Nursery automation and phenotyping: final report

Grant Pearse, Alan Tan, Peter Massam, Honey Estarija, Heidi Dungey

September 2020

Report information sheet

Report title	Nursery automation and phenotyping: final report
Authors	Grant Pearse Alan Tan Peter Massam Honey Estarija
Client	Forest Growers Research
Client contract number	QT 7559 J03303
MBIE contract number	-
PAD output number	18033868
Signed off by	Heidi Dungey
Date	September 2020
Confidentiality requirement	Confidential (for client use only)
Intellectual property	
Disclaimer	The information and opinions provided in the Report have been prepared for the Client and its specified purposes. Accordingly, any person other than the Client uses the information and opinions in this report entirely at its own risk. The Report has been provided in good faith and on the basis that reasonable endeavours have been made to be accurate and not misleading and to exercise reasonable care, skill and judgment in providing such information and opinions. Neither Scion, nor any of its employees, officers, contractors, agents or other persons acting on its behalf or under its control accepts any responsibility or liability in respect of any information or opinions provided in this Report.

Published by: Scion, 49 Sala Street, Private Bag 3020, Rotorua 3046, New Zealand. www.scionresearch.co.nz

Executive Summary

Background

Nursery managers require new methods to automate tasks such as stock assessments and implement in-field phenotyping and trials. Unmanned aerial vehicles have already been shown to offer counts and some indication of height in open-grown tree nurseries but methods have relied on orthorectification of highly overlapping imagery. The inputs are expensive and time-consuming, and the outputs lack detail, scalability and are poorly suited to phenotyping. New methods from modern computer vision could enable rapid, scalable stock assessments and improve the prospects for phenotyping in tree nurseries while new cameras equipped with time-of-flight image ranging could enable rapid counting and growth assessment.

Methods

A trial containing three beds of open-grown *Pinus radiata* D. Don seedlings was established. New sensors were trialled for measuring seedling heights to develop growth curves. Two new methods for imaging long seedling beds using unmanned aerial vehicles (UAVs) were trialled. The first was based on panorama generation and the second used direct georeferencing based on high-precision positional information. Two deep learning algorithms for counting seedling were developed. The first algorithm focused exclusively on accuracy and speed for final seeding counts after topping. The second focused on phenotyping and counting applications by detecting and delineating seedling crowns mid-way through the growth cycle and again shortly after topping. Lastly, a new technique for automatically extracting high-quality imagery of plot locations in large open-grown trial setups was tested for use in forest tree nurseries.

Results

The height measurement trial was inconclusive due to repeated missed measurements during the growth period because of COVID-19 restrictions. Initial testing suggested stereo and time-of-flight sensors were most promising for direct measurement of seedling height. Methods based on panorama generation and direct georeferencing were able to rapidly image long seedling beds using low-overlap (40%) imagery from a UAV flying at 2.5-3.5 meters height. Panorama-based imagery retained higher detail but mosaics from direct georeferencing were geolocated to decimetre accuracy and had consistent scale and perspective. The first deep learning algorithm offered rapid and scalable counting of seedlings, achieving 96% precision and 92% sensitivity. The second deep- learning model delineated the crown area of individual seedlings phenotyping and automation with 86% accuracy in a bed with moderate weed cover and 91% accuracy in topped seedlings. Automatic extraction of micro-plot locations and imagery was successfully demonstrated using a simulated block trial and the georegistered orthomosaic.

Conclusions

UAVs combined with modern computer vision algorithms are powerful tools for automation and phenotyping. Alternative methods allow single-pass, low-overlap imagery to rapidly image long seedling beds. Deep learning algorithms accurately detected seedlings. At the cost of extra training and computation, these models were also able to accurately delineate the crown of individual seedlings – a key requirement for phenotyping. Combined with the demonstrated techniques for micro-plot extraction, UAVs enable both automation and large-scale phenotyping in forest tree nurseries.

Next steps - Industry

For industry adoption, we believe the technology is ready for a commercial pilot. Scion is seeking partners to see at least the counting approach implemented in a commercial nursery. The recommended pathway for a commercial pilot is:

1) Select two ages at which seedlings counts would be useful for nursery managers.
2) Develop a pilot based on capturing new data from a larger number of seedling beds at the target ages.
3) Annotate at least 3000-5000 seedlings for each dataset to increase data diversity and model robustness.
4) Test the system against ground-truth inventory counts performed in seedlings beds that were not used to develop the model.
5) Develop an iterative approach to deploy, assess, improve and redeploy the counting method.
6) Gradually phase-out traditional inventory counts.

Next steps – Phenotyping

The methods developed for segmenting individual seedling crowns are suitable for both in-field and controlled facility phenotyping. The approach is transferable so that seedlings could be segmented in imagery acquired from sensors ranging from hyperspectral cameras through to depth sensing cameras. Although the height estimation could not be definitively tested in this trial, it is likely that time-of-flight cameras combined with segmentation algorithms will enable regular assessments of seedling height. Future goals identified with nursery researchers at Scion envision an automated system to monitor height growth in nurseries. This would make it possible to use growth data to forecast the numbers and timing of seedlings expected to reaching maturity. In containerised nurseries, this concept could be coupled with robotic handling systems to automatically sort and select seedlings that had achieved growth targets for dispatch. This approach could greatly improve efficiency and negate the need for practices such as topping.

Introduction

Production nurseries across New Zealand face labour shortages and pressures to reduce the use of pesticides and herbicides. At the same time, nurseries need to increase total production to meet the increasing demand for new plantings (MPI, 2019). To achieve these objectives, tree seedling nurseries will require automated approaches to crop management and care. In the agricultural context, automation has largely been driven by advances in robotics to tackle in-field tasks such as automated weed control, crop counting and crop health assessment (Champ et al., 2020; Shannon et al., 2020; Skovsen et al., 2019). Algorithms that can detect the crops/seedlings of interest in real-time are key enabling technologies for robotic weed control, automated crop counts and growth monitoring (Dyrmann et al., 2018; Milioto et al., 2018).

The enabling technologies required for automation are closely aligned with those required for crop phenotyping. Progress in non-destructive sensors and methods to measure physiological attributes on large numbers of plants have led to rapid advances in crop phenotyping (Araus & Cairns, 2014; Yang et al., 2020). Combined with advances in genomic methods, these advances are expected to lead to substantial gains in global crop improvement efforts (Yang et al., 2020). The ability to accurately identify and characterise individual plants is a key requirement for plant-level phenotyping. Dedicated high-throughput phenotyping facilities have often addressed the challenges of delineating single plants by housing plants in controlled conditions. In a controlled phenotyping facility, plants can be grown and handled individually in pots or trays with regular spacing. These can be moved beneath dedicated sensor arrays and data such as images and 3D scans can be acquired individually under controlled conditions. While this approach solves many technical challenges, it requires expensive specialised facilities with reduced capacity and planting density compared to open-grown field trials.

High-throughput, field-based phenotyping has attracted growing interest from research and commercial phenotyping programmes. The decreasing cost of off-the-shelf sensors and the much larger number of plants that can be grown in densely occupied rows improves the addressed the critical bottlenecks of scale and cost faced by many phenotyping programmes (Araus & Cairns, 2014). The increased density and size achievable with in-field phenotyping approaches comes with significant challenges. In this setup, the sensors must be brought to the plants to capture data under variable conditions. A variety of solutions have been adopted including large gantry-mounted sensor platforms, mobile ‘phenocarts’ and unmanned aerial vehicles (UAVs) (Crain et al., 2016; Yang et al., 2020). These platforms are used to move a range of sensors such as multispectral or hyperspectral sensors over the plants to enable in situ measurements to be obtained.

Echoing the requirements for automation, a key step in the phenotyping process is the ability to identify and delineate individual plants within rows and/or treatment blocks and automatically extract measurements for quantitative analysis. Traditional image segmentation methods have been successfully used in some contexts, but these approaches have typically struggled with the variability of in-field scenes (Thorp & Tian, 2004). In recent years, deep learning has revolutionised computer vision tasks such as detection, classification and even segmentation of objects in imagery (Y. Guo et al., 2016; Rawat & Wang, 2017). Recent studies have explored the potential of these new methods to overcome the complexity of in-field conditions to assist tasks in phenotyping and automation. Examples include efforts to build weed detection datasets (Skovsen et al., 2019) crop health detection datasets (Parraga-Alava et al., 2019; Singh et al., 2020) and international efforts such as the Global Wheat Head Detection Challenge aimed at developing algorithms to phenotype wheat varieties (David et al., 2020).

