Design and evaluation of a modular robotic plum harvesting system utilizing soft components

Jasper Brown | Salah Sukkarieh

The Australian Centre for Field Robotics, Faculty of Engineering, The University of Sydney, Sydney, Australia

Correspondence
Jasper Brown, The Australian Centre for Field Robotics, Faculty of Engineering, The University of Sydney, Sydney 2006, Australia. Email: j.brown@acfr.usyd.edu.au; jasperebrown@gmail.com

Abstract
The human labor required for tree crop harvesting is a major cost component in fruit production and is increasing. To address this, many existing research works have sought to demonstrate commercially viable robotic harvesting for tree crops, though successful commercial products resulting from these have been few and far between. Systems developed for specific crops such as sweet peppers or apples have shown promise, but the vast majority of cultivar types remain unaddressed, and developing a specific system for each one is inefficient. In this study, an easily modifiable development platform for robotic fruit harvesting is presented, this can be used to test specific design choices on different fruit and growing conditions. The system is evaluated in a commercial plum orchard, with no crop modifications. Both a hard and soft gripper are trialed, along with three object detector approaches and two picking motions. Some existing techniques are found to be counterproductive for plums, while soft robotics and persistent target tracking significantly improve performance. The best harvest success rate of 42%, was observed when using the soft gripper with complex motion. This is lower than expected based on prior testing with apples and indicates the difficulty in moving to new fruit types. Unique challenges specific to the plum type and growing style are examined in the context of system module design choices.

KEYWORDS
agriculture, harvesting, soft robotics, terrestrial robotics, tree crops

1 | INTRODUCTION

Robotic fruit harvesting is a problem that has captured the attention of researchers for over 40 years, who have generated a wealth of publications on the topic. The economic benefits to growers, of reduced reliance on a largely seasonal and often untrained labor force are clear. Despite obvious motivation and extensive research, translating research outcomes into commercial solutions has proven difficult. Those which have recently emerged are limited to specific crop types and the vast majority of tree crop cultivars, such as plums, remain unaddressed. What we present here is a system design that uses a modular software architecture with highly flexible hardware, to allow for rapid development of system components that may be specific to different growing conditions or fruit cultivars. By using this flexible platform for development, we are able to rapidly test multiple approaches for key components and algorithms. This system is validated on a plum crop and studies of the design decisions for various components are carried out. In addition to commonly used techniques, a persistent target detection and filtering module is introduced to overcome the limitations of object detection in highly obscured crops. Also, the use of soft robotics components in the gripper hardware module was shown to have a very large impact on harvest success while eliminating issues associated with collisions and tree damage.
The primary limitation of this study is the small amount of testing data, which is restricted to a single growing season. We plan to address this in the coming year. Gripper force and longevity were also an issue, with further design iterations of the soft gripper geometry and materials required. Assessment of nonpickable fruit to avoid, and autonomous picking motion selection, are identified as areas for possible system performance improvement.

2 | RELATED WORK

There is an extensive body of literature on both systems and techniques for robotic fruit picking. Over 50 papers have been published on the topic, many of these in the past several years (Bac, van Henten, Hemming, & Edan, 2014; Comba, Gay, Piccarolo, & Aimonino, 2010; De-An, Jidong, Wei, Ying, & Yu, 2011). Despite significant research effort, commercial progress has been slow and the field has been hampered by the difficulty of implementing standard testing protocols to allow direct comparison. Many works are specific to one crop type and growing setup, while the difficulty of creating common testing setups makes direct comparisons largely impossible even within a fruit type.

System implementations of robotic harvesters using modern techniques are presented by Arad et al. (2020) for sweet peppers and by Xiong, Peng, Grimstad, From, and Isler (2019) for strawberries. Many additional research efforts have been directed towards autonomous harvesting in protected cultivation systems (Tanigaki, Fujiura, Akase, & Imagawa, 2008; van Henten et al., 2002, 2013). The growing pattern and lighting regularity of these environments, along with typically higher-value crops, makes them ideal candidates for automated harvesting. Crops such as sweet peppers possess soft and flexible stems, in contrast to the hard lignified wood of tree crops, meaning collisions will easily damage the plant but will not often lead to the stoppage of the harvesting process.

Operating on outdoor tree crops brings the opposite challenge, with system component damage more likely than tree damage. It also introduces additional complexities in perception, gripping, and crop variation. An early apple picking system was developed by Baeten, Donné, Boedrij, Beckers, and Claesen (2008) which makes use of an eye-in-hand camera and soft, suction gripper. Apple harvesting with a low-cost 3D printed gripper was demonstrated by Silwal et al. (2017) but occurred on a modified crop. A custom-designed manipulator arm with a prismatic base is used to reduce cost and simplify planning. Rotational motion is applied during picking. Common failure causes were identified as position error, followed by apples on long shoots which behave as pendulums when picked.

Soft robotics is a field that focuses on the use of compliant, biologically inspired embodiments, sensors, and actuators (Verl et al., 2015). Numerous works have identified this as a key technology for robotic agriculture and harvesting. The Silwal et al. (2017) prototype was iterated upon in Hohimer et al. (2019) which performed apple harvesting on an unmodified crop using a soft pneumatic gripper. They identify fruit clustering as the most common failure case, followed by positioning error. Multiple exposures are used to increase image dynamic range and an ensemble of two learned feature detectors are applied to the resulting image. Picking motion is a straight pull and operating with a nonzero pitch angle was not found to improve performance, but did increase collisions. No research activity on robotic plum harvesting was found during a review of the literature, however mechanized plum harvesting is assessed in Mika, Buler, Rabcewicz, Bia Ikowski, and Konopacka (2015).

Commercial systems are beginning to become available for crops such as tomatoes (Panasonic1), strawberries (Agrobot,2 Octinion,3 and CROO Robotics4), raspberries (fieldwork robotics5) and apples (Abundant Robotics6 and FFRobotics7) among others. As of writing, all systems targeted at tree crops remain in the development phase. These solutions target a single type of fruit and require specific growing styles to be effective. Developing harvesting robots for less common fruit and growing systems will likely remain an open problem for many years and creates the need for flexible development platforms such as proposed in this study.

