2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Archie Snr: A Robotic Platform for Autonomous Apple Fruitlet Thinning

Henry Williams[1*], Ans H. Qureshi[1**], David Smith[1], Trevor Gee[1],
Ben McGuinness[2], Rahul Jangali[2], Kale Black[2], Scott Harvey[2], Catherine Downes[2], Hin Lim[2],
Richard Oliver[3], Mike Duke[2], Bruce A. MacDonald[1]

***Abstract*— Apple fruitlet thinning is critical in cultivating high-quality apples, requiring an expert workforce to manage the orchard. The thinning process requires precise mapping of fruitlet clusters across the tree branches to manage the desired load for each tree. This paper presents Archie Snr, which was developed to autonomously assess the current load of the tree and thin the excess apples as an expert thinner would. The platform has been extensively evaluated in a real-world commercial orchard. The results show the platform can generate an average load count accuracy of 82.1% with a recall of 93.3%. The system was then able to successfully thin 66.14% of the fruitlets from the canopy.**

I. INTRODUCTION

In modern agriculture, the cultivation of apple orchards faces the challenge of optimising fruit yield while ensuring the quality and size of each apple. One crucial aspect of this process is fruitlet thinning, which involves removing excess young fruits to promote the development of larger, healthier apples [1].

Apple fruitlet thinning typically occurs shortly after the natural fruit set occurs, when apple trees produce an excess of small fruitlets. If left unmanaged, this abundance of fruit can lead to overcrowding, resource competition, and stunted growth, diminishing overall harvest quality. Thinning involves selectively removing a portion of these fruitlets, allowing the tree to allocate resources more efficiently to the remaining fruits. This not only enhances the size and uniformity of the apples but also prevents biennial bearing, a phenomenon where the tree alternates between heavy and light fruit production years. Fruitlet thinning is a delicate balancing act that aims to achieve an optimal fruit load, ensuring that each remaining apple receives the necessary nutrients, spacing, and sunlight for robust development.

Chemical thinning is the primary method to reduce crop load, but it is unpredictable and susceptible to adverse weather conditions [2]. Traditional manual thinning methods are labour-intensive, time-consuming, and often inconsistent between labourers [1]. As the demand for efficient and sustainable farming practices continues to rise, there is a growing need for automated apple fruitlet thinning technologies. Automation not only addresses the limitations of manual labour but also offers precision, speed, and the potential for increased overall productivity.

Working to meet this demand, the autonomous apple thinning platform Archie Snr (shown in Fig. 1) is being developed by a consortium of research institutes in New Zealand [3]. This paper presents the complete system and real-world evaluation of the platform's ability to generate accurate maps of the apple fruitlets on the tree and its ability to thin the excess fruit reliably within a real-world commercial orchard.

Fig. 1. The apple fruitlet thinning platform Archie Snr over a 2D apple tree row on the left. On the right is the internal view of one of the UR5 robotic arms mounted on the linear rail system used to scan and thin the apple trees. The Charuco board for alignment of the scans between each side is visible on the ground in the right image.

II. VIDEO DEMONSTRATION

A video demonstration of the platform can be found here[1].

III. BACKGROUND

A growing trend in the apple industry is to structure the canopy using a 2D growing system. This approach prunes the apple tree flat against a wire structure with seven to eight branches typically spaced $300\,\mathrm{mm}$ along the tree's height; an example of this is shown in Fig. 2. The reported benefits of this growing structure are improved yield quality through improved light interception across the tree [4], [5]. The primary advantage of this structure is its suitability for automation.

[1]The authors are with the Centre for Automation and Robotic Engineering Science, The University of Auckland, New Zealand.
[2]The authors are with the School of Engineering, University of Waikato, Hamilton, New Zealand.
[3]The author is with Plant and Food Research, Hamilton, New Zealand
Corresponding authors: henry.williams@auckland.ac.nz*, aqur476@aucklanduni.ac.nz**

[1]https://www.youtube.com/watch?v=K9HKhdngv5k

Fig. 2. An example of two apple tree branches following a 2D system approach. This photo is taken during the fruitlet thinning period with the fruitlets sized 5 mm to 40 mm in diameter.

Apple trees inherently produce excess fruits, necessitating proactive crop load management to attain optimal fruit size and quality [1]. An orchardist will determine a tree's total load based on age, variety, and health. The excess number of fruits must be thinned to provide adequate spacing between the remaining fruitlets to prevent overcrowding and ensure sufficient access to sunlight, nutrients, and airflow to produce high-quality apples.

Typically, a bud on an apple tree can produce between three to seven flowers, depending on the variety [6]. These flowers develop into *clusters* of apple fruitlets as shown in Fig. 3. These clusters are typically reduced to one or two fruits for adequate spacing and nutrient flow. Further fruitlets are removed if required to maintain the desired load and spacing across the branch. To automate this process, the robot platform needs to precisely map the clusters and positions of the fruitlets across each branch. This is distinctly different to estimating the yield on the tree in prior literature.

Fig. 3. An example of apple fruitlets growing in clusters together - growing from the same bud along the apple tree's branch. From left to right - a cluster of three, five, and three apples.

IV. RELATED WORK

Mechanised thinning solutions such as limb and drum shakers risk over-thinning and causing irregularities in the thinning pattern [7]. A detailed study of mechanised thinning showed that existing solutions could reduce crop load by an average of 58%, increasing the fruit size by 9% [8]. There was a noticeable decrease in the time needed for hand thinning as the shaking frequency increased. Using mechanical thinning shows the potential to save time, but it is vital to be cautious about excessively high rotor speeds. These speeds might reduce yield, cause fruits to grow too large, and lead to issues such as fruit cracking and split pits [9]. Furthermore, these approaches are non-selective, and larger fruits are more frequently removed than smaller ones. They can also cause damage to the trees. A selective means of thinning the fruitlets like an expert human is desirable.