Scion has recently begun investing in a research nursery to achieve a high level of automation and, in the longer-term, conducting phenotyping research. However, this investment is focused on advancing research capabilities and developing new methods for adoption in commercial nurseries. Research done at Scion’s nursery has highlighted substantial opportunities to apply advanced sensors and methods to automate aspects of crop production in commercial nurseries growing bare-root or containerised seedlings outdoors. This work aims to address two areas of research that have applications for both operational and phenotyping nurseries. Specifically, we aim to:

Objective 1: Automate radiata pine seedling counts at scale using computer vision techniques and extend these beyond counting to segment individual seedling crowns.

Objective 2: Test methods to develop a radiata pine seedling growth/height estimation model using new sensors.

Objective 3: Write a work plan to extend the trial into a commercial setting with support from an industry partner.

Previous work

Previous work at Scion has addressed the topic of counting seedlings from high-resolution RGB orthomosaics acquired from an unmanned aerial vehicle (UAV). A small pilot project tested a method based on developing cross-correlation templates after applying a series of image processing techniques including bilateral filtering followed by morphological erosion to enhance the contrast between seedlings (Pearse 2017). This approach was successful, achieving 90% accuracy where the quality of the final orthomosaic imagery was high enough (Figure 1).

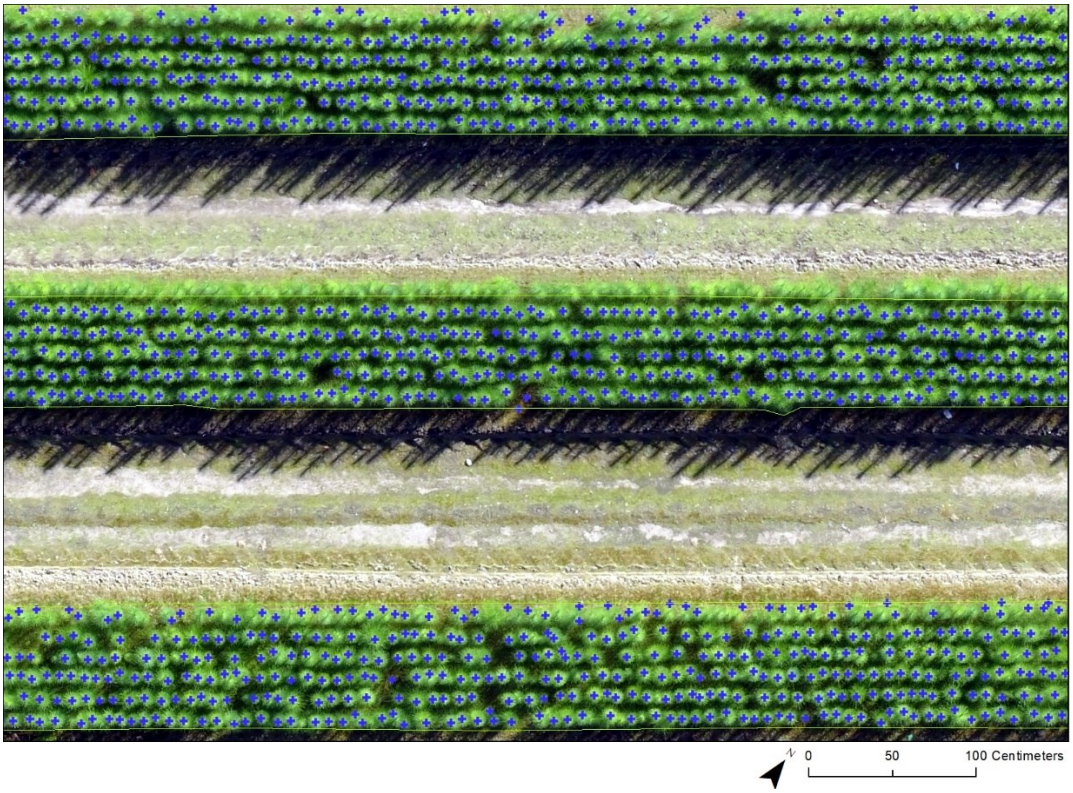

Figure 1: Seedling counting using cross-correlation templates. In this example, the upper row was excluded from analysis due to poor image stitching. Detection accuracy was 90% in the remainder of the areas.

A follow-up study using improved imagery verified these results and demonstrated the potential to extract heights from the photogrammetric point clouds produced during orthorectification of the UAV imagery (Figure 2). This work also confirmed that a wider range of methods such as simple top-hat transforms could be used to count seedlings with relative ease.

These approaches, and to the best of our knowledge all commercially available tools or services, rely on the use of high-resolution orthomosaic imagery acquired from a UAV. The process of orthorectification removes the impact of perspective from the imagery and ensures consistent scale across scenes but it also introduces significant obstacles for use in automation and phenotyping. Firstly, the process typically requires up to 90% overlap between images to generate good orthomosaics over homogenous crops like tree seedlings. This translates to long UAV flight campaigns and significant volumes of raw imagery that must be post-processed into an orthomosaic using compute-intensive software such as Pix4Dmapper (Pix4D, Prilly, Switzerland) or Agisoft Metashape (Agisoft LLC, St. Petersburg, Russia). This poses a significant challenge in tree nurseries where long rows can contain tens of thousands of homogenous seedlings. This approach is also unsuitable for automation where image processing must be done in real-time.

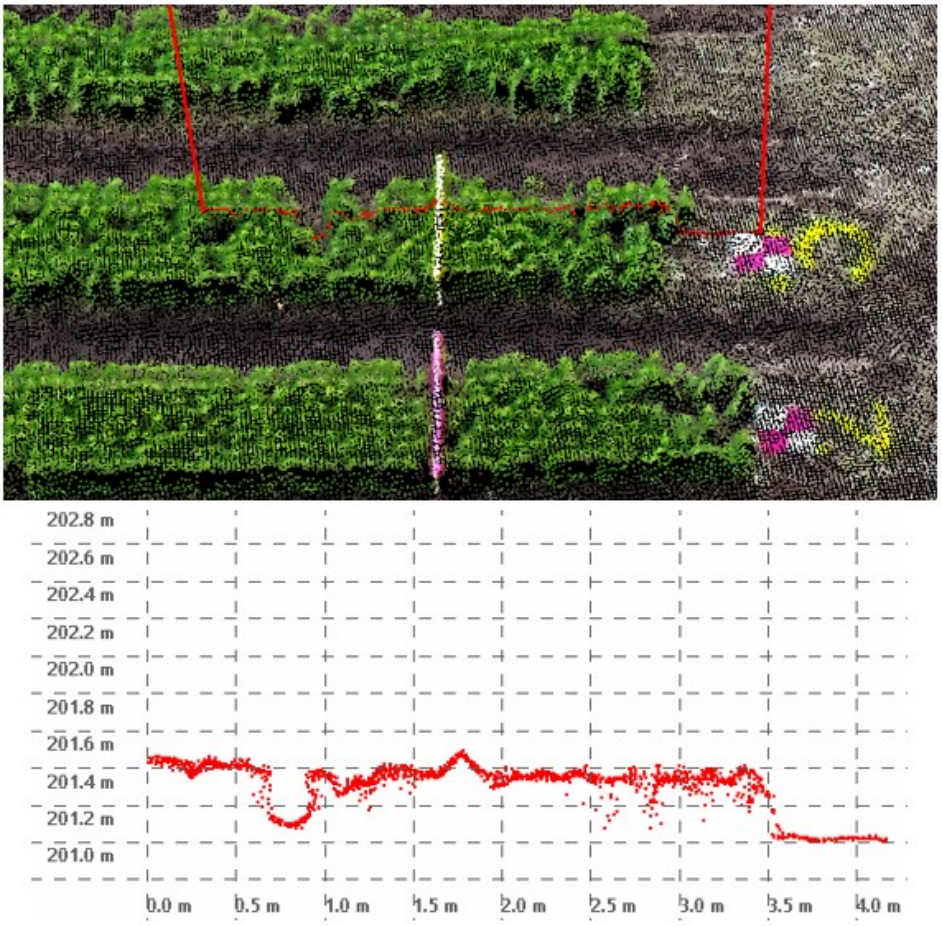

Figure 2: Extraction of height profiles from photogrammetric point clouds acquired over seedlings growing in Scion's nursery.

