Robotic Kiwifruit Harvesting using Machine Vision, Convolutional Neural Networks, and Robotic Arms

Henry Williams[a,*], Mahla Nejati[a], Jamie Bell[a], Nicky Penhall[a], Ho Seok Ahn[a], JongYoon Lim[a], Bruce MacDonald[a], Mark H. Jones[b,**], Matthew Seabright[b], Josh Barnett[b], Mike Duke[b], Alistair Scarfe[c]

[a] *Faculty of Engineering, University of Auckland, Auckland, New Zealand*
[b] *School of Engineering, University of Waikato, Hamilton, New Zealand*
[c] *Robotics Plus Ltd, Newnham Innovation Park, Tauranga, New Zealand*

Abstract

As labor requirements in horticultural increase, so too does the feasibility of increased automation in these industries. This paper presents a performance evaluation of a kiwifruit harvesting robot designed to operate autonomously in pergola style orchards. The robot consists of four harvesting arms, end-effectors designed specifically for kiwifruit detachment, and a machine vision system employing convolution neural networks. Performance evaluations are presented for the harvester as a whole, as well as the machine vision system. We show the system as a whole is capable of harvesting over half of all fruit within three test orchards, equating a substantial reduction in peak harvesting labor requirements.

Key words: Horticulture, Robotics, Neural Networking, Machine Vision, Harvesting, Convolution Neural Networks, Orchard

2017 MSC: 00-01, 99-00

1. Introduction

The New Zealand government is targeting a two-fold increase in primary exports for the thirteen year period ending 2025 [1]. Pressure from this growth

*Henry.Williams@auckland.ac.nz
**markj@waikato.ac.nz

is driving demand for new technologies and approaches to handle the increasing demand on manual labor forces to ensure high quality exports. Research presented here focuses on kiwifruit, which represents the largest share of New Zealand's exports by value and 30% of the market-share globally.

Kiwifruit are harvested in New Zealand from late March to mid-June and currently demand upwards of 2,500 people over a 50 day period. Timely, efficient, and careful harvesting of this fruit is critical to ensure optimum returns with export quality fruit. The kiwifruit industry, like most horticultural sectors, struggles to attract and retain laborers, especially during seasonal high demand periods. To meet peak labor demands of the harvesting and pollination seasons, migrant workers are often sourced through government assisted work programmes. However, relying on these seasonal workers can pose reliability, quality, and socioeconomic challenges.

Development of robotic solutions to assist manual labor is emerging as a strategic necessity to harvest desired yields at a high quality within the short harvesting time-frame. This is reflected by the investment programmes the New Zealand government is engaging in to fund research with the potential to increase those export figures. Furthermore, automated harvesting systems have the potential to provide more reliable and consistent harvest quality over their human counterparts.

Kiwifruit are most commonly grown in pergola style frames with rows separated by 3 m to 5 m. This creates a relatively flat overhead canopy that sits approximately 1.7 m above the ground. Fruit generally hang below the leafy canopy, but are also able to grow within the canopy area where obstructions are common. Such obstructions may be a cane of the plant itself, wires used to tie the canopy down, or the cross-beams that hold the canopy up.

The project to develop the harvester (along with the base platform and pollinator) is jointly funded by the New Zealand Government, and industry investment. It is a collaboration between two universities, a government funded research organisation, and an early stage technology development company. A photo of the developed harvesting system is presented as figure 1.

This system combines Convolutional Neural Networks (CNN), stereo vision, and robotic harvesting arms on a robotic platform as a means of harvesting individual fruit. An end-effector that mimics the harvesting action of a human is used to ensure that fruit detachment minimises force and promotes detachment of the fruit from its stem. Each arm has been designed to pick continuously at a rate of one fruit per second. It is expected that the system will be unable to harvest all fruit in a given canopy area due to obstructions and/or occlusion from the canopy itself.

Figure 1: The robotic harvesting system sitting under a kiwifruit canopy.

2. Related Work

Traditionally, object recognition for machine vision applications is based on combinations of processes such as thresholding, masking, colour segmentation, edge detection, and filtering. These approaches work well when lighting is predictable, the object's appearance is well defined, and occlusions are minimal, such as in consumer-goods factories. However, in the agricultural environment it is often the case that none of these conditions are met. With regards to detecting kiwifruit, cameras face a canopy which is often back-lit by the sun,

kiwifruit grow in a number of shapes, hang on various angles, and be obstructed by canes and other kiwifruit. Harvesting these fruit complicates matters further by adding the need for a robotic positioning system, a suitable end-effector, precise localisation of target fruit, and obstacle detection.

Research toward overcoming these problems is gaining in popularity as demand for agricultural automation increases globally [2, 3, 4, 5, 6]. A comprehensive review on harvesting systems investigated over 50 projects reported in the last three decades [4]. It was found that on average a system would locate 85%, detach 75%, harvest 66%, and damage 5% of the fruit.

A key challenge to convert this research into a commercial product is developing a system with a commercially viable operational time. Currently an average cycle time of 33 s is being reported [4]. This 'cycle time' refers to the time to complete an average full harvest operation. This may include ripeness determination (if used), localization, fruit detachment, transport of detached fruit, wasted time from failed pick attempts, and movements between fruit. This indicator is relevant to determine the economic feasibility of the robot.

A kiwifruit harvesting system intended to be capable of picking four kiwifruit per second has previously been reported [7], but the final system presented in that work was unable to meet the target figures. The work presented in this paper is a continuation of the earlier work by Scarfe et. al that attempts to meet those targets.