Yield estimation attempts to measure the crop yield based on estimates of the fruit in the canopy. Automated solutions utilising advances in computer vision to detect the fruit within the canopy have seen extensive work in recent years [10], [11], [12], [13], [14]. Most work has attempted to measure the yield of a given crop through 2D images. These techniques leverage advances in deep learning for bounding box-based [15] and mask-based [16] detection methods. However, the challenge of occlusions [17] leads to an inability to accurately determine the actual count of fruits. This has led to additional work to observe the tree from multiple angles [18], [19], [20]. Nevertheless, the multi-view method can lead to over-estimation of fruit load due to counting the same fruit at different angles [21]. To address this issue, subsequent studies aimed to incorporate 3D spatial information to register fruits between frames [22], [23], [24], [25].

Work looking to thin or harvest fruit also requires the orientation of the fruit. Specific features of the fruit and the canopy can be used to approximate the orientation of the target fruit [26], [27], [14]. A notable approach involves utilising calyx [28] and stem information [29], [14] to determine the orientation of the fruit. In [14], the stem and calyx were detected in the 2D images to determine the orientation of the apple fruitlets. This approach is effective but is limited by occlusions from the canopy and other fruitlets. The work is also limited to only determining the 2D orientation of the fruitlet relative to the image frame.

A wide range of robotic fruit harvesters have been developed to locate and selectively pick individual fruit from a canopy without damage for several crops [30], [31], [32]. Robotic apple harvesters typically utilise either a grasping [30], [33] or a vacuum-based [32] end-effector to detach the apple from the tree. Each method shows promise, but the gripper-based end-effectors tend to suffer from interference from neighbouring fruit in clusters of apples or from the canopy itself. This is particularly evident for the kiwifruit harvester in [34], [31]. However, the work by [32] also shows that the quality of the canopy can still impact the harvesting success rate when utilising a suction-based end-effector with a difference of 17% between two different orchards. Given the significantly smaller size of the apple fruitlets (10 mm to 30 mm) and larger cluster sizes, we propose that a suction-based approach provides a potentially more precise means of targeting the individual fruit for thinning.

V. ARCHIE SR

Archie Snr (refer to Fig. 1) is an experimental research platform specifically created for the automated thinning of apple trees using the 2D growing system. The platform

arches over the canopy row, simultaneously operating on both sides of the tree. The platform's full dimensions are $3.8\,\mathrm{m}$ tall, $3.2\,\mathrm{m}$ wide, and $4.97\,\mathrm{m}$ long.

On either side of the platform are linear rails to facilitate the precise positioning of UR5 robotic arms at branches along the tree's height. The linear rails can move the robotic arms $1500\,\mathrm{mm}$ horizontally and vertically. This allows for comprehensive scanning and manipulation of fruit within the two-dimensional tree branch-by-branch.

The decision to straddle the apple tree serves three primary purposes. Firstly, it mitigates extreme lighting conditions, especially when the sun is low and directly faces the camera systems, causing glare and adversely affecting detection and stereo-matching performance. Secondly, straddling eliminates the impact of wind, which can disturb the canopy and disrupt the alignment of the fruitlet mapping process, also observed in [35]. Lastly, similar to the work on vines [36], straddling is essential because not all fruitlets are physically accessible or visible from a single side of the canopy.

VI. VISION SYSTEM

The vision system must generate accurate 3D maps of fruitlets along every apple tree branch for precise thinning decisions. This proves challenging owing to the dense foliage, which frequently obscures fruitlets partially or entirely from a singular viewpoint. Additionally, the tendency of fruitlets to grow in clusters further hampers visibility. The platform uses its robotic arms to scan the branch with stereo cameras to provide a comprehensive view of all the fruitlets from both sides of the canopy. This image data is processed through the flow diagram shown in Fig. 4. The full details of the vision system have previously been described in [3]; this section provides an overview of the vision system and highlights new features for the fully integrated system presented in this work, specifically clustering of the fruitlets.

A. Scanning Process

The fruitlet mapping process involves coordinating the UR5 robotic arms to capture the stereo data of the tree from various pre-planned poses. The arm scans each branch following an arc focused on the branch $300\,\mathrm{mm}$ to $400\,\mathrm{mm}$ from its centre. Five points are captured around the branch to observe fruitlets in and around the leaves. The arm scans the branch in nine arcs spaced $100\,\mathrm{mm}$ apart for 45 stereo pair images per branch.

The stereo data obtained from the platform is geometrically aligned through hand-eye calibration, which relies on the precise movements of the UR5 arms and the calibration between the arm and the stereo pair. The stereo-data between each side of the platform is aligned geometrically using a Charuco board placed at the bottom of the tree (in the future, the roof of the platform), shown in Fig. 1. This marker provides a common reference point for each side at the start of each scan.

The cameras of choice are a pair of *Basler acA2440-35uc* USB 3.0 with a resolution of 2056x2464 pixels and high dynamic range ideal for the dynamic lighting conditions in the orchard. To achieve the desired working distance of $300\,\mathrm{mm}$ to $600\,\mathrm{mm}$ for scanning the branch's width and guiding the thinning tool, Kowa lenses (Kowa F1.8, 2/3" format, focal length of $5\,\mathrm{mm}$) were employed. The cameras were positioned with a baseline of $100\,\mathrm{mm}$ to minimise stereo errors, ensuring accurate depth resolution while maintaining the required field of view.

