Autonomous robotic pruning in orchards and vineyards: A review

(Article begins on next page)

Autonomous Robotic Pruning in Orchards and Vineyards: a Review

Alessandro Navone[a,b,*], Mauro Martini[a,b], Marcello Chiaberge[a,b]

[a] *Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, 10129, TO, Italy*
[b] *PIC4SeR, Politecnico di Torino Interdepartmental Center for Service Robotics, Corso Ferrucci 112, Torino, 10141, TO, Italy*

Abstract

Manual pruning is labor intensive and represents up to 25% of annual labor costs in fruit production, notably in apple orchards and vineyards where operational challenges and cost constraints limit the adoption of large-scale machinery. In response, a growing body of research is investigating compact, flexible robotic platforms capable of precise pruning in varied terrains, particularly where traditional mechanization falls short.

This paper reviews recent advances in autonomous robotic pruning for orchards and vineyards, addressing a critical need in precision agriculture. Our review examines literature published between 2014 and 2024, focusing on innovative contributions across key system components. Special attention is given to recent developments in machine vision, perception, plant skeletonization, and control strategies, areas that have experienced significant influence from advancements in artificial intelligence and machine learning. The analysis situates these technological trends within broader agricultural challenges, including rising labor costs, a decline in the number of young farmers, and the diverse pruning requirements of different fruit species such as apple, grapevine, and cherry trees.

[*] corresponding author

Email addresses: alessandro.navone@polito.it (Alessandro Navone), mauro.martini@polito.it (Mauro Martini), marcello.chiaberge@polito.it (Marcello Chiaberge)

[1] This paper is part of the project NODES which has received funding from the MUR – M4C2 1.5 of PNRR funded by the European Union - NextGenerationEU (Grant agreement no. ECS00000036).

By comparing various robotic architectures and methodologies, this survey not only highlights the progress made toward autonomous pruning but also identifies critical open challenges and future research directions. The findings underscore the potential of robotic systems to bridge the gap between manual and mechanized operations, paving the way for more efficient, sustainable, and precise agricultural practices.

Keywords: Pruning, Autonomous Robots, Review, Orchards, Vineyards.

1. Introduction

Every year, fruit orchards need intensive labor operations such as harvesting and pruning, which are still conducted manually by the majority of farmers. Pruning is a repetitive and costly process that aims to ensure the growth of new plant nodes after winter seasons. For instance, recent studies show that the pruning operation covers from 20% to 25% of the overall annual labor cost for an apple farmer [1].

Fruit production is a key agricultural industry in both the European Union (EU) and the United States (US). In 2022, the EU's fruit harvesting sector was valued at €27.3 billion, with Spain, Italy, and Poland leading in production. The EU produced 14.7 million tonnes of pome fruit, 10.5 million tonnes of citrus fruit, 6.3 million tonnes of stone fruit, 2.6 million tonnes of tropical fruit, 1.1 million tonnes of nuts, and 0.7 million tonnes of berries [2]. In the same year, the US recorded $34.2 billion in sales from fruit, tree nuts, and berries, representing 6.3% of total agricultural sales, according to the US Department of Agriculture (USDA) Economic Research Service. Moreover, vines covered 3.2 million hectares (ha) in the EU in 2020, with 2.2 million vineyard holdings. More importantly, 83.3% had less than 1 ha of vineyards, confirming the prevalence of small activities and fields.

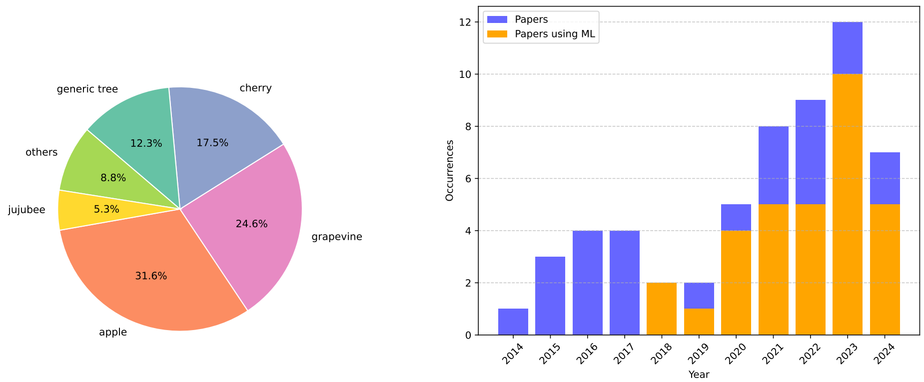
Despite the sector's economic significance, the EU faces demographic and labor-related challenges. The proportion of young farmers has sharply declined, with only 6.5% of EU farm managers under the age of 35 in 2020 [3]. Additionally, the availability of manual labor has decreased in recent years. From 2010 to 2019, labor costs—typically around 30% of total production costs for fruit and nut trees—have increased for several reasons. In the US, a recent USDA Economic Research Service report discusses how producers are responding to these rising costs and labor shortages [4]. Long-term strategies adopted in the US include the use of temporary worker visas such as the

H-2A program and, in some cases, a reduction in domestic production.

Automatic machinery has emerged as a possible support for boosting efficiency in agriculture and reducing costs without requiring intensive manual operations. Several recent studies discussed the feasibility of adopting a mechanized solution for hedging in both vineyards and orchards [5, 1]. Agriculture experts are advancing methods to organize canopies in walls or trellis structures to foster mechanized pruning [6]. However, huge machines have limitations. First, they cannot be applied to slopes and hills, typical of European vineyards and orchards. Moreover, the investment required is not ignorable and affordable for small and medium-sized producers, ranging around €$12k$ for the machine itself to be coupled with a tractor, without considering maintenance and insurance costs. Finally, the fine wine economy requires precise and careful operations to prevent fruit damage that can affect the temporal evolution of the fermentation process. In this context, custom pruning strategies are adopted by different wine producers, raising the need for flexible and customizable solutions.

Apples and vines represent the principal production industry, and, hence, the main focus of research advancement in the direction of technological solutions. Indeed, among papers published from 2014 to 2024, as presented in Fig. 1a, 31.6% of the studies consider the pruning of apple trees and 24.6% consider grapevine as application scenarios. A relevant part of the considered studies covers the pruning of cherry trees (17.5%), while many works target generic trees, preferring not to target a specific fruit tree but addressing part of the methodology or analyzing the performance of a specific type of sensor. Finally, jujubees cover a smaller part of the analysis, while the rest of the considered work, which is grouped as "others", analyzes sparsely other fruit trees such as walnut, avocado, pear, mango, and more.

Robotic pruning represents an optimal trade-off between manual and massively mechanized operations, leading to a flexible solution with small- to medium-sized platforms. A typical robotic platform for pruning is a complex system composed of different modules. In this review, we analyze the innovative contributions to each robotic component. Nonetheless, a special focus is devoted to machine vision, perception, plant skeletonization, reconstruction, and control, considering the huge impact of recent Artificial Intelligence (AI) on these areas.



(a) (b)

Figure 1: (a) Pie chart of the distribution of the addressed crops in the analyzed papers. (b) Distribution of the analyzed papers in the period of time from 2014 to 2024. The orange bars represent the paper using machine learning, showing a large increase in recent years.

1.1. Scope of the review

Currently, autonomous robotic pruning in orchards and vineyards remains a niche area within the broader robotics literature focused on precision agriculture, as evidenced by the data presented in Fig. 1b. A first survey in the field of autonomous pruning was conducted by He and Schupp in 2018 [6], in which sensing and perception techniques are explored solely in the context of apple orchards. It covers mainly the fields of tree training systems, pruning strategies, and plant structure reconstruction. A subsequent study by Tinoco et al. [7] provides a more detailed analysis of the design of a manipulator for pruning and harvesting. The considered pruning manipulators were tested in two distinct settings: vineyards and apple orchards. Finally, Zahid et al. [8] examined the technological advancements pertaining to the primary components of a robotic pruner for apple trees, including machine vision, manipulation, path planning, obstacle avoidance, and the design of the end effector.

In this work, we consider papers from the decade 2014-2024, which, as emerged from an attentive literature review, capture the evolution of trends and the recent advancements in AI that have deeply influenced this research context. For the identification of the relevant publications in this field, we carried out an extensive research using Google Scholar[2] and Scopus[3] in two main phases, the first one in June 2024 and a later one in February 2025 for recent updates and advancements.

The graph shown in Fig. 1b shows clearly that the number of published papers, and therefore the interest towards autonomous pruning of fruit trees and the related tasks, has consistently increased in recent years. We believe that recent advancements in AI and robot autonomy have paved the way for consistent research in robotic pruning, as shown in the graph. Hence, in this review, we discuss the latest contributions, trends, and open challenges related to the main aspects of pruning in orchards and vineyards.

1.2. Paper organization

The paper is organized as follows. Section 2 offers a wide overview of the autonomous pruning pipeline, giving insight into a possible subdivision of the tasks. Section 3 reports visual perception methods for robotic pruning, organizing them in classical computer vision methods, in Subsection 3.1, and deep-learning-based methods, in Subsection 3.2. Section 4 describes skeletonization and tree reconstruction methods, while Section 5 describes the pruning points estimation methods. Section 6 treats simulation environments for autonomous pruning training and development. Section 7 discusses the various hardware parts needed to constitute the different parts of an autonomous robotic pruner, starting from sensors, in Subsection 7.1, then manipulators in Subsection 7.2, and end-effectors in Subsection 7.3, and how to control them, in Subsection 7.4. Then, an overview of works that merge all the different parts is given in Section 8. Finally, Section 9 discusses the challenges and limitations, in Subsection 9.1 and 9.2 , and gives possible future developments in Subsection 9.3. Conclusions are drawn in Section 10.

2. Autonomous Pruning with Robots

Automated robotics solutions are emerging as a significant support for human jobs in various sectors. Precision agriculture is one of them, as it involves many repetitive tasks in the field that require only partial human input and supervision. Plant pruning in orchards and, above all, vineyards represents one of the most attractive tasks in this direction. The intelligence

[2] `https://scholar.google.com/`
[3] `https://scopus.com/`

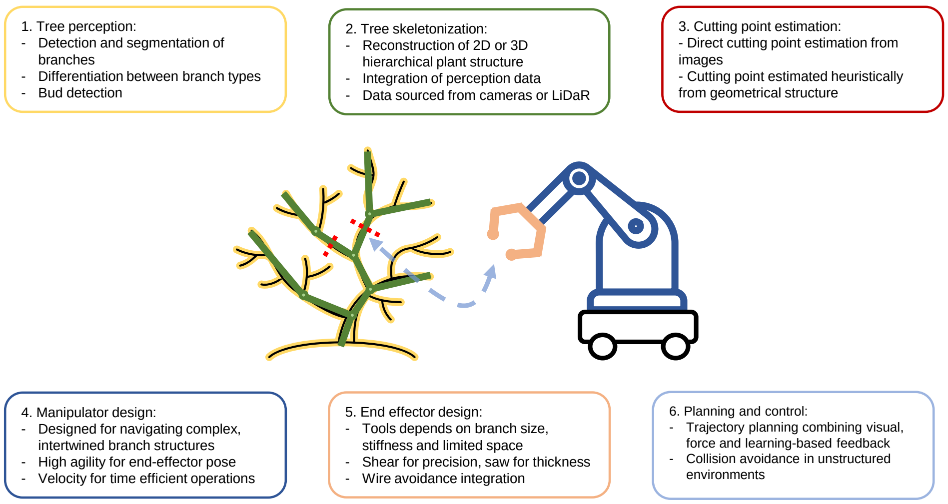


Figure 2: Scheme of the autonomous robotic pruning pipeline. The above part of the scheme depicts the perception-related tasks, while the bottom part shows the tasks related to manipulator design and control.

of a robotic solution resides in the combination of sensors and algorithms. In fact, according to the literature of the last ten years, a robotic pruning system can be divided into a few modules, as shown in the scheme in Fig. 2.

The robotic pruning pipeline can be decomposed in two main aspects. First, a perception pipeline, which aims at sensing and reconstructing the structure of the target plant and its surroundings, as shown in the upper part of Fig. 2. Then, the design of the manipulator, its actuation system, and logic are shown in the bottom part of the figure. The perception part of the pruning task can be decomposed in three main phases:

1. **Tree perception**: sensor data, mostly obtained from cameras or LIDARs, must be interpreted to identify regions of interest. Segmentation and detection algorithms are used to distinguish the plant parts from the background. In some cases, the different plant parts, which are fundamental to have more meaningful information on the pruning action, are differentiated from each other. Some works also target the detection of the single buds lying on the branches.
2. **Tree skeletonization**: after the regions of interest are identified, it is fundamental for the pruning operation to obtain a clear structure of the

plant. For this operation, the previously obtained information about the position of the branches, their type, and some additional parameters such as their width and their orientation, is merged to form a graph-like representation, obtaining a simplified essential structure of the plant, which is more easily interpretable from the following algorithms.

3. **Cutting point estimation**: this phase represents the crucial point of the perception activity, aiming at finally identifying the points on the plants that need to be cut. For this purpose, two main approaches are employed: the first one exploits the previously obtained structure and applies several algorithms or heuristics to find the desired cutting points. The second one consists of learning-based approaches that train algorithms based on the annotations of experts.

Therefore, the output of the perception pipeline is the pose in the 6D space of the desired cutting point, which is used as input to the manipulation part.

Regarding the manipulation, we considered three different macro-areas that were faced separately in the analyzed works:

4. **Manipulator design**: the architecture of the manipulator should be chosen wisely to address the movement in an intricate environment, such as interwined branches. In fact, it should have enough degrees of freedom to make the desired pose of the end-effector possible. Moreover, to be efficient enough, alongside the precision, the manipulator should be fast enough during its movement
5. **End effector design**: the tools should be designed depending on the target branch stiffness and size, considering also the limited space due to the intricate branches and supports. Moreover, depending on the type of cut to be made, shears or saws are considered. In some cases, due to the presence of support wires, the shears can be equipped with a specific solution for their avoidance.
6. **Planning and control**: given the input from the perception pipeline, a trajectory should be planned, considering collision avoidance. Hybrid control strategies can be adopted, combining visual force and learning-based feedback.

The plot in Fig. 3 shows the number of papers analyzed covering the different areas proposed in the previous summarization. As the graph shows, plant perception alone is the most covered task, and it has been strongly pushed by the advancements of ML in the last five years. Right after, the

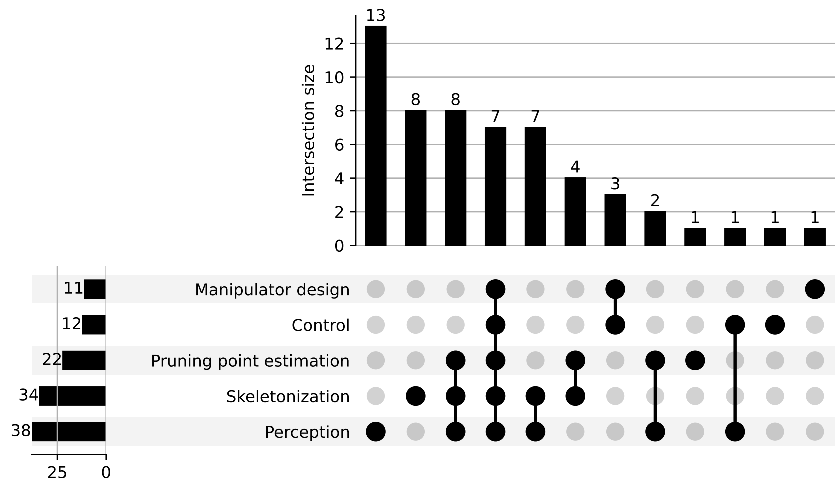


Figure 3: This plot shows the number of papers addressing individual tasks in the autonomous pruning pipeline (bottom left) and combinations of tasks (top). The manipulator and end-effector design tasks are grouped together, as they are consistently paired in the reviewed papers.

skeletonization module was run alone and merged with the perception and the estimation of the pruning point. These trends highlight that the most covered topics are all related to the problem of sensing the plant. Complete research works proposing a full robotic pruning system integrating sensing and manipulation have also been published, with a similar effort in terms of the number of papers. Finally, the design of the manipulators, end effectors, and control has a smaller coverage in terms of published works, probably due to the more general state-of-the-art solutions already available and the recent developments of enhanced commercial platforms that can be deployed directly.