2.1. Fruit Detection

Accurate fruit detection and locating is currently considered the biggest hurdle for development of commercial level harvesting systems [8]. Detection systems should be capable of dealing with variations in fruit shape and lighting, as well coping with clustering and occlusion of fruit.

A wide range of sensor types have been investigated for a variety of fruit types in an attempt to overcome these issues. An extensive review of fruit detection systems can be found in [8].

The most common approach to fruit detection utilises colour cameras for

fruit detection and stereo cameras to determine position. Colour images provide a range of information for detection including colour, geometric, and textural information about the fruit [9]. However, uncontrolled lighting conditions make it difficult to develop robust detection approaches using traditional computing algorithms [8].

Soft computing methods are often used to perform modeling and analysis of complex problems, and to provide solutions, which are tolerant to imprecision, uncertainty, partial truth and approximation [10]. This makes them ideal for the detection of fruit. Work has even showing the feasibility for using soft computing methods for the detection of kiwifruit [5], however it is limited for use in night time conditions where the light can be controlled.

Recent advancements in Convolutional Neural Networks (CNNs) have shown improvements in classification accuracy and robustness [11]. Having been utilised in autonomous cars [12] for detection in uncontrolled environments, they could provide a robust and effective means of fruit detection.

2.1.1. Convolutional Neural Networks

In machine learning, a CNN is a type of feed-forward neural network for analyzing visual imagery developed in the 1990s [13, 14]. Recent developments by Krizhevsky et al. in 2012 [15] have rekindled interest in CNNs by showing substantially higher image classification accuracy on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [11].

A typical CNN is comprised of one or more convolutional layers (often with a subsampling step) and then followed by one or more fully connected layers as in a standard multi-layer neural network [15]. The architecture of a CNN is designed to take advantage of the 2D structure of an input image (or other 2D input such as a speech signal). This is achieved with local connections and tied weights followed by some form of pooling which results in translation invariant features. One benefit of CNNs relative to fully connected networks is a reduced number of input parameters, leading to faster training times.

Another advantage of CNNs, and deep learning in general, is that the self-

generated feature extraction model only requires ground-truth images of the desired operating environment. Given images in a wide range of operating conditions, the system can learn a robust means of classifying objects. One downside however is the time and computing power required to train the model and the gathering of the ground truth data itself.

2.2. Fruit detachment methods

Once the fruit is located in the canopy, the next challenge is picking the fruit without causing damage to the fruit or canopy. Bulk harvesting approaches, such as shaking techniques [16, 17], are already used in the industry. Such techniques are unsuitable for harvesting fruit that are further processed before being sold to customers, such as fresh kiwifruit. In the case of kiwifruit the stem of a kiwifruit can damage other fruit when packed in a bin it is important for the stems to be detached during the harvest process. Furthermore, care must be taken to prevent bruising or piercing to the kiwifruit.

Mechanisms for selectively picking the fruit are required where the fruit must be handled with care to ensure the quality of the fruit. To achieve this, picking arms with gripping hands [18] or even vacuum based methods [19] have been developed. Both approaches have shown the ability to effectively harvest apples [2, 20], pears [20], grapes [21], and kiwifruit [3, 7] without harm.

2.3. Base Robotic Platform

In addition to the kiwifruit harvesting unit, a base robotic platform has been developed that is capable of moving the harvesting unit through an orchard. This platform is a second generation unit of what has previously been published by [7]. Modularity has been increased over the previous design, allowing the unit to be used for both pollination and harvesting. It is electrically driven with four in-wheel hub motors that allow the harvesting unit to be moved with greater precision than a hydraulic or combustion driven system. Between each harvesting cycle the harvester must move forward approximately 300 mm – this equates to a high number of stop-start cycles which the electric drive is

well suited to. This base platform contains all the power conversion necessary to power the harvesting module electrically as well as provide compressed air. Further details of this platform and its specifications will be published separately [22].

3. Harvesting System

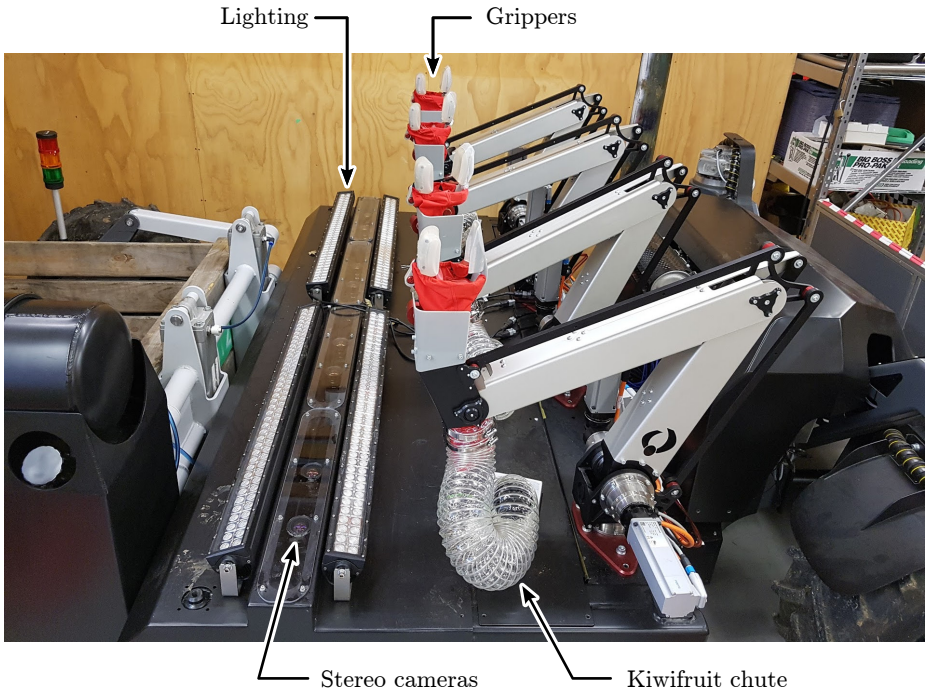

Figure 2: Photo showing the harvester as mounted on the base robotic platform.