Fruit detection is the first step in picking and has been done using both hand engineered and learned approaches (Bargoti & Underwood, 2017; Kapach, Barnea, Mairon, Edan, & Ben-Shahar, 2012; Nguyen et al., 2014; Sa et al., 2016; Vitrabin & Edan, 2016). An apple harvesting system is presented by Onishi et al. (2019) who focus mainly on the detection aspect, and utilize deep learning. Detection for yield mapping in orchards is reviewed in recent work by Hání, Roy, and Isler (2020). Dynamic thresholding in a range of color spaces is applied to detect common crops in Zemmour, Kurtser, and Edan (2019). They report comparable results to deep learning approaches, though with minimal training data. Multispectral sensing has also been explored as an alternative to RGB color channels for fruit detection (Hung, Nieto, Taylor, Underwood, & Sukkarieh, 2013). Synthetic image generation is explored by Barth, Ijsselmuider, Hemming, and van Henten (2018) as a means of overcoming the need to gather large training datasets. A deep network is combined with morphological and color thresholding to yield high frame rates while detecting sweet peppers and fruit in Arad et al. (2020), though this operates on yellow fruit. Lighting control is effective for greenhouse environments through filtering and flash-no-flash image sequences, as demonstrated in Arad et al. (2019). State of the art fruit counting results are generated in Chen et al. (2017) using a multistage deep learning approach optimized for overall fruit quantity estimation, rather than individual instance localization. Multiview detection has also been considered (Hemming, Ruizendaal, Hofstee, & van Henten, 2014).

Target tracking and information fusion over multiple frames is rarely implemented in harvesting system designs. In work by Mehta, Ton, Asundi, and Burks (2017) multiple monocular cameras are used...
with a particle filtering framework to localize fruit in three-dimensional (3D). Spring-mass motion models are developed, though only validated in simulation. Disparity maps are used to detect and measure broccoli seedlings in Ge et al. (2019). Outliers are removed by applying K-nearest neighbors to the point clouds but no filtering occurs on the final detections.

Gripper design is a critical component of autonomous harvesting (Blanes, Mellado, Ortiz, & Valera, 2011; Rodríguez, Moreno, Sánchez, & Berenguel, 2013). In addition to the pneumatic gripper of Hohimer et al. (2019), other experimental techniques include under-actuated cable grippers (Xiong et al., 2019) and integrating tactile feedback (Dimeas, Sako, Moulianitis, & Aspragathos, 2015). The Abundant Robotics commercial platform uses a vacuum type gripper that swallows the fruit. A commercial deformable-finger end-effector design is used by Eizicovits, van Tuijl, Berman, and Edan (2016), where simulation tools are used to map grasp position tolerances to sensing requirements. Crops such as sweet peppers require stem cutting which has been addressed in combination with suction (Bac & Hemming, 2017; Lehnert, English, McCool, Tow, & Perez, 2017) and catching (Arad et al., 2020) type grippers.

Final target approach is commonly done using image-based visual servoing (IBVS) with eye-in-hand cameras (Arad et al., 2020; Barth, Hemming, & van Henten, 2016; Mehta, Mackunis, & Burks, 2016). A review of vision-based control is presented in Zhao, Gong, Huang, and Liu (2016). End effector integrated IR distance sensors are used by Xiong et al. (2019) for final positioning. Multiple sweet pepper approach strategies are assessed under laboratory and greenhouse conditions by Ringdahl, Kurtser, and Edan (2019). Using multiple approach attempts was found to slightly increase the probability of reaching a pepper, at the cost of higher cycle times. Radicchio harvesting is targeted in Foglia and Reina (2006) which requires cutting the plant stem approximately 10 mm below the ground plane using a claw type mechanism. High density two-dimensional (2D) growing systems are well suited to cartesian picking motions with up to 91% of fruit reachable in this manner (Vougioukas, Arikapudi, & Munic, 2016). Picking motions for apples are identified and evaluated in Li, Karkee, Zhang, Xiao, and Feng (2016) with some motions being much more prone to fruit damage.

In this study, two new gripper designs featuring both hard and soft components are compared. Additionally, a persistent target tracking filter is developed to go beyond single-frame perception techniques. Unlike many existing works, significant emphasis is placed on the modularity of hardware and software. This modular approach is tested by harvesting a novel crop type, with multiple module design choices being evaluated.

3 | EXPERIMENTAL DESIGN

Full system evaluation was carried out as part of a week-long plum harvesting trial. The aim of this was to discover systemic strengths and weaknesses of the flexible development platform, and to understand the requirements specific to plum harvesting. Within this trial, three experiments were done to test specific module design choices. In the first, multiple object detector algorithms are tested in both day and night scenes. Next, both simple and complex harvesting motions are assessed for a soft gripper. Finally, hard and soft grippers are compared on criteria of their effectiveness and robustness to collisions.

Target crop and trellis style are essential considerations for autonomous harvesting. Plums were chosen for this study due to their availability in a modern fruiting wall style 2D trellis configuration. This 2D style, shown in Figure 1, is widely employed and well suited to mechanization. It provides for very uniform growing conditions, leading to more predictable and profitable crops. The fruiting wall trellis also provides clear access to the target fruit and is ideal for testing robotic harvesters.

Flower thinning directly determines where fruit will grow in the coming months, so is essential to harvest success. Current best practice thinning, which discourages bunching and matches fruit density to branch carrying capacity, was applied to this crop. A mechanized brush first reduces overall flower load, followed by a manual thinning step.

4 | HARDWARE SYSTEM DESIGN

The hardware system design goals are to realize a fully self-contained development platform that has long-endurance, is easily modified, and can support a variety of payloads. To maximize the deployment options, it is constructed on a trailer base which can be towed by a robotic platform, moved by hand, or attached to existing farm vehicles. Key hardware modules are shown in Figure 2. The requirements and implementation of these is described below.

4.1 | Support

The support hardware module provides 240 V power and compressed air at 60 psi for the soft gripper. An uninterruptible power supply is also included to provide approximately 30 min of battery run time for the compute system while the generator is being refueled. A system power budget is shown in Table 1 which is met by a 3 kW petrol generator.

4.2 | Compute

The compute module provides a high-performance PC for generic tasks and an embedded deep learning computer to offload model inference. Low-level control of the gripper positioning system, a UR5 manipulator arm, is handled by a dedicated controller, in this case, the standard universal robots (UR) control unit. An off the shelf router provides local networking as well as a WiFi link for remote control, monitoring, and visualization.

For the high-performance PC, a Zotac EN72080V mini PC is used, this has an i7 processor, 32 GB of RAM and an NVIDIA
RTX2080 graphics processing unit (GPU). However, this GPU is only used for development and during deployment the object detection models are run on an NVIDIA Xavier embedded deep learning computer. Compute components are mounted to a easily removable tray, shown in Figure 3, for rapid component replacement.