3. Visual Perception for Robotic Pruning

Accurate detection and segmentation of tree structures—such as trunks, branches, and nodes—are critical tasks for enabling automation in horticultural operations like pruning. While early efforts relied on handcrafted features and traditional image processing pipelines, recent advancements in machine learning and deep neural networks have enabled more robust and scalable solutions. The following sections outline this progression, highlighting the key contributions, technologies, and limitations of each approach.

3.1. Classical Computer Vision Algorithms

Early works strongly rely on the employment of classical computer vision algorithms, as deep learning had not yet gained traction, and the hardware lacked sufficient computational power. Initial systems commonly used 2D RGB cameras to detect tree structures. For instance, Qiang et al. [9] employed an RGB camera and morphological features with a multi-class Support Vector Machine (SVM) to detect citrus branches thicker than 5 pixels. Ji et al. [10] applied a Contrast-Limited Adaptive Histogram Equalization (CLAHE) technique to segment apple tree branches, finding it superior to Otsu and basic histogram-based methods in terms of recognition rate. Shalal et al. [11] combined RGB imagery with laser scanning to detect apple trunks. However, these 2D camera systems were insufficient for precise spatial localization [12, 13].

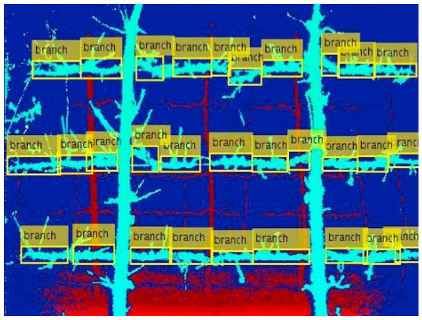
To overcome these limitations, researchers began incorporating 3D imaging technologies, such as stereo vision and Time-of-Flight (ToF) cameras. These systems significantly improved spatial accuracy in detecting trunks, branches, and canopies. For instance, Zhang et al. [14] used a stereo vision camera to capture RGB, depth, and index images. Amatya et al. [15] utilized a Bayesian classifier based on morphological features (e.g., orientation, length, and thickness) to distinguish between cherry branches, leaves, fruit, and background. In a follow up study, the authors [16] modeled branch geometry using linear or logarithmic functions and employed cherry cluster locations to infer occluded branches, achieving enhanced segmentation even under foliage. In a different approach, Botterill et al. [17] reconstructed vine structures in 3D using a triangulation-based feature matching algorithm, followed by a Radial Basis Function SVM (RBF-SVM) classifier trained on hand-labeled 5×5 image patches. Controlled lighting and monochrome backgrounds further improved segmentation performance in their setup.

3.2. Deep Learning Techniques

Adopting deep learning methods has significantly improved the accuracy of tree structure detection and segmentation. In the field of branch detection and segmentation, several machine learning algorithms and deep learning

Table 1: Overview of deep learning-based computer vision methods for various tree crops.

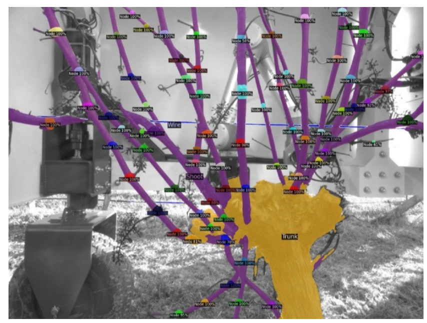
Crops	Method	Performance	Year	Ref.
Citrus	SVM + morphological ops	Branch seg. (diameter >5 px)	2014	[9]
Generic tree	HSV color + edge detection	Seg. acc. = 96.64%	2015	[11]
Apple	CLAHE-based seg.	Acc. = 94%	2016	[10]
Cherry	Bayesian classifier	Acc. = 93.8%	2016	[16]
Cherry	Bayesian classifier	Acc. = 89.2% (w/ foliage)	2017	[15]
Grapevine	RBF-SVM	P = 58%, R = 49% (canes)	2017	[17]
Apple	R-CNN (AlexNet)	PCI: P = 86%, R = 81%; +Depth: P = 92%, R = 86%	2018	[14]
Apple	Depth cut + SegNet (multi-class)	Trunk: Acc. = 0.92, IoU = 0.59; Branch: Acc. = 0.93, IoU = 0.44	2018	[21]
Grapevine	Faster R-CNN (various backbones)	Best (ResNet18): F1 = 0.55, AP = 45.1%	2019	[22]
Apple	Faster R-CNN + U-net (UAV)	Det: IoU = 0.735, F1 = 0.925; Seg: Acc. = 94.68%	2020	[23]
Apple	Depth cut + SegNet (multi-class)	Trunk: Acc. = 0.91; Branch: Acc. = 0.92; Wire: Acc. = 0.97	2020	[24]
Grapevine	CNN (FCN-VGG16 best)	F1: background = 0.97, trunk = 0.94	2020	[25]
Grapevine	Mask R-CNN (ResNet101)	Avg P = 41.1%, R = 48.7%	2021	[26]
Guava	Tiny Mask R-CNN	F1 = 0.518	2021	[27]
Jujubee	SPGNet + DBSCAN	Trunk: IoU = 0.85; Branch: IoU = 0.76	2021	[28]
Apple	Cascade Mask R-CNN (Swin-T)	@IoU=0.5: bboxAP = 9.879, seg. AP = 0.893	2022	[29]
Cherry	FlowNet2 + pix2pix (sim)	IoU = 62.3%, FN = 28.7%, FP = 1.8%	2022	[30]
Cherry	FlowNet2 + pix2pix (sim)	-	2022	[31]
Cherry	Mask R-CNN + postproc.	IoU: 0.78 (light), 0.67 (nat.); P = 0.89/0.81; R = 0.97/0.81	2023	[32]
Grapevine	Faster R-CNN + Mask R-CNN	Node seg.: Acc. = 0.88, R = 0.85; pruning pt. det. = 0.97	2023	[33]
Apple/Grapevine	HOB-CNN (occlusion)	RMSE = 1.63–2.66, Corr. = 0.925–0.948	2023	[34]
Apple	UNet++ (InceptionV3)	IoU = 0.625, F1 = 0.7692	2023	[35]
Apple	SOLOv2 (ResNet50)	AP: front = 0.934, back = 0.917, side = 0.940, cloudy = 0.947	2023	[36]
Grapevine	Hourglass net	Nodes: F1 = 0.92, ADE = 0.66px; Courson: F1 = 0.75	2023	[37]
Grapevine	Detectron2 (panoptic)	PQ: node = 0.6832, wire = 0.674; SQ varies 0.77–0.85	2023	[20]
Grapevine	YOLOv7–10	F1 = 70–86.5%	2024	[19]
Grapevine	PSP-Net + YOLOv5	PSP: mIoU = 83.7%, Acc. = 96.8%; YOLOv5: F1 = 0.72	2024	[38]



(a)

(b)

(c)



(d)

Figure 4: Overview of four different tasks of computer vision for autonomous pruning: (a): detection of branches on pseudo-color image extracted from PCD [14], (b): semantic segmentation of different parts of cherry trees [18], (c): bud detection on grapevine [19], (d): panoptic segmentation of the structure of different parts of grapevine and different buds[20].

architectures have been proposed: a complete overview of the different employed methods is shown in Table 1, and some examples of the possible applications of deep computer vision are shown in Fig. 4.

Region-based CNN (R-CNN) [39] architectures have been widely applied for tree component detection and segmentation. These models rely on region proposals to localize potential objects before classifying them and generating bounding boxes or segmentation masks. *Fast R-CNN* [40] and *Faster R-CNN* [41] enhance computational efficiency by processing the entire image with a CNN to generate feature maps and using a Region Proposal Network (RPN). Mask R-CNN [42] extends this approach by adding a segmentation mask prediction branch. Numerous studies have adopted these architectures for different detection tasks. Tong et al. [29], for instance, evaluated two kinds of R-CNN, namely *Mask R-CNN* and *Cascade Mask R-CNN* [43], one of its variants, to detect and segment the trunk, primary branches, and supports of apple trees. For each model, they tested two backbones: Swin-T [44] and ResNet50. Borrenpohl and Karkee [32] compared two models of a Mask R-CNN with a ResNet101 backbone to detect vigorous leaders from Upright Fruiting Upshots (UFO) in sweet cherry trees, both in artificial and natural light. Images were first segmented using a stereo image and a deep stereo matching [45], to identify the tree of interest, and then manually annotated to identify the trunk and branches. The accuracy was increased by developing two post-processing algorithms: to avoid double detections of the same leader while the second to avoid simultaneous detection of two adjacent leaders. In other cases, such R-CNN are used to segment multiple grapevine branch classes, such as main cordon, canes, and nodes, which facilitates the process of reconstructing the tree structure and the following estimation of pruning points. In this regard, Fernandes et al. [26] compared a Mask R-CNN with several backbone models (R50-FPN, R101-FPN, X101-FPN), demonstrating ResNet101 to be the best backbone. Similarly, Guadagna et al. [33] fine-tuned a Mask R-CNN for grapevine organ segmentation, distinguishing more classes: cordon, arm, spur, cane, and node. The trained model was then tested on images of light-pruned, shoot-thinned, and unthinned control samples. In this case, better performances were achieved on thinned shoots compared to unthinned control samples, underlining the role of canopy management in improving the performance of autonomous solutions. A tiny Mask R-CNN was also used by Lin et al. [27] to simultaneously obtain an instance segmentation of branches and fruits in guava crops for obstacle avoidance during fruit harvesting. Later, they were reconstructed as spheres and cylindrical segments starting from the segmented frames and the point cloud data.

Other works proposed the usage of a detection model, such as the one by Majeed et al. [22], where a Faster R-CNN was employed to detect the visible part of the cordon for green shoot thinning in grapevine. It was deployed through transfer learning and fine-tuning using pre-trained networks such as AlexNet, VGG16, VGG19, and ResNet18, obtaining a higher accuracy with the latter. A Faster R-CNN has also been used by Wu et al. [23] to identify apple trees' bounding boxes from UAV images. Then, a *U-net* model [46] was applied on the previously extracted bounding boxes to segment branch pixels.

In similar applications, R-CNN has been employed with different inputs:

Zhang et al. [14], which addressed the problem of branch detection in apple orchards using a Pseudo-Color depth Image (PCI) obtained by mapping the depth to a certain color value, as shown in Fig. 4a. The adopted architecture consisted of an improved AlexNet [47] network trained and compared on depth images and PCIs and depth images together. The issue of reconstructing the position of occluded branches has been addressed by Chen et al. [34], who proposed a regression-based deep learning model, Hallucination of Occluded Branches CNN (HOB-CNN), to achieve this in apples and vine plants, enabling direct reconstruction of the three structures. Tree branch position prediction is formulated as a regression problem towards the horizontal locations of the branch along the vertical direction or vice versa. Two different versions of U-Net [46] were trained for the reconstruction of the trunk: U-Net Visible is used to detect only visible branch sections, while U-Net Whole is used to hallucinate the whole branch in occlusion conditions. Finally, a curve fitting method is employed to obtain a single-pixel skeleton of the plants. Similarly, Kok et al. [35], used a U-net++ [48] with an InceptionV3 encoder to segment visible apple branches, followed by a skeletonization algorithm in order to reconstruct the obstructed parts of the plant.

Another popular architecture employed in different work is *SegNet* [49], a fully convolutional network for semantic pixel-wise segmentation consisting of an encoder-decoder structure. Majeed et al. (2018) [21] employed RGB images and Pointcloud Data (PCD) to detect one-year-old apple trees. After a preprocessing employing the depth information to cut out the background noise, SegNet was employed for the semantic segmentation of background, trunk, and branches. In a follow-up study [24], they extended the segmentation to trellis wires, confirming improved performance when foreground-RGB images were used. In another work, Majeed et al. [25] segmented the two cordon on grapevines, comparing SegNet with its weight initialized with VGG16 and VGG19 [50], and other fully convolutional networks, such as AlexNet and VGG16 modified for semantic segmentation as in [51]. Moreover, to separate the right and the left cordons, a mathematical model was applied, evaluating first the centroid of the trunk and, later, the two parts of the cordon and fitting a curve on the centroids of each cordon to better fit their shape.

Other segmentation models were successfully employed. Tong et al. [36] used a *SOLOv2* [52] with a Resnet50 backbone to segment RGB images of dormant apple trees, identifying trunks, branches, and supports. It is then compared to A Mask R-CNN and Cascade Mask R-CNN, achieving a better performance, also compared in different illumination weather conditions. Moreover, Chen et al. [38] employed a *PSP-net* [53] to separate branches and the main stem from the background in grapevines. The results show better performance with respect to other considered models such as U-net and DeepLabv3+. Moreover, on the segmentation mask achieved with the first step of segmentation, different versions of YOLO were employed to detect the nodes of the canes. Experimental results show how YOLOv5 achieves better results. Similarly, Oliveira et al. [19] compared four different versions of *YOLO*, such as YOLOv7, YOLOv8, YOLOv9, and YOLOv10, for the detection of nodes in dormant grapevine. They trained the networks using a public dataset presented in [37] with images presenting diverse artificial backgrounds, and validated them on other data, contributing with a new dataset. YOLOv7 was assessed as the most robust model, achieving a trade-off between accuracy and inference speed. The bud detection task can be seen in Fig. 4c. With the same objective, Gentilhomme et al. [37] employed an architecture based on a Stacked Hourglass Network [54], ViNet, which outputs a heatmap with the different kinds of branches of interest such as nodes and branches. It produces a stack of node and branch heatmaps. The position of the nodes is extracted through a post-processing step that localizes local maxima within large enough blobs in the heat maps. Williams et al. [20][55] developed a novel vision system, which aims at reconstructing the whole structure of the plant. The detection system employs a *Detectron2* [56] for panoptic segmentation [57] of the cane information from the 2D images, as seen in Fig. 4d. In this case, panoptic segmentation is used for semantic segmentation of trunk, canes, and wires, while for the instance segmentation of the different nodes, which are fundamental to reconstruct the structure of the plant.

You et al. [30][31] exploit the optical flow technique joined with neural networks to enhance the segmentation of branches, trained entirely in a simulation environment. First, optical flow is estimated with a fully-sized FlowNet2 [58]. Then, alongside the three-channel RGB frame, is passed to a pix2pix [59] *Generative Adversarial Networks* to perform segmentation of the branches. Moreover, a neural network comparison is carried out with different types of inputs, demonstrating the similar accuracy in simulation environment and the real world.