Figure 2 shows the arrangement of arms, cameras, lights and fruit chutes on the harvesting system. The overall system is comprised of four sub-units spaced across the width of the base platform. Each sub-unit consists of a servo driven three-axis articulated arm, servo controllers, computing hardware, and stereo camera pair. The computers in each sub-unit are linked via ethernet connection which allows them to be coordinated with one another. The module as a whole receives electrical power and pneumatic pressure from the base robotic platform, which also provides the ability to cut power in an emergency.

Conceptually, the harvester can be broken down into the following sub-systems: machine vision for object recognition, stereo depth calculation, a pick-order scheduler, arm path-planning and servo control, and fruit grip-and-detach.

Functionally, the harvester takes an image of the canopy which is processed by the vision system to detect fruit in the image and locate their positions in 3D space. The system then determines the order in which it will pick the located fruit. Next, optimised movement profiles are generated and streamed to each of the servo-motor controllers. When in position, the end-effector is actuated so as to detach the target fruit from the canopy. Finally, the harvested fruit roll down a chute into a conveyor contained within the harvester and the arm moves onto the next fruit in the schedule. When the schedule is empty the arms return to a home position and the systems starts again, this process is referred to as a sub-cycle.

The system repeats these sub-cycles until the system determines there are no kiwifruit available to harvest, or is unable to locate any harvestable fruit. The base platform then advances 300 mm completing a full harvesting cycle. A new harvesting cycle then begins and continues until the harvester has completed the entire orchard.

3.1. Machine Vision

The vision system needs to be capable of accurately locating the position of the kiwifruit to correctly place the hand for a successful pick. It may also detect obstructions as a means of protecting itself and reducing damage to the canopy. Furthermore, it is required to operate in a wide range of lighting and weather conditions.

To achieve this the presented system utilities a CNN[1] to perform semantic segmentation [23, 24] on images of the canopy. The network was trained on a total of 63 hand-labeled images collected across a range of conditions and locations. These images varied by time of day, the camera settings used, and

[1] https://devblogs.nvidia.com/parallelforall/image-segmentation-using-digits-5/

location. Each were labeled for kiwifruit calyxes, canes, and guide wires. An example of the network's output is shown in figure 3.

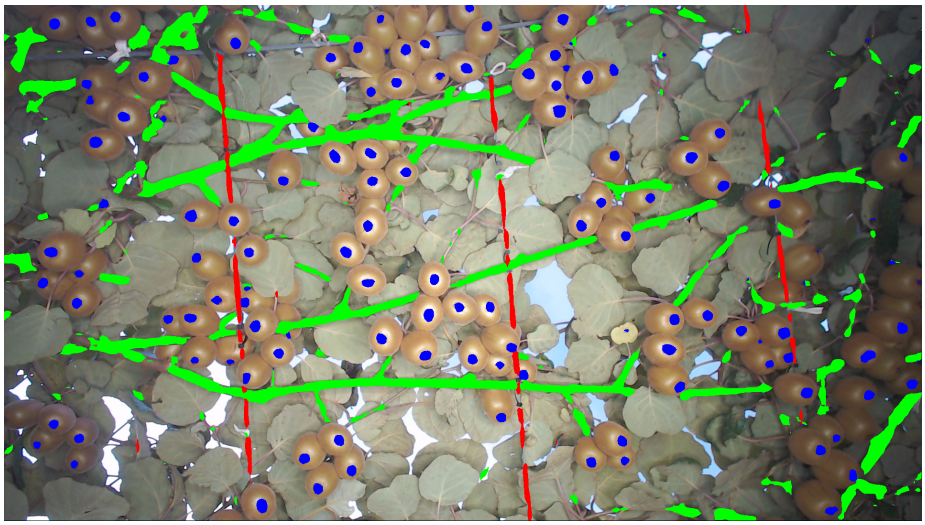

Figure 3: Output from the trained neural network. Blue represents a kiwifruit's calyx, green represents canes, and red shows wires.

The network used here is capable of performing inference on a 500x500 pixel image with a Nvidia GTX-1080 8GB graphics card in 250 ms. As the network is limited in its image input resolution, the input images are fed through in smaller chunks (500x500 pixels) and then reassembled post-inference. The cameras used on the harvester have a resolution of 1900 by 1200 and so each image is broken into 12 pieces to cover the whole image. Therefore, performing inference on a full image takes 3 s, a significant amount of time per sub-cycle.

The position of each kiwifruit is then determined through standard stereo point matching methods. The cameras are spaced with a baseline of 170 mm to minimize the positional error at the working distance of 900 mm. The positions of detected kiwifruit and any obstructions are subsequently passed to the scheduling system.