### 4.3 Sensing

Two Intel Realsense cameras are used for 3D sensing, along with a wide-angle camera to implement the visual servoing and in-gripper visualization. A D435i RGBD camera is mounted above the gripper to provide 3D fruit locations and can also be used for obstacle detection. A T265 simultaneous localization and mapping (SLAM) camera is mounted to the rear corner of the trailer to allow for persistent global target tracking as the trailer moves through the orchard. This global tracking is not required for picking, but allows for tracking fruit during trailer movements and for functionality such as counting and crop density mapping.

The small footprint of the D435i is essential when using the on-gripper mounting configuration and was one of the selection factors when choosing a 3D camera. Many existing approaches use an off-arm camera setup or multiple cameras. The 3D camera is mounted on the end effector to allow for maximum flexibility when positioning it, this can also be used to simulate an off-arm setup by moving the end effector to a fixed position for each new camera frame.

While the D435i provides general fruit localization, the field of view is insufficient for final approach. Most on-gripper sensing will be obscured at very close range, so a 170° diagonal field of view RGB camera is embedded within the soft gripper cup for visual servoing control and grasp visualization.

### 4.4 Actuation

Collaborative robot (cobot) arms are designed to safely work in close proximity to humans by limiting the maximum force and inertia of links through feedback control or compliant mechanisms. These are essential safety requirements when working near untrained farm staff and during platform development. The UR5 CB3 arm was chosen as a well supported and kinematically suitable cobot arm. Workspace and payload limits of the UR5 were sufficient for our development goals.

Gripper design is an essential aspect of successful fruit harvesting and a huge variety of approaches have been proposed. Additionally, the size, cultivar type and growing conditions of fruit will all alter the optimal gripper design. A specialized gripper control unit is used to provide a standardized software and hardware interface to two grippers tested in this study. Both of these take identical serial commands over USB and are supplied by 24 V DC, this commonality allows for rapid integration of existing and future designs.

A tendon driven parallel gripper, shown in Figure 4, was first developed and is actuated by a Dynamixel Mx28T servomotor, controlled using the OpenCM9.04 board. Two parallel plates are mounted on a linear rail with a common tendon-cam design actuating
them simultaneously. By keeping the total tendon loop length constant at all gripper positions, the need for a tensioning mechanism is eliminated allowing for a very simple and low cost gripper that can be made arbitrarily wide. The tested configuration has a minimum closure distance of 15 mm and maximum of 200 mm. All the components for this are commercially available or 3D printed, and the designs are made available online.8

Following preliminary testing on apple crops, the most common failure mode of the system was identified as fruit being grasped with one side of the parallel gripper fingers on the outside of a branch, leading to the fruit not being held on that side. The most obvious post-grasp mechanism to deal with this was identified as having that finger slide off the branch and snap back into contact with the fruit as the arm retracts. Initial design concepts looked at mechanical linkages, however soft robotics was chosen as an approachable and robust solution.

To demonstrate a soft robotics approach, a second gripper was developed with four pneumatic fingers, shown in Figure 5. This uses the finger designs from Sun, Song, and Paik (2013), and Sun et al. (2017), all connected to a common air loop ensuring they close together. Control for this occurs on an Arduino Uno which actuates a

---

8github.com/jaspereb/SimpleSliderHand.
pair of pneumatic solenoids to alternatively pressurize and depressurize the gripper air loop.

Damaging contact with tree branches and fruit is also a risk with any rigid gripper and will preclude the fully autonomous long term deployment of harvesting robots unless extremely reliable perception systems can be developed. By employing soft robotics components, key failure cases can be avoided in hardware rather than software. By having only soft components extending beyond the gripper cup, collisions in this region can be ignored.

### SOFTWARE AND ALGORITHM DESIGN

The software architecture can be divided into five key modules, shown in Figure 6. Sensing and perception form a standalone unit that can run without any additional input, allowing the same platform to be used for tree crop analytics such as fruit counting and yield volume estimation. The robot operating system (ROS) is used for message passing and software bringup. All key tuning values for the system are stored on the ROS parameter server and can be updated live without restarting any nodes. While this induced additional development overhead, the ability to live update all parameters was found to be key for field development.

Dealing with hard and soft obstacles is a primary challenge of fruit harvesting and avoiding damaging collisions requires a combined hardware and software approach. In addition to soft robotics components, a two stage planning mechanism is used. First, an approach pose 150 mm out from the fruit location is planned to while respecting a “hard obstacles” safety constraint plane just beyond that, this prevents the arm from moving through the trellis while reaching that pose. Following this, one of two final approach controllers takes over and moves the arm to the fruit location while not exceeding a second “soft obstacles” safety plane which prevents the hard gripper components from contacting the trellis. With this two stage approach, only the soft gripper components can contact the trellis wires, tree limbs, and posts.

Visualization occurs in Rviz, either onboard the generic PC or off-board via WiFi link. Shown in Figure 7 is an example of the default visualization window showing the arm state, target fruit and planning scene obstacles.

### TABLE 1 System power budget

| Component          | Nominal power (W) | Maximum power (W) | Input voltage (V) |
|--------------------|-------------------|-------------------|-------------------|
| Compressor         | 50                | 1000              | 240 AC            |
| Zotac              | 200               | 330               | 240 AC            |
| Arm                | 150               | 325               | 240 AC            |
| Xavier             | 40                | 75                | 9–20 DC           |
| Router             | 10                | 10                | 12 DC             |
| Arduino            | 1                 | 5                 | 5 DC              |
| D435               | 1                 | 2.5               | 5 DC              |
| T265               | 2                 | 2.5               | 5 DC              |
| Wide angle camera  | 2                 | 2                 | 5 DC              |

Note: Compressor duty cycle is approximately 5% resulting in a low average, but high maximum power draw.
5.1 | Sensing

Framerates are limited to 15 FPS for the D435i and wide-angle camera, including the depth feeds, to reduce computation intensity. Overall CPU load for the system is approximately 60%, of which 20% is due to visualization tools. Odometry transformations from the T265 SLAM camera were found to be unstable under certain circumstances, so an option to disable this was also built in. All cameras operate with automatic exposure turned on.

5.2 | Perception

Object detection is run on both the D435i and wide-angle RGB streams using one of the three models described in Section 6.1.

