.11: Illustration of failure cases: (a) the target pruning point falls into the "failure" region; (b) the target pruning point is blocked by the cordon; (c) the target pruning point is blocked by the cane.

not find a valid path due to unreachable target pose, which means pruning pose generated is not feasible). This kind of situation happens 26 times in the top cane cut cases, pruning the previous year's two-bud spurs in half. It is required to approach pruning points from a feasible orientation. Generating pruning orientation is based on the geometry method, which is sensitive to model parameters and leads to insufficient robustness in practical applications. (ii) the cane is not aligned (e.g., the jaws of the shears do not align properly with the cane, especially top cane cut cases, see Figure 5.11 (a)). This is easy to happen for canes with thicker diameters and when the shears need to be rotated at a large angle, it happened 40 times in total. (iii) occlusion-related failure (e.g., target pruning points were lost from sight during the motion, due to the occlusion of other canes, see Figure 5.11 (b) and (c)). When the distance between the canes is too close, this failure occurred 19 times in total.

The results show that the planning framework is computationally efficient enough to do real-time pruning. They also show that the system can prune with acceptable precision while avoiding unsafe contact with the environment. Our experiments provide useful information

for future system development. The success rate is still the first priority of our research. On the one hand, we can optimize the hardware design, so that the end-effector is able to complete pruning easily, such as designing jaws that are easier to engage. On the other hand, optimizing our planning algorithm improves accuracy by incorporating closed-loop feedback methods such as visual servoing.

5.3   Conclusion

In this chapter, we proposed a grapevine winter pruning planning framework "Three-phase Approach Planning", in which the path of the end-effector to the end target pose is divided into three discrete sets of trajectories. In the first phase, the robot moves from the home pose to the detection pose. During this stage, the main consideration is obstacle avoidance, so we adopt the STOMP planning algorithm. The second phase navigates the robot shears to the pre-pruning pose, in which the shears are about 1.5 cm ahead of the pruning point, by calling our whole body controller, where we make sure the target pruning point is positioned in the "success" region. In the third stage, the robot shears move forward to enclose the pruning location to achieve the desired cut.

Inside this planning framework, we also present a task-priority coordinated whole-body motion controller for a non-holonomic mobile manipulator. The controller can plan and schedule whole-body coordinated motion to complete the grapevine winter pruning task. The top priority task can be executed by employing all capabilities (manipulation and locomotion) of the robotic system. The second (lower) priority task is then projected into the null space of the top priority task, hence they have no impact on its execution. We conducted a grapevine winter pruning experiment to demonstrate the performance of the proposed framework using a two-wheeled mobile base with a 7-DoF robot manipulator.

Our experiments provide useful insights into the further development of the system. We evaluate the performance of an end-to-end pruning system in terms of accuracy reliability and cycle time. Although the three-phase planning framework takes more execution time to complete the task, the planning framework can improve the overall success rate and reliability. Looking into the different reasons causing an approach to fail is important from two aspects. First, it provides insights into the reasons. Second, it gives us a direction on how to improve planning methods.

6 Compliant Control for Grapevine Pruning

In agricultural robotics, harvesting, pruning, and grasping are common applications. Until recently, the majority of agriculture robots employ high-gain (stiff) position feedback control for the above-mentioned tasks. However, there is not much work that incorporates the idea of compliant control into the above application scenarios to improve the compliance and robustness of operations in agriculture applications[143, 144, 145]. Owing to the different biological characteristics of the target plant, there are often situations in which grapevine fails to be pruned resulting in the inability to precisely control the cut during the pruning process. If the pushing force during pruning is too large, it will cause irreversible mechanical damage to the target spur and reduce plant/cordon vitality and fruit quality. While the pushing force is too small, it will cause the spur to slide or even fall off, resulting in failure to enclose the target branch [47, 145]. Therefore, studying the non-destructive pruning technology of the grapevine pruning robot is of great significance for reducing the mechanical damage to the grapevine and improving the success rate of spur pruning. Impedance control is a useful framework to allow the robot to follow reference trajectories and, simultaneously, handle external disturbances which are always present in dynamic manipulation.

Once the robot moves to the approach position pre-pruning pose, the final step in the pruning process is to move the shears toward the target pruning point and execute the cut.

At this point, we make an assumption that the initial estimation of the target pruning point is good enough so that, when the arm is moved to approach the pre-pruning pose, the target pruning point is inside the "success" region. We then execute a compliant controller that uses a human-like manner to guide the shears to enclose the target pruning point, at which point we execute the cut. The use of a Cartesian impedance controller, which will be introduced in the next subsection is more effective than a pure position controller [146] since the impedance parameters, especially the stiffness parameter, which regulates the relationship between the interaction force and the tracking error, can be tuned in arbitrary directions to implement a trade-off between position accuracy and the "softness" of the response.
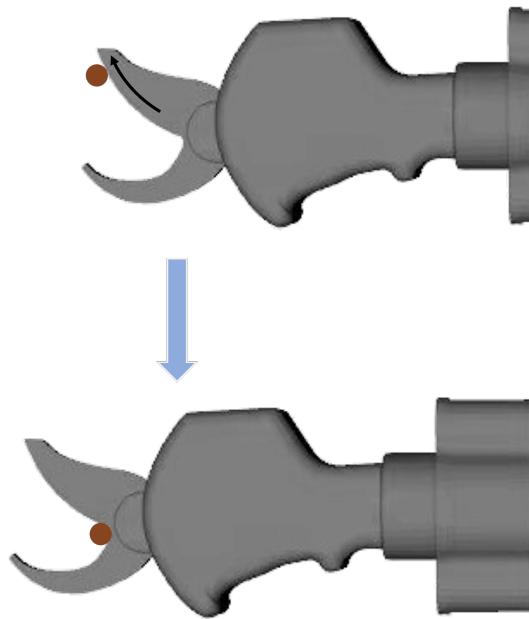
Figure 6.1: Schematic diagram of shears enclosing grapevine target pruning point. Up: a possible contact point; Down: the cane at the pivot point, ideal for executing a cut.

6.1 Compliant Control Strategy for Grapevine Pruning

6.1.1 Impedance Control for Grapevine Pruning

The Cartesian impedance control depends on the combined model of the damping-spring-mass system and first-order spring system. The first-order spring model is used to simulate the grapevine with complex biological characteristics. Figure 6.2 describes the three stages of the shears and the grapevine from never contacting $x < x_{envi}$ to stable engagement $x > x_{envi}$. When the pruning process is in an ideal state (the environmental stiffness and the environmental position remain unchanged), it means that the biological characteristics of the vine are not changed or the shears do not slide. Otherwise, it means that the control precision of pruning force is not enough due to some other reasons, for example, inappropriate pruning point.
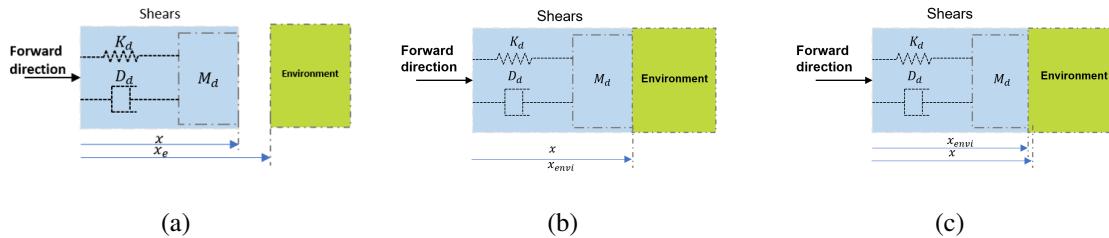
Figure 6.2: Pruning model between shears and grapevine for impedance control: (a) before contacting $(x < x_{envi})$; (b) during contacting $(x = x_{envi})$; (c) stable pruning $(x > x_{envi})$.

As it is well known, a robotic manipulator is a coupled, time-varying and nonlinear system. The general dynamic model of a manipulator with joint coordinates $q_n \in \mathbb{R}^n$ can be written as

$$M\left(q_n\right)\ddot{q}_n + C\left(q_n, \dot{q}_n\right) + G\left(q_n\right) = \tau_n + \tau_n^{ext}, \tag{6.1}$$

where $M \in \mathbb{R}^{n \times n}$ is the symmetric and positive definite inertial matrix of the manipulator. $C \in \mathbb{R}^n$ is the Coriolis and centrifugal force, $G \in \mathbb{R}^n$ is the gravity vector, $\tau_n \in \mathbb{R}^n$ and

$\tau_n^{ext} \in \mathbb{R}^n$ are the commanded torque vector and external torque vector, respectively. In Equation (6.1), it is maintained the property that $\dot{M} - 2C \in \mathbb{R}^{n \times n}$ is skew symmetric.