Secondly, the resulting orthomosaic imagery is severely degraded in terms of detail relative to the raw input imagery (Figure 3A & 3B). Uncertainty and errors arising from the bundle adjustment process used to reconstruct scene geometry during rectification, mosaic blending and histogram matching all contribute to lost detail and alteration of raw pixel values. This also makes orthomosaic imagery poorly suited to phenotyping. Sensors such as hyperspectral cameras or multispectral cameras acquire frame or line-scanned images that must be carefully processed to enable retrieval of physiological parameters using the spectral reflectance recorded by each pixel (Figure 3B). Likewise, automation requires both real-time detection and fine-grained detail for tasks such as weed detection, localisation and control (Figure 3C).

This problem has been widely acknowledged by researchers working with in-field phenotyping data. Several approaches have been proposed to avoid orthorectification while still being able to work at the level of individual plants (Roth et al., 2018; Tresch et al., 2019). One of the most promising is to use advanced positioning techniques to precisely locate imagery and other observations (J. Guo et al., 2018). Combined with image analysis methods, this could allow long rows of crops or seedlings to be imaged in a single pass. These methods are being integrated into advanced mobile phenotyping systems developed for row crop analysis, but little research has explored applying this approach to data captured from UAVs.

This report presents findings from a project to develop UAV-based methods that can be used in both research and commercial tree seedling nurseries. The goal of this work was to develop systems and methods that can support automated crop management such as counting and weed control as well as in-field phenotyping for tree breeding programmes.

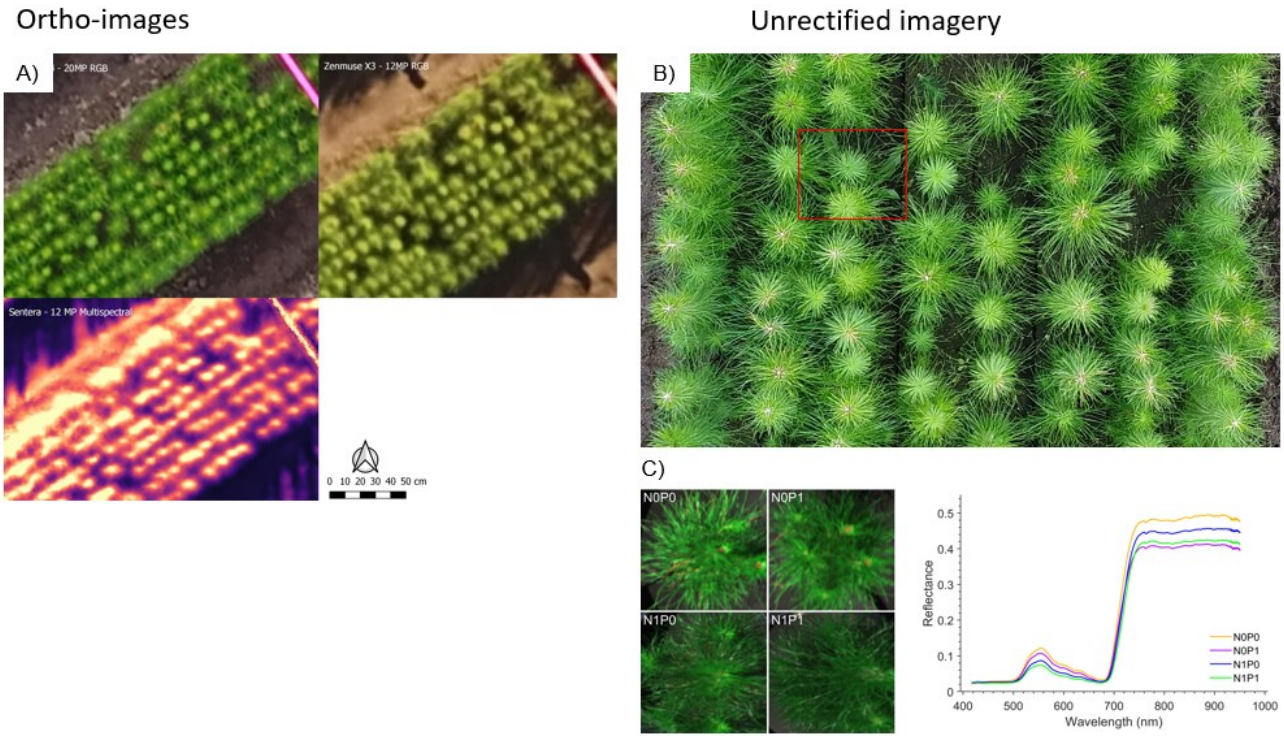

Figure 3: Panel A shows examples of orthorectified imagery acquired from high-overlap UAV flights using RGB colour and multispectral cameras over Scion's nursery. Panel B shows the higher level of detail available from unprocessed RGB imagery from a comparable camera. The improved detail in B) is required for tasks such as detecting and treating weeds as shown in the red box. Panel C illustrates line-scanned imagery from a hyperspectral camera. The pixel values record spectral reflectance and must be preserved to perform phenotyping at the seedling level.

Objective 1: Height estimation.

Nursery managers currently top seedlings to ensure they obtain a uniform crop that is suitable for packing and transport. This damages the seedlings and creates infection sites for pathogens. For research purposes, height estimation from point clouds generated alongside UAV orthomosaic imagery gave only general indications of height and was prone to noise and imprecision that made it unsuitable for growth monitoring (Pearse et al. 2017).

This objective tested the hypothesis that orthomosaic generation for height estimation is less efficient and accurate than methods based on new technologies such as depth-sensing cameras. In this objective, we aimed to test a range of depth-sensing camera systems covering the leading technologies. The objective was to test the ability of these sensors to acquire height measurements with reduced cost and improved accuracy.

Objective 2: Scalable, automated nursery seedling counts.

While nursery stocks can be estimated at the time of establishment, conditioning treatments, mortality, and poor growth may reduce the final number of seedlings viable for sale or planting of trials. At present, historical data or manual inventory are used to estimate the number of seedlings available. Manual approaches can be labour intensive and usually require the use of sample-based methods for larger crops, introducing a level of uncertainty into inventory figures. Nursery managers have expressed an interest in automating the inventory process using imagery acquired from a UAV. Industry providers and Scion's remote sensing team have both demonstrated methods that achieve suitable accuracy of around 90% (Pearse, 2017) but these approaches are based on generating high-resolution orthomosaics that are both inefficient and unsuitable for use in automation or phenotyping.

The goal of this objective was to test the hypothesis that a combination of computer vision principles and techniques and/or high precision position technologies applied to UAV imagery could allow us to generate seedling counts without the need for orthomosaic generation. If successful, this approach would allow a

sequence of images or video recordings to be used to generate accurate seedling counts over potentially very-large nursery beds. Ideally, these approaches should be suitable for both phenotyping and automation. We also aimed to test methods to move from simple detection to segmentation (delineation) of individual seedling crowns – a key requirement for phenotyping work where the reflectance spectra from the photosynthetic plant elements alone (the crown and the bulk of the needle mass in seedlings) can be used to estimate e.g. growth, health and nutrient status.

Objective 3: Micro-plot extraction

The final objective sought to trial a recently developed technique called micro-plot extraction (MPE) (also referred to as automatic field plot extraction) for use in forest tree nurseries. MPE allows large, open-grown trials to be phenotyped using UAV-borne sensors (Khan & Miklavcic, 2019; Tresch et al., 2019). The method leverages the precision and spatial consistency of orthomosaic imagery while still allowing rapid, automatic extraction of raw imagery covering treatment blocks for use in phenotyping.

Methods

Trial location and layout

Scion established a test crop of seedlings for this study. Cuttings were set in the 2019 season and covered in frost cloth until late October of the same year. The layout and management regime used on these beds was similar to that used in most bare-root production nurseries - 8 evenly spaced rows per bed set out after mechanical cultivation (Figure 4).

Figure 4: Trial beds containing cuttings of *P. radiata* for sensor testing at Scion's nursery.

Testing of the in-field sensors and initial development of the methods for the three objectives were carried out within these trial beds. A total of 36 permanent measurement locations were established at fixed intervals along the beds. Each measurement location was marked and assigned a row code that was also printed on a stake that would be visible within any imagery captured during the trial. At each measurement location, the eight seedlings within each row were marked for re-identification (Figure 5). A custom measuring board was fabricated and used to rapidly measure the heights of seedlings within measurement rows by placing the board behind each row and reading off heights. Photographs of the seedlings in front of the height board were also recorded. Using this approach, monthly measurements of the trial were scheduled to commence at the end of January and run until the crop was lifted at the end of June 2020. These measurements formed the ground-truth data that the sensor-based height measurements were compared to.

Figure 5: Measurement board for rapid, accurate measurement of all seedlings in a row.