FIGURE 4 The parallel tendon drive gripper and D435i Realsense. Not rendered are the flexible tendons that run from each sliding car through the mid pulley and onto the actuation pulley. These form a closed loop with the tendon end pulleys to keep constant tension in the drive system as the cars slide along the linear rail [Color figure can be viewed at wileyonlinelibrary.com]

FIGURE 5 The soft robotic gripper showing the two zones of hard and soft components, as well as three commonly observed stable modes. Objects fully within the gripper cup usually resulted in mode (a), while larger or further away objects would typically put the gripper in either mode (b) or (c). Different air pressures also impact gripper geometry and mode [Color figure can be viewed at wileyonlinelibrary.com]
A single model is used for both camera streams. For all operations except the detector type study, the YoloV3 detector is used. This is an embedded version and is deployed on the NVIDIA Xavier, with image data being passed over the LAN. This reduces load on the primary computer and allows for very low power (30 W) inference at over 15 FPS. Detector training labels are constructed to capture the full target extents, even where these are obscured, this allows for better estimation of fruit size using just the bounding box.

While deep learning-based object detection does have much larger training and inference burdens than hand-engineered features such as hue-saturation-value (HSV) thresholding, by using a commercial labeling service it was possible to gather data, label and train a model overnight for a 24 h turn around. Rapid detector training is important for crops that may change appearance each season.

To deproject a D435i detection into the world frame, the pixel coordinates \((u, v)\) are required along with a depth value \((d)\). Calculation of the \((d)\) value is done by extracting the image patch corresponding to the bounding box extents, then applying a HSV filter to that patch and keeping only fruit-colored depth pixels. This is only applied to the local image patch, and is distinct from the HSV filter described in Section 6.1. The depth and image feeds to the patch HSV extractor are downscaled from \(640 \times 480\) to \(320 \times 240\) for...
efficiency. The median depth value of fruit-colored pixels is then returned. By using a patch which is known to contain a fruit, a highly tolerant HSV threshold filter can be used. In the case of no HSV thresholded pixels being present, the depth value at the bounding box centroid is used. Where the centroid depth value is also missing, that target is skipped. If all pixels fall within the threshold this approach reduces to the case of taking the bounding box median depth value. This masking technique is specifically required for instances where the fruit is more than 50% obscured, otherwise the depth will correspond to the obscuring objects and be artificially low. For a spherical fruit with no obscurations, a maximum ratio of 1\(\pi\) accurate bounding box depth pixels will come from the fruit.

A 3D position extraction module takes the \((u, v)\) bounding box centroid position and calculated \((d)\) value, and uses the camera intrinsics and frame transformation tree to deproject each detection into the world frame. Checks on minimum and maximum depth and size are performed here. A motion filter is also applied to ignore detections while the camera is moving above a certain linear or angular velocity, this eliminates blurred detections. A fruit size property is set for each fruit object, which is the mean bounding box dimension projected to the depth value. Additional properties such as ripeness, pickability, health, and visibility can also be set here to allow for future functionality. This module publishes a list of all currently detected fruit objects, which is used by the persistent target tracking module.

This target tracking module is used to track all currently and previously seen fruit within the world frame. All observed fruit are assigned a unique ID and tracked within the state vector of an Extended Kalman Filter (EKF). For each currently observed fruit, a euclidean distance threshold is applied to associate it with an existing target. If there is no existing target, it is added to the state matrix, otherwise the EKF update step is run for reobserved targets. At each filter iteration a static state transition model is applied, though this could easily be modified to allow target motion models, such as for wind. When a current detector frame is received, all elements (tracked targets) in the EKF state matrix are projected into the current camera frame. If a target should be observed but is not, a counter is incremented for that target. At the end of each update loop, targets with counters above a set threshold are removed as false-positive detections.

The EKF state vector, default noise matrices and initial covariance matrix are given by

\[
\begin{bmatrix}
    x_1 \\
    y_1 \\
    z_1 \\
    x_2 \\
    y_2 \\
    z_2 \\
    \vdots \\
    x_n \\
    y_n \\
    z_n
\end{bmatrix},
\]

\[
Q = 0.01 \times I_{n,3}, \quad R = 0.02 \times I_{n,3},
\]

\[
P = \begin{bmatrix}
    0.05 & 0 & 0 \\
    0 & 0.05 & 0 \\
    0 & 0 & 0.1
\end{bmatrix},
\] (1)

where \(I\) is the identity matrix and \(n\) is the number of fruit currently being tracked. A method for estimating these noise covariance matrices is presented in our previous work (Brown, Su, Kong, Sukkarieh, & Kerrigan, 2019). All units are in meters and the third element of \(P\) is aligned with the camera depth axis \((d)\) under our frame definitions, so has greater variance than the \((u, v)\) pixel readings. An association distance of 0.03 m was used.

5.3 | Planning

Target filtering is the first planning step, this takes the tracking module output and applies a tag to fruit objects which are either within the arm workspace or within a more restrictive region of interest (RoI). By using a highly restrictive RoI it is possible to largely eliminate arm singularities within the workspace, essential for reliable functioning of the approach controllers. For all tests, picks were attempted within a 0.5 \(\times\) 0.5 \(\times\) 0.8 m RoI window directly beside the trailer, with a mobile platform, this did not slow down the overall pick rate significantly.

Core picking functionality is constructed as a state machine, illustrated in Figure 8, which controls the current arm goal, gripper...
actuation and approach controller status. Five poses are defined for the UR5 arm at any given time; a home position for convenience, a drop position where the fruit are deposited, a look position that gives a view of the entire RoI, a pick approach position 150 mm offset from the fruit and the actual fruit position for harvesting. The first three of these are user defined, while the last two are dynamically updated by the approach controller. Operator input consists of setting or moving to the above poses, manually actuating the gripper, listing current target poses or starting automatic picking. For automatic picking, the state machine caches the most recent list of tracked and filtered fruit, all targets within the RoI are then worked through in order of highest to lowest on the trellis. We note that many existing works have applied traveling-salesperson solutions or similar to harvest order planning. This is only beneficial if the drop position varies, in our case the total tour length only varies with the choice of first fruit target after leaving the home position. This may be optimally chosen as the fruit that projects furthest along the drop to home position vector.