For a manipulator with joint coordinates $q \in \mathbb{R}^n$, the desired dynamic relationship between Cartesian error $\tilde{x} \in \mathbb{R}^6$ and external force $F_{ext} \in \mathbb{R}^6$ in the Cartesian impedance controller is given by [146]:

$$\Lambda(x)\ddot{\tilde{x}} + (\psi(x, \dot{x}) + D_d)\dot{\tilde{x}} + K_d\tilde{x} = F_{ext}, \tag{6.2}$$

where $K_d \in \mathbb{R}^{6 \times 6}$ and $D_d \in \mathbb{R}^{6 \times 6}$ are the desired Cartesian stiffness and damping respectively. $\Lambda(x) \in \mathbb{R}^{6 \times 6}$ represents Cartesian inertial and $\psi(x, \dot{x}) \in \mathbb{R}^{6 \times 6}$ represents Cartesian Coriolis and centrifugal matrix. They can be computed respectively by Equation (6.3):

$$\begin{aligned} \Lambda(x) &= J(q)^{-T} M(q) J(q)^{-1} \\ \psi(x, \dot{x}) &= J(q)^{-T} \left(C(q, \dot{q}) - M(q) J(q)^{-1} \dot{J}(q) \right) J(q)^{-1} \end{aligned} \tag{6.3}$$

where $M(q) \in \mathbb{R}^{n \times n}$ represents the inertial matrix, $C(q, \dot{q}) \in \mathbb{R}^{n \times n}$ is the Coriolis and centrifugal matrix and $J(q) \in \mathbb{R}^{6 \times n}$ represents the Jacobian matrix.

The Cartesian impedance controller input for the final push stage of the pruning task is as follows:

$$\tau_{imp} = G(q) + J(q)^\top \left(\Lambda(x)\ddot{x}_d + \mu(x, \dot{x})\dot{x}_d - K_d\tilde{x} - D_d\dot{\tilde{x}} \right) \tag{6.4}$$

where $G(q) \in \mathbb{R}^n$ denotes the gravity force and $x_d \in \mathbb{R}^6$ is the desired Cartesian position and orientation.

6.1.2 Impedance Regulation for Human-Like Pruning

Imitation learning has been extensively applied in the kinematics-level modeling of task movement, while there are very few works about learning dynamics [147, 148]. The barrier is the difficulty in extracting physically meaningful dynamic parameters from human demonstrations. Aiming at demonstrating desired end-effector trajectories and interaction wrenches

Figure 6.3: Demonstration of pruning robot, in which Cartesian position trajectories of end-effector and force profile are recorded. The red arrow indicates the movement process of the robot.

to shears, we integrated kinesthetic teaching with the FCI[1] interface of Franka Emika Panda. The robotic manipulator used is torque controlled, and the interaction wrenches of the robot end-effector with the environment $\hat{F}_{ext}$ can be estimated through the joint torque measurements of the manipulator. Hence, during the demonstration, the human instructor can guide the robot through the Franka FCI interface while the robot manipulator estimates its interaction with the environment.

The desired force profiles learned from demonstrations are generated using Gaussian mixture modeling (GMM) and Gaussian mixture regression (GMR) [149]. The parameters of GMM can be trained by the Expectation-Maximization (EM) algorithm [150] with an offline training process. The demonstration data η is collected, $\eta^{\mathcal{I}}$ and $\eta^{\mathcal{O}}$ respectively denote the input and output variables in demonstration data, where the subscripts $\mathcal{I}$ and $\mathcal{O}$ stand

---

[1]https://frankaemika.github.io/docs/

for their dimensions, respectively. The GMM is trained from the data obtained from the demonstrations, in which the input variable is time $\eta^{\mathcal{I}} = t$ and output variables are position and force $\eta^{\mathcal{O}} = [\hat{x}^{\top}, \hat{F}^{\top}]^{\top}$. Given an input variable $\eta^{\mathcal{I}}$, the best estimation of output $\hat{\eta}^{\mathcal{O}}$ is computed by GMR, the mean $\hat{\mu}$ of the conditional probability distribution $\hat{\eta}^{\mathcal{O}} \mid \eta^{\mathcal{I}} \sim \mathcal{N}(\hat{\mu}, \hat{\Sigma})$, and $\hat{\Sigma}$ is covariance matrix.

Impedance-based skill refers to different stiffness levels that a robot needs to accomplish a given task. Then we need to estimate from the collected data the stiffness and damping parameters K_d, D_d to formulate control input Equation (6.4) of robots. Based on the demonstrated force, the desired impedance parameters are computed using a quadratic program (QP) from the point of view of optimal control. A stiffness matrix $K_{d,i}$ is estimated for each time step i, by assuming that the robot behavior is driven by a set of virtual springs:

$$
\min_{K_{d,i}} \frac{1}{2} \sum_{i=1}^{N} \left\| K_{d,i} \left(x_{d,i} - \hat{x}_i \right) - \hat{F}_i \right\|_W^2
\tag{6.5}
$$
$$
\text{s.t.} \quad K_{\min} \leq K_{d,i} \leq K_{\max} \quad i = 1, \ldots, N
$$

where $x_{d,i}$ is the predefined reference trajectory to push shears forward, $\hat{x}_i$ and $\hat{F}_i$ are the output of the GMR at time step i, N is the length of the time window, and W is weighting matrix $W = \mathrm{diag}([\omega_1, \omega_2, \ldots, \omega_N])$. $K_{\min}$ and $K_{\max} \in \mathbb{R}^{6 \times 6}$ are respectively minimum and maximum allowed stiffness. The constraint inequality between vectors is element-wise.

Some simplifying hypotheses have been assumed in the implemented method. First of all, $K_{d,i}$ and $D_{d,i}$ are assumed diagonal. Moreover, only the translational part of the interaction model is considered while the rotational part is kept constant. The value of $D_{d,i}$ at the time step i control loop is computed at the previous optimization step $K_{d,i-1}$, with critical damping factor [146]:

$$
D_{d,i} = 2 \times 0.707 \times \sqrt{K_{d,i-1}}
\tag{6.6}
$$

6.2   Experiment
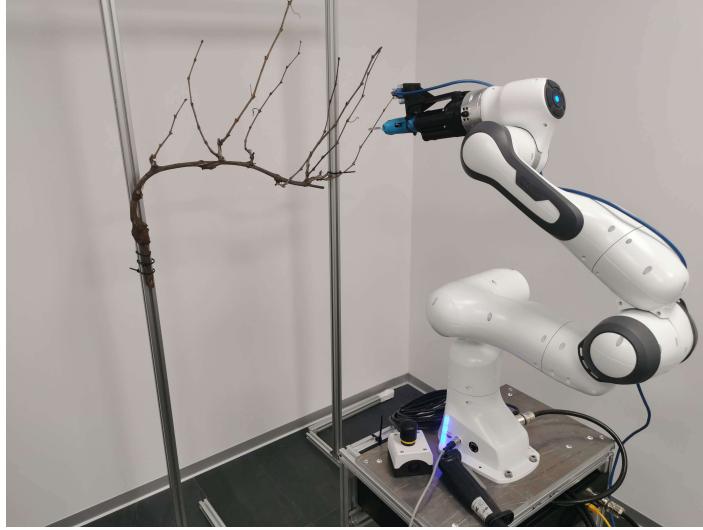
Figure 6.4: Experimental setup: mock trellis system.

We set up a mock trellis system (as shown in Figure 6.4) to compare our controller with a baseline controller [47] in a physical environment. The experimental setup for the human demonstrations is shown in Figure 6.5 along with the path followed by the human. The human instructor grasps Franka demostrator to teach the grapevine winter pruning task. The demonstrated trajectory starts from the pre-pruning pose. Then, the human guides the robot end-effector to enclose the grapevine, moving blades so that canes engage into the pivot points of shears. A shears tool is attached to the end-effector of the robot. The end-effector forces $\hat{F}_{ext}$ are estimated thanks to the torque sensors integrated into the robot joints, which will be used for training desired impedance parameters.