Most of the aforementioned models operate on 2D RGB data. A limited number of models directly exploit 3D data such as depth images or point clouds, which, in the case of robotic pruning, may offer some additional information that is fundamental for a better 3D reconstruction of the considered plant. For instance, Ma et al. [28] utilized two cameras to obtain a high-quality 3D point cloud in the field to create a comprehensive 3D model. Then, they used SPGNet [60] to automatically segment trunks and branches.

As shown in Table 1, there has been a clear progression from traditional image processing and shallow machine learning methods (e.g., SVM, Bayesian classifiers) toward deep learning-based approaches. In particular, CNNs have demonstrated strong performance by learning visual features such as edges, textures, and shapes that are critical for detecting trunks and branches. Their translation invariance and scalability make them well-suited to handling variations in branch positions and lighting conditions. This evolution is accompanied by a notable increase in segmentation and detection performance. Early methods achieved moderate accuracy (e.g., Accuracy = 89–94%), while more recent deep learning models report higher and more consistent metrics such as F1 scores exceeding 0.90, IoU values above 0.80, and Precision/Recall rates $>$90% in some cases. For instance, the use of Faster R-CNN, Mask R-CNN, U-Net++, and YOLOv7–10 has led to metrics like IoU = 0.85, Precision = 0.89, Recall = 0.97, and mIoU = 83.7%, even in challenging natural conditions. Additionally, models incorporating depth or 3D data further improve performance, as seen with HOB-CNN (RMSE = 1.63–2.66, Correlation = 0.925–0.948) and SPGNet + DBSCAN (IoU = 0.85 / 0.76 for trunk and branches, respectively). These results underscore the effectiveness of deep learning in handling occlusions, variable lighting, and complex tree architectures.

4. Skeletonization and Tree Structure Reconstruction

After the tree is detected, the data obtained from the detection phase must be interpreted and exploited to reconstruct its structure. The reconstruction of the tree's shape can be performed in 2D or 3D, merging depth data from stereo cameras or other sources. A broad overview about all the analyzed methods is shown in Table 2 For this scope, Dukić et al. [69], presented the BRANCH dataset, a multiple-viewpoint dataset of 70 images which covers pear trees, both before and after pruning. They proposed a pointcloud extraction for the multiple POV images, then the 3D final model is obtained by using the RANSAC and Iterative Closest Point (ICP) [70] algorithms. In addition, by superimposing the pointcloud before and after the pruning, they obtained the position of the pruned branches. A similar multi-POV approach

Table 2: Overview of skeletonization and thinning methods.

Crop	Method	Performances	Year	References
Generic tree	SbG + RANSAC	Branch Id. Acc. = 96.0%, Branch modeling Acc. = 92.2%, Radius Acc. = 78.0%	2015	[61]
Apple tree	Semicircle fitting + Center point optimization	Diameter Acc. = 89.4% (±5mm), Branch ID = 100%	2016	[62]
Grapevine	Stochastic Grammar + RDP line detection	Overlap with groundtruth: 95%	2017	[17]
Apple tree	Split-and-merge + robust cylinder fitting	Branch ID Acc. = 96%, Diameter error = 0.6 cm	2017	[63]
Apple tree	Polynomial interpolation	Corr. Coeff. r = 0.91	2018	[14]
Apple tree	Harris corner detection + AHCA	-	2019	[64]
Grapevine	6-th degree polynomial interpolation	80% $R^2 > 0.98$	2020	[25]
Grapevine	Auxiliary matrix	-	2021	[26]
Guava tree	Point cloud + RANSAC cylinders	2D F1 = 0.394, 3D F1 = 0.451	2021	[27]
Apple tree	2D junction point estimation	F1 = 91.06%	2022	[29]
Cherry	DBSCAN + Space Colonization	Angle RMSE = 3.309°, Trunk RMSE = 0.069 m	2022	[65]
Cherry	Topo/Geom. Constraints	Median Acc. = 70%	2022	[18]
Apple tree	Point2Skeleton + OBR	Total err. = 18.68 ± 14.3 mm, Obscured = 38.11 ± 32.64 mm	2023	[35]
Apple tree	3D junction estimation	Diam. MAE = 1.10 mm, Spacing MAE = 16.06 mm	2023	[36]
Grapevine	Shortest-path on resistivity graph	-	2023	[37]
Generic tree	DBSCAN + slicing	Pos. dev. = 0.418 cm, Dir. dev. = 8.47°	2023	[66]
Jujubee	ICP + SCA	Reg. err. = 0.66–0.91 cm	2023	[67]
Grapevine	Zhang–Suen + YOLOv5	Thinning = 95.44%, Sens. = 90.87%	2024	[38]
Walnut	Delaunay + Dijkstra + MST + LM Fit	Acc. dev. < 7%	2024	[68]
Grapevine	2D/3D node constraints	Conn. Acc.: 67.05% (1-side), 78.4% (2-side)	2023	[20]

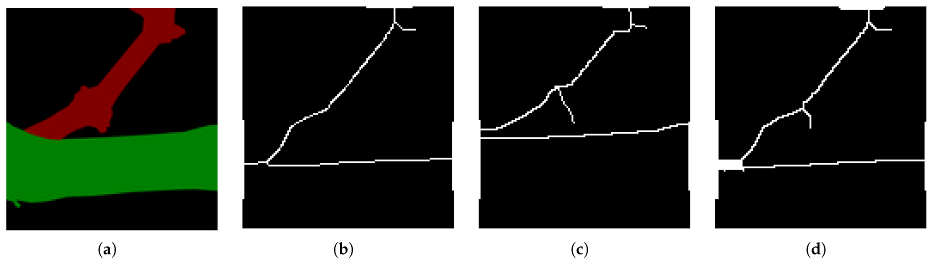


Figure 5: Comparison with Zhang and Suen with other methods [38]. (a) RGB image. (b) The result image of the Zhang–Suen thinning algorithm. (c) The result image of the Hilditch algorithm. (d) The result image of the Stentiford algorithm.

has been adopted by Li et al. [68], where the PCD of a ToF sensor have been merged and filtered to reconstruct the full structure of a walnut tree. Moreover, the IWOA-RANSAC-NDT algorithms are used for the 3D registration. Then, Delaunay triangulation [71] and Dijkstra shortest path [72] algorithms are employed to evaluate the minimum spanning tree, a concept originating from the nutrient transport paths in ecology. Then, the Kd-tree data structure is used for branch segmentation and, finally, the Levenberg-Marquardt algorithm [73] is used to obtain the full skeleton model.

An important part of the tree reconstruction process consists of the estimation of the parameters of the tree, such as in the work by Tabb and Medeiros [74], where parameters such as branch diameters, branch length, and branch angles are estimated. Another more recent study, by Dong et al. [75], introduces a novel method for three-dimensional trunk and branch volume calculation of individual apple trees, enabling the creation of pre- and post-pruning maps using a voxel-based algorithm.

Some studies use simple solutions, such as Zhang et al. [14], where a polynomial fitting is used as a skeletonization algorithm after the segmentation mask is obtained. Majeed et al. [25] reconstructed the structure of the cordon in a vine plant by fitting its centroids, obtained through applying an 8-connectivity matrix to the segmentation mask, with several methods such as polynomial, Gaussian, Fourier and sum of sines. Also, Borrenphol et al. [32] employed a high-degree polynomial to fit each prediction of branches of cherry tree to estimate each trajectory, differentiating one from another. Another study by Botterill et al. [17] starts from the a-priori knowledge that each cane is a smooth curve with approximately uniform thickness and they are all connected to a tree. For this sake, they employed a Stochastic Image

Grammar [76] to represent this knowledge, which is described in [77]. First, the cane edge segments are extracted by taking the outline of the foreground and approximating them as a set of straight lines using the Ramer-Douglas-Peucker [78] algorithm. The nodes are joined through an SVM considering lengths, thickness, curvature, and distances.

The approach proposed by Zhang and Suen [79] found several applications for skeletonizing tree structures. It consists of a fast parallel thinning algorithm consisting of two subiterations: the first targets the removal of southeast boundary points and northwest corner points, while the second focuses on eliminating northwest boundary points and southeast corner points. This approach keeps endpoint and pixel connectivity unchanged, thinning each pattern to a skeleton of unitary thickness. In fact, a comparison with other skeletonization algorithms for the sake of plant structure reconstruction has been carried out by Cuevas et al. for rose structure reconstruction [80], demonstrating its superiority. For instance, Tong et al. [29][36] were able to obtain the skeleton structure of the apple trees once they evaluated the 2D segmentation mask, reconstructing the junction points between the trunk and the primary branches. In this case, the junction point between the trunk and the branches has been evaluated at the intersection of the main trunk and branches in the skeleton, taking into account the diameter of the branches. Also Chen et al. [38] uses the same algorithm for plant skeletonization reconstruction, after it emerged to be the best one when compared with the Hilitch thinning algorithm [81] and the Stentifod thinning algorithm [82], joined with buds identification with a YOLOv5 network: the final obtained data is added the depth information, which results in the 3D reconstruction of grapevine plants. The comparison can be seen in Fig. 5.

To reconstruct the skeleton line, Ma et al. [28] employed a Laplacian algorithm [83] and, after downsampling, calculated their eigenvectors by a Fat Point Feature Histogram (FPFH). Then, a Sampling Consistency Algorithm (SAC-IA) was used to achieve the mapping relationship between the skeleton points. Finally, the classical ICP algorithm was applied to refine the initial matrix. After the application of a network for branch-type segmentation. A clustering algorithm, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), was used to estimate the branch count. You et al. [66] employed Terrestrial Laser Scanning (TLS) to reconstruct the tree skeleton. First, DBSCAN is employed to remove all the outliers of the obtained point. Then, the point cloud was divided into slices employing contour planes, obtaining several tree segments by applying DBSCAN. After the point-inversion transformation, the tree skeleton points were obtained from each tree segment. The relationship between skeleton points and the weight of each part was evaluated based on the point-traversal order, organizing the branch hierarchy. Angle consistency was used to optimize the skeleton structure, employing positional and directivity deviations. Similarly, Xu et al. [65] employed terrestrial laser scanning o and Space Colonization Algorithm (SCA) on cherries and begonia trees to obtain their skeletal structure. Since the SCA algorithm alone leads to over-extension of the skeletal trunk of the trees, improvements are made, preprocessing the data with noise filtering and identifying the number of bifurcations using the DBSCAN algorithm.

Liu et al. [64] reconstructed a graph representing the spatial structure of apple branches, and employed a Harris' corner detection algorithm [84] to process binary images to detect the initial candidate point. Then, a 2-level Aggregation Hierarchy Clustering (2HAC) is used to address invalid points in the initial candidates, and finally, skeleton points are obtained. An E-line converging algorithm is proposed for processing to obtain an adjacency matrix. Also, Fu et al. [67] employed a 3D pointcloud, reconstructed from frames taken from different points and merged with the ICP algorithm, as in [85]. Once the full point cloud has been reconstructed, the space colonization algorithm [86] has been employed to reconstruct the structure of the plant, where neighboring space points of the skeleton influence the structure to be formed. The results of this process consist of a directed graph, whose nodes are later classified heuristically to separate the trunk from primary and secondary branches. The obtained model consists of multi-segment cylindrical-connection graphs, considering the decreasing diameters of branches. You et al. [18] aim to reconstruct the skeleton of a UFO cherry tree from a point cloud in the form of a graph defined by edges, nodes, and labels, depending on the type of branches. In this process, the point cloud is first pre-processed, downsampling it into superpoints. A set of graph edges is identified using nearest neighbors, and each edge's likelihood to belong to the skeleton is evaluated. Finally, branch tips are identified. Then, to identify the best skeleton candidate, some a-priori knowledge about the structure of the trees is employed to formulate some constraints used to score the best candidates: topological constraints, label progression, label linearity, trunk-support split, and label-based geometric constraints, which are based on turning angles and growth direction. Also, Fernandes et. al. [26] represent the structure of a grapevine as a graph. Starting with the multiclass segmentation of cordon, canes, and nodes, they identify three sets of connections in the graph structure: the one connecting cordon to canes, the interconnection between elements of the same canes, and the connections between nodes and canes. The algorithm employed is based on an auxiliary matrix containing all main cordon instances. The algorithm identifies the main cordon overlapping a cane mask by checking non-zero values and unique overlaps. If no overlap is found, the cane mask is dilated incrementally until an overlap is detected or a limit is reached.

Several applications have been exploiting several geometry fitting algorithms, approximating the trunk and branches to geometric figures such as cylinders. Elfiky et al. [61] investigated the use of Skeleton-based Geometric (SbG) to reconstruct the 3D skeleton of plants. A point cloud is obtained by merging different perspectives. First, a Laplacian smoothing is performed to extract the skeleton of the tree. Once the graph is obtained, an algorithm based on Geometric Feature Extraction is presented to assign the nodes respectively to the trunk or to the branches. Then, circles are recursively fitted to the point cloud, estimating the best locations for the centers of the circles and the radii, depending also on adjacent circles, using a Random Sample Consensus-based (RANSAC) algorithm and fitting the circles on the 2D plane when the pointcloud is split on horizontal planes. Primary branches are then modeled as cylinders using RANSAC, determining their growing direction and segmenting them to differentiate them from the other branches.

Akbar et al. [87] and Chattopadhyay et al. [62] built a 3D reconstruction of apple trees exploiting the predominantly cylindrical structure of the branches, introducing a new public dataset for benchmarking. In fact, the diameters and the centers of the branches are estimated from a single depth image, filtering lens distortion, and background. Then, the skeleton of the plant is reconstructed [83], the neighboring points of a skeleton segment are found, the dominant direction is found, the neighboring points are projected into the cross section, and a semicircle is fit on it, optimizing the center points [88]. Meideiros et al. [63] employed a laser sensor from a different perspective to measure and model the fruit tree. At first, points are classified in three sets (trunk, junction point, and branches) by a split-and-merge clustering. Then, trunk candidates and junction points are refined by a robust fitting algorithm that models trunks and branches as cylinders, measuring the diameters of the trunk and of the primary branches. In another case, Lin et al. [27] reconstructed the shape of branches and fruits in guava plants under the hypothesis where guava fruit detections could be approximated to a sphere and irregular branches to a finite sequence of 3D cylindrical segments. Within this hypothesis, a RANSAC sphere fitting method [89] and a Principal Component Analysis (PCA) were investigated to reconstruct fruit and branches from the point cloud, once the region of interest is identified via segmentation. Kok et al. [35] obtained the a representation made by a series of skeletal points and the radius by employing a PointNet++ [90] with a shared multilayer perception network [91], an unsupervised method that learns the skeletal representation by utilizing insights from the Medial Axis Transform (MAT) to capture the intrinsic geometric and topological characteristics of the point cloud. To connect skeletal points, a greedy algorithm inspired by [92], connecting the points hierarchically in the bottom-up direction. Moreover, this work also introduces an algorithm to recover branches obscured by vegetation, namely Obscured Branch Recovery (ORB).