B. Fruitlet Mapping

Fruitlet mapping processes the scan data view by view, as shown in Fig. 4. The image and depth data are used to extract the position, orientation, and size of the fruitlets from the immediate view. The fruitlet is presumed to be a 3D sphere with the position measured as the sphere's centre point. With ideal data, this should mean the centre point is the same when measured from any viewpoint. The calibrated position of the camera is then used to integrate this information between frames.

A two-stage detection process determines the fruitlet poses within an individual scan. The Detectron2[2][37] network is first used to detect instances of fruitlets in the 2D image data (Fig. 4c). The fruitlet instances are then cropped into higher-resolution sub-images for detecting the calyx with a separately trained Detectron2 network. This step produces the instance masks for the fruitlet and calyx for each apple in the images. The instance masks are then used to crop the point cloud data for each fruitlet from the depth images generated by HSMnet [38] for the stereo image pair. Fruitlets with less than 1000 points are filtered out as empirically there is too little data to measure the size of the fruitlet reliably.

The second stage first uses Random Sample Consensus (RANSAC) [39] to fit a sphere to the fruitlet depth data (Fig. 4f). In prior work [3], the sphere diameter was estimated using the average point distance approach. The radius was estimated by finding the average sum of the distance between the centre of the fruitlet point cloud and each point of the fruitlet point cloud. This size was then used as a threshold for RANSAC to fit maximum points. However, we found this approach led to overestimating the size of the smaller fruitlets in the range of $20\,\mathrm{mm}$ to $30\,\mathrm{mm}$ [3]. The approach employed in this paper utilises the distance between the two furthest points in the point cloud data. The RANSAC centroid is also adjusted to ensure that the sphere is encapsulated behind the surface of the point cloud.

The orientation of the fruitlet is then measured as the 3D gradient between the sphere's centre point and the calyx (bottom of the fruitlet), as shown in Fig. 4g. Where the position of the calyx is the average 3D position of the calyx instance. If the calyx is not detected, then the orientation of the fruitlet is not measured in this viewpoint.

Once the position, orientation, and size of the fruitlets in the current view are measured, they are integrated into the overall map of the fruitlets. Fruitlets from previous views that overlap based on their position and size are associated and placed into the overall map. At the end of the scanning

[2]https://github.com/facebookresearch/detectron2

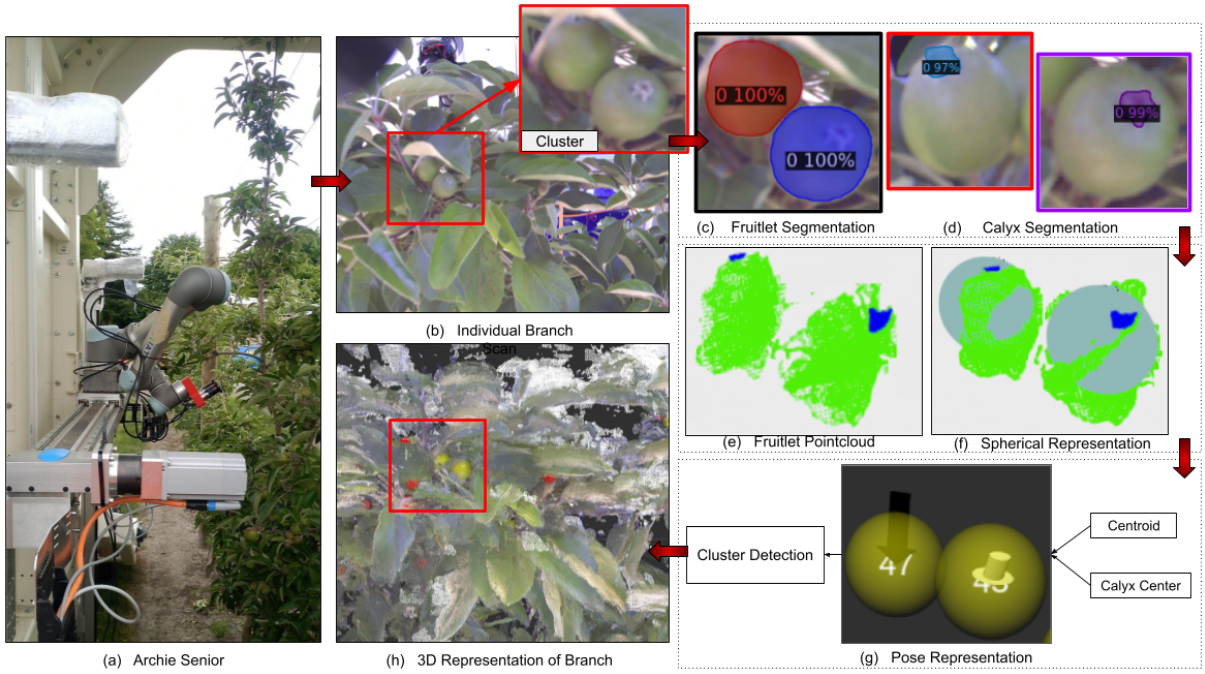

Fig. 4. Visualisation of the Fruitlet Mapping Process step by step: (a) Step 1: Archie Snr scans the branch on either side, (b) Example snapshot of a branch from one side with a single cluster to highlight following stages, (c) Step 2: Segmentation of the fruitlets from the 2D image, (d) Step 3: Calyx segmentation from the individual fruitlet segmentation, (e) Step 4: Point cloud extraction for individual fruitlets, (f) Step 5: Sphere fitting to find the size and centre of the fruitlet, (g) Step 6: Orientation measurement using centroid and calyx information, and (h) Step 7: Mapping together into a 3D map of the fruitlets.

phase, the associated fruitlets are merged, and each fruitlet's final position, size, and orientation are found by averaging the associated views together.