Three demonstrations are performed, all from the same human instructor, where desired position and interaction force of the end-effector with the environment are recorded and are later used to train the impedance regulator method. The control algorithm ran on a laptop with a core-i7 processor $1.8\ GHz$ with $32\ GB$ RAM.

<div align="center">(a)   (b)</div>
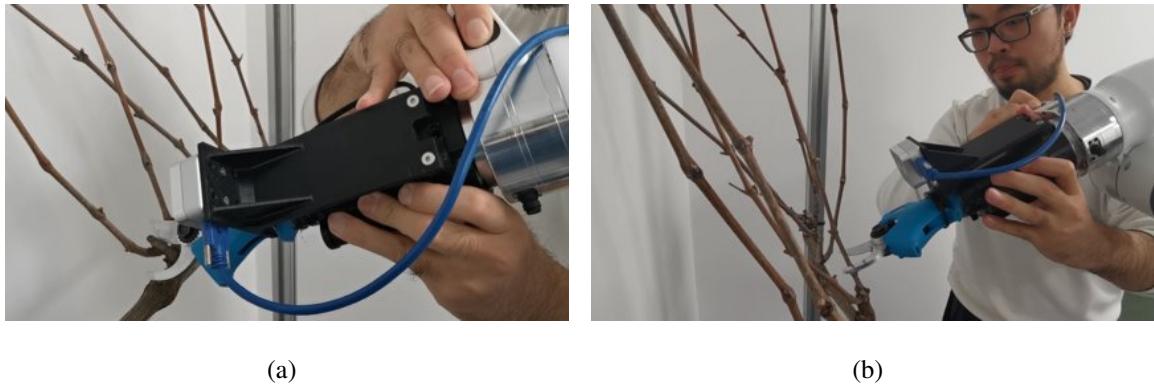
Figure 6.5: Experimental setup. The instructor demonstrates the grapevine pruning task: (a) top cane pruning; (b) basal cane pruning.

In order to assess the contribution of each component of our impedance controller, we define three controllers that remove or alter certain aspects of the hybrid controller:

- **learned-based impedance control (LIC):** Given a target, the LIC uses the impedance profile learned from demonstrations. The controller stops when the shears push forward 8 cm.

- **pre-difined impedance control (PIC):** Given a target, the PIC uses a predefined impedance parameter, using the average of the learned impedance files. The controller stops when the shears push forward 8 cm.

- **position-based closed-loop controller (PC):** A closed-loop position controller sets a virtual target 8 cm past the input target and moves towards it.

6.2.1 Experimental Trials and Evaluation Criteria

To verify our controller, we selected 5 pruning regions on a grapevine, including 7 pruning points, as shown in Figure 6.6. For each trial, we first selected one of the seven target pruning points. After driving the robot back to the home pose, which provided a distinct view of the
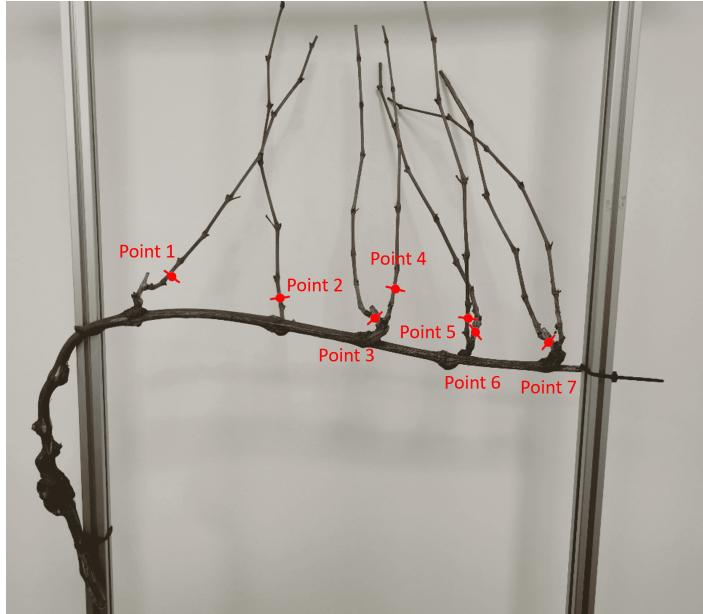
Figure 6.6: Seven target pruning points are marked with red dots.

grapevine, we selected the one of points from the point cloud corresponding to the target based on grapevine pruning rules, giving us a 3D estimate of the desired pruning point. We then planned a path so that the shears would reach the pre-pruning pose, where the shears were 1.5 cm away from the estimated target. In total, for each of the three controllers, we ran 63 trials, corresponding to 3 trials each for 7 target points for each controller. At this point, we ran each of the three controllers in succession and recorded the following metrics to evaluate the performance of each cut:

- **Accuracy:** Is the targeted cane inside the shears jaw.

- **Pivot offset length:** Distance of the cane to the shears pivot, measured by hand.

- **Max force:** The maximum force magnitude from Franka's force-torque sensor during the execution.

Table 6.1: Averaged summary stats for each controller.

	Accuracy	Pivot Offset (cm)	Max Force (N)
learning-based impedance controller	76.19%	1.8 ± 1.2	4.6 ± 1.7
predefined impedance controller	66.67%	2.1 ± 1.5	4.5 ± 2.6
position based controller	57.14%	2.7 ± 1.7	4.9 ± 3.1

(a)   (b)
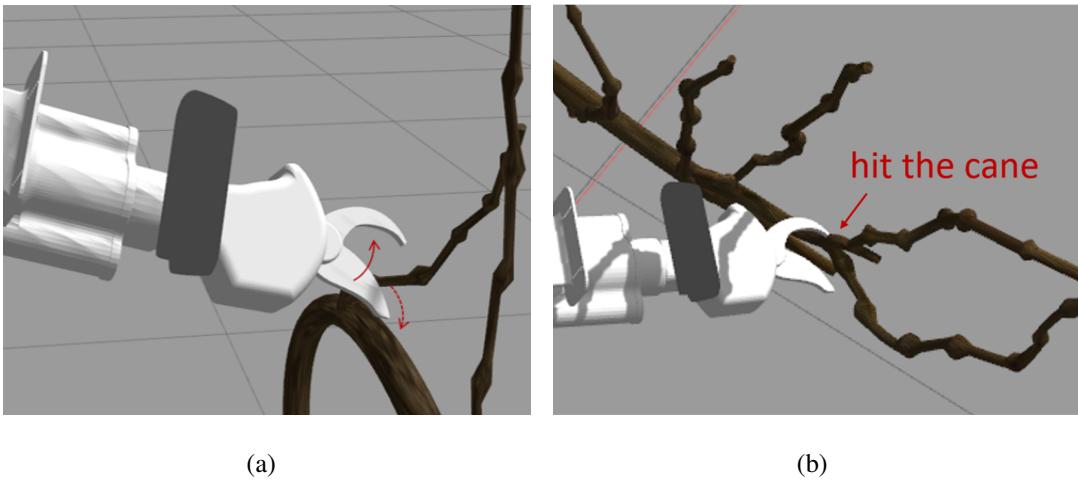
Figure 6.7: Illustration of failure cases: (a) the target cane is not aligned to the pivot point; (b) the fixed blade hit the thick cane.