3.2. Scheduling

The kiwifruit are first filtered to remove kiwifruit not within reach of the arms, based on the arm's kinematic profile. The remaining kiwifruit are then ordered into a picking list that represents the order each kiwifruit will be picked in. That list is ordered in such a way that the risk of damaging adjacent fruit from actuation of harvesting mechanism is minimised. Figure 6 shows the harvesting action of the end-effector as a sequence of images. The image shows how parts of the mechanism protrude from the initial profile during the harvesting action, which could damage neighbouring fruit.

An example of fruit clustering and the resulting order in which they are picked is shown as figure 4. A cluster is picked in the following order: lowest hanging fruit that is furtherest away and to the right relative to the picking arm, progressing toward the closest left. The picking order between clusters happens from the closest to the furtherest away from the arm.

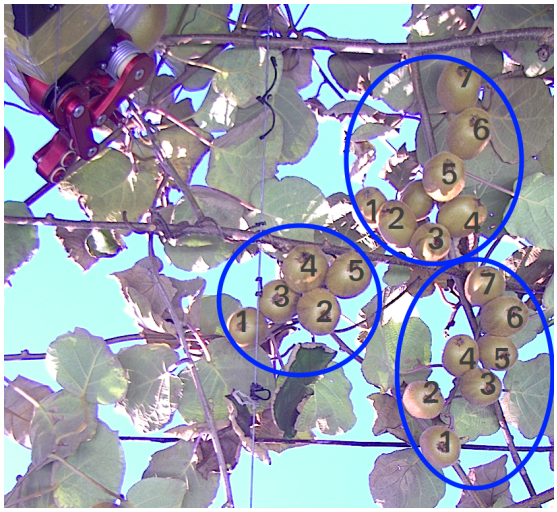

Figure 4: Three clusters of kiwifruit and the order in which the arm will harvest the kiwifruit.

If the system locates a kiwifruit numerous times it will only attempt to pick it twice before marking it as not pickable and removing it from the pick list. This happens if the arm is repeatedly prevented from picking the fruit, due to obstructions, or the system has detected a false positive. This prevents the

system from repeatedly attempting to pick false positives or fruit placed behind obstructions.

3.3. Path Planning and Control

Each arm is controlled using three Festo EMME-AS series servo motors and three CMMP-AS series motor controllers. Theses controllers support the CiA 402 CANopen device profile which is used as the control interface. To ensure each axis remains synchronised during movements, the controllers are set to operate using an interpolated position mode. This allows the controlling computer to stream sets of axis positions to the controllers in near real-time.

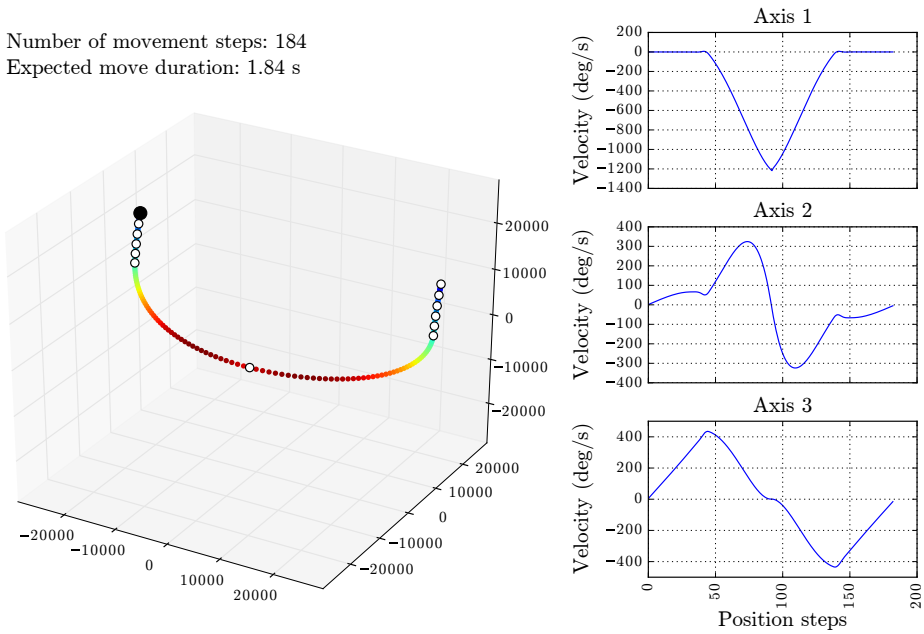

Figure 5: Plot showing a generated 'U-move' path (left) and its associated motor velocities (right). The visualised path is represented in units of a millidegree.

The design of the fruit detachment mechanism necessitates vertical movement both when approaching and detaching a fruit. Figure 5 represents a typical instance of this movement. The end-effector begins at the black circle (top left) where it first retracts from the canopy before moving under the next fruit, then

moves vertically onto the fruit. Fruit detachment is performed each end of the move.

Each of the white circles on the path represent points that the arm must pass through as part of the 'U-move'. Between those points a series of positions are generated that represent a time-optimal path based on velocity and acceleration limits for each axis. This path is generated using the Time-Optimal Path Parameterization library (TOPP) [25]. This ensures the arm moves smoothly and reduces the amount of time spent moving between fruit.

During the move, positions and torques applied by each axis are monitored. If excessive torque is being applied, the scheduler will be notified so it can abort the picking operation and move on to the next fruit. The fruit that caused the excessive torque is then marked as unsafe and removed from the pick list. This reduces damage to the arm, end-effector, and the canopy.