In prior work [3], the fruitlet size from each viewpoint was equally weighted when merging the associated fruitlets. It was observed in further investigations that viewpoints that only had a limited view of the fruitlet consequently had less point cloud data to fit the sphere and thus tended to provide less accurate size estimates. This paper introduces a weighted average based on the point cloud size representing the fruitlet in each view. This approach biases towards views with more direct views of the fruitlets while not entirely discarding size estimates of the fruitlets from indirect views. The weighting factor is the normalised size for each point cloud.

Clustering is the final step once all the views have been merged and is a new feature of the system. A naive approach based on just the position of the fruitlet is insufficient, and direct observation of the bud the fruitlet is growing on is a significant challenge, even with a multi-view approach. The position of the bud each fruitlet is connected to is estimated based on their orientation and the average length of the stem for the given variety. A straight line of half the fruitlets' size length plus the average length of the stem ($20\,\mathrm{mm}$) is drawn from the centre point along the fruitlets' orientation. Fruitlets with bud predictions within $20\,\mathrm{mm}$ of each other are considered to be growing from the same bud. This process only works for fruitlets detected with an orientation. Fruitlets that have failed to be mapped with an orientation are associated with fruitlet clusters if they are within $30\,\mathrm{mm}$ of a predicted bud location. A fruitlet lacks an orientation when the system does not view the fruitlets' calyx.

VII. Automated Thinning

An orchardist will specify the number of fruitlets that must be removed from all the trees to meet a desired load. The human thinners will then follow specific rules based on the orchardist's preference for the given variety to determine which fruitlet to remove. Typically, they remove damaged, smaller, or less coloured fruitlets while retaining the larger and prominently coloured fruitlets within each cluster. On a 2D apple tree, this process is applied branch by branch with the assumption this will meet the desired load for each tree. The decision-making process is intended to follow this process; for this paper, the system targeted the smallest fruitlet in each cluster.

The robotic arms are equipped with custom end-effectors, shown in Fig. 5, designed to remove an individual fruitlet from a cluster. The end-effector latches onto the fruit through suction and twists it off its stem. Suction is achieved using a 1250W Ryobi RVC-1220I-G[3] vacuum cleaner with the nozzle rotation controlled by a DC motor operating at $21.0\,\mathrm{rad\,s^{-1}}$ (200 RPM).

Taking advantage of the relatively flat 2D structure of the apple tree, the planning presumes the arm can move directly towards the selected fruitlet. During the thinning process, the thinning tool is first placed $200\,\mathrm{mm}$ from the fruitlet. The

[3]https://www.ryobi.co.nz/products/details/1250w-20l-wet-dry-vacuum-rvc-1220i-g

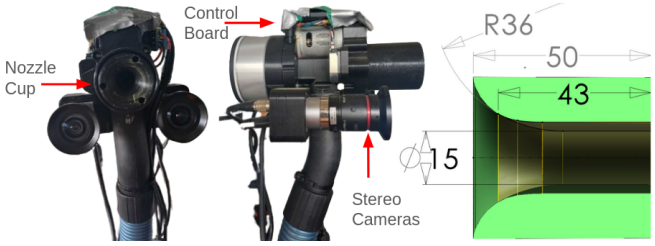

Fig. 5. The custom end-effector for apple thinning utilises a rotating suction nozzle to latch onto and twist fruitlets off the canopy. The left shows a front view of the end-effector, while the middle image shows the side view. The dimensions of the inner rotating nozzle cup are shown on the right, with dimensions in millimetres.

thinning tool is then moved to the desired fruitlet. During this movement, the suction and rotation are activated. The thinning tool is then retracted back 200 mm from the tree at an angle of 45 degrees to drop the fruitlet away from the tree by deactivating the suction and rotation. This process is then repeated for the next fruitlet. The MoveIt[4] library is utilised for planning the pruning tool between poses.

VIII. RESULTS

The platform was evaluated in an active real-world commercial apple orchard in Hastings, New Zealand, during the 2023 thinning season. The evaluation encompassed a comprehensive and realistic field trial, comparing the system's performance metrics to ground truth measurements. Before our trials, the trees were not thinned, trimmed, or altered in any way. Data collection took place over two weeks, with the robot operating during daylight hours under varying lighting conditions.

Fruitlet mapping was conducted by manually driving the platform over a tree and placing the scanning arms at the centre of the branches. The tree branches were scanned individually to produce a map of their fruitlets. Once completed, the robot was driven forward to the next tree in the row, repeating this process as often as possible.

A. Fruitlet Counting

The accuracy of the fruitlet count was evaluated by manually comparing the generated map of the fruitlets against the individual branches. The fruitlets in the map produced by the robot were marked as true or false positives by directly comparing the 3D map on a laptop screen against the branch. False negatives were manually added to the map where a fruitlet was not correctly identified. The size and orientation of the fruitlet and its associated cluster were also evaluated. An example of this evaluation method is shown in Fig. 6. This approach has the potential for human error in associating the digital map and 2D images with the real tree. However, it is the best means of validating the system, and a team of four put extensive effort into validating the maps over two weeks.