Once a target is popped from this list the arm is planned to the approach pose for that target and one of the two approach controllers is then activated. If the IBVS controller is used and fails, the direct approach controller is attempted. If both fail, likely due to a collision or dangerous singularity, the pick is aborted and the next target chosen. If either controller succeeds the gripper is closed and the arm moved backward 150 mm using the direct approach controller. From there, the arm is planned to the drop pose and the next target chosen. Feedback from the parallel gripper is used to determine when nothing has been contacted during a grab, in that case the IBVS controller is run again once from that position and a grasp reattempted.

Arm motion planning and inverse kinematics (IK) is handled using the RRTConnect planner from the Open Motion Planning Library (Sucan, Moll, & Kavraki, 2012) and KDL-Kinematics IK solver within the Moveit ROS framework. This is run in the threaded mode for concurrent planning and movement. Known paths, such as the drop to home motion, are precalculated and stored, these are also executed at a higher speed. For all tests, a 1 s planning timeout was used, though this can be reduced in most cases.

5.4 | Approach controller

During final approach linear motion of the end effector is essential to avoid tree or hardware damage. One goal of this system is to assess the viability of a cartesian manipulator, so end effector orientation is restricted to be perpendicular to the trellis. Only cartesian motions are used during final approach, with the exception of rotational motion described in Section 6.2.

The direct servo controller uses the UR script interface of the UR5 driver to send direct servo commands. This uses the instantaneous joint motion planner running on the UR control box and ensures the end effector moves exactly as desired. But software limitations prevent collision and singularity checking in this mode, thus the importance of a restricted RoI. Software joint limits are checked within both approach controllers by comparing current angles to the configured limits. If the joints are within 2 degrees of any limit, the controller is stopped and returns a failure condition. If the joints are over any limit, the previous velocity vector command is reversed and sent for one time step to bring them back within bounds, allowing the planner to work.
Image-based visual servoing is used to provide low level feedback control of the gripper final positioning. This was found to be essential in lab testing where contacting leaves and stems would cause the fruit to move during approach, it also makes the system robust to wind movement and perception errors. The nearest in-hand camera detection to the target point is tracked when the filter initializes, then subsequent frames attempt to associate a current detection with this target using a 2D distance. If association fails, the previous motion command is run again and a counter of missed frames is incremented. If this counter reaches a threshold, 15 by default, it triggers a failure status and the controller returns. Each successful association is used to calculate a pixel space error

\[ \delta_u = u_{\text{target}} - u_{\text{detection}}, \delta_v = v_{\text{target}} - v_{\text{detection}} \]  

(2)

where \( u \) and \( v \) are the vertical and horizontal image pixel coordinates of the detection centroid or target point, this has pixel space gains \( G_u \) and \( G_v \) applied to generate a velocity and is then used to scale the end effector depth axis velocity \( d_{\text{vel}} \) against a preset value \( ee_{\text{vel}} \)

\[ u_{\text{vel}} = \delta_u G_u, \quad v_{\text{vel}} = \delta_v G_v \]  

(3)

\[ d_{\text{vel}} = ee_{\text{vel}} - \max(\text{abs}(u_{\text{vel}}, v_{\text{vel}})) \]  

(4)

the three velocity components are combined into a normalized velocity vector which is scaled by the preset end effector velocity \( V_{\text{cmd}} \) and then transformed from the camera optical frame to the arm controller frame. The UR5 controller handles low level joint control to track the commanded end effector velocity \( V_{\text{cmd}} \).

\[ \vec{V} = \begin{bmatrix} u_{\text{vel}} \\ v_{\text{vel}} \\ d_{\text{vel}} \end{bmatrix}, \]  

(5)

\[ V_{\text{cmd}} = \left\| \vec{V} \right\|_2 \times ee_{\text{vel}}. \]  

(6)

While this results in a nonlinear controller form, it was found to be stable and effective over a wide range of gains. The final velocity vector is sent directly to the arm controller as a commanded end effector motion. Detection size above a threshold is used as the successful stopping condition. Checks on failure conditions include singularities detected by high joint velocity, minimum and maximum bounding box sizes, traversal distance limit exceeded and controller timeout reached.

5.5 Actuation

Direct arm actuation is handled by the UR controller box which implements PID joint control. Both velocity and torque limits can trigger a emergency stop on the arm, usually caused by a servoing singularity or collision respectively. These can be detected and reset in software, and cause the state machine to select the next target.

Soft gripper actuation is binary and open loop. The parallel gripper provides torque sensing which is used to set a binary feedback flag indicating if something has been grasped. Both grippers are controlled using a ROS action service that sends serial commands via a common ASCII protocol to an Arduino for the soft gripper, or OpenCM9.04 for the parallel gripper. When a close command is received by the OpenCM9.04 it runs the servomotor until a torque or width limit is reached. The minimum closing width and torque both reside on the ROS parameter server and can be dynamically set using the global fruit object diameter if desired.

6 TESTING PROCESS

The prototype system was first tested on an apple crop, to confirm all the components were working properly, before being evaluated on a commercial plum crop. Harvesting time windows in different crops make testing under identical conditions difficult, this is one motivating factor behind building a flexible and modular development platform. Some unexpected changes were required when moving to the plum crop, which are detailed in Section 8.1.

System evaluation occurred on two rows of a commercial red plum cultivar known as “Late Scheffer” grown in the 2D fruiting wall style of trellis. Fruit bunching was observed to be common, which also reduces the fruit sale quality where they are in contact. The grower indicated that he is trying to reduce bunching through more targeted thinning in the future, so this is not considered a major problem for harvest performance. Pick attempts were made on all detected fruit within the RoI box, the RoI height corresponds roughly to the space between the middle two trellis wires in all photos. After all fruit in the box were attempted, the platform was moved forward and stopped, leaving a 100 mm horizontal RoI overlap with the last pick attempt. Fruit falling within this overlap are automatically excluded from a second pick attempt. No crop modifications or target exclusions were applied. Detected targets behind trunks, trellis wires and other fruit were still attempted.

System tuning and evaluation took place over a week under conditions of rain, high wind, and darkness. The platform is able to operate in light rain by placing hook and loop secured flaps over the lower level, as in Figure 9, and utilizing water-resistant components on the top level. Wind caused a small amount of fruit movement, primarily in the depth axis which the gripper design is naturally tolerant of. Most fruits was within the typical picking ripeness window. Logs of all sensor feeds and system modules are made available for direct inspection of growing conditions.9

Many different system architectures and algorithms have been suggested in the literature and without existing results on

9data.acfr.usyd.edu.au/Agriculture/PlumHarvesting.
plums testing multiple module design choices was necessary. Specifically the choice of gripper, detector, and picking motion were all assessed.