6.2.2   Results and Discussion

During recent trials, the robot successfully encloses 43 of 63 target canes. Of the 20 failures, eighteen were due to a failure to engage canes into the blade pivot point, and the remaining two were due to the robot missing the canes as it attempted to cut. For the cases of "basal cane cut" (pints 1,2,4,5) only position-based control failed once because the approach angle is relatively large and the targeted cane is lost during the approach. The "top cane cut" condition, on the other hand, is more challenging. This is because the correct pruning orientation is required to enclose the pruning point on a spur. In addition, the spur has a relatively large diameter and is easy to get stuck on the blade of the shears (as shown in Figure 6.7 (b)). A summary of the average results is shown in Table 6.1. Our learning-based

impedance controller (LIC) performed better than the predefined impedance controller (PIC) and closed-loop position controller (PC) in terms of cutting accuracy and precision, where the LIC has a 9.52 percentage point increase in accuracy over the predefined impedance controller and a 19.05 percentage point increase in accuracy over the position based controller. Both impedance-based controllers performed better than position-based controllers. That is because the compliant controller reduces the pressure between the blade and the cane while touching the cane and then reduces the friction so that the cane can be engaged into the pivot point along the arc of the blade. Finally, it can guide the shears as close to the blade pivot point as possible to enclose the cane completely, where the cutting blade is able to deliver maximum cutting force. For high-stiffness position controllers, canes are easier to get stuck by the blade or draw the edge of the shears, leading to a failed cut. Compared with LIC, although the PIC shows compliance, it needs adaptive stiffness performance during the whole pruning point engagement process to avoid the blade being stuck by the cane when lower stiffness is required and enforce engagement to deliver maximum cutting force when higher stiffness is required to stabilize cane at the pivot point. Furthermore, LIC managed to avoid bumping into the canes on every trial and achieved relatively consistent branch remnant lengths averaging 1.8 cm with a standard deviation of just 1.2 cm. This is in comparison to the position-based closed-loop controller which had nearly 1.5 times the standard deviation at 1.7 cm and sometimes drove straight into cane or spur. LIC also more consistently ended up with low pivot offset lengths. This was largely a result of situations in which the shears were initially placed in front of a cane: our controller would move the shears away from the cane along the curve of the blade, while the closed-loop position controller would move straight into them. Moreover, LIC also more consistently ended up with a smaller force, casing by impedance profile learned from the demonstration.

6.3   Conclusion

The grapevine winter pruning task is a typical representation, which shares some crucial requirements with other complex interactive tasks, such as the need for continuous regulation of the task stiffness geometry to comply with environmental constraints. In this chapter, we proposed a learning framework to address the autonomous impedance regulation problem of robots for grapevine pruning tasks, where we encode and reproduce impedance behaviors using learning from demonstration. Our method allows encoding behaviors that rely on task variables, yielding only one model to encode the whole task. In contrast to previous approaches where robot impedance is learned from position variability, our QP-based framework extracts the impedance behavior from the manner in which the instructor behaves during the task, relying on the recorded force profile. We use demonstrated data not only to encode the skill but also to estimate the stiffness of virtual springs. Indeed, the strength of the proposed approach is the ability to learn, replicate, and optimize the interaction behavior of the human arm that can deal with different manipulation tasks. Through our experiments, we demonstrated that our system is robust to a number of issues that have traditionally been associated with the geometry of shears and the diameter of the canes. While we conduct our experiments in an indoor laboratory setting, our framework represents a step towards moving robotic precision pruning to the outdoor vineyard environment. This set of experiments opens the road for various further investigations, the human-like skill transfer controller design can be extended to similar plant pruning or harvesting applications.

7   General Conclusion and Future Work

7.1   Conclusion

The main contribution of this work is to develop an autonomous robot system to realize the perceptual identification and inference of grapevines and then perform the pruning manipulation with the aim to automate grapevine winter pruning. The following goals have been reached to meet the final objective:
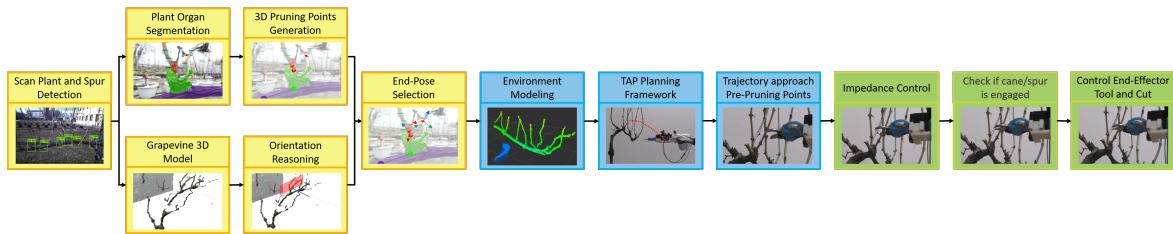

Figure 7.1: Grapevine winter pruning pipeline.

- The detailed models of the robotic pruning platform *Rolling Panda* and *HyQReal-Kinova* have been designed and analyzed. Both prototypes are composed of a mobile platform (wheeled and legged base), manipulator, camera, and shears, which enable loco-manipulation ability.

- Large datasets to train machine learning models for deep learning-based computer vision are a bottleneck in the specialty crop industry and agriculture, in general. To develop a visual perception system, we present hierarchical guidelines for a translation of the human cognition process into algorithms for automated grapevine recognition, and robotized pruning. Multiple combinations of integrated general rules driving the winter pruning process with specifically conceived case studies will compose a dedicated database for developing a perception system for the recognition of grapevine architecture and identification of target regions and pruning points.

- To autonomously perform the pruning point approach from the output of the perception system. The proposed "Three-phase Approach Planning" motion planner is highly manipulative to avoid collisions with the grapevine cordon and canes, while a high pruning success rate is achieved.

- To finish the final step in the pruning process moves the shears toward the target pruning point and execute the cut. We then design a compliant controller that uses a human-like manner to guide the shears to enclose the target pruning point. The designed learning-based Cartesian impedance controller, which learns stiffness parameters from human demonstration, improves the robustness of the pruning point.

7.2 Future Work

For future directions, most of the problems presented in this thesis are to be resolved with the goal of creating an effective pruning robot system with performance similar to a human. These consist of improving the precision of perception, finding the best loco-manipulation strategy, in order to have more human-like behavior, and controlling algorithms/architectures for different kinds of grapevine pruning variation. Future work will focus on improving

robustness and throughput as well as generalizing our approach to other field systems with more complicated pruning rules.

The most researched perception solutions are those based on vision and cloud point sensors, often combined with machine learning approaches to interpreting the collected data. To improve and make our results more consistent, in future work, we plan to increase the dataset with more frames and use patches instead of the entire image. Our hypothesis is that with this approach, the model will learn how to segment grapevine at different crop scales. In addition, a more complete expert knowledge base needs to be established to guide how to convert agronomic workflows into algorithms

We plan in future work to provide the robot with a more active role. We devise to exploit the visual information more efficiently by using active perception, which is a motion perception coupling method, to improve the accuracy of pruning points generation. Moreover, in order to achieve an agricultural manipulator capable of operating in more broad scenarios and applications, as opposed to their current applications of only one type of product to harvest or plant to prune, we plan to integrate learning-based control algorithms into our proposed grapevine pruning pipeline to improve the generalization of manipulation.

Bibliography

[1] J. M. Alston and O. Sambucci, "Grapes in the world economy," in *The grape genome*, pp. 1–24, Springer, 2019.

[2] D. Romano, B. Rocchi, A. Sadiddin, G. Stefani, R. Zucaro, and V. Manganiello, "A sam-based analysis of the economic impact of frauds in the italian wine value chain," *Italian Economic Journal*, vol. 7, no. 2, pp. 297–321, 2021.

[3] T. Galpin, "Pruning the grapevine," *Training & Development*, vol. 49, no. 4, pp. 28–34, 1995.

[4] S. Poni, S. Tombesi, A. Palliotti, V. Ughini, and M. Gatti, "Mechanical winter pruning of grapevine: Physiological bases and applications," *Scientia Horticulturae*, vol. 204, pp. 88–98, 2016.

[5] M. Brown and G. Gao, "Basic principles of pruning backyard grapevines," *Ohio State University Extension Fact Sheet. HYG-1428*, 2004.

[6] S. Poni, M. Gatti, A. Palliotti, Z. Dai, E. Duchêne, T.-T. Truong, G. Ferrara, A. M. S. Matarrese, A. Gallotta, A. Bellincontro, *et al.*, "Grapevine quality: A multiple choice issue," *Scientia horticulturae*, vol. 234, pp. 445–462, 2018.

[7] A. Palliotti, S. Tombesi, O. Silvestroni, V. Lanari, M. Gatti, and S. Poni, "Changes in vineyard establishment and canopy management urged by earlier climate-related grape ripening: A review," *Scientia Horticulturae*, vol. 178, pp. 43–54, 2014.

[8] M. Gatti, S. Civardi, F. Bernizzoni, and S. Poni, "Long-term effects of mechanical winter pruning on growth, yield, and grape composition of barbera grapevines," *American journal of enology and viticulture*, vol. 62, no. 2, pp. 199–206, 2011.