[4]https://moveit.ros.org/

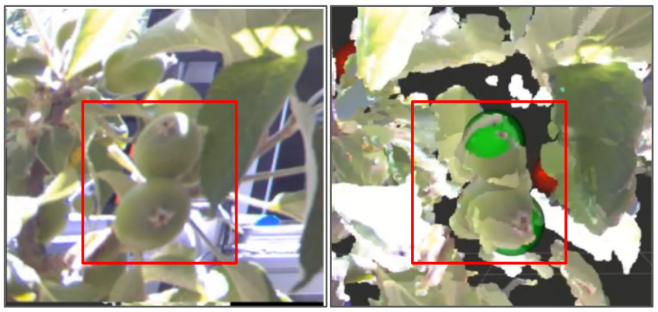

Fig. 6. Example section of a 3D map of fruitlets represented by green and red spheres with the point cloud of the canopy generated by the robot (right) compared against the real-world image of the fruitlets (left). The red box highlights the detected fruitlets with the real fruitlets. Note: The 3D map lets the user look in and around the point cloud. The point cloud can also be removed to make it easier to view the detected fruitlets as shown in the video provided in Section II.

In total, 30 individual branches were scanned on both sides by the Archie Snr platform over 10 trees with 1891 fruitlets mapped. The performance of the mapping system is shown in table I. There were only 127 false negatives and 257 false positives throughout the 30 branches. This demonstrates the effectiveness of the system in re-associating the fruitlets. The source of mapping errors arises from inaccuracies in several places: stereo calibration, hand-eye calibration, marker alignment, and depth inference. Improvements to these areas are non-trivial but will improve the accuracy of aligning the scan data from each side of the platform. Furthermore, investigating additional methods of local feature matching may help improve associations throughout the scan data.

TABLE I
FRUITLET LOAD RESULTS

Fruitlets	Accuracy	Precision	Recall
1891	82.1%	87.3%	93.3%

B. Size Measurements

The diameter of 157 fruitlets randomly spread throughout the scanned trees was manually measured with callipers and compared to the system's estimate. The fruitlets averaged 26 mm in diameter. The system showed a Mean Average Error (MAE) of only 4.4 mm across the 157 fruitlets. This is higher than the previous method's MAE results at 3.3 mm in the previous evaluation [3]. However, applying the new method on the prior years' scans shows an improved MAE of 2.1 mm.

The time between the platform scanning the tree and the physical measurement of the fruitlets in the latest trial is a factor. The physical measurements were taken a day after the scanning. This was to associate the fruitlet with its mapped position in the map and to limit access to the row while the robot operated for health and safety. The fruitlets continued to grow in that period, up to a potential increase of 1 mm in optimal conditions [40], [41]. Of the measured fruitlets, 93.0% had their diameter underestimated to their measured

size, supporting this conclusion. Despite this, the approach would still tend to underestimate the size of the fruitlets. It is suspected the primary cause of the bias is the sphere fitting method being used, but further investigation is required as there are numerous sources of potential error. Overall, the accuracy of the size measurements is consistent with recent work using stereo cameras to estimate the size of apples at various stages of maturity [41] and would be sufficient for distinguishing the relatively smaller fruitlets for thinning.

C. Orientation Measurements

The orientation was evaluated by manually validating the fruitlets' measured 3D orientation aligned with the calyx of the fruitlet. This is a binary measurement of correct or incorrect. Fig. 7 shows an example of the measured orientation compared to the orientation of the real fruitlet.

The results of the orientation performance are shown in table II. Overall, 80.8% of the measured orientations correctly aligned with the fruitlet. The issue is that only 56.1% of the evaluated fruitlets had an orientation estimate. This was caused by the system not detecting an associated calyx for those fruitlets. The primary reason for this, when examining the scan data, was that the calyx was never visible in any of the views of the fruitlet. Exploring denser scan patterns or intelligently placing the camera to find the calyx for each fruitlet may help reduce this issue. Additionally, future work can extract additional features from the canopy itself, such as the branches, to assist with determining the orientation of the fruitlets [26], [29]. The errors when a calyx was detected were typically from inaccuracies in the placement of the centroid of the fruitlet with fruitlets with limited point cloud data to fit the spheres accurately.

TABLE II
ORIENTATION ESTIMATION RESULTS

Fruitlets	892	100.0%
Orientation Measured	500	56.1%
Correct Orientation	404	80.8%

D. Clustering Measurements

The clustering performance was evaluated by manually validating that a fruitlet was associated with the correct cluster. Fig. 7 shows an example of the measured clustering association compared to the cluster of the fruitlets. The results of the ground-truth clustering performance are shown in table III. The results are divided to show the orientation accuracy for only fruitlets with a measured orientation and all fruitlets mapped by the system. This provides insight into whether the naive clustering association negatively impacts the overall performance. The system has a clustering precision of 84.1% when assigning all fruitlets into clusters with a recall of 93.4%. Assigning the fruitlets without position increases the precision but does allow all fruitlets to be clustered, which reduces the recall value. The primary issue is the limited number of fruitlets captured with an accurate orientation estimate, as shown in table II. Improvements to the orientation and position estimates will enable better prediction of the position of the bud locations to associate the fruitlets into clusters correctly.