Height Estimation

In this objective, we aimed to develop a means of rapidly and repeatedly measuring individual seedling heights within entire beds of seedlings. This would allow growth curves to be captured with a view to removing the need for topping to achieve final crop uniformity. These measurements could also be used to assess the performance of different crops and regimes within nurseries. Scion has demonstrated that methods based on photogrammetry can be used to assess height at the bed level, but this approach requires relatively expensive input data (high-resolution, high-overlap UAV orthomosaics) and suffered from errors in the input orthomosaics. New methods based on advances in sensors and computer vision offer the potential to perform rapid, in-field assessment of crop height. Moreover, many new sensors are capable of simultaneously capturing RGB colour imagery and depth (RGB+D) at the same time. These sensors potentially offer counting, weed detection and height estimation from a single source and could be well suited for use in robotic automation.

For this project, Scion identified and purchased three affordable camera systems that offer direct depth estimation alongside RGB imagery capture (RGB+D). The first camera was the Intel RealSense D435i depth camera (Intel Corporation, Santa Clara CA, USA) (Figure 6, A). The Intel camera allows high-resolution imagery to be captured alongside accurate depth estimates at up to 10 m using stereo-enabled vision. The

second camera is the Stereolabs ZED camera (Stereolabs, San Francisco CA, USA). This system uses a different method for depth estimation based on visual-inertial SLAM (Figure 6, B).

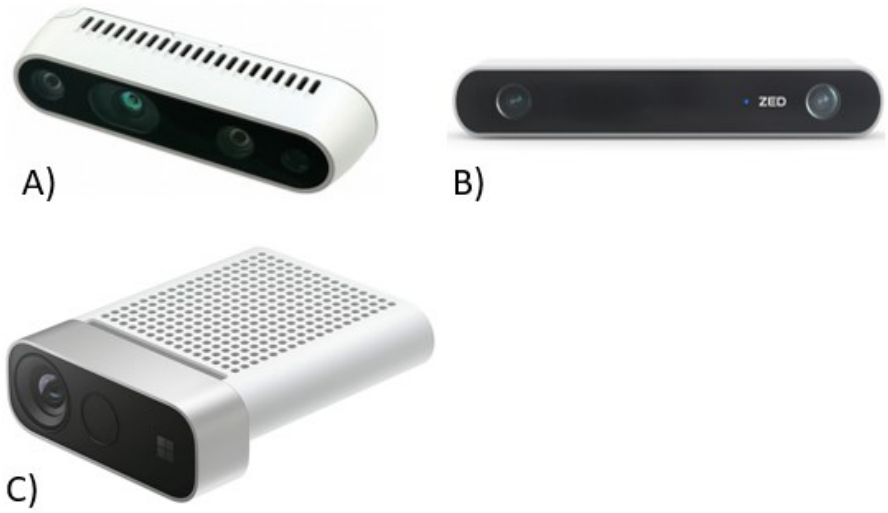

Figure 6: A) Intel RealSense depth camera based on infra-red stereo vision. B) Stereolabs ZED depth camera based on visual-inertial SLAM and C) Microsoft Azure Kinect with time-of-flight sensor.

The final camera purchased was a Microsoft Azure Kinect (Microsoft, Redmond WA, USA) (Figure 6, C). This system was released after the project commenced but was added to the project outline because the sensor offered class-leading depth accuracy and resolution using an advanced time-of-flight (ToF) range imaging camera to measure the distance to objects. The full specifications for the three cameras used in the study are shown in Table 1.

Table 1: Summary of depth cameras selected to measure seedling growth.

	Intel RealSense D435i	Stereolabs ZED	Microsoft Azure Kinect
Depth range	0.11 – 10 m	0.5 – 20 m	0.5 – 3.8 m
Max resolution	1920x1080	4416x1242	3840x2160
Max frame rate	30 fps	100 fps	30 fps
Technology	Active IR stereo	Visual Inertial SLAM	Time-of-flight
Output	RGB+D+ point cloud	RGB+D + point cloud	RGB+D+ point cloud
Est. Price	NZ$350-400	NZ$800	NZ$800

The cameras were calibrated and configured for use in the field. Measurements were scheduled to occur as close as possible to the monthly ground-truth measurements using the measuring board. Although some of the cameras can be adapted to work from a UAV, a mobile platform that could carry a laptop computer was needed during the testing phase. A ‘Phenocart’ was fabricated using additional funding provided by Scion to facilitate rapid data capture using the depth cameras and to facilitate future nursery phenotyping work at Scion (Figure 7). The phenocart was used to mount and operate the depth cameras during data acquisition. This allowed the camera settings to be adjusted and monitored during data capture.

Figure 7: Phenocart used to collect RGB+D data over seedling beds.

Scalable counting and segmenting of seedlings in rows

Radiata seedlings are typically grown in long, dense seedling beds that can hold several thousand seedlings. Identifying and segmenting the crowns of every seedling in these beds depends on having efficient methods to acquire high-resolution imagery over these long seedling beds. In practice, this means moving away from high overlap orthomosaics generated from UAV imagery. Two new methods were investigated for rapidly imaging seedling beds while preserving detail and image properties.

Method 1 for acquiring imagery - panoramas from undistorted imagery

A DJI Phantom 4 Pro UAV (DJI, Shenzhen, China) was used to capture a sequence of images over the bed. The intrinsic properties of the UAV camera were obtained through professional camera calibration; however, similar results can be obtained from open source software or commercial solutions such as Pix4Dmapper or Agisoft Metashape. The camera parameters allowed the radial distortion of the camera to be modelled and removed from the imagery. The camera properties also allowed the field of view for the camera to be calculated at different heights. This information was used to set a flight height that would allow the target seedling bed to occupy the portion of the camera's field while minimizing barrel and perspective distortion as the UAV flew down each row in a single-pass flight line (Figure 8). The barrel distortion on this wide-angled lens was substantial but this is a common property of consumer UAVs and our objective was to develop methods using limited specialist equipment.

Figure 8: Image with radial distortion removed and target seedling bed located in the central portion of the image to minimize barrel and perspective distortion. The impact of barrel distortion can be seen on the seedlings in the left-hand side bed. The imagery was further split into two sub-frames (outlined in yellow) for use in the deep learning models.

The scale invariant feature transform (SIFT) was used to find common features in overlapping images along the flight path. Random sample consensus (RANSAC) was then used to remove erroneous matches before using the remaining image tie points to create a seamless mosaic of the seedling beds. This process was done using the Python OpenCV 3 (https://opencv.org/) computer vision library. Once the general approach was validated, we sought to identify a software tool that could achieve the same outcome with little or no code to facilitate uptake by industry. The PTGui (https://www.ptgui.com/) software tool was identified as offering similar capabilities and Scion purchased and tested the programme for creating seamless panoramas from overlapping UAV imagery. Options such as Microsoft's Image Composite Editor offer similar capabilities but were not tested in this project.

Method 2 for acquiring imagery - direct georeferencing with high-precision sensors

The panorama-based approach is simple; however, it does rely on accurate stitching of images using common features which can be problematic for homogenous crops that offer few distinctive features. In addition, the panorama image does not have consistent scale or perspective and coordinates are not georeferenced. This does not prevent seedling counting, but it does make it harder to link the imagery to real-world locations. An alternative approach combines calibrated cameras with imagery acquired using high-precision positioning systems such as real-time kinematic (RTK) or post-processing kinematic (PPK) GNSS (GPS) data. In this approach, the position of the UAV and camera is known to sub-decimetre accuracy and the properties of the camera are also known with a high degree of certainty. The high-precision positional information greatly reduces the post-processing uncertainty and allows overlapping imagery to be combined into an orthomosaic with far less distortion and fewer artefacts. Crucially, the precise positional information allows the overlap required between images to be decreased from 80-95% down to 40-50% - greatly reducing cost and flight time. Although more technically demanding, this approach is highly suited for automation and large-scale deployment as it offers precise geolocation of the mosaiced imagery without the need for ground control points and greatly reduced capture costs due to lower overlap. The approach can also be generalised to sensors such as hyperspectral cameras for phenotyping and crop monitoring. For this project, Scion used the same DJI Phantom 4 Pro UAV and calibrated camera but a Klau Geomatics PPK solution (Klau Geomatics, Nowra NSW, Australia) was added to the configuration. The imagery was acquired using a single flight over each bed with just 40% average forward overlap between successive images. The images were then geotagged using the post-processed PPK data from the Klau system before being mosaiced in Pix4Dmapper.