3.4. Fruit Grip and Detachment

Simply pulling a kiwifruit away from the canopy places excessive force on the fruit and causes unnecessary shaking of the canopy. Excessive force can lead to bruising or piercing of the fruit. Shaking causes other fruit to swing and often leads to fruit dropping out of the canopy. Minimising dropped fruit is important for the economic operation of the harvester as it is a source of unrecoverable loss. Additionally, the fruit locations previously determined by the machine vision system are only valid as long as the fruit remain still.

One way to reduce the amount of force needed to detach a kiwifruit is to rotate the fruit about its stem before pulling it away form the canopy. This concentrates the force at the fruit-stem interface into a smaller region, which promotes tearing between the fruit and stem. That tearing reduces the peak force required to detach the fruit and ensures the fruit is separated from its stem. As kiwifruit stems can damage other fruit post-harvest, e.g., when placed in bins, ensuring they are detached is important to the design of the end-effector. A custom end-effector has been developed that produces this rotation using an asymmetrical four bar linkage. This end-effector and its harvesting action is

shown in figure 6, where a kiwifruit like object is detached from a magnetic rod and dropped through the hand mechanism.

The gripping mechanism is made of food-grade silicon molded around a 3D-printed digit. The molded silicon sections contain channeled air pockets that allow the silicon to conform to the shape of the kiwifruit. By conforming to the shape of the fruit, the contact area between the fruit and gripper is increased, thus reducing the peak pressure applied, thereby reducing the chance of bruising. A single pneumatic cylinder actuates both the clasp and rotate actions of the hand. Adjusting the pressure applied to that cylinder therefore controls the total force being applied to the fruit.

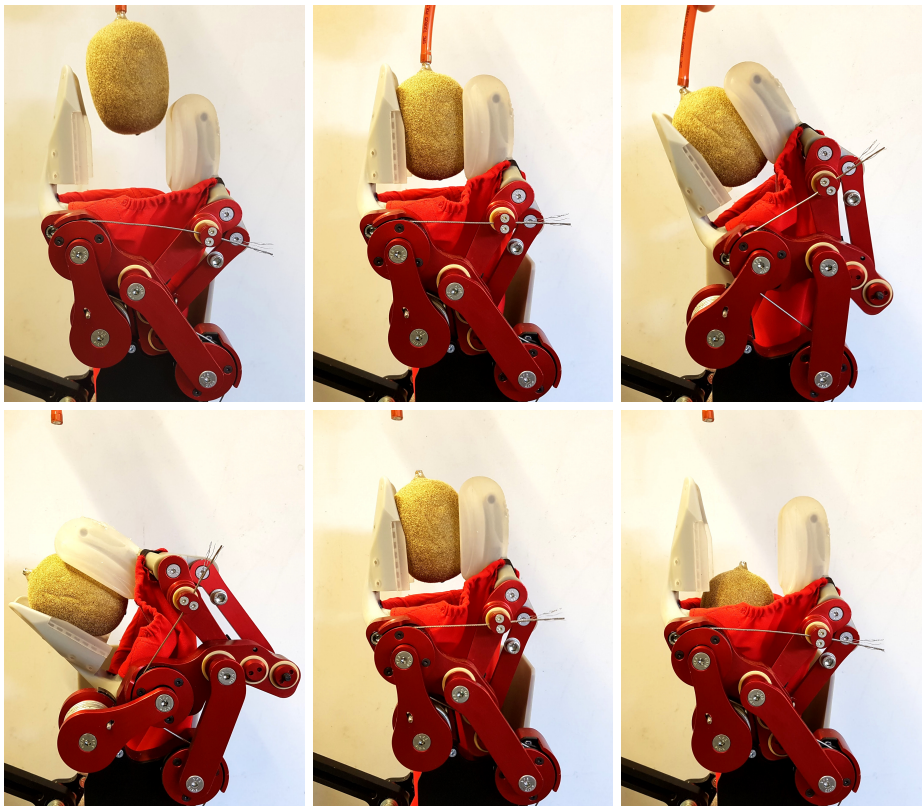

Figure 6: Picking action of harvesting hand, viewed from left to right. The hand grips the fruit and then rotates to gently snap the kiwifruit off the stem before releasing the fruit down the chute.

4. Evaluation

The harvesting system is evaluated in two parts: the machine vision system, and the system as a whole. As results of system as a whole include those of the machine vision system, simple subtraction yields an evaluation of the hardware. Firstly, an evaluation of the vision system is presented, followed by the remainder of the system.

4.1. Machine Vision

The performance of the machine vision system was evaluated over three static tests in separate locations within a single orchard. The tests were static in that the harvesting arms were disengaged and the platform remained stationary. During the test the harvesting unit captured images of the canopy, which were fed through the detection and localisation system. The source and output images from the system were then analysed offline to assess performance. An example of this is shown in figure 7, where the kiwifruit are labeled with colored circles to represent whether the kiwifruit was detected, matched, and/or reachable.

Prior to capturing data, the canopy area visible to the four stereo camera pairs was marked out with red tape. Kiwifruit outside the visible area were manually picked before the image data was captured to make a final count easier to measure. Once images of the canopy were captured, all fruit within the region were hand-picked and counted.