[9] G. S. Howell, "Sustainable grape productivity and the growth-yield relationship: A review," *American Journal of Enology and Viticulture*, vol. 52, no. 3, pp. 165–174, 2001.

[10] W. M. Kliewer and N. Dokoozlian, "Leaf area/crop weight ratios of grapevines: Influence on fruit composition and wine quality," *American Journal of Enology and Viticulture*, vol. 56, pp. 170–181, 2001.

[11] M. C. Trought, J. Bennett, and H. Boldingh, "Influence of retained cane number and pruning time on grapevine yield components, fruit composition and vine phenology of sauvignon blanc vines," *Australian Journal of Grape and Wine Research*, vol. 17, no. 2, pp. 258–262, 2011.

[12] T. Bates, "Mechanical crop control in new york'concord'vineyards target desirable crop load levels," in *International Symposium on Physiological Principles and Their Application to Fruit Production 1177*, pp. 259–264, 2014.

[13] J. R. Morris, "Development and commercialization of a complete vineyard mechanization system," *HortTechnology*, vol. 17, no. 4, pp. 411–420, 2007.

[14] T. Bates, "Pruning level affects growth and yield of new york concord on two training systems," *American Journal of Enology and Viticulture*, vol. 59, no. 3, pp. 276–286, 2008.

[15] T. Bates and J. Morris, "Mechanical cane pruning and crop adjustment decreases labor costs and maintains fruit quality in new york 'concord'grape production," *HortTechnology*, vol. 19, no. 2, pp. 247–253, 2009.

[16] A. Zahid, M. S. Mahmud, L. He, P. Heinemann, D. Choi, and J. Schupp, "Technological advancements towards developing a robotic pruner for apple trees: A review," *Computers and Electronics in Agriculture*, vol. 189, p. 106383, 2021.

[17] S. Mehta and T. Burks, "Vision-based control of robotic manipulator for citrus harvesting," *Computers and Electronics in Agriculture*, vol. 102, pp. 146–158, 2014.

[18] F. A. A. Cheein and R. Carelli, "Agricultural robotics: Unmanned robotic service units in agricultural tasks," *IEEE industrial electronics magazine*, vol. 7, no. 3, pp. 48–58, 2013.

[19] T. Teng, M. Fernandes, M. Gatti, S. Poni, C. Semini, D. Caldwell, and F. Chen, "Whole-body control on non-holonomic mobile manipulation for grapevine winter pruning automation," in *2021 6th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM)*, pp. 37–42, IEEE, 2021.

[20] M. Gao and T.-F. Lu, "Image processing and analysis for autonomous grapevine pruning," in *2006 International Conference on Mechatronics and Automation*, pp. 922–927, IEEE, 2006.

[21] N. S. Naik, V. V. Shete, and S. R. Danve, "Precision agriculture robot for seeding function," in *2016 international conference on inventive computation technologies (ICICT)*, vol. 2, pp. 1–3, IEEE, 2016.

[22] G. Kootstra, X. Wang, P. M. Blok, J. Hemming, and E. Van Henten, "Selective harvesting robotics: current research, trends, and future directions," *Current Robotics Reports*, vol. 2, pp. 95–104, 2021.

[23] A. T. Meshram, A. V. Vanalkar, K. B. Kalambe, and A. M. Badar, "Pesticide spraying robot for precision agriculture: A categorical literature review and future trends," *Journal of Field Robotics*, vol. 39, no. 2, pp. 153–171, 2022.

[24] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, 2019.

[25] S. Paulin, T. Botterill, J. Lin, X. Chen, and R. Green, "A comparison of sampling-based path planners for a grape vine pruning robot arm," in *2015 6th International Conference on Automation, Robotics and Applications (ICARA)*, pp. 98–103, IEEE, 2015.

[26] V. Tinoco, M. F. Silva, F. N. Santos, L. F. Rocha, S. Magalhães, and L. C. Santos, "A review of pruning and harvesting manipulators," in *2021 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pp. 155–160, IEEE, 2021.

[27] S. A. Magalhães, F. N. d. Santos, R. C. Martins, L. F. Rocha, and J. Brito, "Path planning algorithms benchmarking for grapevines pruning and monitoring," in *EPIA Conference on Artificial Intelligence*, pp. 295–306, Springer, 2019.

[28] R. Barth, J. Hemming, and E. J. van Henten, "Design of an eye-in-hand sensing and servo control framework for harvesting robotics in dense vegetation," *Biosystems Engineering*, vol. 146, pp. 71–84, 2016.

[29] K.-S. Han, S.-C. Kim, Y.-B. Lee, S.-C. Kim, D.-H. Im, H.-K. Choi, and H. Hwang, "Strawberry harvesting robot for bench-type cultivation," *Journal of Biosystems Engineering*, vol. 37, no. 1, pp. 65–74, 2012.

[30] S. Paulin, T. Botterill, X. Chen, and R. Green, "A specialised collision detector for grape vines," in *Proceedings of the Australasian Conference on Robotics and Automation*, pp. 1–5, 2015.

[31] M. Fernandes, A. Scaldaferri, G. Fiameni, T. Teng, M. Gatti, S. Poni, C. Semini, D. Caldwell, and F. Chen, "Grapevine winter pruning automation: On potential pruning points detection through 2d plant modeling using grapevine segmentation," in

2021 IEEE 11th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pp. 13–18, IEEE, 2021.

[32] Y.-R. Chen, K. Chao, and M. S. Kim, "Machine vision technology for agricultural applications," *Computers and electronics in Agriculture*, vol. 36, no. 2-3, pp. 173–191, 2002.

[33] E. Davies, "The application of machine vision to food and agriculture: a review," *The Imaging Science Journal*, vol. 57, no. 4, pp. 197–217, 2009.

[34] J. Xue, L. Zhang, and T. E. Grift, "Variable field-of-view machine vision based row guidance of an agricultural robot," *Computers and Electronics in Agriculture*, vol. 84, pp. 85–91, 2012.

[35] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, "Machine learning applications for precision agriculture: A comprehensive review," *IEEE Access*, vol. 9, pp. 4843–4873, 2020.

[36] H. Tian, T. Wang, Y. Liu, X. Qiao, and Y. Li, "Computer vision technology in agricultural automation—a review," *Information Processing in Agriculture*, vol. 7, no. 1, pp. 1–19, 2020.

[37] C. Zheng, A. Abd-Elrahman, and V. Whitaker, "Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming," *Remote Sensing*, vol. 13, no. 3, p. 531, 2021.

[38] W. Kazmi, S. Foix, G. Alenyà, and H. J. Andersen, "Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison," *ISPRS journal of photogrammetry and remote sensing*, vol. 88, pp. 128–146, 2014.

[39] D. Houle, D. R. Govindaraju, and S. Omholt, "Phenomics: the next challenge," *Nature reviews genetics*, vol. 11, no. 12, pp. 855–866, 2010.

[40] S. Tu, J. Pang, H. Liu, N. Zhuang, Y. Chen, C. Zheng, H. Wan, and Y. Xue, "Passion fruit detection and counting based on multiple scale faster r-cnn using rgb-d images," *Precision Agriculture*, vol. 21, pp. 1072–1091, 2020.

[41] R. Goel and P. Gupta, "Robotics and industry 4.0," *A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable Development*, pp. 157–169, 2020.

[42] C. W. Bac, E. J. Van Henten, J. Hemming, and Y. Edan, "Harvesting robots for high-value crops: State-of-the-art review and challenges ahead," *Journal of Field Robotics*, vol. 31, no. 6, pp. 888–911, 2014.

[43] A. Botta, P. Cavallone, L. Baglieri, G. Colucci, L. Tagliavini, and G. Quaglia, "A review of robots, perception, and tasks in precision agriculture," *Applied Mechanics*, vol. 3, no. 3, pp. 830–854, 2022.

[44] E. Vrochidou, K. Tziridis, A. Nikolaou, T. Kalampokas, G. A. Papakostas, T. P. Pachidis, S. Mamalis, S. Koundouras, and V. G. Kaburlasos, "An autonomous grape-harvester robot: integrated system architecture," *Electronics*, vol. 10, no. 9, p. 1056, 2021.