TABLE III
CLUSTERING RESULTS

Evaluation	Clusters	Fruitlets	Precision	Recall
Orientation Only	74	213	88.3%	85.5%
Orientation and Position	74	251	84.1%	93.4%

E. Automated Thinning

Automated thinning tests were conducted by mapping a branch on the tree and then manually selecting the smallest fruitlet of each cluster in the map. The robot then attempted to thin each fruitlet one after the other. The outcome of the thinning attempt was then recorded as a success or failure, with the reason for failure being categorised as either:

- Incorrect Pose: the end-effector was not correctly aligned to the centre of the fruitlet due to errors in its pose estimate.
- Obstruction: the end-effector was obstructed by the canopy (e.g. leaf or branch) or another fruitlet.
- Failed to Attach: the suction was insufficient to latch onto the fruitlet to remove it from the canopy.
- Failed to Detach: the fruitlet was sucked into the end-effector but was unable to twist off the stem.

In total, 14 branches across six trees were thinned by the robotic platform attempting to remove 127 fruitlets. The outcome of these attempts is shown in Table IV. The video in Section II shows the end-effector operating as desired and examples of the categorised failure cases.

TABLE IV
AUTOMATED THINNING RESULTS

Success Total	84	66.14%
Fail Total	43	33.86%
Incorrect Pose	19	44.19%
Obstruction	12	27.91%
Failed to Attach	10	23.26%
Failed to Detach	2	4.65%

The primary cause of failure was placing the end-effector slightly off the fruitlets' centre. This offset error is a product of the errors in the calibration of the stereo cameras, the hand-eye calibration, and the estimated position of the fruitlet. Qualitatively, the offset was typically only $10\,\mathrm{mm}$ to $15\,\mathrm{mm}$ from being placed at the centre of mass of the fruitlet. The suction from the end-effector could account for positional errors less than this but could not complete a seal around fruitlets too far off the centre of the cup. Neighbouring fruitlets (typically in the same cluster) were the primary cause of obstructions and failures to attach. The end-effector would push the obstructing fruitlet, moving the target fruitlet out of the way or preventing the end-effector from forming a proper seal. Overall, the thinning results also exemplify the validity of the vision system as the end-effector was able to be correctly placed on 66.14% of the fruitlets.

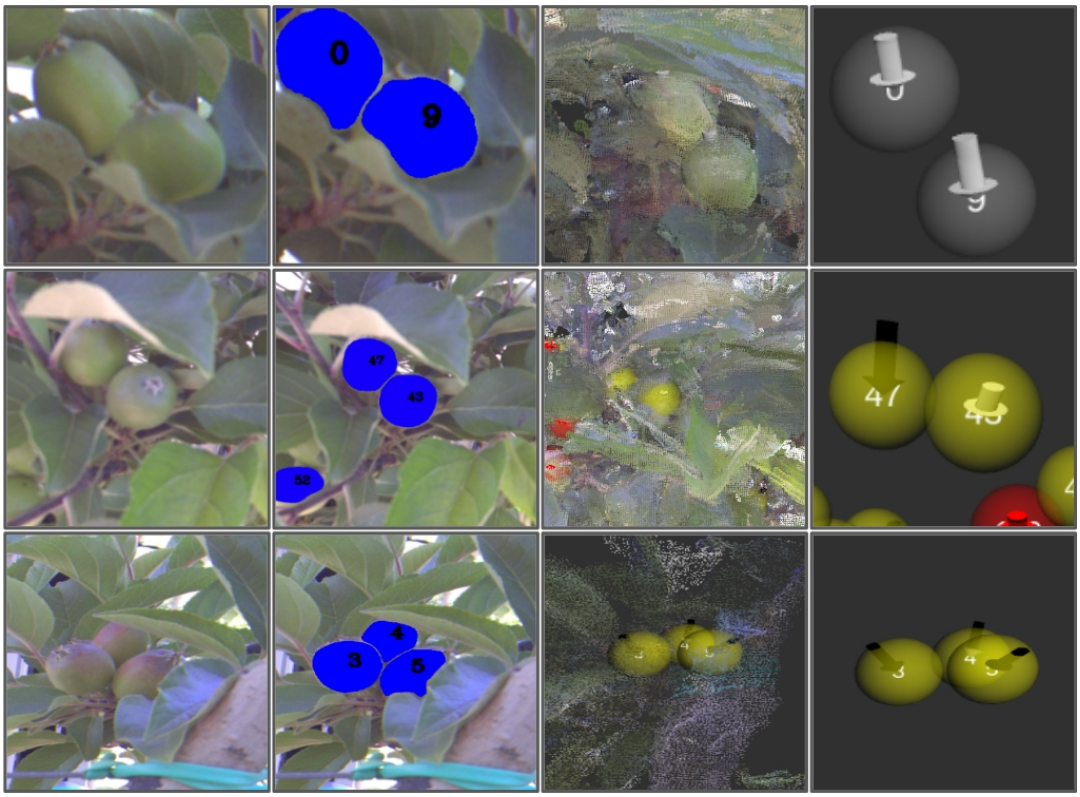

Fig. 7. From left to right: Raw RGB image of a single apple fruitlet cluster (cropped from the full image), semantic detection and identification of the fruitlets in the RGB image, fruitlets as represented in the 3D map of the branch (yellow spheres), and the isolated fruitlet cluster with orientation measured (arrows represent the direction of the stem).

IX. DISCUSSION

The evaluation of the autonomous apple thinning robot Archie Snr demonstrates the potential for automation of this critical task. The large-scale field trial mapping over 1800 fruitlets has exhibited the approach's effectiveness in determining the position, orientation, and clustering of fruitlets along the tree branches. The results presented in this paper align with prior evaluations [3] even on trees in different locations and years. This signifies that the approach can scale and generalise across orchards. However, further work needs to refine the quality of the depth information to improve the estimates of the fruitlet centre of mass.