Table 1 shows the progression of of kiwifruit identification rates. It shows an average of 96.7 % of kiwifruit can be seen by the harvester's cameras; a relatively high figure considering the frequency of occlusion in individual images. Because the cameras have a wide field of view, with overlap between adjacent camera pairs, fruit occluded in one image are generally visible in other images.

The detection system is then shown to be capable of detecting 76.3% of all the kiwifruit in the region, including kiwifruit not visible in the images. Of those kiwifruit visible in the images (by human inspection) the system was able

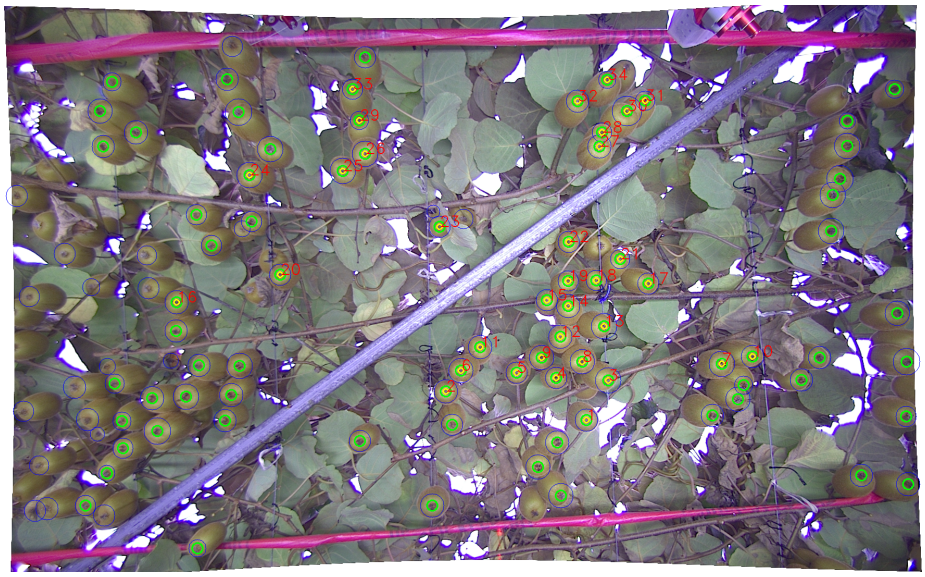

Figure 7: Resulting photo after being processed on-line by the harvesting vision system. Blue circles mark detected kiwifruit, green circles show fruit which have been matched between stereo images, and yellow circles indicate that the fruit's position lies within the picking area of the respective harvesting arm.

to detect 79.0% of the visible fruit. Of those kiwifruit that are detected in the left and right camera images, 99.7% were correctly stereo matched to generate a location for the arm to pick. Leading to a final localisation rate of 76.1% of all kiwifruit in the canopy within the field of view of the cameras.

As can be seen in figure 7, the lighting intensity decreases towards the edge of the image. This reduction in lighting quality tends to result in a lower detection rate in these areas using the current network. Increasing the detection rate should be possible with further training of the detection networks as more training data in these conditions are acquired.

The false positive count of the system adds another 1.2% of false kiwifruit to be harvested. This does effect the picking time as time is wasted picking non-existent fruit, but it is substantially lower than the real value and does not significantly effective the overall time to harvest.

These figures were obtained from a single capture of the canopy images. Dur-

Trial	Total	Visible	Detected	Localised	False Positive
1	70	70	59	59	1
2	182	176	137	136	1
3	162	154	120	120	3
Total	414	400	316	315	5
Cumulative	100 %	96.7 %	76.3 %	76.1 %	1.2 %
Step loss	0 %	3.3 %	21.0 %	0.3 %	

Table 1: Breakdown of identification and steps from processed canopy images. 'Total' refers to the number of hand counted fruit, 'visible' refers to the number of human identified fruit from the images, 'detected' refers to the number of fruit identified by the system, and 'localised' refers to the number of fruit matched between stereo images to give a position. 'False positive' refers to the number of non-kiwifruit objects localised as kiwifruit by the vision system.

ing operation, images are captured and processed at the start of every harvest sub-cycle. Fruit not detected in one sub-cycle have a chance of being detected on subsequent sub-cycles due small changes in the image, e.g. clusters thin out and reveal more fruit. Additionally, between each harvest cycle the base platform advances by approximately one third of the viewable area. Any fruit missed in the first round are likely to be seen in the next due to the change in viewpoint and lighting.

Finally, a manual detection approach was used to determine the localisation rate of the reachable fruit in the canopy, these results are shown in table 2. It was found that 55.0% of all the kiwifruit visible in the field of view of the cameras were considered reachable by the arm system. Of those that are reachable 89.6% were correctly Localised by the vision system. Effectively this means that the cameras have a much wider field of view than the reachable area of the harvesting arms.

4.2. Harvesting System

The harvester was evaluated in three orchards near Tauranga, in the Bay of Plenty region of New Zealand. Each contained plantations of the Hayward variety, commonly known as 'green' kiwifruit. Two of the orchards were main-

Trial	Total fruit	Reachable fruit	Localised reachable fruit
1	70	59	57
2	182	97	84
3	162	75	66
Total	414	231	207
Cumulative	100 %	55.8 %	89.6 %

Table 2: Comparison between the number of fruit in the harvester's viewing window (hand counted), the number determined to be reachable by the arms, and the number of reachable fruit correctly localised

tained to a standard commonly found in the kiwifruit industry, while the third was particularly well maintained in terms of its canopy.