A fast algorithm for counting seedlings

Several computer vision methods have been shown to produce accurate seedling counts from RGB imagery using e.g. correlation templates and image transformations (Pearse, 2017). These methods, and to the best our knowledge current commercial methods, are all based on the use of orthorectified imagery. To overcome the cost and quality limitations of traditional orthoimagery, we attempted to develop counting algorithms that could work on the imagery obtained from the panorama and/or direct georeferencing methods trialled in this project. The first model developed focused only on counting seedlings after topping. We chose a deep learning object detection model based on the 'Darknet' architecture (Bochkovskiy et al., 2020). This neural network architecture is both accurate and markedly fast at prediction time, often enabling real-time prediction using e.g. consumer smartphones (Redmon et al., 2015; Redmon & Farhadi, 2017). This was an important consideration given the scale of many tree seedling nurseries.

The training data for the darknet-based algorithm was obtained from single-pass UAV flights carried out at 2.5-3.5 m height above the three seedling beds. The target forward/backward overlap between successive images was 40%. The images were cropped to include only the seedling bed of interest and then split into two sub-frames of 1300 x 800 pixels (Figure 8). This image size was small enough to train the algorithm on without any input resizing – retaining the full resolution of the seedlings and detail present in the imagery. The images were then hand-labelled by marking the location of each seedling in the imagery using the VGG VIA image labelling software (Dutta & Zisserman, 2019). Standard data augmentation processes were applied to the training images, including flipping, rotations, shearing, gaussian blurring and variation of image contrast, brightness and colour balance. The data from two of the three seedling beds were split to allocate 127 images to the training set and a further 5 images to the validation set for hyperparameter tuning and to detect overfitting. The test set for this model was imagery from the third seedling bed that was completely withheld from training and used only for testing. The panorama imagery was used to test the fast counting algorithm as these techniques could be combined to demonstrate rapid acquisition and processing at scale.

Counting and segmenting seedlings for phenotyping

For automation and phenotyping, it is valuable to have an image mask that outlines the needles and crown of each seedling in the bed. In multispectral or hyperspectral analysis, these pixels represent reflectance spectra that in-turn capture physiological properties such as health and nutrient status. To achieve this, we selected a deep learning model capable of performing 'instance segmentation'. Instance segmentation models learn to distinguish individual objects of the same class and assign masks to the pixels predicted to belong to each object. The instances can be tallied to provide seedling counts while the masks allow the spectral information from the pixels belonging to each seedling to be extracted and analysed in phenotyping. In the case of the previously described height detection cameras, the segmentation mask from the RGB colour channel can be adapted to the depth channel and used to automatically extract the height of each seedling. This is also useful for robotic vision systems to 'understand' the extent and location of seedlings or weeds.

We were also interested to know if seedlings could be detected and segmented before topping. This could provide early counts for nursery managers and enhance precision management such as targeted weed control or chemical application. To test this, herbicide was withheld from one of the two beds used for model development to allow the weed level to increase. The same UAV system was used to capture video footage of this bed from an altitude of 4 m during February 2020. Non-overlapping image frames were extracted from the video and cropped to the area covering the seedling bed before manually defining polygons around each seedling crown (Figure 9). A total of 25 annotated images were used for training and validation while 5 images were withheld for testing. For clarity, this model is hereafter referred to as the 'early-crop' detection and segmentation model.

Figure 9: Examples of imagery used to train segmentation and detection model on early-crop seedlings. Moderate levels of weed cover were present within the beds. The left panel shows how seedling crowns were manually delineated in the images to train the segmentation model.

For detecting and segmenting topped seedlings, we used the same images and data splits as the Darknet algorithm. However, the two beds used for training data were re-annotated so that a polygon outlined the inner portion of the crown containing the topped leader as well as the bulk of the needles. This was a labour-intensive process with each image containing around 100 seedlings to be annotated.

The final model for detecting and segmenting topped seedlings as well as the model testing early-crop detection was implemented using the Mask-RCNN algorithm within the Detectron2 deep learning framework (Wu et al., 2019). The model used a ResNet-50 backbone model pre-trained on the COCO dataset as provided in the Detectron2 model repository. All models were trained on a cloud hosted Nvidia K80 GPU with 12GB of RAM.

Image preparation for the counting and segmentation algorithms

The fast darknet-based counting model and the two instance segmentation models were then applied to the imagery of the third seedling bed withheld for testing purposes. The test data was in the form of a long, high-resolution image generating using the panorama or direct georeferencing techniques previously described. None of the deep learning models trained to detect seedlings could ingest such large images. Therefore, a small Python script was developed to perform the following steps:

1. For the panorama imagery - register the full image into a local coordinate frame. For the geolocated imagery, use the NZTM2000 coordinate projection system.
2. Tile the mosaic into smaller images (1000 x 1000 pixels) for the chosen deep learning model while recording the origin and size of each tiled image within the larger image.
3. Apply the model to each tile and record the predictions.
4. Collect local coordinates of detected seedlings with a confidence score >= 0.5 and translate the location back to positions in the imagery of the full seedling bed.

This process is summarised in Figure 10. The output of this script is a seamless image of the seedling beds with the detected seedling locations overlaid onto the image for visualisation. In the case of the georeferenced imagery, the locations correspond to seedling locations in the real-world (+- 10 cm).

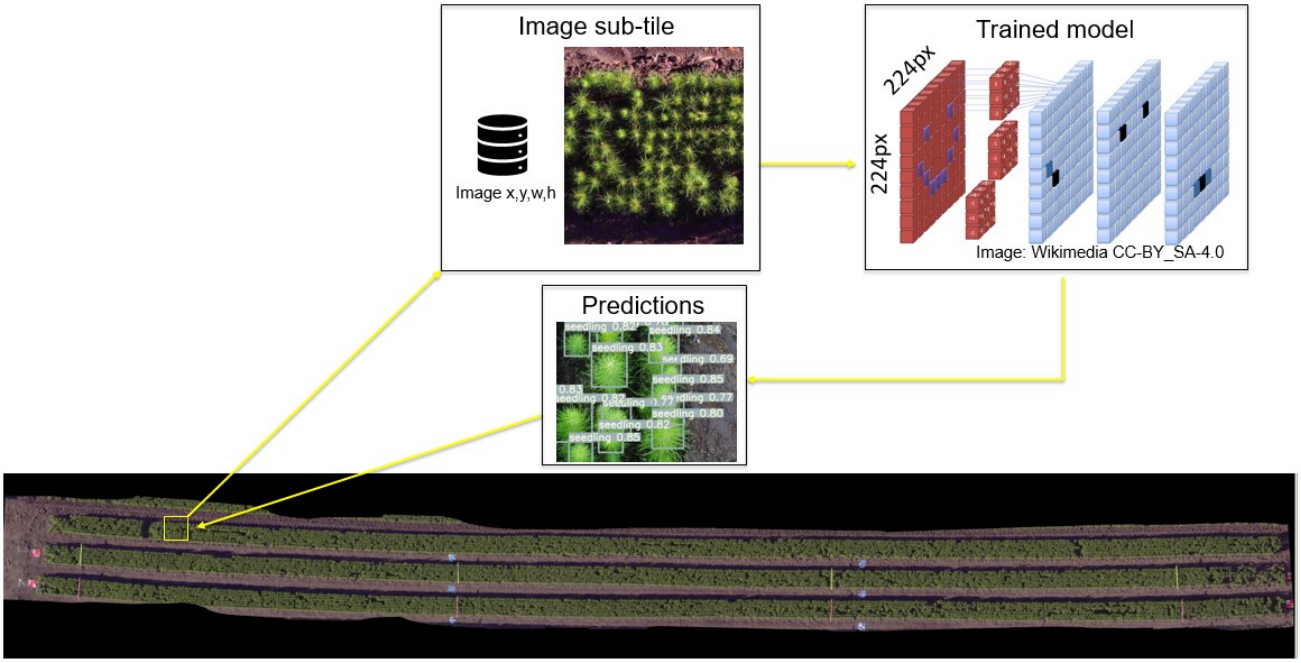

Figure 10: Visual representation of the approach used to divide long panoramas or georeferenced mosaics into small sub-tiles (chips) for input into deep learning models. The results are combined with information on the sub tile's position within the larger image to translate predictions back their locations in the overall scene. This is achieved by recording the tile's upper left corner coordinates (x,y) relative to the larger image's upper left corner and the width (w) and height (h) of the sub tile.