6.1 | Detector type study

Effective object detection for harvesting is typically measured using recall and precision, which are essential for being able to harvest a large portion of the crop. However, commercial operation also places requirements on frame rates and power consumption, something not always achieved by high accuracy deep learning approaches. To compare a range of options, a basic HSV detector is compared to the low power embedded YoloV3 model and a state of the art Retinanet model. All of these are able to run at 15 FPS.

An HSV detector is built by manually tuning the 3 HSV channel thresholds on a representative daytime image to yield a binary mask of targets, which region size thresholds are applied to. Contiguous connected regions are considered a single object and returned as a positive detection.

Both deep learning models are trained on 492 images, representing over 3000 bounding box examples from both day and night scenes. Default hyper-parameter values are used for these with training data augmentation turned on. The embedded YoloV3 model is tuned to preference very low false-positive rates above high recall. The Retinanet model uses a Keras implementation \(^1\) running on the Zotac RTX2080, and a confidence cutoff of 0.7 to consider an object a true detection.

Sixteen frames exhibiting no camera motion were randomly selected from both the day and night data. The true-positive, false-positive, false-negative, and misseparation rates are calculated for these images. An intersection over union (IOU) threshold of 0.5 is used to define true-positives. False-negatives are missed fruit, while misseparation errors occur where a single bounding box has stretched to include multiple fruit, or multiple boxes exist for one fruit. False-positives occur where predicted boxes do not meet the IOU or misseparation criteria. Truncated bounding boxes and fruit are ignored. Bounding box accuracy was not found to be a limiting factor so is not further quantified.

6.2 | Picking motion study

Numerous complex motion strategies have been applied in existing works. Following discussions with human pickers and the grower, two motion strategies were proposed. The first is a simple straight-in, straight-out approach which is easily implemented on a cartesian system. The second was straight-in, rotate, angled-out which mimics the motion of human pickers, this is denoted “complex motion” and is shown in Figure 10.

6.3 | Gripper type study

Use of soft components was primarily chosen to reduce the impact of collisions but will also alter harvesting performance. To assess this, the parallel gripper was manually positioned for a number of pick attempts in either a vertical or horizontal orientation. Rotational motion with this gripper was rarely possible due to collisions, so a straight-in, straight-out motion was used.

7 | RESULTS

Overall harvesting success was far below that required for commercial viability, but represents a reasonable first step for this platform and numerous system components show clear opportunity

\(^1\)github.com/fizyr/keras-retinanet.
for improvement. Time per pick was not a major goal during testing, but should be optimized in future. The majority of time elapsed is due to actuation, and arm speed is constrained by worker safety and platform stability, rather than actuator limits. Lab testing, without the risk of damaging collisions, indicated a possible per pick time of 12 s, including all steps except trailer platform repositioning.

7.1 | Detector type

HSV thresholding with region size filtering was found to be not at all effective. Although it performed well on red apples in the Sydney region during preliminary testing, the HSV channels in red plums could not be well separated from the red soil in the Victoria region, even under artificial light. The bounding boxes produced were also of lower quality, as in Figure 11.

The embedded YoloV3 model achieved a very low false-positive and misseparation rate, at the cost of low recall. Practically the detection rate of this model is insufficient, even with the target tracking module able to compensate for detections that are missed in some frames.

Table 2 shows the day and night performance of the three detectors. Moving to artificial lighting improved the HSV detector, despite it being tuned for daytime operation. Retinanet also performed slightly better at night while YoloV3 was less effective.
Retinanet significantly outperformed the embedded YoloV3 model but recorded an average GPU power draw of 168 W, over four times that of the Xavier embedded computer. The low recall rate of YoloV3 was unexpected and may be due to the conversion process to an embedded model type, a more detailed comparative study of current deep learning object detectors for fruit will be performed in future work.

7.2 | Picking motion

Soft gripper failure modes and rates are enumerated in Table 3. The simple straight-in, straight-out motion was not very effective and frequently resulted in insufficient gripper force to detach the fruit. There are two reasons for this, the first being that only a subset of the fingers would frequently be in contact with the fruit leading to lower overall force transferred to the target. The second is due to fruit stems having much higher tensile strength, as opposed to shear strength. Mature plums detach at the abscission layer, which is the natural process by which they fall from the tree. However, pulling directly along the stem axis often resulted in stem pull-out which damages the fruit and requires a much higher amount of grip force. Rotational motion while pulling, or applying force at an angle to the stem, resulted in more detachments occurring at the abscission layer with lower required force. By applying the rotation and angled pull back motion, the fingers also fall into a more tightly closed stable mode, shown in Figure 5 bottom left, where the target is fully within the rigid gripper cup.

Bad positioning refers to any case where it was judged that more accurate positioning may have led to a successful pick. Often this was the result of fruit falling through the finger gaps in the gripper as it retracted. Stated failure modes are overlapping and must be considered in order, for instance a greater grip force may correct for a badly positioned pick. Likewise an attempt that fails due to positioning may also have subsequently failed due to grip force even with perfect positioning (Table 4).

A large portion of the "other" failure types for the soft gripper were fruit in difficult positions, such as on the other side of a trellis wire or trunk. Use of the more complex motion also resulted in many more fruit being knocked off, typically a missed pick during simple motion would leave the target fruit on the tree. When gripper rotation is introduced the fruit are much more effectively detached, but if not well gripped will then fall to the ground.

7.3 | Gripper type

Both gripper type and retraction motion were found to have a large effect on picking success rate. Specifically, the hard parallel gripper effectiveness tripled when used in a vertical rather than horizontal

### Table 2

| Detector   | True-positives | False-positives | False negatives | Misseparation | Recall (%) | Precision (%) | Bad box rate (%) |
|------------|----------------|-----------------|-----------------|---------------|------------|---------------|------------------|
| HSV Day    | 57             | 14              | 170             | 24            | 25.1       | 80.3          | 42.1             |
| YoloV3 Day | 73             | 0               | 154             | 2             | 32.2       | 100.0         | 2.7              |
| Retinanet Day | 185       | 7               | 42              | 8             | 81.5       | 96.4          | 4.3              |
| HSV Night  | 30             | 6               | 40              | 4             | 42.9       | 83.3          | 13.3             |
| YoloV3 Night | 13          | 0               | 57              | 1             | 18.6       | 100.0         | 7.7              |
| Retinanet Night | 61        | 0               | 9               | 4             | 87.1       | 100.0         | 6.6              |

Abbreviation: HSV, hue-saturation-value.