Evaluation was by way of operating the harvester through a pre-defined part of each orchard and recording its performance. These regions were chosen based on how well they appeared to represent the orchard as a whole. Each were marked out with tape prior to testing and the kiwifruit within the region counted by hand multiple times as a check.

The test regions spanned the full width of a row, which were between 4.0 m and 4.5 m. Because the harvester is roughly 2.4 m wide, it was driven through each area twice; once along the left-hand edge, then along the right. Doing so allowed it, in theory, to reach all points within the designated area, allowing the possibility of a 100 % pick. A photo of the unit at the beginning of an evaluation run is presented as figure 8.

Primarily, evaluation of the system is based on the number of fruit picked versus those lost or left in the canopy. Lost fruit includes fruit successfully detached but subsequently dropped, and non-targeted fruit which were knocked from the canopy. These numbers were determined by in-field counts of kiwifruit retrieved from the harvester, found on the ground, or left in the canopy. Secondly, the use of video analysis provided a means of quantifying other performance related metrics such as: the ratio of fruit picked then dropped, versus those knocked directly from the canopy; or the frequency of obstacles being encountered during

Figure 8: Photo showing the harvesting unit and base robotic platform before performing an evaluation run. Red tape has been used to mark the evaluated region.

harvest; or the frequency of the gripper failing to detach fruit.

The primary evaluation results, as shown in table 3, are relatively consistent between each trial. They show that approximately one half of fruit are successfully harvested, a quarter are left on the canopy, and a quarter are lost. Calculation based on the total duration, fruit count, and percentage harvested, yields an average cycle time of 5.5 s/fruit across all four arms. This figure also takes into account the time taken for the platform to move forward between harvest cycles; estimated to be around 3 s at a speed of $100\,\mathrm{mm\,s^{-1}}$.

Secondary results, taken from video analysis and presented as table 4, show a much greater variation between orchards. 'Grip failure' occurs when the harvester attempts to pick kiwifruit but fails to detach the fruit from the canopy. Possible causes are interference with obstacles, positioning error of the end-effector, or a lack of friction between the gripper and the fruit. A lack of friction was responsible for most of the grip failures during the first trial due to the addition of plastic skins over the silicone grippers. The skins were intended

Orchard	Kiwifruit (total)	Harvested (%)	Lost (%)	Remaining (%)	Duration (s)
1	340	48.5	28.0	23.5	1155
1	285	53.0	22.0	25.0	669
2	444	55.0	20.0	25.0	1140
3	387	47.5	28.0	24.5	1098
Total	1456	51.0	24.6	24.5	4062

Table 3: Primary evaluation results from in-orchard harvesting trials. Total duration: 67 min and 42 s. Lost includes target kiwifruit dropped by the harvester and non-target fruit being knocked out of the canopy.

Orchard	Grip Failure	Obstacle	Drops/Lost	Cycles	Sub-cycles
1	138	48	71 %	17	68
1	14	29	59 %	12	41
2	52	102	70 %	13	57
3	54	31	66 %	14	67

Table 4: Secondary evaluation results from in-orchard harvesting trials. A sub-cycle refers to an image acquisition, detection, and harvest sequence. A cycle refers to a set of sub-cycles being carried out until no reachable fruit remain followed by the system moving forward.

to allow the gripper to slip into clusters of fruit without gripping onto non-target fruit and knocking them from the canopy. The reduction of grip failures in subsequent trials is due in-part to adjustment of those skins.

'Obstacle' represents a count of the times the end-effector came into contact with anything that limited its ability to move. In some cases encountering an obstacle had no negative effect, i.e. the target fruit was still harvested, but it often means the harvest move was aborted. While this count offers little insight into the performance of the harvester, it does give some idea of an orchard's suitability for automated harvesting.

Across all four trials the average cycle time for the harvester is 5.5 s/fruit with all four arms. The biggest contributor to the cycle time is currently the detection system, at 3.0 s/image at each detection step. As the robot harvests

the system checks if any fruit have become visible or are still in the canopy after a picking step. With the vision system being a constant 3.0 s, there is a diminishing returns on the time spent in the same location checking for viable fruit after each attempt at picking.

5. Discussion

The vision system has been demonstrated to be sufficient to detect up to 89.6% of the pick-able kiwifruit, based on the current harvesting arm, and detecting up to 76.0% of all the kiwifruit in the canopy in the field of view of the cameras. Further improvements are required to reduce processing time of the system, currently contributing a significant detriment to the overall cycle time. With decreasing costs in GPU memory this problem could be solved with more advanced GPU's in the future, however it is not guaranteed. A less complex model targeted at kiwifruit may be capable of the same performance with less memory requirements.

The overall system can currently pick on average 51.0% of all the kiwifruit successfully. The results and observations of the system indicate that a pick rate above 70.0% may be feasible if the drop rate of 15.7% and knock off rate of 7.7% can be resolved. This may be achievable with small improvements to the positional accuracy of the vision system, and design of the hand used to pick the fruit.

To achieve higher than 70% pick rates, further considerations about the design of the harvesting arm appear to be required. Adding more degrees of freedom to the hand may allow the system to pick around obstacles, for example rotating the hand via a wrist joint. However, increased complexity to the overall system may lower the cycle time of the system.

The cycle time of 5.5 s/fruit appears to rate highly compared to other reported systems [4]. Overall, the biggest time constraint is processing time of the CNN in the detection system. However, the high detection performance indicates it is a viable means of fruit and object detection.