Gentilhomme et al. [37] start from a set of nodes of different types and belonging to different branch categories, and connect them to build the grapevine structure. They rely on the predicted branch vector fields to derive the spatial dependencies between nodes. It is highlighted how finding an optimal association by analyzing all the possible connections is an NP-hard problem. Studies about human pose estimation [83] optimize it by creating a spanning tree skeleton and matching adjacent body parts, assuming known node patterns. Here, with unknown nodes and patterns, the approach instead uses shortest-path optimization on a resistivity graph to link nodes to parent nodes, forming a tree structure that's robust and adaptable to varying patterns and missing data. An example of the desired output is shown in Fig. 6. Williams et al. [20] exploit the geometric feature of the grapevine, asserting that the internode cane growth occurs similarly to a straight line. Taking advantage of this, several images from different perspectives were taken, using a Charuco board as a background for calibration, and the 3D structure was recreated by imposing the following constraint. First, the lines between two nodes fitting 95% of their length in cane segmentation were kept. Then, depth data is used to determine if the lines are contiguous in 3D space. Discontinuities are assessed when neighbouring pixels on an image are more than 2 mm apart. Finally, canes were linked to complete the final structure of the vine, eliminating invalid connections.

The performance metrics reported in Table 2 vary widely across methods and datasets, reflecting the diversity of objectives and experimental setups in tree skeletonization and thinning research. While some works report classic geometric accuracy measures such as RMSE, MAE, or diameter estimation errors, others adopt more task-specific metrics such as junction estimation

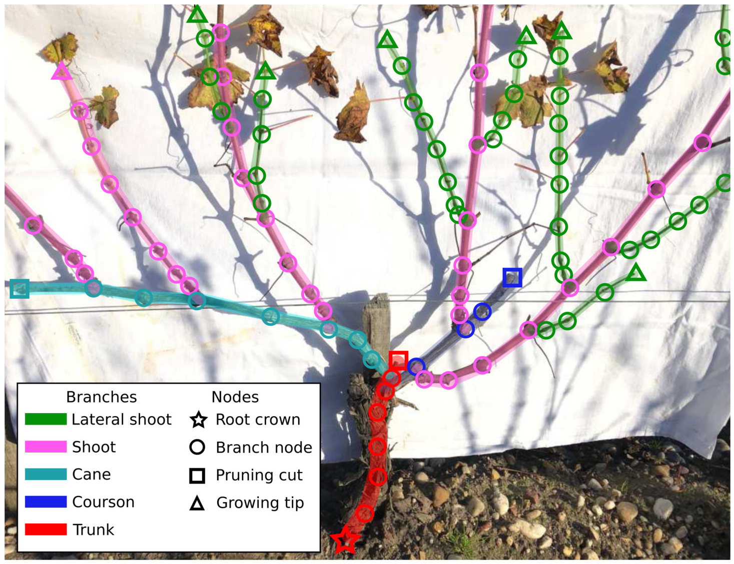


Figure 6: Example of target grapevine plant structure. The types of branches are shown, such as branches and nodes, including different branch types (depicted in different colors), as well as the structure. The dots represent the plant nodes [37].

accuracy, F1 scores in 2D/3D space, or overlap percentages with ground truth. This heterogeneity underscores the need for standardized benchmarks and evaluation protocols to enable a more objective comparison of methods across species and reconstruction goals.

5. Pruning point estimation

In the previous sections, several aspects of the process of identifying and reconstructing a tree's structure have been discussed. In fact, the robustness and reliability of the tree detection and structure estimation process are fundamental to accurately estimate pruning points. Analyzing the works presented in this section, two primary methodologies emerge: the first one relies on the previously obtained structure. In contrast, the second one is based on training algorithms directly on examples provided by experts. Fig. 8 provides an example of the correct estimation of pruning points in different trees.

Generally, the localization of pruning points can be seen as a combinatorial optimization problem, as claimed by the work of Strnad et Kohek [96]. Here, two novel discrete differential evolution (DDE) methods are proposed

Table 3: Overview of pruning point estimation methods.

Crop	Method	Performances	Year	Ref.
Grapevine	Cost function depending on parameters (length, position, angle from head, etc)	-	2017	[17]
Grapevine	Midpoint between nodes. The final point is located above the second node of a cane.	-	2021	[26]
Grapevine	Faster R-CNN for detecting pruning regions,	Recall = (76±15) %, Precision = (82±12)%, F1-score = (78±11)%	2021	[93]
Apple tree	Estimation of branches diameter	Accuracy = 87.2%	2023	[36]
Jujubee	Skeleton analysis: 1/3 of the lenght of the primary branch	-	2023	[67]
Grapevine	Faster R-CNN for detecting pruning regions,	Highest pruning point detection (depending on visibility conditions) = 0.97	2023	[33]
Grapevine	Faster R-CNN for detecting pruning regions (depending on wood type, orientation and visibility)	Highest detection rate: visible intermediate complex spurs = 0.97	2023	[33]
Grapevine	Analysis of the skeleton and the buds position	Accuracy = 82.35%	2024	[38]

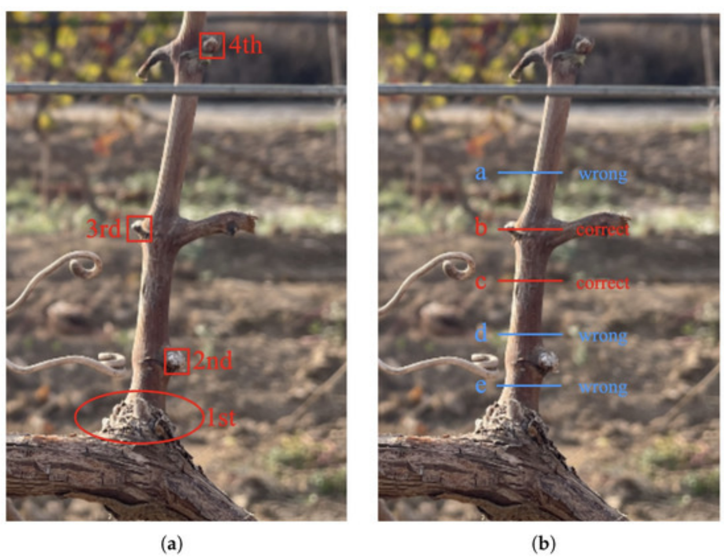
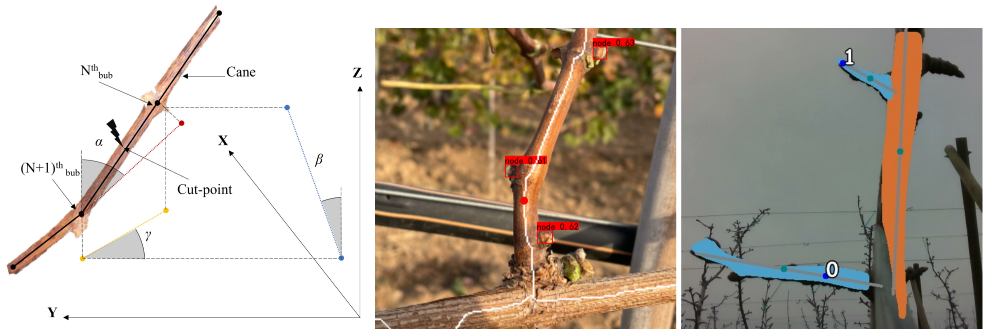


Figure 7: Example of how the pruning point is evaluated depending on the position of the buds on a grapevine. [38]

to optimize simulated tree pruning. Within this problem configuration, the main issue is defining a quantifiable set of metrics to be evaluated and optimized. For this purpose, Williams et al. [20] proposed a new set of quantifiable metrics, named DOLPHIN, to evaluate pruning on grapevine. The set



(a) (b) (c)

Figure 8: (a): cut-point detection: three-dimensional vector projection of a cane section in the YZ (red line – roll angle α), ZX (blue line – pitch angle β) and XY axis (yellow line – yaw angle γ) [94]. (b): correct identification of pruning points on grapevine [38]. (c): correct identification of pruning points on cherry tree [95].

of metrics includes the diameter of the cane, measured between the second and third nodes, the orientation of the cane with respect to the center of the vine's head, length, position on the head, health of the cane, internode length, and node count along the cane. In other cases, the policy for the decision of the branches to be pruned can be more shallow: for instance, Fu et al. [67] considered that in the jujube plant, the primary branches should be shortened each year by approximately 1/3, to reduce nutrient consumption. Therefore, given the previously obtained plant graph, the evaluation of the pruning point is based only on the branch length. Other works consider a more complex vector of features (such as length, position, angle from head, distance below wires, growth position with respect to the trunk) to decide which grapevines' canes to keep. A cost function, represented by a simple linear combination of features, was parametrized using vines labeled by experts measuring pruning quality [97].

In other plants, such as the grapevine, it is necessary to consider a more complex structure, differentiating the various types of branches. In this case, a previous semantic segmentation and branch type identification, with a successive reconstruction, is fundamental. Moreover, bud detection also plays a significant role [98][99], since the position of the cutting point depends strongly on their position, as shown in Fig. 7. Fernandes et al. [26], after reconstructing the structure of the plant, obtain the potential pruning points between two nodes on the same cane or between the bases of two canes growing from the same cane, between the base of a cane and its first node. Similarly, Chen et al. [38], after reconstructing the structure of the grapevine, identify the position of the buds, merged with the depth data, to better identify the pruning point. In fact, the step followed to achieve this result consists of a connection check, where bud pixel coordinates are linked via the mask map to group buds on the same branch, and a key point selection, where the bud with the highest vertical coordinate is chosen as the key point. Starting from the key point, the pruning point is identified on the skeleton image, and its 3D coordinates are recorded. In other works, the pruning points are directly estimated without previously reconstructing the structure of the plant, training the estimation algorithms directly on experts' annotations. For instance, Guadagna et al. [93] trained a Faster R-CNN [41] model for identifying pruning regions in vineyards on a hand-labeled dataset. Then, the obtained result is analyzed according to the type of wood, orientation, and visibility of the pruning point. In a later work, Guadagna et al. [33] fine-tuned a Faster R-CNN with images collected in various vineyards and categorized according by wood type, orientation and visibility. It was emphasized that the accuracy of detecting pruning regions was strongly influenced by visibility.

Table 3 summarizes different approaches to pruning point estimation across several crops, highlighting both qualitative strategies (e.g., structural rules) and quantitative evaluations (e.g., accuracy, recall, F1-score). The reported metrics reflect the diversity in methodologies. While early approaches rely on structural heuristics—such as fixed-length reductions or geometric relationships—more recent methods leverage deep learning models like Faster R-CNN, allowing for the computation of standard performance metrics (e.g., recall, precision, F1-score). However, the table also illustrates a lack of standardized evaluation protocols across studies. Some report raw detection accuracy, others rely on domain-specific heuristics, and still others present detailed classification metrics. This diversity complicates direct comparisons and highlights the need for unified benchmarks and annotated datasets to ensure fair evaluation and reproducibility in pruning detection research.

6. Simulation Environments

In the field of autonomous pruning system research, the use of simulation is of great importance for several reasons. First, it provides a safe environment for evaluating the algorithms' effectiveness, given their destructive nature and the potential risk of compromising the health of the plants. In addition, simulation allows training data generation for the proposed algorithms, which is often time-consuming and difficult to obtain with real-world data. Nevertheless, the generation of a realistic three-dimensional structure of branches is a significant challenge in the field of three-dimensional reconstruction [100][101]. In fact, in order to obtain a more accurate model, the processes influencing the evolution of plants should be taken into account [97], such as the growth against gravity (gravitropism) [102], the search for light (phototropism) [103] and the competition with other plants [104].

One of the earliest simulated environments was created by Corbett-Davies et al. [105], where a system for making pruning decisions using simulated vines is presented. It employs a vector space model where decisions are made based on a vector of characteristics of a cane. The realism of plant models is essential in simulation environments: the accuracy directly affects the system's ability to scale to real-world applications, directly improving the transferability of skills or algorithms to real-world environments. In addition, simulations must provide a variety of scenarios to increase the robustness of training. This diversity helps expose models to various conditions, improving their adaptability and reducing the risk of overfitting, ultimately leading to more reliable real-world performance. To this end, Zhao and Wang [100] used a Space Colonization Algorithm to generate 3D skeletons of cherry trees [106]. The generated models were then segmented with different labels (trunk and branches), and the annotated point cloud was exported for training neural networks for tree structure detection and pruning, and point estimation. Some constraints inherent to SCA were used to obtain more realistic plant structures, including potential growth directions. In addition, Bryson et al. [107] illustrate that artificially generated point clouds from a custom tree simulator can achieve comparable results when used to train detection algorithms. In particular, a PointNet++ network trained on the synthetic dataset outperformed a network trained on a real dataset acquired with a LiDAR, with an increase in IoU of approximately $1-7\%$.

Kolmanivc et al. [97] demonstrated the possibility of a simulation environment by developing an algorithm for automated apple tree pruning in simulation, based on creating artificial tree structures, using EduAPPLE [108] on hierarchical modularity, as many other simulators [109, 110]. In addition, phenomena such as phototropism and competitive growth are included in the simulation,evaluating the light received by buds, as demonstrated in [104], evaluating the impact of pruning in subsequent years, as shown in Fig. 9a. In

(a) (b)

Figure 9: Two possible applications of simulation environments. (a): evolution of the growth of the plant after pruning during several years in apple trees[97]. (b): estimation of the tree crown eliminated given a pruning point on a mango tree [112].

this work, a novel algorithm based on differential evolution [111] optimizes the light received by the inner parts of the plant and minimizes the shading to achieve the optimal result. Subsequently, two pruning templates, namely cylinder and cone, were used to compare the proposed method and expert pruning.

Simulation is also a valuable tool for visualizing immediately the effects of pruning on the structure of the plant, as highlighted by Westling et al. [112], who proposed a LiDaR-based method for suggesting pruning strategies on avocado and mango plants. This method is based on light availability on different plant parts and uses ray tracing to create a voxelized tree structure. Next, the tree was reconstructed as a graph and then explored using the A* algorithm [113]. Using a structural analysis of the tree, the simulation environment determined which portion of the point cloud, corresponding to a plant part, is pruned, based on a single, specific cut point, as shown in Fig. 9b. Finally, novel cut points were proposed by identifying tree components that negatively affect the light distribution. In addition, simulation environments can help improve the understanding of plant structure, as in the work of Qiu et al. [114], where the incompleteness and discontinuity of point clouds were improved using a closed-loop framework [115]. In fact, the issue of PCD incompleteness was observed that modern orchards with high tree density result in high occlusion and low sensor visibility. In addition, adverse weather conditions, such as variable sunlight and strong winds, introduce noise and ghost points. To tackle this problem, a closed-loop framework is proposed, consisting of integrating a data generation process based on real-world observations (Real2Sim) and a subsequent application of these data in a simulation environment (Sim2Real). Three-dimensional apple tree models were generated through a characterization pipeline [115], which was used to train a skeletonization model. Subsequently, a transformer encoder [116] and a joint decoder using a Generative Adversarial Network provide point feature embedding and coarse completion of the point cloud.

7. Building a robotic platform for autonomous pruning

The advancement of autonomous pruning systems for orchards and vineyards requires carefully selecting and designing the manipulator and end effector. It must possess the necessary degrees of freedom, strength, and dexterity to navigate complex and variable environments, such as intertwined branches and irregular plant structures [7]. Furthermore, it must be capable of sophisticated obstacle avoidance, avoiding causing damage to the surrounding vegetation [117]. It is equally crucial to ensure the end effector is designed with the highest level of accuracy to perform precise, clean cuts that promote healthy regrowth. There is a clear gap in the literature regarding the employment of a mechanical actuator and the development of a control system that can plan an efficient path, avoid obstacles, and move the end effector toward the goal. Indeed, numerous similar works concentrate on the autonomous harvesting of fruits in orchards and vineyards [118][119][120][121][122]. In this section we will analyze in detail the study of manipulators, end-effectors and, path planning and control regarding the development of an autonomous pruner.