Micro-plot extraction

In this approach, crop rows or beds are assigned to experimental treatments according to a pre-determined trial design. The spatial coordinates of all treatment blocks are recorded using high-grade GPS equipment or extracted from precisely geo-registered orthoimagery covering the area. UAV-borne sensors are then flown over the trials to capture imagery for phenotyping. In the first processing stage, MPE accepts the high-overlap and loss of image quality to generate orthomosaics with consistent scale, perspective and accurate geo-registration. This process also precisely locates the camera position for each input image. In the second phase, a software script is used to extract and combine camera calibration parameters with position and orientation information generated during the orthorectification process. Ray tracing from the camera position to the digital surface model generated during the orthomosaic generation defines the footprint of every image used to generate the mosaic. This allows 3D real-world coordinates of the treatment blocks to be converted to 2D coordinates defining the block locations in the raw input imagery. This can be used in phenotyping workflows to automatically extract the raw, unadjusted pixels values for the crops/seedlings in each treatment block. MPE accepts the cost and time penalty of acquiring high-overlap UAV imagery in exchange for the ability to perform repeatable and automatic extraction of the raw image values from e.g. multispectral sensors flown regularly over the crop. This is an essential workflow for large-scale, in-field phenotyping and we are unaware of previous work trialling these methods in tree seedling crops.

The trial beds in this study did not include randomized treatments. Instead, a grid of 10 trial blocks was simulated in a GIS. The grid was overlaid onto an orthomosaic generated from the imagery collected to test direct georeferencing over the third bed of seedlings. This was the same imagery used to test the counting and segmentation models. A script to perform MPE was developed based on the methods described in Tresch et al. (2019) & Khan & Miklavcic (2019) and the Pix4Dmapper documentation. The script took as input the raw input images used to produce the direct georeferencing test mosaic, metadata produced by Pix4Dmapper during orthomosaic generation and the 3D coordinates of the simulated treatment blocks. The output of the script was a list of the raw input images that contained the simulated treatment blocks and the 2D pixel coordinates describing the footprint of the trial block in the raw, unprocessed imagery.

Results

Height estimation

An initial calibration run for height estimation was completed in December 2019. Based on the results of this capture, the Intel RealSense and Microsoft Azure Kinect were best suited to height measurements. Two measurement datasets were acquired over the height trial on January 23rd and again on February 23rd. Ground measurements and UAV flights to capture co-incident imagery were also carried out. The next capture was scheduled for March 23rd; however, the move to alert level 3 in response to COVID-19 prevented the capture of this dataset. The April data capture also had to be forfeit due to Alert Level 4 restrictions. Efforts were made to revisit the trial at the end of May, but the trial beds had suffered lack of maintenance and with two important height measurements missed the decision was made to abandon efforts to recover the trial.

Preliminary analysis was carried out on the datasets captured at the start of the trial. Seedlings within each of the measured rows were identified in the RGB imagery from the RealSense and Azure Kinect imagery and the lowest value from the depth channel (representing the closest distance between the seedling and camera) was recorded. This value was subtracted from the distance to a sample of adjacent ground pixels to compute seedling height. The best RMSE of 8 cm was obtained from the Intel RealSense camera. This value was high compared to the mean height of seedlings (15 cm). Bright, sunlit conditions appeared to interfere with the ability of both sensors to accurately range the top of each seedling. However, the interruption of the trial prevented further testing under different conditions. There were also numerous data enhancement methods exposed through the software API for both cameras that we were unable to test in the field. Work was then refocused on the detection and segmentation models as well as micro-plot extraction methods for phenotyping.

Results generating panoramas from undistorted imagery

The goal of this approach was to use low-cost methods based on panorama generation to seamlessly combine overlapping images of a seedling bed into a continuous image. The results from the manual OpenCV code-based approach generally worked well. Most image sequences captured over the beds could be combined into a panorama with few mismatches. It was clear that most mismatched tie points were in the actual seedling beds, with positive matching features often detected in the imagery of the ground surrounding the beds. The PTGui software produced panoramas with the same or higher quality than the code-based approach. The software was intuitive and easy to use, allowing the user to refine mismatched points if necessary. Figure 11 shows an example of a panorama produced by PTGui with the seamlines and areas of uncertainty highlighted in red along with examples of poor image matching. Because of the differences in approach to traditional orthorectification, the panoramas lacked geolocation and the scale and perspective were not consistent across the image. Despite these limitations, this approach appeared to offer a rapid and accessible method for constructing seamless panoramas of long seedling beds. The outputs were primarily suitable for use in the counting algorithms.

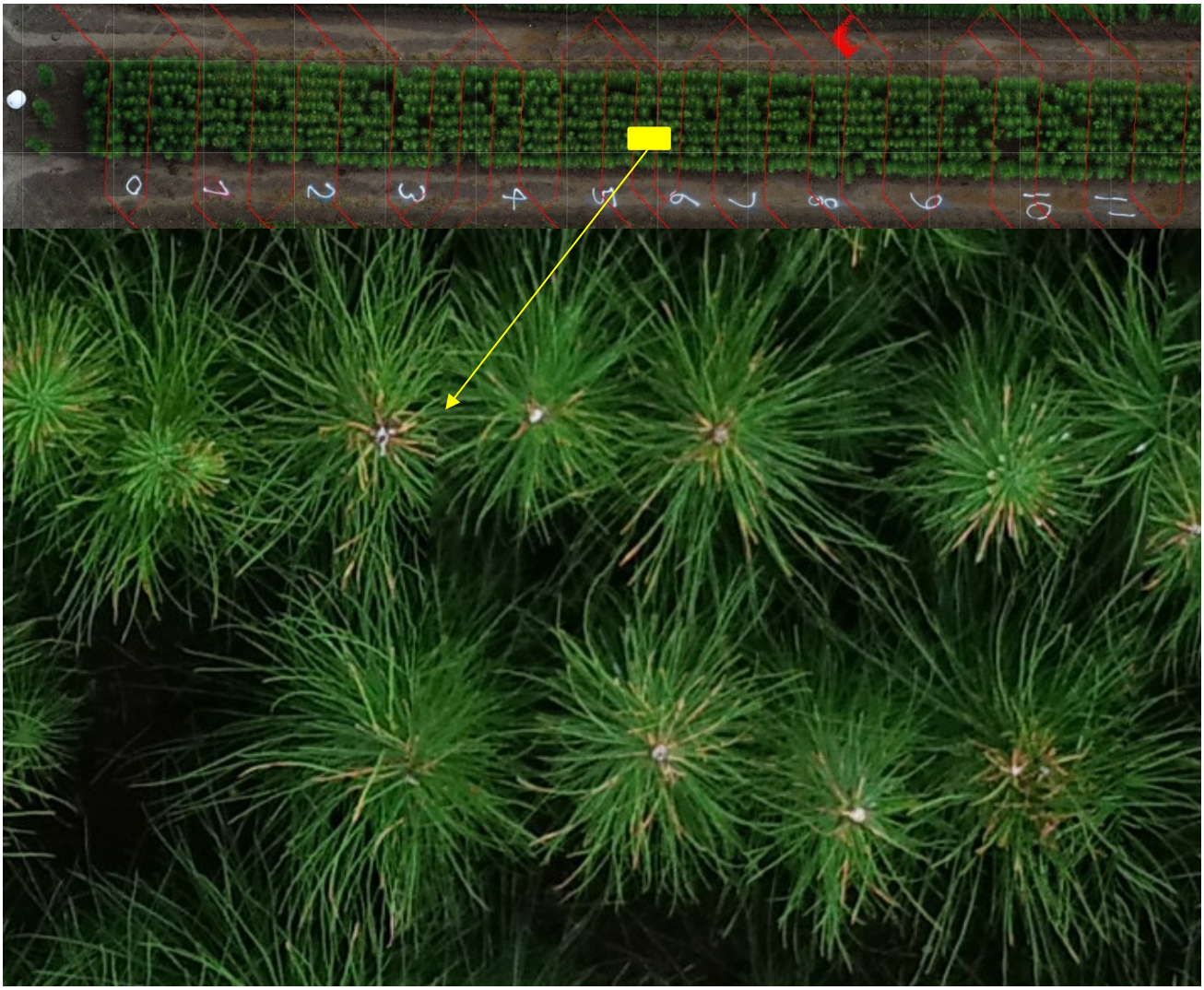

Figure 11: Panorama output from PTGui with seamlines between merged, overlapping images shown in red (top). The quality of the panorama was high and retained the detail of the original images (bottom). The yellow arrow highlights minor artefacts encountered at the seamlines that caused occasional loss of detail and continuity.