The automated thinning results denote the successful integration of the end-effector with the vision system. The failure cases indicate clear areas for future improvements through modifications to the design of the end-effector and refined calibration methods. Arguably, a more complex planning algorithm may be able to approach the fruitlets from different angles to avoid obstructions from the canopy or neighbouring fruitlets. However, the space constraints around the clusters in the canopy severely limit the number of alternative approach angles. This would also slow the system down, requiring complex motion planning and actuation. We argue that a more precise end-effector design is the better solution to this problem but would require more accurate pose estimates. A larger opening of the suction area may allow for improved compliance around positional errors but may increase failures around obstructions with the canopy. Additional work will investigate a soft end-effector with a smaller surface area that can conform around the fruitlet to account for such positional errors with a lower surface area.

X. CONCLUSIONS

This paper presents Archie Snr, the autonomous apple thinning robotic platform. The platform has been extensively evaluated in a real-world commercial orchard without any modifications or pre-thinning of the trees. The results have shown the effectiveness of the vision system for generating accurate maps of the apple fruitlets (82.1% over 1891 fruitlets), including measuring their orientation (80.8%) and associated clusters (84.1%). Using the custom vacuum-based end-effector, the platform also successfully thinned 66.14% of the apple fruitlets. Future work will further refine the accuracy of the vision system and explore compliant end-effectors to improve the tolerance to positional errors in the measured position of the fruitlets.

ACKNOWLEDGMENT

This research was supported by the New Zealand Ministry for Business, Innovation and Employment (MBIE) on contract UOAX1810.

REFERENCES

[1] G. Costa, M. Blanke, and A. Widmer, "Principles of thinning in fruit tree crops-needs and novelties," in *EUFRIN Thinning Working Group Symposia 998*, 2012, pp. 17–26.

[2] T. Robinson, A. Lakso, D. Greene, G. Reginato, and A. De R. Rufato, "Managing fruit abscission in apple," in *XXIX International Horticultural Congress on Horticulture: Sustaining Lives, Livelihoods and Landscapes (IHC2014): 1119*, 2014, pp. 1–14.

[3] A. Qureshi, D. Smith, T. Gee, M. Nejati, J. Shahabi, J. Lim, H. S. Ahn, B. McGuinness, C. Downes, R. Jangali, *et al.*, "Seeing the fruit for the leaves: Robotically mapping apple fruitlets in a commercial orchard," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 3234–3239.

[4] D. Tustin, "Orchard systems for the 21st century: perspectives, considerations and critique," in *XII International Symposium on Integrating Canopy, Rootstock and Environmental Physiology in Orchard Systems 1346*, 2021, pp. 195–206.

[5] D. Tustin, K. Breen, and B. van Hooijdonk, "Light utilisation, leaf canopy properties and fruiting responses of narrow-row, planar cordon apple orchard planting systems—a study of the productivity of apple," *Scientia Horticulturae*, vol. 294, p. 110778, 2022.

[6] J. Jakopic, A. Zupan, K. Eler, V. Schmitzer, F. Stampar, and R. Veberic, "It's great to be the king: Apple fruit development affected by the position in the cluster," *Scientia Horticulturae*, vol. 194, pp. 18–25, 2015.

[7] M. Lopes, P. D. Gaspar, and M. P. Simões, "Current status and future trends of mechanized fruit thinning devices and sensor technology," *Proceedings of the World Academy of Science, Engineering and Technology*, 2019.

[8] J. Schupp, T. A. Baugher, S. Miller, R. Harsh, and K. Lesser, "Mechanical thinning of peach and apple trees reduces labor input and increases fruit size," *HortTechnology*, vol. 18, no. 4, pp. 660–670, 2008.

[9] W. Steyn, K. Theron, and M. De Villiers, "The effect of mechanical bloom thinning with the darwin 300tm, on the hand thinning time, yield, and fruit quality in'zephyr'nectarine," in *XXIX International Horticultural Congress on Horticulture: Sustaining Lives, Livelihoods and Landscapes (IHC2014): 1130*, 2014, pp. 631–638.

[10] B. Sagar and N. Cauvery, "Agriculture data analytics in crop yield estimation: a critical review," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 12, no. 3, pp. 1087–1093, 2018.

[11] A. Koirala, K. B. Walsh, Z. Wang, and C. McCarthy, "Deep learning–method overview and review of use for fruit detection and yield estimation," *Computers and electronics in agriculture*, vol. 162, pp. 219–234, 2019.

[12] T. Van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020.

[13] L. Droukas, Z. Doulgeri, N. L. Tsakiridis, D. Triantafyllou, I. Kleitsiotis, I. Mariolis, D. Giakoumis, D. Tzovaras, D. Kateris, and D. Bochtis, "A survey of robotic harvesting systems and enabling technologies," *Journal of Intelligent & Robotic Systems*, vol. 107, no. 2, p. 21, 2023.

[14] M. Hussain, L. He, J. Schupp, D. Lyons, and P. Heinemann, "Green fruit segmentation and orientation estimation for robotic green fruit thinning of apples," *Computers and Electronics in Agriculture*, vol. 207, p. 107734, 2023.

[15] D. Wang and D. He, "Channel pruned yolo v5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning," *Biosystems Engineering*, vol. 210, pp. 271–281, 2021.

[16] J. Gené-Mola, M. Ferrer-Ferrer, E. Gregorio, P. M. Blok, J. Hemming, J.-R. Morros, J. R. Rosell-Polo, V. Vilaplana, and J. Ruiz-Hidalgo, "Looking behind occlusions: A study on amodal segmentation for robust on-tree apple fruit size estimation," *Computers and Electronics in Agriculture*, vol. 209, p. 107854, 2023.