7.1. Sensors for Robotic Pruning

Autonomous pruning of orchards and vineyards relies on a variety of advanced sensors to ensure precision and efficiency. Different types of cameras have been used to gather datasets of RGB frames for branch detection and segmentation. A consistent group of works utilizes a reflex and high-resolution RGB camera to capture images of branches. This sensor's main

Figure 10: Comparison of natural and artificial light [32].

advantage is its high resolution, which is ideal for identifying branches, even the smaller secondary ones. However, high-resolution images will inevitably result in a lower frame rate in some algorithms. Some studies highlight how the illumination influences the performance of the considered algorithms, comparing the orientation of the sun and the weather conditions. Moreover, in some cases, as shown in Fig. 10, artificial light is employed to obtain a uniform and controlled illumination.

Many works utilize RGB-D cameras, such as the Intel RealSense or Kinect family. These cameras provide an estimation of the depth of the frame, despite the slightly lower adjustable resolution. They are stereo cameras, which is why they are the best choice for this task. This conversion to a point cloud provides meaningful information for segmentation, reconstruction of the branches, identification of the pruning points, and the control system. However, it has some limitations. The limited measured interval and sparse depth resolution may pose a significant problem at certain thresholds due to the invisibility of the smallest branches and buds. Obtaining accurate depth information for small objects, such as branches, presents a considerable challenge for stereo vision systems. In fact, depth images derive from a disparity map, used to infer depth information from two slightly different images (stereo images) captured by cameras placed side by side. The limited number of pixels representing these objects makes it difficult to perform the necessary comparisons between stereo images for reliable depth calculation.

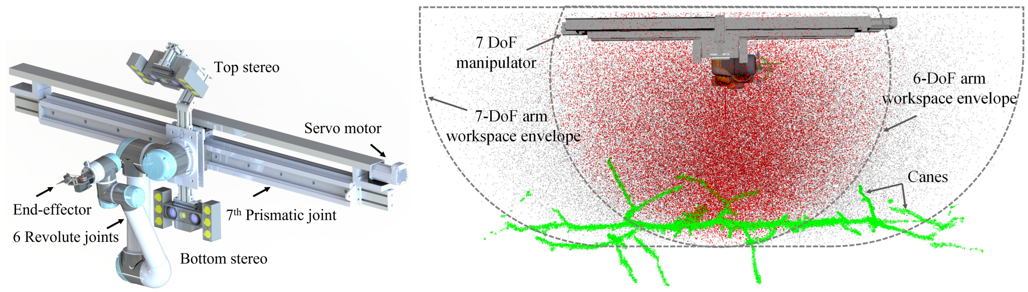
In some cases, traditional stereo vision systems are replaced or complemented with structured light or time-of-flight (ToF) sensors, which are the optimal choice for providing more reliable depth information for small and challenging objects. In other works, such as [28], multiple cameras are used to obtain a more accurate point cloud reconstruction. Specifically, they positioned two Azure Kinect DK cameras at a fixed distance and used an ArUco calibration board to register the point cloud. Another work by Li et al. [123] analyzed the planting mode to better define the way of image acquisition. They identified the challenges of matching images of dormant jujube trees, specifically the lack of local feature regions, weak continuity, and variations in individual morphological structures. They developed a bilateral image-matching method based on three-view geometry constraints to address this. EXIF tags enhanced reconstruction efficiency by implementing a camera self-calibration algorithm based on images. This enabled them to create a 3D point cloud using a Structure from Motion (SfM) reconstruction strategy, combined with a Patch-based Multi-view (PMVS) algorithm.

7.2. Manipulators

The classification of manipulators depends on the number of joints, which corresponds to the number of degrees of freedom (DoF), and the joint type, which is either revolute (R) or prismatic (P). These two variables have a significant impact on the manipulator's key properties, including dexterity, obstacle avoidance, spatial requirements, and the final orientation of the end-effector, which is essential for this task. The literature provides only a limited number of examples of manipulators used for robotic pruning. Table 4 shows a summary of the different manipulators taken into consideration in this paper.

The Universal Robots UR5 manipulator [4], which is a 6 DoF manipulator with 6 prismatic joints, has been successfully employed in several works. For instance, You et al. [124] proved that it can be used to prune cherry trees when equipped with a cutter and an eye-in-hand camera. A multi-pose approach was used to achieve the target cut points successfully. However, it should be noted that the study was carried out in a simplified laboratory setting. The authors were clear that increasing the number of degrees of freedom would be key to improving timing within a sequence of goal poses. Furthermore, placing the robot on a gantry will effectively resolve the visual serving issue. Botterill et al. [17] employed the same manipulator with 6 joints for pruning vine plants. The robot arm is considered the main cause of

[4]https://universal-robots.com/products/ur5-robot/



(a) (b)

Figure 11: (a) CAD rendering of the 7 DoF robot with its components. (b) Top view of the work volume of the 6 and 7 DoF arms. [94]

slow path execution times. They also emphasized that an arm with more degrees of freedom or a different joint configuration will enable fast movements. Furthermore, Silwal et al. [94] demonstrate the limitations of the same 6 DoF robotic arm, which are determined by several factors, including singularities, self-collision, and collision models of the environment. They added a prismatic joint to the base to overcome the abovementioned limitations. This allows the end-effector to achieve any combination of orientation required to reach pruning locations. The result, showing the design of the 7 DoF manipulator and the improved workspace is shown in Fig. 11. Teng et al. [125] employed another commercial manipulator, using a velocity-controlled two-wheel non-holonomic mobile robot, MP-500 (Neobotix GmbH. Co.)[5], and a 7-DoFs robot arm (Franka Emika. Co.)[6]. Each had its own controller. The base localization algorithm employed odometry and twist information of the base central frame.

In other works, the manipulator is designed directly, such as in Zhang et al. [126], where a 5-DoF arm is designed by studying the movement of the human body when pruning. The machine is composed of three main parts: a foundation support with a rotary joint, the machine body with a prismatic joint, and the large arm, which includes the elbow joint, the rotary joint of the forearm, and the forearm. Furthermore, a three-dimensional mathematical simulation model has been created and its workspace was calculated using

[5]`https://neobotix-roboter.de/`
[6]`https://franka.de/`

Table 4: The different configurations of pruning manipulators and relative performances.

Crop	Configuration	Performance	Year	Refe.
Grapevine	6 DoF: 6R	Cut success rate: 66%, cut time 12s per vine plant	2017	[17]
Cherry	6 DoF 6R	Cut success rate: 75%, cut time: 13 s, laboratory setup.	2020	[124]
Apple	6 DoF: 3P + 3R	Cut success rate: 65%,	2020	[127]
Grapevine	7 DoF: 7R	Cut time: 28 s, laboratory setup. The time includes the approach of a non-holonomic robot	2021	[125]
-	2 arms, 6 DoF: 6R	The design has been tested only in simulation.	2021	[128]
Grapevine	7 DoF: 1P + 6R	Pruning accuracy: 87%, cut time: 213 s per vine plant	2022	[94]
Jujube	5 DoF: 4R, 1P	Max position error: 10 mm, success rate: 85.16%, cut time: 29.3 min per Jujubee tree	2022	[126]

the Monte Carlo Method. In this case, the pruning points of the jujube were identified manually by an operator and reproduced by the arm. Zahid et al. [127] developed a 3 DoF end effector, which led to the design of a 3-prismatic DoF Cartesian manipulator, with the first three joints along the x, y, and z-axis respectively. Furthermore, a linear rigid arm was attached to the z-axis, enabling the Cartesian manipulator to be placed outside the tree canopy without interfering with the branches. As an alternative, Lu [128] proposed a dual-arm pruning robot design, providing the 3D model and the Denavit-Hartenberg kinematic model,analyzing, as well as, the trajectory planning in the joint space.

Table 4 shows that manipulator performance varies widely across setups, crops, and testing conditions. Success rates range from 65% to 87%, with higher values often achieved in controlled environments. Cut times also vary significantly—from 12 seconds per vine to over 29 minutes per tree—highlighting the trade-off between speed and precision. Only a few works report positional accuracy, which limits direct comparison. Overall, inconsistent metrics and testing conditions make it difficult to benchmark systems objectively, underlining the need for standardized performance evaluation.

7.3. End-effector Design

Selecting the appropriate end effector is a crucial step in designing an autonomous pruning tool. The design process must address both mechanical and spatial factors, such as size, shape, weight, and maneuverability, alongside the specific requirements of the branches themselves, considering their physical structure, horticultural characteristics, and biological properties [129]. The most common pruning techniques for fruit trees are shear cutting and saw cutting. A complete overview of the methods proposed in the literature is summarized in Table 5.

Given the structural characteristics of the articulated manipulator, Zhang et al. [126] chose a shear mechanism as the end-effector for the jujube pruning manipulator. The robotic arm moves the end-effector mounted on the forearm to the target branch. Once the diagonal photoelectric sensor detects the branch in the scissor opening, the executive motor activates, closing the moving cutter. To cut fruit trees such as cherries, You et al. [124] developed a pneumatically-actuated four-bar linkage with custom ground blades. In initial tests, the cutter proved capable of consistently cutting branches up to 10 mm in diameter near the pivot point of the blades. In a later work [130], they employed an end-effector made up of an integrated battery-operated electric bypass pruner and an eye-in-hand Intel RealSense D435 RGB-D camera, ensuring the top blade is always visible.

The end effector must take into account the variability of the mechanical properties of the canes in vine plants, according to Silwal et al. [94]: vines exhibit a wide variation in the length and diameter of dormant canes depending on nutrients and water. Furthermore, dead samples of vines required a higher force to cut. The end-effector designed in this work fixes one of the handles to a rigid surface, while the weight needed to cut canes of different diameters was evaluated by applying different weights to the other handle. To address the problem of obtaining all the degrees of freedom needed for robotic pruning, Zahid et al. [127] designed a 3 revolute DoF end effector, as the continuation of a Cartesian 3P manipulator, ending with a shear cutter. The diameter of apple branches was considered: usually less than 25 mm. Afront opening of 60 mm was used. The problem of metal support wires was solved by Williams et al. [55], who designed a custom cutting system, called Barracuda, based on small recesses on the bigger teeth of the cutter. It allows the wires not to be cut and to be pushed into the small recesses between teeth, while the thicker cane, which does not fit into the narrow gaps, is cut by the blades.

As an alternative, saw cutting was explored by Botterill et al. [17], where a CNC router mill-end attached to a brushless DC motor was employed. To cut a cane, the robot arm sweeps the cutter through the cane. However, as the cut motion requires a considerable collision-free space, the cutter sometimes fails to make separation cuts.

Table 5 highlights various end-effector designs for robotic pruning, mostly focusing on shear cutting due to its clean and precise cuts. Pneumatic shears

Table 5: The different designs of the end-effectors and the respective performances.

Crop	Mechanism and actuation	Performances	Year	Ref.
Apple	Shear, electromechanical	-	2020	[127]
Cherry	Shear, pneumatic	92% of cutting success	2020	[124]
Grape	Saw (mill), electromechanical	66% of cutting success	2020	[127]
Jujubee	Shear, electromechanical	Max cutting diameter: 12 mm, time per cut: 1 s	2022	[126]
Grape	Shear, electromechanical	-	2022	[94]
Cherry	Shear, electromechanical	-	2023	[130]
Grape	Shears with gaps for wires, electromechanical	97.0% cutting success	2024	[55]

achieve high success rates (e.g., 92% for cherries), while electromechanical shears show good versatility, with quick cut times and effective performance on different crops. Saw-based cutters generally have lower success ($\tilde{6}6\%$) due to collision challenges. Innovations like wire-friendly shears improve cutting success up to 97%, addressing specific vineyard needs. However, many studies lack detailed metrics, and factors like branch variability and environment impact performance. Overall, shear mechanisms remain the preferred choice, but customization is key to improving efficiency and reliability.

7.4. Path planning and control

Once a pruning point is detected and localized, the final step is planning the end effector trajectory to achieve a precise, collision-free cut amid environmental constraints. Teng et al. [125] proposed a hierarchical control strategy for highly redundant robots that uses self-motion redundancy and task-space segmentation to prioritize core tasks while managing secondary constraints; joint velocities are computed incrementally via the Jacobian [131]. Pruning points are identified with a visual perception system [26], and quintic polynomial interpolation [132] ensures smooth paths. The robot's non-holonomic base helps navigate joint limits and singularities.

For motion planning, Rapidly-exploring Random Tree (RRT)-Connect [133] is widely used. Chen et al. [134] introduced goal-biased sampling, adaptive step size, and cubic B-spline smoothing. You et al. [130] combined RRT-Connect with a stop-and-go approach where a mobile base reduces joint motion near pruning points, and a hybrid visual-force controller aligns for minimal disturbance. Silwal et al. [120] implemented a two-phase planner,

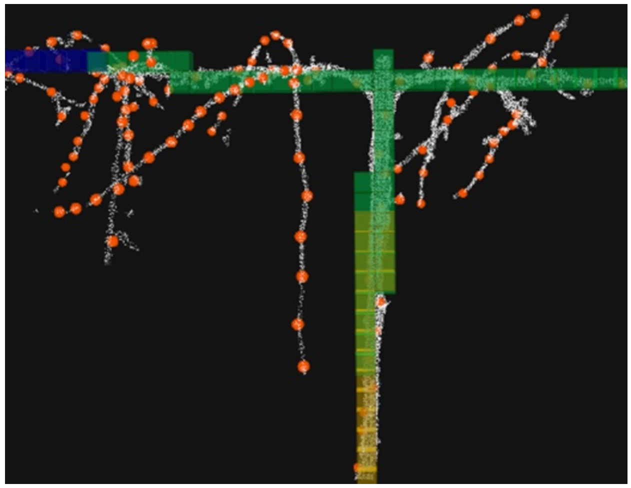


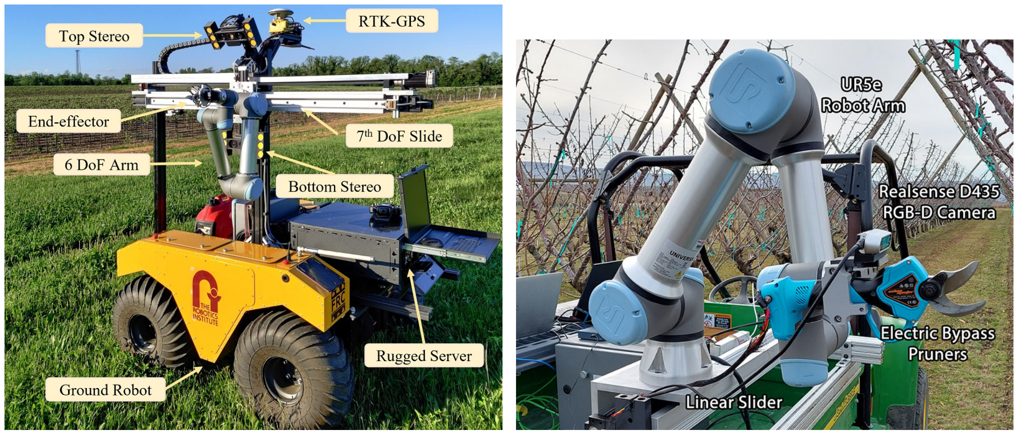
Figure 12: Cordon and trellis wire are considered hard obstacles for motion planning [94].

using RRT-Connect for coarse approach and Cartesian planning for final alignment, with retract-and-reset routines to enhance collision avoidance, as shown in Fig. 12. Botterill et al. [17] dynamically adjusted safety margins and used geometric primitives for robust collision detection.