Results generating seamless mosaics using direct geo-referencing

This method required precision GNSS equipment to reduce the required overlap and improve the final quality of the orthomosaic. The data required some post-processing to attach the high-grade GNSS positions to the imagery. Image mosaicking was trialled using both Pix4Dmapper and Agisoft Metashape. In both cases, the software was instructed to consider camera positions as being highly accurate - constraining the alignment and bundle adjustment process. In both cases, high-quality orthomosaics with cm-level precision were produced by the software from low-overlap imagery (~40%) acquired in a single flight. The distortions in the imagery were greater than in the panorama but still improved from the traditional orthorectification method without precise positional information (Figure 12). The mosaic imagery had the advantage of being accurately georeferenced, making it suitable for micro-plot extraction. The scale and perspective were also consistent across the scene, but some detail was lost from the individual seedlings.

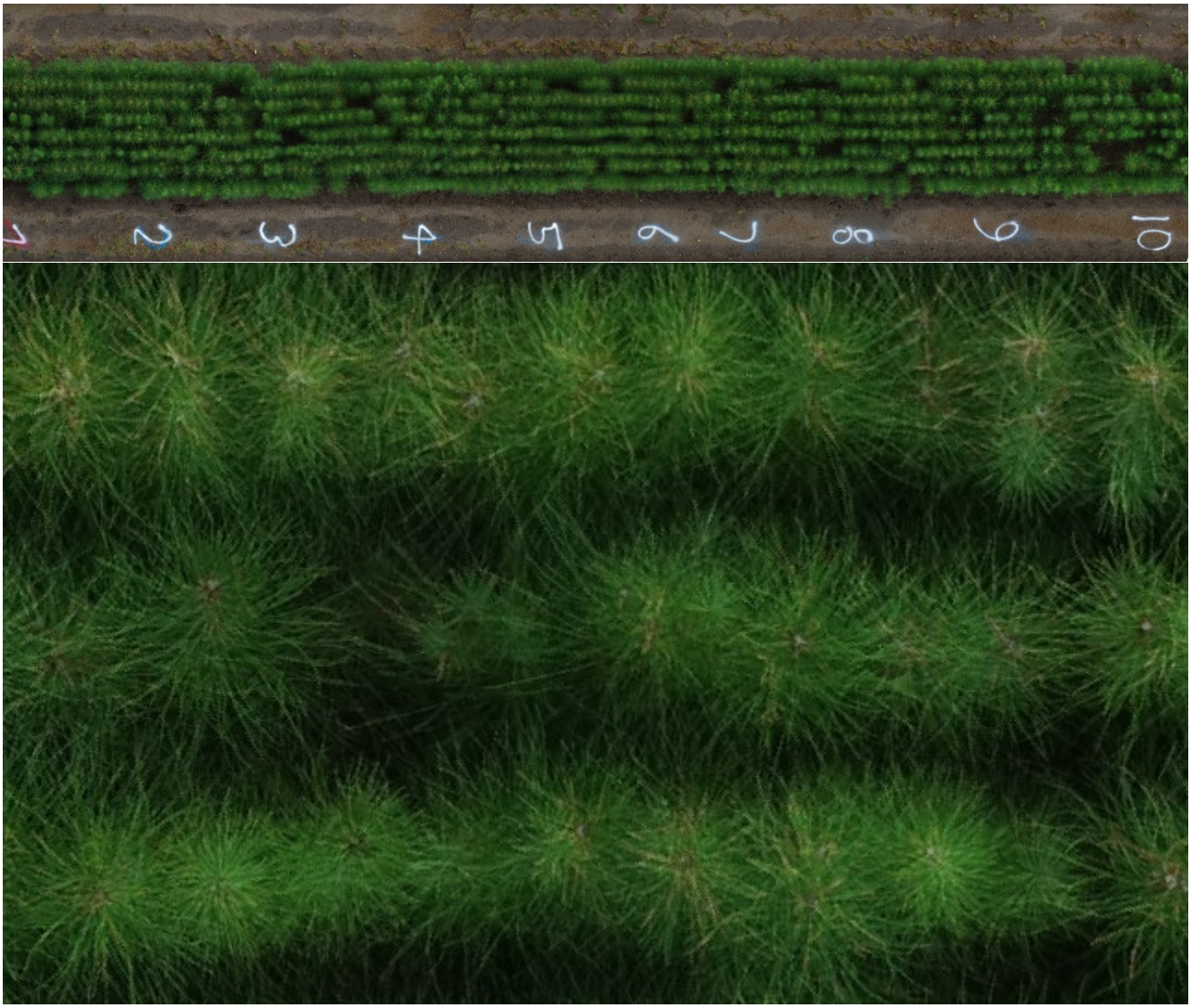

Figure 12: Orthomosaic imagery generated using direct georeferencing. The full mosaic (top) was geo-registered to cm-level accuracy without the need for ground control points. The mosaic was obtained with ~40% overlap and the quality of the final imagery was improved from orthorectification without direct georeferencing but showed some loss of detail (bottom). The imagery was well suited to micro-plot extraction where the 3D real-world coordinates could be used to extract undistorted 2D imagery for phenotyping.

In summary, a single flight at 2.5 to 3.5 m capturing imagery with a target overlap of 40% was successfully used to generate both panoramas and directly georeferenced orthomosaics of the seedling beds.

Results from Darknet-based counting algorithm

The Darknet-based algorithm focused on rapid counting of seedlings in the test bed. The model took ~6 hours to reach convergence – indicating adequate training time. The tiling methodology described earlier allowed the trained model to be applied to sub-tiles without image resizing. Making predictions on all the tiles from the entire test bed took around 3 minutes. Inspecting the results on the full test bed showed a high level of detection accuracy (Figure 13). The model had a precision of 96% and sensitivity of 92%. The lower sensitivity reflected a higher rate of false negatives associated with seedlings at the edge of the bed and shorter seedlings that were not topped. The training set overwhelmingly included topped seedlings and the model likely did not have enough examples of un-topped and edge seedlings to learn from. The false positives were most often caused by seedlings with multiple leaders. Overall, the Darknet-based counting algorithm provided fast and accurate detection based on a relatively small training set.

Figure 13: Results from Darknet-based counting algorithm. False negatives can be seen along the edge of the bed while some seedlings with more than one leader led to false positives.

Results from instance segmentation models

The first instance segmentation model tested counting and segmentation of seedlings approximately mid-way through the crop cycle in a single bed with moderate weed cover. The Mask-RCNN model achieved a detection accuracy of 86% on the withheld test images. The model generally provided accurate segmentation masks for the individual seedlings (Figure 14). Most errors were associated with false negatives (missed seedlings) and no weeds were incorrectly detected as seedlings. These results were a positive indication given the small training set.

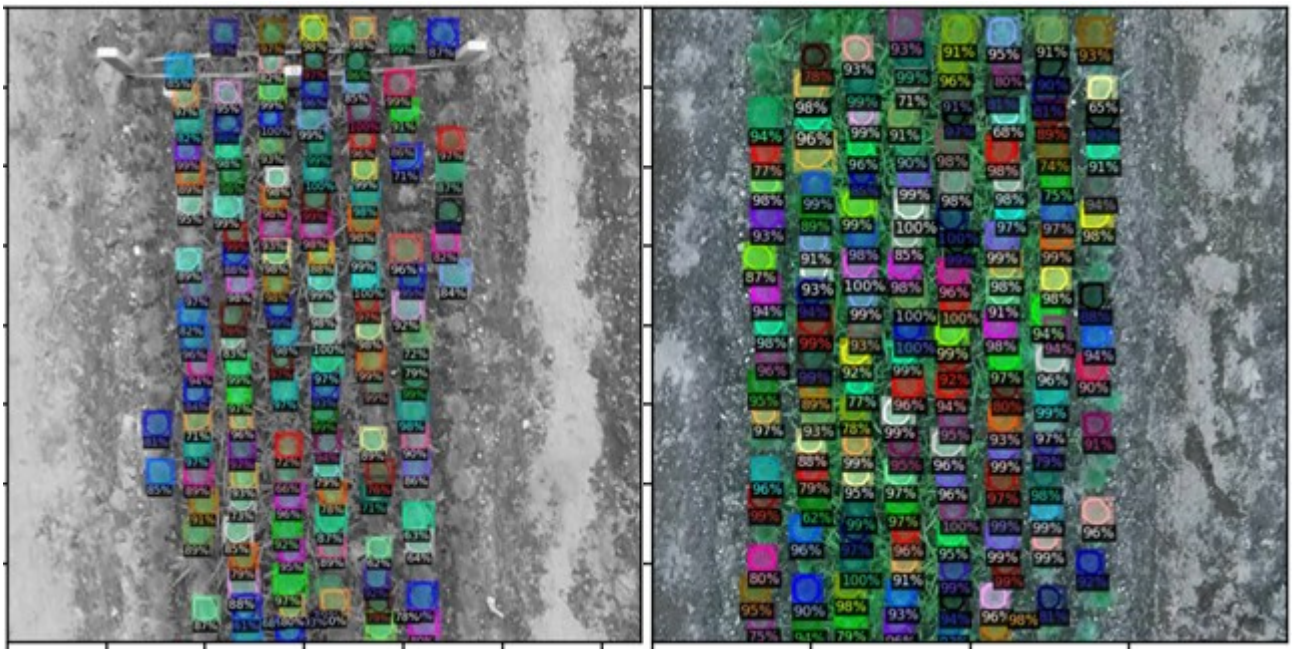

Figure 14: Results from instance segmentation model on seedlings mid-way through the crop cycle. Segmentation masks were generally accurate and captured the crown area, but a moderate number of false negatives were evident towards the edge of the seedling beds.