[45] Y. Peng, J. Liu, B. Xie, H. Shan, M. He, G. Hou, and Y. Jin, "Research progress of urban dual-arm humanoid grape harvesting robot," in *2021 IEEE 11th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 879–885, IEEE, 2021.

[46] A. Silwal, F. Yandun, A. Nellithimaru, T. Bates, and G. Kantor, "Bumblebee: A path towards fully autonomous robotic vine pruning," *arXiv preprint arXiv:2112.00291*, 2021.

[47] A. You, F. Sukkar, R. Fitch, M. Karkee, and J. R. Davidson, "An efficient planning and control framework for pruning fruit trees," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3930–3936, IEEE, 2020.

[48] L. He and J. Schupp, "Sensing and automation in pruning of apple trees: A review," *Agronomy*, vol. 8, no. 10, p. 211, 2018.

[49] J. R. Clark and N. Matheny, "The research foundation to tree pruning: a review of the literature," *Arboriculture & Urban Forestry*, vol. 36, no. 3, pp. 110–120, 2010.

[50] T. Botterill, S. Paulin, R. Green, S. Williams, J. Lin, V. Saxton, S. Mills, X. Chen, and S. Corbett-Davies, "A robot system for pruning grape vines," *Journal of Field Robotics*, vol. 34, no. 6, pp. 1100–1122, 2017.

[51] A. You, N. Parayil, J. G. Krishna, U. Bhattarai, R. Sapkota, D. Ahmed, M. Whiting, M. Karkee, C. M. Grimm, and J. R. Davidson, "An autonomous robot for pruning modern, planar fruit trees," *arXiv preprint arXiv:2206.07201*, 2022.

[52] B. Ma, J. Du, L. Wang, H. Jiang, and M. Zhou, "Automatic branch detection of jujube trees based on 3d reconstruction for dormant pruning using the deep learning-based method," *Computers and Electronics in Agriculture*, vol. 190, p. 106484, 2021.

[53] M. Fernandes, A. Scaldaferri, P. Guadagna, G. Fiameni, T. Teng, M. Gatti, S. Poni, C. Semini, D. Caldwell, and F. Chen, "Towards precise pruning points detection using semantic-instance-aware plant models for grapevine winter pruning automation," *arXiv preprint arXiv:2109.07247*, 2021.

[54] A. Zahid, M. S. Mahmud, L. He, D. Choi, P. Heinemann, and J. Schupp, "Development of an integrated 3r end-effector with a cartesian manipulator for pruning apple trees," *Computers and Electronics in Agriculture*, vol. 179, p. 105837, 2020.

[55] B. Zhang, X. Chen, H. Zhang, C. Shen, and W. Fu, "Design and performance test of a jujube pruning manipulator," *Agriculture*, vol. 12, no. 4, p. 552, 2022.

[56] F. Molaei and S. Ghatrehsamani, "Kinematic-based multi-objective design optimization of a grapevine pruning robotic manipulator," *AgriEngineering*, vol. 4, no. 3, pp. 606–625, 2022.

[57] A. Zahid, L. He, D. D. Choi, J. Schupp, and P. Heinemann, "Collision free path planning of a robotic manipulator for pruning apple trees," in *2020 ASABE Annual International Virtual Meeting*, p. 1, American Society of Agricultural and Biological Engineers, 2020.

[58] F. Yandun, T. Parhar, A. Silwal, D. Clifford, Z. Yuan, G. Levine, S. Yaroshenko, and G. Kantor, "Reaching pruning locations in a vine using a deep reinforcement learning policy," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2400–2406, IEEE, 2021.

[59] N. Strisciuglio, R. Tylecek, M. Blaich, N. Petkov, P. Biber, J. Hemming, E. van Henten, T. Sattler, M. Pollefeys, T. Gevers, *et al.*, "Trimbot2020: An outdoor robot for automatic gardening," in *ISR 2018; 50th International Symposium on Robotics*, pp. 1–6, VDE, 2018.

[60] A. H. Qureshi and Y. Ayaz, "Intelligent bidirectional rapidly-exploring random trees for optimal motion planning in complex cluttered environments," *Robotics and Autonomous Systems*, vol. 68, pp. 1–11, 2015.

[61] S. Corbett-Davies, T. Botterill, R. Green, and V. Saxton, "An expert system for automatically pruning vines," in *Proceedings of the 27th Conference on Image and Vision Computing New Zealand*, pp. 55–60, 2012.

[62] Y. Nakamura, H. Hanafusa, and T. Yoshikawa, "Task-priority based redundancy control of robot manipulators," *The International Journal of Robotics Research*, vol. 6, no. 2, pp. 3–15, 1987.

[63] P.-B. Wieber, R. Tedrake, and S. Kuindersma, "Modeling and control of legged robots," in *Springer handbook of robotics*, pp. 1203–1234, Springer, 2016.

[64] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, *et al.*, "Ros: an open-source robot operating system," in *IEEE International Conference on Robotics and Automation Workshop on Open Source Software*, 2009.

[65] M. L. Felis, "Rbdl: an efficient rigid-body dynamics library using recursive algorithms," *Autonomous Robots*, vol. 41, no. 2, pp. 495–511, 2017.

[66] B. Siciliano, O. Khatib, and T. Kröger, *Springer handbook of robotics*, vol. 200. Springer, 2008.

[67] Y. Liu, Z. Li, H. Su, and C.-y. Su, "Whole body control of an autonomous mobile manipulator using series elastic actuators," *IEEE/ASME Transactions on Mechatronics*, 2021.

[68] S. Kim, K. Jang, S. Park, Y. Lee, S. Y. Lee, and J. Park, "Whole-body control of non-holonomic mobile manipulator based on hierarchical quadratic programming and continuous task transition," in *2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM)*, pp. 414–419, IEEE, 2019.

[69] F. Rubio, F. Valero, and C. Llopis-Albert, "A review of mobile robots: Concepts, methods, theoretical framework, and applications," *International Journal of Advanced Robotic Systems*, vol. 16, no. 2, p. 1729881419839596, 2019.

[70] A. Roshanianfard, N. Noguchi, H. Okamoto, and K. Ishii, "A review of autonomous agricultural vehicles (the experience of hokkaido university)," *Journal of Terramechanics*, vol. 91, pp. 155–183, 2020.

[71] C. Semini, V. Barasuol, M. Focchi, C. Boelens, M. Emara, S. Casella, O. Villarreal, R. Orsolino, G. Fink, S. Fahmi, *et al.*, "Brief introduction to the quadruped robot hyqreal," *Istituto di Robotica e Macchine Intelligenti (I-RIM)*, 2019.

[72] P. Baranyi, "Kinova robotics: Kinova gen3.." `https://www.kinovarobotics.com/en/products/gen3-robot`.

[73] S. Amatya, M. Karkee, A. Gongal, Q. Zhang, and M. D. Whiting, "Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting," *Biosystems engineering*, vol. 146, pp. 3–15, 2016.

[74] S. Amatya and M. Karkee, "Integration of visible branch sections and cherry clusters for detecting cherry tree branches in dense foliage canopies," *Biosystems Engineering*, vol. 149, pp. 72–81, 2016.

[75] J. Zhang, L. He, M. Karkee, Q. Zhang, X. Zhang, and Z. Gao, "Branch detection for apple trees trained in fruiting wall architecture using depth features and regions-convolutional neural network (r-cnn)," *Computers and Electronics in Agriculture*, vol. 155, pp. 386–393, 2018.

[76] C. Yang, L. Xiong, Z. Wang, Y. Wang, G. Shi, T. Kuremot, W. Zhao, and Y. Yang, "Integrated detection of citrus fruits and branches using a convolutional neural network," *Computers and Electronics in Agriculture*, vol. 174, p. 105469, 2020.

[77] L. Qiang, C. Jianrong, L. Bin, D. Lie, and Z. Yajing, "Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine," *International Journal of Agricultural and Biological Engineering*, vol. 7, no. 2, pp. 115–121, 2014.

[78] C. A. Díaz, D. S. Pérez, H. Miatello, and F. Bromberg, "Grapevine buds detection and localization in 3d space based on structure from motion and 2d image classification," *Computers in Industry*, vol. 99, pp. 303–312, 2018.