### Table 3

| Outcome                           | Straight | Angled | Straight percentage (%) | Angled percentage (%) |
|-----------------------------------|----------|--------|-------------------------|-----------------------|
| Success                           | 4        | 27     | 20.0                    | 42.2                  |
| Grip force failure                | 7        | 5      | 35.0                    | 7.8                   |
| Bad positioning failure           | 6        | 9      | 30.0                    | 14.0                  |
| Knocked off target failure        | 1        | 9      | 5.0                     | 14.0                  |
| Gripper failure                   | 0        | 2      | 0.0                     | 3.1                   |
| Other failure                     | 2        | 12     | 10.0                    | 18.8                  |
| Total                             | 20       | 64     |                         |                       |

Note: Straight and angled percentages refer to the relative rate of that outcome for the given motion type.
orienation. It was observed that more fruit have hard obstacles on their sides than above or below them and the majority of hard gripper failures were a result of grasping with an obstacle between the finger and fruit.

No fruit damage was observed with either gripper, although the parallel gripper resulted in minor tree damage where collisions occurred. Zero emergency stops due to collisions were registered with the soft gripper in place.

8 | DISCUSSION

Overall goals of testing system modularity, module choice comparisons and evaluation on a plum crop were met. Picking performance was lower than anticipated but comparisons to existing work will not be very informative for a range of reasons, including the crop type and testing conditions. The modularity criteria was well met, with minimal issues when swapping over components for module testing.

Several lessons around system design and development were learned. One clear lesson was the importance of rigorous and frequent testing. This is made difficult by the short harvesting window for most crops, uniqueness of each plant and trellis style, and the difficulty of creating accurate lab-based testing setups. Many existing harvesting systems use a visual servoing approach controller, whereas this was found to be worse than going directly to the estimated fruit position for plum crops. This led to wasted development effort and unnecessary sensing hardware, which could have been reduced with earlier testing on plums.

Target tracking and filtering is an essential tool if less than perfect object detectors are used, such as in highly obscured crops. This not only compensates for some detector failures, but also for additional fruit properties such as ripeness, to be estimated from multiple sensor frames or modalities.

Compute capacity of the current system was sufficient but not excessive. Utilizing an embedded solution for deep object detection saved energy while providing a stable and predictable frame rate. Bandwidth overall system topics was approximately 40 MB/s and the current software is not optimized for efficiency. This demonstrates that computing requirements can be easily met with commercial systems and off the shelf software frameworks, such as ROS.

Despite the water-resistant properties of the platform, heavy rain caused the loss of several testing days. While building water-proof hardware is not difficult, the performance of most perception and gripping systems remains untested for heavy rain and could pose unforeseen issues. Night operation, under artificial light was successful with little drop in detector or depth accuracy. Human pickers typically harvest at night due to better weather conditions and lower wind, day harvesting of this cultivar is not possible during very hot weather as the fruit bruise when stacked in collector bins. Controlled lighting can also be applied to improve day time machine vision performance, but requires high-intensity lighting or artificial shading.

The key software assumptions around soft and hard obstacle collisions occurring in the soft obstacle area and no hard components colliding with the trellis or trunk. This was in part due to careful calibration of these planes, which could in future be done online using sensor data, and also due to the uniform and well-kept trellis structure. We stress that no modifications were made to the crop for this study work, it is a commercial crop undergoing an otherwise normal harvest process. Fruiting wall plum crops do appear to be slightly more difficult, or at least different, problem than apple harvesting. Primarily due to fruit proximity to branches and trunks, though thinning protocol also impacts this.

8.1 | Crop specific observations

When the prototype system was previously tested on apples it performed subjectively well. The additional size and separation of these from the growing trunk led to the current soft gripper design. HSV detection also worked very well on red apples and the bunching was less severe, leading to less collateral damage fruit. The evaluation plum crop grows extremely close to woody trunks and branches, while being a lot smaller in size. This meant the soft gripper was oversized and needed to be modified with a soft skirt to prevent fruit from escaping between the fingers. While growing systems will vary greatly, even within plum crops, the proximity of target plums to branches caused many collisions and made the use of an angled motion critical to pick success. Images of the growing conditions can be found in the online data logs.

While the visual servoing approach controller was not originally intended to be one of the module design choice tests, it quickly became apparent that it was not contributing to picking success. Lab testing on fake orange trees and pre-evaluation tests on apples had indicated that the IBVS controller was essential to compensating for fruit motion. Unlike these crops, the tested plum crop has extremely short stems and exhibited almost no motion during picking, making the IBVS controller counter productive, even with the very small number of failures attributable to that module.

Some fruit and market combinations require the stem to be present, while for others it must be removed. In this plum crop either was acceptable, but stem removal often led to stem pull out where the plum skin is broken, rendering it worthless. Where stems are
required for commercial sale, the picking motion strategy must reflect this.

Crop style, gripper design, and motion choice are all closely interlinked. These cannot be separated and will pose a major problem for future development. The seasonal availability of testing crops compounds these difficulties, though some work has been done to address this using simulation tools (Wang et al., 2018).

8.2 Module design choices

Choice of detector architecture had a surprising impact on performance, with a significant difference in precision and recall between the two deep networks. The optimal balance of these two objectives will depend on the cost of attempting bad picks due to false-positives, weighed against the missed picks due to false-negatives. Detector limitations were partially compensated for by the persistent target tracking and filtering. The detector type study is intended to be comparative only and careful further tuning could marginally improve the performance of all three detector types.

Picking motion was also important, with even minor changes to the soft gripper motion resulting in different pick success figures. This does come at the cost of increased actuation time and additional gripper wear and tear. Longevity was one major downside of the soft gripper, with two catastrophic and two minor finger failures observed even for the small number of picks, approximately 300, carried out over the week. All motion strategies resulted in some fingers being caught between hard components and tree branches leading to external fabric and internal structure damage. Contact with rough old growth wood caused abrasions to the external silicone, introducing failure points when this was deformed by the finger actuation. More robust material selection should significantly improve soft gripper lifespan.