Recent research explores greater adaptability via model-free control. RL-based approaches [31] use vision to select cut points, optimizing for visibility and alignment, with force-based admittance finalizing the cut. Database-driven strategies [124] discretize the workspace and use TSP solvers with CHOMP [135], STOMP [136], and BIT* [137] for efficient sequencing and online replanning.

8. Complete Robotic Pruning Systems

Several studies have proposed comprehensive end-to-end pruning pipelines that integrate perception, control, and the design of both the robot and its end-effector. These integrated systems are designed to operate autonomously and efficiently across diverse environments. The perception modules typically employ advanced sensors and algorithms to detect and analyze environmental features, allowing for accurate target identification. Control systems rely on sophisticated motion planning and execution software to ensure precise and coordinated movements. Meanwhile, the robot and end-effector designs are tailored specifically for pruning tasks, incorporating characteristics like dexterity, reach, and durability. These complete systems are validated through



(a) (b)

Figure 13: Integrated robotic systems: (a): the system comprises a rover, a 7 DoF robot arm, cutting end-effector, dual stereo cameras and on-board computers [94]; (b): the system comprises a 6 DoF robot arm, a cutting end-effector, a stereo camera and on-board computer [95].

extensive field trials to ensure reliability and effectiveness under real-world conditions.

Silwal et al. [94] introduced a fully autonomous system for dormant season grapevine pruning, as shown in Fig. 13a. Their solution combines a vision system, a ground robot, a kinematically redundant 7-DoF manipulator—achieved by adding a prismatic joint to a standard 6-DoF UR5—and purpose-built pruning algorithms. The end-effector was specifically designed to handle the cutting forces correlated with branch diameter. To reduce the impact of uneven illumination, the robot employs an active light stereo camera system [138] that reconstructs dual point clouds. These are registered using transformation matrices and refined with the ICP algorithm. Bud detection—a critical step in grapevine pruning—is carried out with a Faster R-CNN model. To avoid collisions with trellis structures and wires, a RANSAC algorithm identifies vertical and horizontal line constraints in the 3D vine model.

Botterill et al. [17] presented a mobile robotic platform that straddles vine rows and captures images using trinocular stereo cameras. A 3D model of each vine is constructed via feature matching, triangulation, and incremental bundle adjustment. A 6-DoF robotic arm, guided by an AI system, executes pruning actions based on the reconstructed structure. The robot arm's trajectory is planned in real-time using a rapidly-exploring random tree algorithm, achieving collision-free paths within 1.5 seconds per vine. Although pruning each vine takes around 2 minutes—comparable to human performance—system reliability is limited by the interdependencies of perception, planning, and actuation components.

In a different crop domain, You et al. [95] designed a pruning system for sweet cherry trees grown under the upright fruiting offshoot (UFO) architecture as shown in Fig. 13b. Their fully autonomous platform integrates prior research in perception and manipulation, achieving a 58% cutting success rate during field trials. While not yet robust enough for commercial deployment, this is a pioneering effort in robotic pruning for fruit trees and sets a solid groundwork for future improvements. To address the reality gap between simulation and field conditions, You et al. [31] also developed a simulated pruning environment featuring planar cherry trees and pre-labeled cutting points. Due to the noise in depth data, they opted for a 2D image-based control system. Using a 6-DoF manipulator with an Intel D435 camera, they trained a Generative Adversarial Network (GAN) to segment branches, enabling the visual controller to function reliably across simulated and real images. An actor-critic reinforcement learning algorithm then generated control signals for branch alignment. Upon contact, control transitions to an interaction-based controller that adjusts the cutter's pivot to minimize forces during cutting.

Lastly, Williams et al. [55] developed Archie Jnr, a fully operational pruning platform designed for a controlled environment under an arched structure (see Fig. 14). The system uses two UR5 robotic arms to scan plants and conduct panoptic segmentation of nodes and canes for precise 3D reconstruction. A Charuco board is employed for accurate image-depth registration. The manipulators incorporate custom mechanisms that enhance agility, making this system especially suitable for precision pruning in high-control settings.

Collectively, these studies highlight the critical role of system integration in achieving autonomous pruning. Robust perception methods—particularly 3D reconstruction and deep learning-based segmentation—enable accurate identification of pruning targets. Control strategies range from traditional trajectory planning to learning-based policies and hybrid force-vision controllers. Manipulator designs are adapted to the physical demands of the pruning task, while platform-level decisions reflect a balance between robustness in the field and operational simplification in controlled environments.

The comparison in Table 6 highlights the progress and remaining chal-

Figure 14: Archie Jnr robotic platform, proposed by [20].

Table 6: Comparison of the key metrics of complete autonomous pruning systems.

Crop	Field trial realisms	Cut success rate	Cut time	Year	Ref.
Vineyard	Real field	-	115 sec/vine	2017	[17]
Vineyard	Real field	87 % success	213 sec/vine	2022	[94]
Cherry tree	Sim + Real field	58 % success	284 s scan + 35.1 s/cut	2022	[95]
Vineyard	Real field	71% success	-	2024	[55]

lenges in developing fully autonomous pruning systems. While vineyard-focused systems, such as those by Silwal et al.[94] and Williams et al.[55], demonstrate relatively high success rates, cutting speed remains slower than that of human labor. The cherry tree system by You et al. represents a pioneering effort in fruit-tree pruning but still suffers from limited success rates, underscoring the difficulty of generalizing perception and control strategies across diverse crop architectures. Overall, these results emphasize the trade-off between field robustness and operational efficiency, pointing to the need for further improvements in perception reliability, motion planning speed, and manipulator design.

9. Discussion: challenges, solutions and future scenarios

The previous chapters have provided a broad overview of the existing solutions for developing an autonomous pruning system. Given the diversity of tasks involved and the range of expertise required to develop each component of such a system, it has become evident that a modular approach to robot design is fundamental for achieving optimal results. As demonstrated throughout this paper, most existing works focus on narrow aspects of the problem, such as branch detection, skeletonization, pruning point estimation, or end-effector design. Each component demands highly specialized skills spanning computer vision, agronomy, mechanical engineering, and control systems. Moreover, a modular platform can be repurposed for various agricultural tasks, such as health monitoring, fruit counting, and harvesting. For instance, a module designed for branch detection can also be used for obstacle avoidance during harvesting. At the same time, the manipulator structure may remain the same, with only the end-effector swapped for the target task.

However, several limitations have emerged from the current state of the art. A primary barrier, even before economic considerations, is the lack of deployment-ready systems. Most solutions remain at the prototype stage, tested only in controlled environments or on a limited number of plants, with robustness to real-world variability—such as plant morphology, illumination, weather, or soil conditions—still largely unproven. This hinders commercial adoption. Reliable performance under continuous operation and the high cost of robotic platforms further limit large-scale deployment. Sensors remain a major cost driver, although recent advances in AI suggest that similar performance may be achieved with fewer, lower-cost sensors. Long-term economic viability will also depend on standardization, shared maintenance infrastructures, and farmer acceptance. In regions with lower labor costs, adoption may remain limited unless robots provide added value, for example, by combining pruning with monitoring or harvesting tasks.

9.1. Challenges and limitations of the perception pipeline

Most of the proposed algorithms for branch detection, segmentation, 3D reconstruction, and pruning point estimation rely on learning-based methods. From the literature analysis of Section 3, learning-based approaches represent the most promising due to their generalization capability in different conditions. In fact, classical computer vision methods appear to be excessively scenario-dependent, without being able to generalize even in different light settings without proper tuning. Among the proposed solutions, convolutional neural networks stand as the best solution available: thanks to their spatial-invariant feature extraction and their capability of learning invariant features in different conditions. While this trend is globally evident, a significant drawback of such approaches is the requirement for large, labeled datasets for effective training. Constructing a structured and diverse dataset is labor-intensive, particularly due to the complexity of manually labeling pruning-relevant features. The lack of shared public datasets and standardized annotation protocols further limits reproducibility and benchmarking. Collaborative initiatives to build open datasets and simulation environments could significantly accelerate progress.

Several works have attempted to address this challenge by developing realistic simulation environments to generate synthetic training data. These environments can simulate plant structures and growth patterns based on agronomic models, allowing for the automatic generation of diverse and labeled training data. This approach not only facilitates training of detection and decision algorithms but also enables analysis of the short- and long-term consequences of different pruning strategies. Models trained in simulation and then fine-tuned with real-world data have shown comparable performance to those trained entirely on real datasets. This is a significant advantage in pruning applications, which are inherently destructive, making real-world testing both costly and risky in terms of yield loss.

The estimation of the pruning point is a crucial step in the perception pipeline. Most studies first reconstruct the tree structure and then estimate pruning points using heuristics based on branch characteristics (e.g., diameter, orientation, bud position). An alternative is direct cut estimation from sensor data (images or point clouds), but both approaches suffer from limited generalization. Accurate pruning requires an understanding of plant structure, as human pruners infer cuts from structural relationships; correlating reconstructed structure with additional features can improve estimation but adds complexity. Direct estimation with deep vision algorithms also faces scalability issues. Moreover, perception and manipulation are often treated separately, causing error propagation—small reconstruction errors can significantly affect cut placement.

9.2. Challenges and limitation of manipulations

Another critical limitation is operational speed. Current systems require one minute per branch to two minutes per vine plant (typically about eight cuts), which is an order of magnitude slower than manual pruning. As noted by [17], adopting faster robotic arms may offer a short-term solution, but real progress depends on improving detection and control algorithms, which remain computationally demanding. Two strategies are typically pursued: (i) optimizing algorithms through lighter-weight neural networks and more efficient pipelines to reduce computational overhead, and (ii) adopting more powerful on-board hardware to enable real-time inference and faster actuation without compromising accuracy.

Mechanical design also plays a crucial role. Integrated systems, such as [94], combine custom manipulators and end-effectors with autonomous navigation (GPS-RTK) and active lighting to enhance perception accuracy and motion planning. Similarly,[31] incorporates a sim-to-real approach using GAN-generated synthetic data[139], reducing the performance gap between controlled and real-world conditions. Despite these advances, manipulators still face challenges related to robustness under outdoor variability, interaction with deformable branches, and continuous operation in harsh agricultural environments. Future work should also explore modular manipulator designs with interchangeable end-effectors, allowing the same robotic platform to perform multiple tasks such as pruning, health monitoring, and harvesting, improving overall economic viability.

9.3. Future Scenarios

One promising direction involves developing autonomous pruning machines prioritizing robustness and precision over compactness. For instance, the system proposed in [17] operates in a partially enclosed, controlled environment, where uniform lighting and backgrounds enhance segmentation accuracy. This setup also supports the use of high-performance sensors, which can significantly improve depth perception—a key requirement for precise control of robotic arms and end-effectors. While lower-cost sensors are attractive from an economic standpoint, they may compromise accuracy, potentially leading to suboptimal cuts or even crop loss. Future research could focus on hybrid solutions that balance hardware cost with intelligent compensation through software. Equally important will be the development of standardized hardware and software platforms, shared datasets, and benchmarking protocols, which could accelerate the transition from research prototypes to commercially deployable systems.

Emerging machine learning paradigms offer new approaches that could bypass traditional step-wise pipelines. Instead of detecting, segmenting, and analyzing plant structure in sequence, it may be possible to learn a direct mapping from sensory input to pruning actions. Reinforcement Learning (RL), for instance, allows an agent to learn optimal pruning policies through trial and error in a simulation environment, where pruning consequences can be integrated into the reward function. While promising for structured crops like apples, mangoes, and cherries, RL alone may be insufficient for grapevines due to the complex, region-specific pruning rules they require. Hybrid approaches that combine RL with agronomic rule-based constraints or active perception strategies could improve reliability, especially in unstructured outdoor conditions.

To address domain-specific nuances, Learning from Demonstration (LfD) provides a powerful paradigm. By observing expert behavior, a pruning policy can be fine-tuned for both structural accuracy and agronomic validity. For instance, [140] shows how demonstration-based approaches can avoid motion discontinuities in complex tasks. Complementary technologies like augmented and virtual reality could support expert-guided training, allowing humans to annotate or demonstrate pruning strategies in an immersive, simulated environment. An example is [141], where an AR mobile app provides pruning suggestions to non-experts using AI-generated recommendations.

Another promising avenue is imitation learning, which has succeeded in tasks like pepper harvesting [142]. There, a visuomotor policy trained via demonstrations enabled a handheld shear-gripper to adapt to variable field conditions and crop types. Applying similar strategies to pruning could enhance adaptability and robustness, especially in unstructured outdoor environments. However, application in pruning remains rare, representing an untapped area for future investigation. In the longer term, multi-robot systems or human-robot collaborative teams could enable large-scale operations, while adaptive manipulators capable of modifying cut strategies in real time may further increase efficiency.

While the potential of AI-driven pruning is evident, future systems must integrate precision hardware, data-efficient learning, and domain-specific agronomic knowledge. Developing scalable, adaptable, and accurate solutions will likely require multi-disciplinary collaboration spanning robotics, machine learning, agronomy, and human-computer interaction. Equally crucial will be considering socio-economic factors, including ease of integration into existing agricultural workflows, training support for operators, and economic models (e.g., leasing or multi-tasking robots) that make adoption viable for farmers.

10. Conclusions

This work has illustrated that autonomous pruning remains an open and highly relevant research challenge, demanding progress in robotics, perception, machine learning, and agronomic integration. While modern orchard training systems and agricultural machinery have streamlined many tasks, such as spraying, fertilization, and green pruning, dormant and selective pruning remain among the most labor-intensive and skill-dependent operations. These operations are crucial, as they directly affect the yield quality and quantity of the following season.

A core requirement for reliable autonomous pruning is the accurate understanding of plant structure, which is inherently unstructured, variable, and occluded. Traditional image processing approaches have shown limited performance in complex outdoor scenes. In contrast, deep learning-based methods, particularly convolutional neural networks, have significantly improved the performance of branch detection and segmentation. These networks allow not only for detecting branches but also for classifying them hierarchically, which is critical for subsequent 3D reconstruction and pruning logic. Detecting buds is another essential capability, as pruning rules often depend on the position and condition of buds. However, their small size and inconsistent visibility make detection challenging, typically requiring high-resolution cameras and precise segmentation. The next logical step after segmentation is skeletonization, which helps define the structural hierarchy and spatial configuration of the plant. This, in turn, supports pruning point estimation, for which two main methodological categories exist: rule-based methods that extract pruning points from a reconstructed structure using deterministic logic learning-based approaches that estimate pruning points directly from raw data.