[79] M. Annalisa, "The vineyard balance always stars with a good pruning." https://www.guadoalmelo.it/en/the-vineyard-balance-always-stars-with-a-good-pruning/.

[80] C. Steger, M. Ulrich, and C. Wiedemann, *Machine vision algorithms and applications*. John Wiley & Sons, 2018.

[81] D. Zapata, M. Salazar, B. Chaves, M. Keller, and G. Hoogenboom, "Estimation of the base temperature and growth phase duration in terms of thermal time for four grapevine cultivars," *International Journal of Biometeorology*, vol. 59, pp. 1771–1781, 2015.

[82] C. Intrieri and S. Poni, "Integrated evolution of trellis training systems and machines to improve grape quality and vintage quality of mechanized italian vineyards," *American Journal of Enology and Viticulture*, vol. 46, no. 1, pp. 116–127, 1995.

[83] S. Poni, C. Intrieri, E. Magnanini, *et al.*, "Seasonal growth and gas exchange of conventionally and minimally pruned chardonnay canopies," *VITIS-GEILWEILERHOF-*, vol. 39, no. 1, pp. 13–18, 2000.

[84] B. G. Coombe and P. Dry, "Viticulture. volume 2. practices," *Winetitles. Adelaide, Australia*, 1992.

[85] S. Poni, F. Bernizzoni, G. Briola, and A. Cenni, "Effects of early leaf removal on cluster morphology, shoot efficiency and grape quality in two vitis vinifera cultivars," in *VII International Symposium on Grapevine Physiology and Biotechnology 689*, pp. 217–226, 2004.

[86] H. Studer, G. Di Collalto, and H. Olmo, "Observations on the agronomic effects of hedge pruning on grapevines trained as horizontal cordons in california," *Riv. Vitic. Enol*, vol. 33, pp. 459–465, 1980.

[87] P. Clingeleffer, "Vine response to modified pruning practices," in *Pruning mechanization and crop control : proceedings of the Second Nelson J. Shaulis Grape Symposium*, Geneva, NY, New York Agricultural Experiment Station, 1993.

[88] J. Possingham, "New concepts in pruning grapevines," *Horticultural Reviews*, vol. 16, pp. 235–254, 1994.

[89] C. Intrieri, O. Silvestroni, and S. Poni, "Long-term trials on winter mechanical pruning of grapes," in *Proceedings of the Second International Seminar on Mechanical Pruning of Vineyards*, pp. 168–173, Edagricole Bologna, Italy, 1988.

[90] M. Keller, L. J. Mills, R. L. Wample, and S. E. Spayd, "Crop load management in concord grapes using different pruning techniques," *American Journal of Enology and Viticulture*, vol. 55, no. 1, pp. 35–50, 2004.

[91] T. J. Zabadal, G. Vanee, T. Dittmer, and R. Ledebuhr, "Evaluation of strategies for pruning and crop control of concord grapevines in southwest michigan," *American Journal of Enology and Viticulture*, vol. 53, no. 3, pp. 204–209, 2002.

[92] F. Bernizzoni, M. Gatti, S. Civardi, and S. Poni, "Long-term performance of barbera grown under different training systems and within-row vine spacings," *American Journal of Enology and Viticulture*, vol. 60, no. 3, pp. 339–348, 2009.

[93] M. Greven, J. Bennett, and S. Neal, "Influence of retained node number on s auvignon b lanc grapevine vegetative growth and yield," *Australian Journal of Grape and Wine Research*, vol. 20, no. 2, pp. 263–271, 2014.

[94] G. L. Creasy and L. L. Creasy, *Grapes*, vol. 27. Cabi, 2018.

[95] S. Poni, M. Gatti, A. Palliotti, Z. Dai, E. Duchêne, T.-T. Truong, G. Ferrara, A. M. S. Matarrese, A. Gallotta, A. Bellincontro, F. Mencarelli, and S. Tombesi, "Grapevine quality: A multiple choice issue," *Scientia Horticulturae*, vol. 234, pp. 445–462, 2018.

[96] F. Bernizzoni, S. Civardi, M. Van Zeller, M. Gatti, and S. Poni, "Shoot thinning effects on seasonal whole-canopy photosynthesis and vine performance in vitis vinifera l. cv. barbera," *Australian Journal of Grape and Wine Research*, vol. 17, no. 3, pp. 351–357, 2011.

[97] J. Brooks, "COCO Annotator." https://github.com/jsbroks/coco-annotator/, 2019.

[98] M. Hofer, M. Maurer, and H. Bischof, "Improving sparse 3d models for man-made environments using line-based 3d reconstruction," in *2014 2nd International Conference on 3D Vision*, vol. 1, pp. 535–542, IEEE, 2014.

[99] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[100] J. Dai, Y. Li, K. He, and J. Sun, "R-fcn: Object detection via region-based fully convolutional networks," *Advances in neural information processing systems*, vol. 29, 2016.

[101] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Springer, 2016.

[102] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron2." https://github.com/facebookresearch/detectron2, 2019.

[103] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1492–1500, 2017.

[104] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pp. 740–755, Springer, 2014.

[105] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 1, pp. 142–158, 2015.

[106] G. Lin, Y. Tang, X. Zou, J. Xiong, and Y. Fang, "Color-, depth-, and shape-based 3d fruit detection," *Precision Agriculture*, vol. 21, no. 1, pp. 1–17, 2020.

[107] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and electronics in agriculture*, vol. 147, pp. 70–90, 2018.

[108] Q. Zhang, Y. Liu, C. Gong, Y. Chen, and H. Yu, "Applications of deep learning for dense scenes analysis in agriculture: A review," *Sensors*, vol. 20, no. 5, p. 1520, 2020.

[109] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.

[110] L. Fu, Y. Majeed, X. Zhang, M. Karkee, and Q. Zhang, "Faster r–cnn–based apple detection in dense-foliage fruiting-wall trees using rgb and depth features for robotic harvesting," *Biosystems Engineering*, vol. 197, pp. 245–256, 2020.

[111] V. Bloch, A. Degani, and A. Bechar, "A methodology of orchard architecture design for an optimal harvesting robot," *Biosystems Engineering*, vol. 166, pp. 126–137, 2018.

[112] R. Verbiest, K. Ruysen, T. Vanwalleghem, E. Demeester, and K. Kellens, "Automation and robotics in the cultivation of pome fruit: Where do we stand today?," *Journal of Field Robotics*, vol. 38, no. 4, pp. 513–531, 2021.

[113] M. Keller, *The science of grapevines*. Academic press, 2020.

[114] M. I. Ferreira Fernandes, A. Scaldaferri, G. Fiameni, T. Teng, M. Gatti, S. Poni, C. Semini, D. Caldwell, and F. Chen, "Grapevine winter pruning automation: On potential pruning points detection through 2d plant modeling using grapevine segmentation," *11th IEEE International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, 2021*, in press, 2021.

[115] L. Devroye and T. J. Wagner, "8 nearest neighbor methods in discrimination," *Handbook of Statistics*, vol. 2, pp. 193–197, 1982.

[116] S. Lipovetsky, "Pca and svd with nonnegative loadings," *PATTERN RECOGNITION*, vol. 42, no. 1, pp. 68–76, 2009.

[117] G. Sánchez and J.-C. Latombe, "On delaying collision checking in prm planning: Application to multi-robot coordination," *The International Journal of Robotics Research*, vol. 21, no. 1, pp. 5–26, 2002.

[118] S. M. LaValle and J. J. Kuffner, "Rapidly-exploring random trees: Progress and prospects: Steven m. lavalle, iowa state university, a james j. kuffner, jr., university of tokyo, tokyo, japan," *Algorithmic and computational robotics*, pp. 303–307, 2001.

[119] B. Li and B. Chen, "An adaptive rapidly-exploring random tree," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 2, pp. 283–294, 2021.

[120] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal, "Stomp: Stochastic trajectory optimization for motion planning," in *2011 IEEE international conference on robotics and automation*, pp. 4569–4574, IEEE, 2011.

[121] J. J. Kuffner and S. M. LaValle, "Rrt-connect: An efficient approach to single-query path planning," in *Proceedings 2000 ICRA. Millennium Conference. IEEE Interna-*

tional Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065), vol. 2, pp. 995–1001, IEEE, 2000.