While the parallel gripper vertical performance was not vastly worse than the soft gripper, it resulted in many emergency stops due to detected collisions. The small number of attempted picks is due to this. Collision emergency stops are not captured in the success rate figures but makes the deployment of a hard gripper infeasible without much more advanced perception algorithms. Some of our previous work (Hung et al., 2013), has addressed trunk detection in figures but makes the deployment of a hard gripper infeasible without much more advanced perception algorithms. Some of our previous work (Hung et al., 2013), has addressed trunk detection in

9 CONCLUSIONS

In this study, we lay out the design and evaluation of a flexible research and development platform for robotic plum harvesting. An RGBD camera mounted to a UR5 cobot arm is used with an object detection, filtering, and tracking framework to locate and pick red plums. Both a hard and soft gripper are tested, with the soft gripper performing better at picking while resulting in less damaging collisions. For object detection, an HSV filter and two deep learning architectures are trialed. HSV is ineffective and RetinaNet outperforms a low power embedded YoloV3 model, but requires higher compute resources. A complex picking motion significantly outperforms a simple motion when using a soft gripper on 2D trellised plum crops. The growing conditions, primarily fruit proximity to branches, led to unanticipated results and a lower than expected success rate highlighting the importance of testing on additional tree crops to those presently in the literature. Visual servoing was found to be unnecessary, but target tracking and filtering was essential. Small changes to picking motion had critical impacts on harvesting success and must be considered in the context of gripper design.

Growers in Australia and overseas have demonstrated their willingness to adopt growing systems suitable for mechanization and eventual automation. Fruit walls appear to be a promising candidate based on this study and current results on apples. But flower thinning practices are also essential to crop success and are rarely assessed in harvesting tests, beyond the impact of bunched fruit, which remains a major problem for robotic harvesters.

Soft robotics has shown good results in building effective harvest systems that minimize fruit, tree and robot damage resulting from collisions. Long term reliability is an issue for the design presented here and will require thorough reliability testing for any commercial systems seeking to use this technology.

Next year we hope to return to the same crop to compare inter-year performance and further develop system modules specific to plum harvesting, before testing on additional fruit types. The hard-and-soft gripper distinction used here for the soft gripper, goes a long way towards addressing collisions, but a unified framework for detecting, categorizing and recovering from all collision cases in robotic harvesting is still lacking. We also hope to improve soft gripper force and robustness in further design iterations. An open problem of this system and others is the autonomous detection of difficult fruit which should be avoided in picking attempts.

ACKNOWLEDGEMENTS

The authors would like to thank Tom Ingram, Eric Dreischer, and Khalid Rafique for their tireless help on this project. Also, Yi Sun for his soft robotics expertise and assistance.

ORCID

Jasper Brown http://orcid.org/0000-0002-0843-3885

REFERENCES

Arad, B., Balendonck, J., Barth, R., Ben-Shahar, O., Edan, Y., Hellström, T., ... van Tuijl, B. (2020). Development of a sweet pepper harvesting robot. Journal of Field Robotics, 37(6), 1027–1039. https://doi.org/10.1002/rob.21937

Arad, B., Kurtser, P., Barnea, E., Harel, B., Edan, Y., & Ben-Shahar, O. (2019). Lighting and illumination independent target detection for real-time cost-efficient applications. The case study of sweet pepper robotic harvesting. Sensors, 19(6), 1390. https://doi.org/10.3390/s19061390

Bac, C., Hemming, J., van Tuijl, B. A. J. v., Barth, R., Wais, E., & van Henten, E. J. (2017). Performance evaluation of a harvesting robot for sweet pepper. Journal of Field Robotics, 34(6), 1123–1139. https://doi.org/10.1002/rob.21709
modulation of silicone-based soft pneumatic actuators. *Soft Robotics*, 4(3), 251–260. https://doi.org/10.1089/soro.2016.0047

Tanigaki, K., Fujiura, T., Akase, A., & Imagawa, J. (2008). Cherry-harvesting robot. *Computers and Electronics in Agriculture*, 63(1), 65–72. https://doi.org/10.1016/j.compag.2008.01.018

van Henten, E. J., Bac, C. W., Hemming, J., & Edan, Y. (2013). Robotics in protected cultivation. *IFAC Proceedings Volumes*, 46(18), 170–177. https://doi.org/10.3182/20130828-2-SF-3019.00070

van Henten, E. J., Hemming, J., van Tuijl, B., Kornet, J., Meuleman, J., Bontsema, J., & van Os, E. (2002). An autonomous robot for harvesting cucumbers in greenhouses. *Autonomous Robots*, 13(3), 241–258. https://doi.org/10.1023/A:1020568125418

Verl, A., Albu-Schäffer, A., Brock, O., & Raatz, A. (Eds.). (2015). *Soft robotics: Transferring theory to application*. Berlin Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-662-44506-8

Vitzrabin, E., & Edan, Y. (2016). Adaptive thresholding with fusion using a RGBD sensor for red sweetpepper detection. *Biosystems Engineering*, 146, 45–56. https://doi.org/10.1016/jbiosystemseng.2015.12.002

Vougioukas, S. G., Arikapudi, R., & Munic, J. (2016). A study of fruit reachability in orchard trees by linear over motion. *IFAC-PapersOnLine*, 49(16), 277–280. https://doi.org/10.1016/j.ifacol.2016.10.051

Wang, H., Hohimer, C. J., Bhusal, S., Karkee, M., Mo, C., & Miller, J. H. (2018). Simulation as a tool in designing and evaluating a robotic apple harvesting system. *IFAC-PapersOnLine*, 51(17), 135–140. https://doi.org/10.1016/j.ifacol.2018.08.076

Xiong, Y., Peng, C., Grimstad, L., From, P. J., & Isler, V. (2019). Development and field evaluation of a strawberry harvesting robot with a cable-driven gripper. *Computers and Electronics in Agriculture*, 157, 392–402. https://doi.org/10.1016/j.compag.2019.01.009

Zemmour, E., Kurtser, P., & Edan, Y. (2019). Automatic parameter tuning for adaptive thresholding in fruit detection. *Sensors*, 19(9) https://doi.org/10.3390/s19092130

Zhao, Y., Gong, L., Huang, Y., & Liu, C. (2016). A review of key techniques of vision-based control for harvesting robot. *Computers and Electronics in Agriculture*, 127, 311–323. https://doi.org/10.1016/j.compag.2016.06.022

How to cite this article: Brown J, Sukkarieh S. Design and evaluation of a modular robotic plum harvesting system utilising soft components. *J Field Robotics*. 2021;38:289–306. https://doi.org/10.1002/rob.21987