Beyond perception, the success of autonomous pruning also hinges on the robot's physical hardware, including the manipulator and end-effector. While perception determines where to prune, the robot must also be capable of acting safely and precisely. Recent studies have developed custom electromechanical end-effectors capable of adapting to varying branch sizes and minimizing damage. Similarly, the choice of manipulator, such as its reach, degrees of freedom, and control workspace, directly affects system flexibility and responsiveness in field conditions. Speed remains a bottleneck; current systems are often an order of magnitude slower than manual pruning, highlighting the need for faster arms and more efficient planning algorithms. Finally, the integration of all these components—perception, planning, actuation, and platform mobility—into fully functional, field-ready pruning robots is still in early stages. Some examples demonstrate the feasibility of such systems in specific scenarios, combining semantic segmentation, GPS navigation, and simulation-based training. These systems highlight both the potential and the complexity of real-world deployment. Notably, the modular design of these robots allows them to be adapted for other agricultural tasks, such as health monitoring or harvesting, thereby improving their return on investment.

In summary, although many technical components have shown promise in isolation, fully integrated, generalizable autonomous pruning systems require further development. Future research must focus on bridging the gap between lab prototypes and robust, scalable platforms, emphasizing generalization, simulation-based training, and human-in-the-loop learning methods. Ultimately, the path forward lies in a multidisciplinary collaboration that combines agronomic expertise with cutting-edge robotics and AI, unlocking a new era of intelligent, sustainable agriculture.

11. Acknowledgements

This work has been developed with the contribution of the Politecnico di Torino Interdepartmental Centre for Service Robotics (PIC4SeR[7]).

References

[1] S. P. Galinato, A. Kendall, C. A. Miles, Costs and profitability for mechanized pruning and harvest in two cider apple orchard systems, HortTechnology 32 (3) (2022) 275–287.

[2] European Commission, The fruit and vegetable sector in the eu - a statistical overview, Eurostat (2024).

[7] `https://pic4ser.polito.it/`

[3] European Commission, Farmers and the agricultural labour force - statistics, Eurostat (2022).

[4] L. Calvin, P. Martin, S. Simnitt, Adjusting to higher labor costs in selected us fresh fruit and vegetable industries (2022).

[5] G. Allegro, R. Martelli, G. Valentini, C. Pastore, R. Mazzoleni, F. Pezzi, I. Filippetti, Effects of mechanical winter pruning on vine performances and management costs in a trebbiano romagnolo vineyard: A five-year study, Horticulturae 9 (1) (2022) 21.

[6] L. He, J. Schupp, Sensing and Automation in Pruning of Apple Trees: A Review, Agronomy 8 (10) (2018) 211. doi:10.3390/agronomy8100211.
URL http://www.mdpi.com/2073-4395/8/10/211

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

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

[9] L. Qiang, C. Jianrong, L. Bin, D. Lie, 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 7 (2) (2014) 115–121.

[10] W. Ji, Z. Qian, B. Xu, Y. Tao, D. Zhao, S. Ding, Apple tree branch segmentation from images with small gray-level difference for agricultural harvesting robot, Optik 127 (23) (2016) 11173–11182.

[11] N. Shalal, T. Low, C. McCarthy, N. Hancock, Orchard mapping and mobile robot localisation using on-board camera and laser scanner data fusion–part a: Tree detection, Computers and Electronics in Agriculture 119 (2015) 254–266.

[12] M. Gao, T.-F. Lu, Image processing and analysis for autonomous grapevine pruning, in: 2006 international conference on mechatronics and automation, IEEE, 2006, pp. 922–927.

[13] N. McFarlane, B. Tisseyre, C. Sinfort, R. Tillett, F. Sevila, Image analysis for pruning of long wood grape vines, Journal of agricultural engineering research 66 (2) (1997) 111–119.

[14] J. Zhang, L. He, M. Karkee, Q. Zhang, X. Zhang, 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 155 (2018) 386–393.

[15] S. Amatya, M. Karkee, Q. Zhang, M. D. Whiting, Automated Detection of Branch Shaking Locations for Robotic Cherry Harvesting Using Machine Vision, Robotics 6 (4) (2017) 31. doi:10.3390/robotics6040031.
URL http://www.mdpi.com/2218-6581/6/4/31

[16] S. Amatya, M. Karkee, Integration of visible branch sections and cherry clusters for detecting cherry tree branches in dense foliage canopies, Biosystems Engineering 149 (2016) 72–81.

[17] T. Botterill, S. Paulin, R. Green, S. Williams, J. Lin, V. Saxton, S. Mills, X. Chen, S. Corbett-Davies, A Robot System for Pruning Grape Vines, Journal of Field Robotics 34 (6) (2017) 1100–1122, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rob.21680. doi:10.1002/rob.21680.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21680

[18] A. You, C. Grimm, A. Silwal, J. R. Davidson, Semantics-guided skeletonization of upright fruiting offshoot trees for robotic pruning, Computers and Electronics in Agriculture 192 (2022) 106622.

[19] F. Oliveira, D. Q. da Silva, V. Filipe, T. M. Pinho, M. Cunha, J. B. Cunha, F. N. Dos Santos, Enhancing grapevine node detection to support pruning automation: Leveraging state-of-the-art yolo detection models for 2d image analysis, Sensors 24 (21) (2024) 6774.

[20] H. Williams, D. Smith, J. Shahabi, T. Gee, M. Nejati, B. McGuinness, K. Black, J. Tobias, R. Jangali, H. Lim, et al., Modelling wine grapevines for autonomous robotic cane pruning, biosystems engineering 235 (2023) 31–49.

[21] Y. Majeed, J. Zhang, X. Zhang, L. Fu, M. Karkee, Q. Zhang, M. D. Whiting, Apple Tree Trunk and Branch Segmentation for Automatic Trellis Training Using Convolutional Neural Network Based Semantic Segmentation, IFAC-PapersOnLine 51 (17) (2018) 75–80. doi:10.1016/j.ifacol.2018.08.064.
URL https://linkinghub.elsevier.com/retrieve/pii/S2405896318311807

[22] Y. Majeed, M. Karkee, Q. Zhang, L. Fu, M. D. Whiting, A study on the detection of visible parts of cordons using deep learning networks for automated green shoot thinning in vineyards, IFAC-PapersOnLine 52 (30) (2019) 82–86, 6th IFAC Conference on Sensing, Control and Automation Technologies for Agriculture AGRICONTROL 2019. doi:https://doi.org/10.1016/j.ifacol.2019.12.501.
URL https://www.sciencedirect.com/science/article/pii/S2405896319324176

[23] J. Wu, G. Yang, H. Yang, Y. Zhu, Z. Li, L. Lei, C. Zhao, Extracting apple tree crown information from remote imagery using deep learning, Computers and Electronics in Agriculture 174 (2020) 105504. doi:10.1016/j.compag.2020.105504.
URL https://linkinghub.elsevier.com/retrieve/pii/S0168169920301605

[24] Y. Majeed, J. Zhang, X. Zhang, L. Fu, M. Karkee, Q. Zhang, M. D. Whiting, Deep learning based segmentation for automated training of apple trees on trellis wires, Computers and Electronics in Agriculture 170 (2020) 105277.

[25] Y. Majeed, M. Karkee, Q. Zhang, L. Fu, M. D. Whiting, Determining grapevine cordon shape for automated green shoot thinning using semantic segmentation-based deep learning networks, Computers and Electronics in Agriculture 171 (2020) 105308.

[26] M. Fernandes, A. Scaldaferri, G. Fiameni, T. Teng, M. Gatti, S. Poni, C. Semini, D. Caldwell, 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), 2021, pp. 13–18, iSSN: 2642-6633. doi:10.1109/CYBER53097.2021.9588303.

[27] G. Lin, Y. Tang, X. Zou, C. Wang, Three-dimensional reconstruction of guava fruits and branches using instance segmentation and geometry analysis, Computers and Electronics in Agriculture 184 (2021) 106107. doi:10.1016/j.compag.2021.106107.
URL https://linkinghub.elsevier.com/retrieve/pii/S0168169921001253

[28] B. Ma, J. Du, L. Wang, H. Jiang, 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 190 (2021) 106484. doi:10.1016/j.compag.2021.106484.
URL https://linkinghub.elsevier.com/retrieve/pii/S0168169921005019

[29] S. Tong, Y. Yue, W. Li, Y. Wang, F. Kang, C. Feng, Branch Identification and Junction Points Location for Apple Trees Based on Deep Learning, Remote Sensing 14 (18) (2022) 4495. doi:10.3390/rs14184495.
URL https://www.mdpi.com/2072-4292/14/18/4495

[30] A. You, C. Grimm, J. R. Davidson, Optical flow-based branch segmentation for complex orchard environments, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2022, pp. 9180–9186.

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

[32] D. Borrenpohl, M. Karkee, Automated pruning decisions in dormant

sweet cherry canopies using instance segmentation, Computers and Electronics in Agriculture 207 (2023) 107716.

[33] P. Guadagna, M. Fernandes, F. Chen, A. Santamaria, T. Teng, T. Frioni, D. Caldwell, S. Poni, C. Semini, M. Gatti, Using deep learning for pruning region detection and plant organ segmentation in dormant spur-pruned grapevines, Precision Agriculture 24 (4) (2023) 1547–1569.

[34] Z. Chen, K. Granland, R. Newbury, C. Chen, Hob-cnn: Hallucination of occluded branches with a convolutional neural network for 2d fruit trees, Smart Agricultural Technology 3 (2023) 100096.

[35] E. Kok, X. Wang, C. Chen, Obscured tree branches segmentation and 3d reconstruction using deep learning and geometrical constraints, Computers and electronics in agriculture 210 (2023) 107884.

[36] S. Tong, J. Zhang, W. Li, Y. Wang, F. Kang, An image-based system for locating pruning points in apple trees using instance segmentation and rgb-d images, Biosystems Engineering 236 (2023) 277–286.

[37] T. Gentilhomme, M. Villamizar, J. Corre, J.-M. Odobez, Towards smart pruning: Vinet, a deep-learning approach for grapevine structure estimation, Computers and Electronics in Agriculture 207 (2023) 107736. `doi:https://doi.org/10.1016/j.compag.2023.107736`.
URL `https://www.sciencedirect.com/science/article/pii/S0168169923001242`

[38] Z. Chen, Y. Wang, S. Tong, C. Chen, F. Kang, Grapevine Branch Recognition and Pruning Point Localization Technology Based on Image Processing, Applied Sciences 14 (8) (2024) 3327, number: 8 Publisher: Multidisciplinary Digital Publishing Institute. `doi:10.3390/app14083327`.
URL `https://www.mdpi.com/2076-3417/14/8/3327`

[39] R. Girshick, J. Donahue, T. Darrell, J. Malik, Region-based convolutional networks for accurate object detection and segmentation, IEEE transactions on pattern analysis and machine intelligence 38 (1) (2015) 142–158.

[40] R. Girshick, Fast r-cnn, in: Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448.

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

[42] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask r-cnn, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980–2988. doi:10.1109/ICCV.2017.322.

[43] Z. Cai, N. Vasconcelos, Cascade r-cnn: High quality object detection and instance segmentation, IEEE transactions on pattern analysis and machine intelligence 43 (5) (2019) 1483–1498.

[44] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 10012–10022.

[45] X. Cheng, Y. Zhong, M. Harandi, Y. Dai, X. Chang, H. Li, T. Drummond, Z. Ge, Hierarchical neural architecture search for deep stereo matching, Advances in neural information processing systems 33 (2020) 22158–22169.

[46] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing, Cham, 2015, pp. 234–241.

[47] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems 25 (2012).

[48] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: Redesigning skip connections to exploit multiscale features in image segmentation, IEEE transactions on medical imaging 39 (6) (2019) 1856–1867.

[49] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE transactions on pattern analysis and machine intelligence 39 (12) (2017) 2481–2495.

[50] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).

[51] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.

[52] X. Wang, R. Zhang, T. Kong, L. Li, C. Shen, Solov2: Dynamic and fast instance segmentation, Advances in Neural information processing systems 33 (2020) 17721–17732.

[53] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid scene parsing network, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2881–2890.

[54] A. Newell, K. Yang, J. Deng, Stacked hourglass networks for human pose estimation, in: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VIII 14, Springer, 2016, pp. 483–499.

[55] H. Williams, D. Smith, J. Shahabi, T. Gee, A. Qureshi, B. McGuinness, S. Harvey, C. Downes, R. Jangali, K. Black, et al., Archie jnr: A robotic platform for autonomous cane pruning of grapevines, in: 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2024, pp. 11736–11743.

[56] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, R. Girshick, Detectron2, `https://github.com/facebookresearch/detectron2` (2019).

[57] A. Kirillov, K. He, R. Girshick, C. Rother, P. Dollár, Panoptic segmentation, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 9404–9413.

[58] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, T. Brox, Flownet 2.0: Evolution of optical flow estimation with deep networks,

in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2462–2470.

[59] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.

[60] L. Landrieu, M. Simonovsky, Large-scale point cloud semantic segmentation with superpoint graphs, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4558–4567.

[61] N. M. Elfiky, S. A. Akbar, J. Sun, J. Park, A. Kak, Automation of dormant pruning in specialty crop production: An adaptive framework for automatic reconstruction and modeling of apple trees, in: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2015, pp. 65–73.

[62] S. Chattopadhyay, S. A. Akbar, N. M. Elfiky, H. Medeiros, A. Kak, Measuring and modeling apple trees using time-of-flight data for automation of dormant pruning applications, in: 2016 IEEE Winter conference on applications of computer vision (WACV), IEEE, 2016, pp. 1–9.

[63] H. Medeiros, D. Kim, J. Sun, H. Seshadri, S. A. Akbar, N. M. Elfiky, J. Park, Modeling Dormant Fruit Trees for Agricultural Automation, Journal of Field Robotics 34 (7) (2017) 1203–1224. doi:10.1002/rob.21679.
URL https://onlinelibrary.wiley.com/doi/10.1002/rob.21679

[64] S. Liu, J. Yao, H. Li, C. Qiu, R. Liu, Research on 3D skeletal model extraction algorithm of branch based on SR4000, Journal of Physics: Conference Series 1237 (2) (2019) 022059. doi:10.1088/1742-6596/1237/2/022059.
URL https://iopscience.iop.org/article/10.1088/1742-6596/1237/2/022059

[65] Y. Xu, C. Hu, Y. Xie, An improved space colonization algorithm with dbscan clustering for a single tree skeleton extraction, International Journal of Remote Sensing 43 (10) (2022) 3692–3713.

[66] L. You, Y. Sun, Y. Liu, X. Chang, J. Jiang, Y. Feng, X. Song, Tree skeletonization with dbscan clustering using terrestrial laser scanning data, Forests 14 (8) (2023) 1525.

[67] Y. Fu, Y. Xia, H. Zhang, M. Fu, Y. Wang, W. Fu, C. Shen, Skeleton extraction and pruning point identification of jujube tree for dormant pruning using space colonization algorithm, Frontiers in Plant Science 13 (2023) 1103794.

[68] X. Li, B. Liu, Y. Shi, M. Xiong, D. Ren, L. Wu, X. Zou, Efficient three-dimensional reconstruction and skeleton extraction for intelligent pruning of fruit trees, Computers and Electronics in Agriculture 227 (2024) 109554.

[69] J. Dukić, P. Pejić, A. Bošnjak, E. K. Nyarko, Branch-a labeled dataset of rgb-d images and 3d models for autonomous tree pruning, in: 2024 International Conference on Smart Systems and Technologies (SST), IEEE, 2024, pp. 57–64.

[70] P. J. Besl, N. D. McKay, Method for registration of 3-d shapes, in: Sensor fusion IV: control paradigms and data structures, Vol. 1611, Spie, 1992, pp. 586–606.