[122] I. A. Sucan, M. Moll, and L. E. Kavraki, "The open motion planning library," *IEEE Robotics & Automation Magazine*, vol. 19, no. 4, pp. 72–82, 2012.

[123] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: An efficient probabilistic 3d mapping framework based on octrees," *Autonomous robots*, vol. 34, no. 3, pp. 189–206, 2013.

[124] D. Meagher, "Geometric modeling using octree encoding," *Computer graphics and image processing*, vol. 19, no. 2, pp. 129–147, 1982.

[125] R. Raja, A. Dutta, and B. Dasgupta, "Learning framework for inverse kinematics of a highly redundant mobile manipulator," *Robotics and Autonomous Systems*, vol. 120, p. 103245, 2019.

[126] T. J. Tsay, Y.-F. Lai, and Y.-L. Hsiao, "Material handling of a mobile manipulator using an eye-in-hand vision system," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4743–4748, IEEE, 2010.

[127] F. Chen, M. Selvaggio, and D. G. Caldwell, "Dexterous grasping by manipulability selection for mobile manipulator with visual guidance," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 1202–1210, 2018.

[128] S. Katyara, F. Ficuciello, D. G. Caldwell, F. Chen, and B. Siciliano, "Reproducible pruning system on dynamic natural plants for field agricultural robots," in *Human-Friendly Robotics 2020*, (Cham), pp. 1–15, Springer International Publishing, 2021.

[129] M. Wrock and S. B. Nokleby, "Decoupled teleoperation of a holonomic mobile-manipulator system using automatic switching," in *2011 24th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 001164–001168, IEEE, 2011.

[130] B. Hamner, S. Koterba, J. Shi, R. Simmons, and S. Singh, "An autonomous mobile manipulator for assembly tasks," *Autonomous Robots*, vol. 28, no. 1, pp. 131–149, 2010.

[131] A. De Luca, G. Oriolo, and P. R. Giordano, "Kinematic modeling and redundancy resolution for nonholonomic mobile manipulators," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, pp. 1867–1873, IEEE, 2006.

[132] M. Li, Z. Yang, F. Zha, X. Wang, P. Wang, P. Li, Q. Ren, and F. Chen, "Design and analysis of a whole-body controller for a velocity controlled robot mobile manipulator," *SCIENCE CHINA Information Sciences*, vol. 63, no. 7, p. 170204, 2020.

[133] A. Roberto, "Redundancy modelling and resolution for robotic mobile manipulators: a general approach," *Advanced Robotics*, vol. 31, no. 13, pp. 706–715, 2017.

[134] S. Katyara, F. Ficuciello, T. Teng, F. Chen, D. G. Caldwell, and B. Siciliano, "Formulating intuitive stack-of-tasks with visuo-tactile perception for collaborative human-robot fine manipulation," *arXiv preprint arXiv:2103.05676*, 2021.

[135] A. Rocchi, E. M. Hoffman, D. G. Caldwell, and N. G. Tsagarakis, "Opensot: a whole-body control library for the compliant humanoid robot coman," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6248–6253, IEEE, 2015.

[136] S. Katyara, F. Ficuciello, T. Teng, F. Chen, B. Siciliano, and D. G. Caldwell, "Intuitive tasks planning using visuo-tactile perception for human robot cooperation," *arXiv preprint arXiv:2104.00342*, 2021.

[137] L. Baron, "A joint-limits avoidance strategy for arc-welding robots," in *Int. Conf. on Integrated Design and Manufacturing in Mech. Eng*, pp. 16–19, 2000.

[138] T. Yoshikawa, "Manipulability of robotic mechanisms," *The international journal of Robotics Research*, vol. 4, no. 2, pp. 3–9, 1985.

[139] B. Bayle, J.-Y. Fourquet, and M. Renaud, "Manipulability of wheeled mobile manipulators: Application to motion generation," *The International Journal of Robotics Research*, vol. 22, no. 7-8, pp. 565–581, 2003.

[140] S. B. Slotine and B. Siciliano, "A general framework for managing multiple tasks in highly redundant robotic systems," in *proceeding of 5th International Conference on Advanced Robotics*, vol. 2, pp. 1211–1216, 1991.

[141] Y. Guan, K. Yokoi, O. Stasse, and A. Kheddar, "On robotic trajectory planning using polynomial interpolations," in *2005 IEEE International Conference on Robotics and Biomimetics-ROBIO*, pp. 111–116, IEEE, 2005.

[142] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," in *Proceedings. 1985 IEEE International Conference on Robotics and Automation*, vol. 2, pp. 500–505, IEEE, 1985.

[143] X. Chen, Z. Li, Y. Wang, and J. Liu, "Effect of fruit and hand characteristics on thumb–index finger power-grasp stability during manual fruit sorting," *Computers and electronics in agriculture*, vol. 157, pp. 479–487, 2019.

[144] J. Peng, H. Xie, Y. Feng, L. Fu, S. Sun, and Y. Cui, "Simulation study of vibratory harvesting of chinese winter jujube (zizyphus jujuba mill. cv. dongzao)," *Computers and Electronics in Agriculture*, vol. 143, pp. 57–65, 2017.

[145] A. You, H. Kolano, N. Parayil, C. Grimm, and J. R. Davidson, "Precision fruit tree pruning using a learned hybrid vision/interaction controller," in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 2280–2286, IEEE, 2022.

[146] C. Ott, *Cartesian impedance control of redundant and flexible-joint robots*. Springer, 2008.

[147] Y. Wu, F. Zhao, T. Tao, and A. Ajoudani, "A framework for autonomous impedance regulation of robots based on imitation learning and optimal control," *IEEE Robotics and Automation Letters*, vol. 6, no. 1, pp. 127–134, 2020.

[148] J. Zhao, A. Giammarino, E. Lamon, J. M. Gandarias, E. De Momi, and A. Ajoudani, "A hybrid learning and optimization framework to achieve physically interactive tasks with mobile manipulators," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 8036–8043, 2022.

[149] S. Calinon and D. Lee, "Learning control," in *Humanoid robotics: A reference*, pp. 1–52, Springer Netherlands, 2017.

[150] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal processing magazine*, vol. 13, no. 6, pp. 47–60, 1996.

A   List of Publications

- Teng, T., Fernandes, M., Gatti, M., Poni, S., Semini, C., Caldwell, D. and Chen, F., 2021, July. Whole-body control on non-holonomic mobile manipulation for grapevine winter pruning automation. In 2021 6th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM) (pp. 37-42). IEEE.

- Teng, T., Gatti, M., Poni, S., Caldwell, D. and Chen, F., 2023. Fuzzy dynamical system for robot learning motion skills from human demonstration. Robotics and Autonomous Systems, p.104406.

- Fernandes, M., Scaldaferri, A., Fiameni, G., Teng, T., Gatti, M., Poni, S., Semini, C., Caldwell, D. and Chen, F., 2021, July. Grapevine winter pruning automation: On potential pruning points detection through 2d plant modeling using grapevine segmentation. In 2021 IEEE 11th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER) (pp. 13-18). IEEE.

- Guadagna, P., Frioni, T., Chen, F., Delmonte, A.I., Teng, T., Fernandes, M., Scaldaferri, A., Semini, C., Poni, S. and Gatti, M., 2021. Fine-tuning and testing of a deep learning algorithm for pruning regions detection in spur-pruned grapevines. In Precision agriculture'21 (pp. 3930-3936). Wageningen Academic Publishers.

- Guadagna, P., Fernandes, M., Chen, F., Santamaria1, A., Teng, T., Frioni1, T., Caldwell, D. G., Poni, S., Semini, C., Gatti, M., 2023. Using Deep Learning for Pruning Region Detection and Plant Organ Segmentation in Dormant Spur-Pruned Grapevines, Precision Agriculture (2023): 1-23.

- Katyara, S., Deshpande, N., Ficuciello, F., Teng, T., Siciliano, B., Caldwell, D.G. and Chen, F., 2022. Formulating Intuitive Stack-of-Tasks using Visuo-Tactile Perception for Collaborative Human-Robot Fine Manipulation. IEEE transactions on human-machine systems, under revision (Major revision).

B Awards

Some prizes have been obtained with the project:

- Best Conference Paper Finalist, 2021 6th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM).

- Best Student Paper Award, 2021 IEEE 11th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)