6. Conclusion

This paper has reported the design and performance of a novel robotic kiwifruit harvesting system. Measurement of its in-orchard performance shows that it is capable of picking 51.0% of kiwifruit in the three test orchards. During those tests the unit harvested fruit with an average cycle time of 5.5 s/fruit. Currently about 25 % of fruit are lost during harvesting. These drops are a form of unrecoverable loss for the harvesting system and future work should prioritise reducing this number. We estimate that with further development the unit may be capable of harvesting 70.0% of kiwifruit grown in orchards similar to the orchards used in this work.

References

[1] Ministry for Primary Industries, Our Strategy 2030 - growing and protecting New Zealand - The export goal (2015).
URL https://www.mpi.govt.nz/about-mpi/our-strategy-2030-growing-and-protecting-new-zealand/the-export-goal

[2] D. Font, T. Pallejà, M. Tresanchez, D. Runcan, J. Moreno, D. Martínez, M. Teixidó, J. Palacín, A proposal for automatic fruit harvesting by combining a low cost stereovision camera and a robotic arm, Sensors 14 (7) (2014) 11557–11579.

[3] L. Fu, F. Zhang, Y. Gejima, Z. Li, B. Wang, Y. Cui, Development and experiment of end-effector for kiwifruit harvesting robot, Trans Chin Soc Agric Mach 46 (3) (2015) 1–8.

[4] C. W. Bac, E. J. Henten, J. Hemming, Y. Edan, Harvesting robots for high-value crops: State-of-the-art review and challenges ahead, Journal of Field Robotics 31 (6) (2014) 888–911.

[5] F. Longsheng, W. Bin, C. Yongjie, S. Shuai, Y. Gejima, T. Kobayashi, Kiwifruit recognition at nighttime using artificial lighting based on machine vision, International Journal of Agricultural and Biological Engineering 8 (4) (2015) 52.

[6] R. Oberti, A. Shapiro, Advances in robotic agriculture for crops, Biosystems Engineering 146 (2016) 1 – 2, special Issue: Advances in Robotic Agriculture for Crops. doi:http://dx.doi.org/10.1016/j.biosystemseng.2016.05.010.
URL http://www.sciencedirect.com/science/article/pii/S1537511016302628

[7] A. J. Scarfe, R. C. Flemmer, H. Bakker, C. L. Flemmer, Development of an autonomous kiwifruit picking robot, in: Autonomous Robots and Agents, 2009. ICARA 2009. 4th International Conference on, IEEE, 2009, pp. 380–384.

[8] A. Gongal, S. Amatya, M. Karkee, Q. Zhang, K. Lewis, Sensors and systems for fruit detection and localization: A review, Computers and Electronics in Agriculture 116 (2015) 8–19.

[9] F. Kurtulmus, W. S. Lee, A. Vardar, Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network, Precision Agriculture 15 (1) (2014) 57.

[10] Y. Huang, Y. Lan, S. J. Thomson, A. Fang, W. C. Hoffmann, R. E. Lacey, Development of soft computing and applications in agricultural and biological engineering, Computers and Electronics in Agriculture 71 (2) (2010) 107–127.

[11] J. Deng, A. Berg, S. Satheesh, H. Su, A. Khosla, L. Fei-Fei, Imagenet large scale visual recognition competition 2012 (ilsvrc2012) (2012).

[12] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal,

L. D. Jackel, M. Monfort, U. Muller, J. Zhang, et al., End to end learning for self-driving cars, arXiv preprint arXiv:1604.07316.

[13] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, Backpropagation applied to handwritten zip code recognition, Neural computation 1 (4) (1989) 541–551.

[14] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 86 (11) (1998) 2278–2324.

[15] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

[16] F. P. Orlando, D. G. Mortensen, Harvester with selective force balanced shaking mechanism, uS Patent 4,418,521 (Dec. 6 1983).

[17] D. Rust, Berry harvesting apparatus, uS Patent 3,184,908 (May 25 1965).

[18] Y. F. Wang, J. J. Yin, S. F. Chen, Path planning of avoiding obstacle for fruit-picking manipulator based on position information, in: Advanced Materials Research, Vol. 383, Trans Tech Publ, 2012, pp. 1318–1325.

[19] R. Luo, K. M. Lewis, Q. Zhang, S. Wang, Assessment of bruise damage by vacuum apple harvester using an impact recording device, in: 2012 Dallas, Texas, July 29-August 1, 2012, American Society of Agricultural and Biological Engineers, 2012, p. 1.

[20] J. R. Davidson, A. Silwal, C. J. Hohimer, M. Karkee, C. Mo, Q. Zhang, Proof-of-concept of a robotic apple harvester, in: Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on, IEEE, 2016, pp. 634–639.

[21] M. Monta, N. Kondo, Y. Shibano, Agricultural robot in grape production system, in: Proceedings of 1995 IEEE International Conference on Robotics

and Automation, Vol. 3, 1995, pp. 2504–2509 vol.3. doi:10.1109/ROBOT.1995.525635.

[22] M. H. Jones, J. Bell, M. Seabright, J. Barnett, A. Scarfe, B. MacDonald, M. Duke, An autonomous platform for use in kiwifruit orchards, Biosystems Engineering 1 (1) (2017) 1–2.

[23] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[24] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.

[25] Q.-C. Pham, A general, fast, and robust implementation of the time-optimal path parameterization algorithm, IEEE Transactions on Robotics 30 (6) (2014) 1533–1540.