[17] S. Bargoti and J. P. Underwood, "Image segmentation for fruit detection and yield estimation in apple orchards," *Journal of Field Robotics*, vol. 34, no. 6, pp. 1039–1060, 2017.

[18] J. Moonrinta, S. Chaivivatrakul, M. N. Dailey, and M. Ekpanyapong, "Fruit detection, tracking, and 3d reconstruction for crop mapping and yield estimation," in *2010 11th International Conference on Control Automation Robotics & Vision*. IEEE, 2010, pp. 1181–1186.

[19] H. Mirhaji, M. Soleymani, A. Asakereh, and S. A. Mehdizadeh, "Fruit detection and load estimation of an orange orchard using the yolo models through simple approaches in different imaging and illumination conditions," *Computers and Electronics in Agriculture*, vol. 191, p. 106533, 2021.

[20] R. R. D. Abeyrathna, V. M. Nakaguchi, A. Minn, and T. Ahamed, "Recognition and counting of apples in a dynamic state using a 3d camera and deep learning algorithms for robotic harvesting systems," *Sensors*, vol. 23, no. 8, p. 3810, 2023.

[21] A. B. Payne, K. B. Walsh, P. Subedi, and D. Jarvis, "Estimation of mango crop yield using image analysis–segmentation method," *Computers and electronics in agriculture*, vol. 91, pp. 57–64, 2013.

[22] Q. Wang, S. Nuske, M. Bergerman, and S. Singh, "Automated crop yield estimation for apple orchards," in *Experimental robotics*. Springer, 2013, pp. 745–758.

[23] Y. Song, C. Glasbey, G. Horgan, G. Polder, J. Dieleman, and G. Van der Heijden, "Automatic fruit recognition and counting from multiple images," *Biosystems Engineering*, vol. 118, pp. 203–215, 2014.

[24] M. Stein, S. Bargoti, and J. Underwood, "Image based mango fruit detection, localisation and yield estimation using multiple view geometry," *Sensors*, vol. 16, no. 11, p. 1915, 2016.

[25] T. Li, Q. Feng, Q. Qiu, F. Xie, and C. Zhao, "Occluded apple fruit detection and localization with a frustum-based point-cloud-processing approach for robotic harvesting," *Remote Sensing*, vol. 14, no. 3, p. 482, 2022.

[26] G. Lin, Y. Tang, X. Zou, J. Xiong, and J. Li, "Guava detection and pose estimation using a low-cost rgb-d sensor in the field," *Sensors*, vol. 19, no. 2, p. 428, 2019.

[27] R. Gao, Q. Zhou, S. Cao, and Q. Jiang, "An algorithm for calculating apple picking direction based on 3d vision," *Agriculture*, vol. 12, no. 8, p. 1170, 2022.

[28] J. Zhou, G. Zhang, R. Liu, Y. Jin, *et al.*, "Apple attitude estimation based on particle filter for harvesting robot." *Nongye Jixie Xuebao= Transactions of the Chinese Society for Agricultural Machinery*, vol. 42, no. 3, pp. 161–165, 2011.

[29] R. Barth, J. Hemming, and E. J. Van Henten, "Angle estimation between plant parts for grasp optimisation in harvest robots," *Biosystems Engineering*, vol. 183, pp. 26–46, 2019.

[30] A. Silwal, J. R. Davidson, M. Karkee, C. Mo, Q. Zhang, and K. Lewis, "Design, integration, and field evaluation of a robotic apple harvester," *Journal of Field Robotics*, vol. 34, no. 6, pp. 1140–1159, 2017.

[31] H. Williams, C. Ting, M. Nejati, M. H. Jones, N. Penhall, J. Lim, M. Seabright, J. Bell, H. S. Ahn, A. Scarfe, *et al.*, "Improvements to and large-scale evaluation of a robotic kiwifruit harvester," *Journal of Field Robotics*, vol. 37, no. 2, pp. 187–201, 2020.

[32] K. Zhang, K. Lammers, P. Chu, Z. Li, and R. Lu, "An automated apple harvesting robot—from system design to field evaluation," *Journal of Field Robotics*, 2023.

[33] L. Bu, C. Chen, G. Hu, A. Sugirbay, H. Sun, and J. Chen, "Design and evaluation of a robotic apple harvester using optimized picking patterns," *Computers and Electronics in Agriculture*, vol. 198, p. 107092, 2022.

[34] H. A. Williams, M. H. Jones, M. Nejati, M. J. Seabright, J. Bell, N. D. Penhall, J. J. Barnett, M. D. Duke, A. J. Scarfe, H. S. Ahn, *et al.*, "Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms," *biosystems engineering*, vol. 181, pp. 140–156, 2019.

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

[36] 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*, vol. 235, pp. 31–49, 2023.

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

[38] G. Yang, J. Manela, M. Happold, and D. Ramanan, "Hierarchical deep stereo matching on high-resolution images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5515–5524.

[39] H. Cantzler, "Random sample consensus (ransac)," *Institute for Perception, Action and Behaviour, Division of Informatics, University of Edinburgh*, vol. 3, 1981.

[40] M. P. DENNE, "The growth of apple fruitlets, and the effect of early thinning on fruit development," *Annals of Botany*, pp. 397–406, 1960.

[41] O. Mirbod, D. Choi, P. H. Heinemann, R. P. Marini, and L. He, "On-tree apple fruit size estimation using stereo vision with deep learning-based occlusion handling," *Biosystems Engineering*, vol. 226, pp. 27–42, 2023.