[71] T. Su, W. Wang, H. Liu, Z. Liu, X. Li, Z. Jia, L. Zhou, Z. Song, M. Ding, A. Cui, An adaptive and rapid 3d delaunay triangulation for randomly distributed point cloud data, The Visual Computer (2022) 1–25.

[72] J. Straub, D. Reiser, N. Lüling, A. Stana, H. W. Griepentrog, Approach for graph-based individual branch modelling of meadow orchard trees with 3d point clouds, Precision Agriculture 23 (6) (2022) 1967–1982.

[73] G. Hu, Z. Zhou, J. Cao, H. Huang, Non-linear calibration optimisation based on the levenberg–marquardt algorithm, IET Image Processing 14 (7) (2020) 1402–1414.

[74] A. Tabb, H. Medeiros, A robotic vision system to measure tree traits, in: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2017, pp. 6005–6012.

[75] X. Dong, W.-Y. Kim, Z. Yu, J.-Y. Oh, R. Ehsani, K.-H. Lee, Improved voxel-based volume estimation and pruning severity mapping of apple trees during the pruning period, Computers and Electronics in Agriculture 219 (2024) 108834.

[76] S.-C. Zhu, D. Mumford, et al., A stochastic grammar of images, Foundations and Trends® in Computer Graphics and Vision 2 (4) (2007) 259–362.

[77] T. Botterill, R. Green, S. Mills, Finding a vine's structure by bottom-up parsing of cane edges, in: 2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013), IEEE, 2013, pp. 112–117.

[78] U. Ramer, An iterative procedure for the polygonal approximation of plane curves, Computer graphics and image processing 1 (3) (1972) 244–256.

[79] T. Y. Zhang, C. Y. Suen, A fast parallel algorithm for thinning digital patterns, Commun. ACM 27 (3) (1984) 236–239. `doi:10.1145/357994.358023`.
URL `https://doi.org/10.1145/357994.358023`

[80] H. Cuevas-Velasquez, A.-J. Gallego, R. B. Fisher, Segmentation and 3d reconstruction of rose plants from stereoscopic images, Computers and Electronics in Agriculture 171 (2020) 105296. `doi:https://doi.org/10.1016/j.compag.2020.105296`.
URL `https://www.sciencedirect.com/science/article/pii/S0168169919323919`

[81] C. J. Hilitch, Linear skeletons from square cupboards (1969).

[82] A. Martin, S. Tosunoglu, Image processing techniques for machine vision, Miami, Florida (2000) 1–9.

[83] J. Cao, A. Tagliasacchi, M. Olson, H. Zhang, Z. Su, Point cloud skeletons via laplacian based contraction, in: 2010 Shape Modeling International Conference, 2010, pp. 187–197. `doi:10.1109/SMI.2010.25`.

[84] C. Harris, M. Stephens, et al., A combined corner and edge detector, in: Alvey vision conference, Vol. 15, Citeseer, 1988, pp. 10–5244.

[85] Y. Fu, C. Li, J. Zhu, B. Wang, B. Zhang, W. Fu, Three-dimensional model construction method and experiment of jujube tree point cloud using alpha-shape algorithm, Trans. Chin. Soc Agric. Eng 36 (22) (2020) 214–221.

[86] A. Runions, M. Fuhrer, B. Lane, P. Federl, A.-G. Rolland-Lagan, P. Prusinkiewicz, Modeling and visualization of leaf venation patterns, in: ACM SIGGRAPH 2005 Papers, 2005, pp. 702–711.

[87] S. A. Akbar, S. Chattopadhyay, N. M. Elfiky, A. Kak, A novel benchmark rgbd dataset for dormant apple trees and its application to automatic pruning, in: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2016, pp. 81–88.

[88] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, Journal of the society for Industrial and Applied Mathematics 11 (2) (1963) 431–441.

[89] R. Schnabel, R. Wahl, R. Klein, Efficient ransac for point-cloud shape detection, in: Computer graphics forum, Vol. 26, Wiley Online Library, 2007, pp. 214–226.

[90] C. R. Qi, L. Yi, H. Su, L. J. Guibas, Pointnet++: Deep hierarchical feature learning on point sets in a metric space, Advances in neural information processing systems 30 (2017).

[91] C. Lin, C. Li, Y. Liu, N. Chen, Y.-K. Choi, W. Wang, Point2skeleton: Learning skeletal representations from point clouds, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 4277–4286.

[92] F. Yandun, A. Silwal, G. Kantor, Visual 3d reconstruction and dynamic simulation of fruit trees for robotic manipulation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 54–55.

[93] P. Guadagna, T. Frioni, F. Chen, A. I. Delmonte, T. Teng, M. Fernandes, A. Scaldaferri, C. Semini, S. Poni, M. Gatti, 16. fine-tuning and testing of a deep learning algorithm for pruning regions detection in spur-pruned grapevines, in: Precision agriculture'21, Wageningen Academic, 2021, pp. 147–153.

[94] A. Silwal, F. Yandun, A. K. Nellithimaru, T. Bates, G. Kantor, Bumblebee: A path towards fully autonomous robotic vine pruning., Field Robotics 2 (1) (2022) 1661–1696.

[95] A. You, N. Parayil, J. G. Krishna, U. Bhattarai, R. Sapkota, D. Ahmed, M. Whiting, M. Karkee, C. M. Grimm, J. R. Davidson, An autonomous robot for pruning modern, planar fruit trees, arXiv:2206.07201 [cs] (Jun. 2022).
URL `http://arxiv.org/abs/2206.07201`

[96] D. Strnad, Š. Kohek, Novel discrete differential evolution methods for virtual tree pruning optimization, Soft Computing 21 (4) (2017) 981–993.

[97] S. Kolmanič, D. Strnad, Š. Kohek, B. Benes, P. Hirst, B. Žalik, An algorithm for automatic dormant tree pruning, Applied Soft Computing 99 (2021) 106931.

[98] W. V. Marset, D. S. Pérez, C. A. Díaz, F. Bromberg, Towards practical 2d grapevine bud detection with fully convolutional networks, Computers and Electronics in Agriculture 182 (2021) 105947.

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

[100] G. Zhao, D. Wang, A Multiple Criteria Decision-Making Method Generated by the Space Colonization Algorithm for Automated Pruning Strategies of Trees, AgriEngineering 6 (1) (2024) 539–554. `doi:10.3390/agriengineering6010033`.
URL `https://www.mdpi.com/2624-7402/6/1/33`

[101] L. Tang, C. Chen, H. Huang, D. Lin, An integrated system for 3d tree modeling and growth simulation, Environmental Earth Sciences 74 (2015) 7015–7028.

[102] A. Bérut, H. Chauvet, V. Legué, B. Moulia, O. Pouliquen, Y. Forterre, Gravisensors in plant cells behave like an active granular liquid, Proceedings of the National Academy of Sciences 115 (20) (2018) 5123–5128.

[103] R. Měch, P. Prusinkiewicz, Visual models of plants interacting with their environment, in: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 1996, pp. 397–410.

[104] W. Palubicki, K. Horel, S. Longay, A. Runions, B. Lane, R. Měch, P. Prusinkiewicz, Self-organizing tree models for image synthesis, ACM Transactions On Graphics (TOG) 28 (3) (2009) 1–10.

[105] S. Corbett-Davies, T. Botterill, R. Green, V. Saxton, An expert system for automatically pruning vines, in: Proceedings of the 27th Conference on Image and Vision Computing New Zealand, ACM, Dunedin New Zealand, 2012, pp. 55–60. `doi:10.1145/2425836.2425849`.
URL `https://dl.acm.org/doi/10.1145/2425836.2425849`

[106] A. Runions, B. Lane, P. Prusinkiewicz, Modeling trees with a space colonization algorithm., Nph 7 (63-70) (2007) 6.

[107] M. Bryson, F. Wang, J. Allworth, Using synthetic tree data in deep learning-based tree segmentation using lidar point clouds, Remote Sensing 15 (9) (2023) 2380.

[108] Š. Kohek, N. Guid, S. Tojnko, T. Unuk, S. Kolmanič, Eduapple: Interactive teaching tool for apple tree crown formation, HortTechnology 25 (2) (2015) 238–246.

[109] H. Kang, M. Fiser, B. Shi, F. Sheibani, P. Hirst, B. Benes, Imapple—functional structural model of apple trees, in: 2016 IEEE international conference on functional-structural plant growth modeling, simulation, visualization and applications (FSPMA), IEEE, 2016, pp. 90–97.

[110] P. De Reffye, C. Edelin, J. Françon, M. Jaeger, C. Puech, Plant models faithful to botanical structure and development, ACM Siggraph Computer Graphics 22 (4) (1988) 151–158.

[111] R. Storn, K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, Journal of global optimization 11 (1997) 341–359.

[112] F. Westling, J. Underwood, M. Bryson, A procedure for automated tree pruning suggestion using LiDAR scans of fruit trees, arXiv:2102.03700

[cs, eess] (Feb. 2021).
URL http://arxiv.org/abs/2102.03700

[113] P. E. Hart, N. J. Nilsson, B. Raphael, A formal basis for the heuristic determination of minimum cost paths, IEEE transactions on Systems Science and Cybernetics 4 (2) (1968) 100–107.

[114] T. Qiu, A. Zoubi, L. Cheng, Y. Jiang, 3d branch point cloud completion for robotic pruning in apple orchards, arXiv preprint arXiv:2404.05953 (2024).

[115] T. Qiu, L. Cheng, Y. Jiang, 3d characterization of tree architecture for apple crop load estimation, in: 2022 ASABE Annual International Meeting, American Society of Agricultural and Biological Engineers, 2022, p. 1.

[116] X. Yu, Y. Rao, Z. Wang, J. Lu, J. Zhou, Adapointr: Diverse point cloud completion with adaptive geometry-aware transformers, IEEE Transactions on Pattern Analysis and Machine Intelligence (2023).

[117] H. Yang, L. Li, Z. Gao, Obstacle avoidance path planning of hybrid harvesting manipulator based on joint configuration space, Transactions of the Chinese Society of Agricultural Engineering 33 (4) (2017) 55–62.

[118] C. W. Bac, J. Hemming, B. Van Tuijl, R. Barth, E. Wais, E. J. van Henten, Performance evaluation of a harvesting robot for sweet pepper, Journal of Field Robotics 34 (6) (2017) 1123–1139.

[119] R. Harrell, P. D. Adsit, R. Munilla, D. Slaughter, Robotic picking of citrus, Robotica 8 (4) (1990) 269–278.

[120] A. Silwal, J. R. Davidson, M. Karkee, C. Mo, Q. Zhang, K. Lewis, Design, integration, and field evaluation of a robotic apple harvester, Journal of Field Robotics 34 (6) (2017) 1140–1159.

[121] K. Tanigaki, T. Fujiura, A. Akase, J. Imagawa, Cherry-harvesting robot, Computers and electronics in agriculture 63 (1) (2008) 65–72.

[122] J. Zhang, J. K. Schueller, Kinematics and dynamics of a fruit picking robotic manipulator, in: 2015 ASABE Annual International Meeting, American Society of Agricultural and Biological Engineers, 2015, p. 1.

[123] Y. Li, Z. Zhang, X. Wang, W. Fu, J. Li, Automatic reconstruction and modeling of dormant jujube trees using three-view image constraints for intelligent pruning applications, Computers and Electronics in Agriculture 212 (2023) 108149.

[124] A. You, F. Sukkar, R. Fitch, M. Karkee, J. R. Davidson, An Efficient Planning and Control Framework for Pruning Fruit Trees, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), 2020, pp. 3930–3936, iSSN: 2577-087X. doi:10.1109/ICRA40945.2020.9197551.

[125] T. Teng, M. Fernandes, M. Gatti, S. Poni, C. Semini, D. Caldwell, 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), 2021, pp. 37–42. doi:10.1109/ICARM52023.2021.9536083.

[126] B. Zhang, X. Chen, H. Zhang, C. Shen, W. Fu, Design and performance test of a jujube pruning manipulator, Agriculture 12 (4) (2022) 552.

[127] A. Zahid, L. He, L. Zeng, D. Choi, J. Schupp, P. Heinemann, Development of a robotic end-effector for apple tree pruning, Transactions of the ASABE 63 (4) (2020) 847–856.

[128] Y. Lu, Kinematics Analysis and Trajectory Planning of Dual-arm Pruning Robot, IOP Conference Series: Earth and Environmental Science 769 (4) (2021) 042067. doi:10.1088/1755-1315/769/4/042067. URL https://iopscience.iop.org/article/10.1088/1755-1315/769/4/042067

[129] N. Kondo, K. Ting*, Robotics for plant production, Artificial intelligence review 12 (1998) 227–243.

[130] A. You, N. Parayil, J. G. Krishna, U. Bhattarai, R. Sapkota, D. Ahmed, M. Whiting, M. Karkee, C. M. Grimm, J. R. Davidson, Semiautonomous precision pruning of upright fruiting offshoot orchard systems: An integrated approach, IEEE Robotics & Automation Magazine (2023).

[131] S. B. Slotine, 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, 1991, pp. 1211–1216.

[132] Y. Guan, K. Yokoi, O. Stasse, A. Kheddar, On robotic trajectory planning using polynomial interpolations, in: 2005 IEEE international conference on robotics and biomimetics-ROBIO, IEEE, 2005, pp. 111–116.

[133] J. J. Kuffner, S. M. LaValle, Rrt-connect: An efficient approach to single-query path planning, in: Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065), Vol. 2, IEEE, 2000, pp. 995–1001.

[134] Y. Chen, Y. Fu, B. Zhang, W. Fu, C. Shen, Path planning of the fruit tree pruning manipulator based on improved rrt-connect algorithm, International Journal of Agricultural and Biological Engineering 15 (2) (2022) 177–188.

[135] M. Zucker, N. Ratliff, A. D. Dragan, M. Pivtoraiko, M. Klingensmith, C. M. Dellin, J. A. Bagnell, S. S. Srinivasa, Chomp: Covariant hamiltonian optimization for motion planning, The International journal of robotics research 32 (9-10) (2013) 1164–1193.

[136] J. Schulman, J. Ho, A. X. Lee, I. Awwal, H. Bradlow, P. Abbeel, Finding locally optimal, collision-free trajectories with sequential convex optimization., in: Robotics: science and systems, Vol. 9, Berlin, Germany, 2013, pp. 1–10.

[137] J. D. Gammell, S. S. Srinivasa, T. D. Barfoot, Batch informed trees (bit*): Sampling-based optimal planning via the heuristically guided search of implicit random geometric graphs, in: 2015 IEEE international conference on robotics and automation (ICRA), IEEE, 2015, pp. 3067–3074.

[138] A. Silwal, T. Parhar, F. Yandun, H. Baweja, G. Kantor, A robust illumination-invariant camera system for agricultural applications, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2021, pp. 3292–3298.

[139] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Advances in neural information processing systems 27 (2014).

[140] C. Lauretti, C. Tamantini, H. Tomè, L. Zollo, Robot learning by demonstration with dynamic parameterization of the orientation: an application to agricultural activities, Robotics 12 (6) (2023) 166.

[141] S. Häring, S. Folawiyo, M. Podguzova, D. Stricker, et al., Vid2cuts: A framework for enabling ai-guided grapevine pruning, IEEE Access 12 (2024) 5814–5836.

[142] C. H. Kim, A. Silwal, G. Kantor, Autonomous robotic pepper harvesting: Imitation learning in unstructured agricultural environments, IEEE Robotics and Automation Letters (2025).