The segmentation model trained on two beds of topped seedlings had an accuracy of 91% when applied to the third, withheld test bed. The required threshold between the ground-truth and model segmentation was set at 50% for true positives. The segmentation masks exceeding this threshold produced accurate delineations of the bulk of the needles and central crown area (Figure 15). Small seedlings were sometimes missed but the rate of false positives due to multi-leader seedlings was lower than the Darknet model.

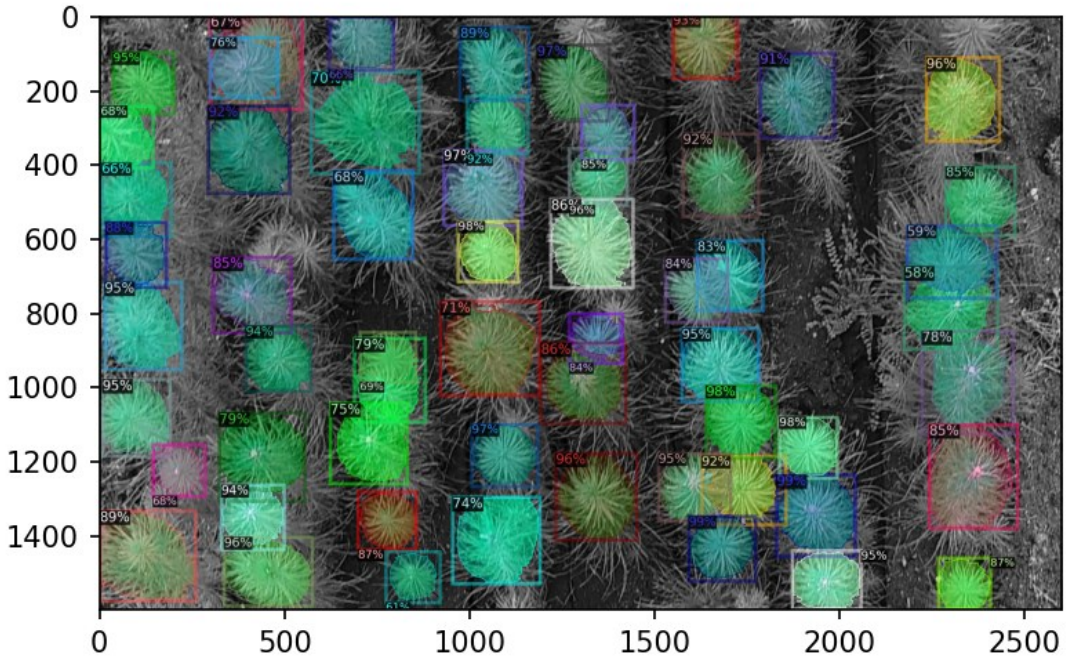

Figure 15: Instance segmentation results on topped seedlings. Segmentation masks for individual seedlings are shown as coloured areas within the bounding boxes.

Micro-plot extraction

The MPE script was tested on the simulated block-design layout. The method was able to correctly identify the raw input images covering the simulated block locations and the process allowed rapid conversion of 3D real-world coordinates to 2D coordinates in images covering the same area. The quality of the digital surface model was important to the process and required full orthorectification to generate a dense surface model. This was likely due to the taller height of the seedlings. This outcome demonstrated that MPE from relatively low overlap UAV imagery could be applied to tree seedlings for in-field phenotyping. This also confirmed the benefit of direct georeferencing for reducing the required image overlap over dense, homogenous seedling beds.

Discussion

The results from the trial confirmed that UAVs are both robust and practical for large-scale phenotyping and automation in tree crop nurseries. Both the panorama and direct georeferencing approaches provided a way to rapidly capture imagery over long seedling beds. Both methods required substantially less overlap and UAV flight time than typical orthomosaic generation methods. The panorama method had the advantage of producing nearly seamless imagery of the beds using only open source or low-cost software but was not suitable for measurement and lacked geolocation information. This worked well when counting was the main objective.

The Darknet-based seedling counting model performed very well on this imagery and the high level of individual detail preserved in this imagery improved detection. The combination of fast inference from the Darknet framework and low-cost, fast imaging based on panoramas is likely to offer a scalable solution for large-scale nursery stock counting. The training set was small by comparison to many production models and the accuracy of the Darknet model could be further improved with additional training data. Additional data could also improve the detection of the seedlings at the edge of beds. Scaling of this method could be further

improved by batching the images together and using a more powerful GPU. This would reduce or remove constraints that may be encountered in very large nurseries. The panorama-based method was sensitive to poor image matching – this is a common challenge over homogenous vegetation. In our testing, the images could be matched with some quality checking by an operator to remove incorrect automatic tie points. Another possible solution is to avoid the use of tie points within the beds and enhance image matching by placing temporary fiducial markers alongside the beds. Many custom markers specifically designed to improve image-to-image matching are freely available and could easily be printed and laid out alongside the beds at regularly spaced intervals.

The accurate positional information added to the imagery captured using the direct georeferencing approach greatly improved the quality of the mosaiced imagery, although the detail was still poorer than the panorama imagery. The advantages of this approach were that high-quality mosaics could still be obtained with only around 40% image overlap – less than half the overlap typically required for orthomosaic generation over seedling beds. This imagery was also geo-registered to within a few centimetres without ground control points. The ability to precisely locate spatial coordinates in the georeferenced imagery made it suitable for micro-plot extraction methods. Automatic extracting of raw imagery overlapping the trial blocks using MPE is the first step in an automated analysis pipeline of (for example) multispectral data to detect treatment effects. Combined with the ability to segment individual seedlings, these tools form the basis for phenotyping research in large-scale, open-grown trials.

The segmentation models required substantially more effort to develop due to the need to carefully outline individual seedlings. The training set was small by the standards of these models, but accurate results were still obtained on topped seedlings as well as on younger ‘early-crop’ seedlings growing in weedy beds. The primary application of these models would not be counting but rather for use in the phenotyping pipeline described above. Accurately delineating the central crown and surrounding needles in multispectral or hyperspectral imagery would facilitate automated detection of treatment effects, health status and nutrient deficiencies at the seedling level. Outside of phenotyping, segmentation models are important for use in automation where robotic systems might employ machine vision systems that must be able to identify and target actions e.g. spraying herbicide only on weeds or fungicide only on seedlings. The results of this work suggest these methods are likely to work well in beds of radiata seedlings.

The loss of the trial designed to extract height information prevented conclusive analysis of the depth camera systems. Although the initial results showed only modest accuracy, further data may have allowed us to refine the capture and/or analysis pipelines to improve the results. Time-of-flight sensors warrant further investigation as these sensors are increasingly common in consumer electronics, including many of the latest smartphones. This widespread adoption will drive developments in both the technology and algorithms for capturing accurate short-range depth data from low-cost sensors. Future experiments with new sensors may well provide sensors that allow simultaneous counting of seedlings and height measurements from a single device flown along seedling beds under a UAV.

Conclusions

Labour shortages and restrictions on chemical use are driving increased demand for automation techniques in the nursery, with rapid stock assessment being of interest to nursery managers. This study demonstrated that existing techniques for counting seedlings can be improved upon by using a combination of computer vision techniques and new deep learning algorithms. Methods based on building panoramas and direct georeferencing of UAV imagery allow the required imagery overlap to be more than halved to as low as 40% forward and backward overlap while still producing high-quality imagery of long seedling beds. Fast, scalable counting algorithms based on both detection and segmentation achieved accuracies over 90% and showed promise for counting at earlier ages as well. The models accurately distinguished between seedlings and weeds and the instance segmentation models accurately delineated individual seedling crowns. This study also demonstrated the potential for UAVs to be used to carry out phenotyping at scale in open-grown seedling beds using a combination of deep learning and micro-plot extraction to enable rapid, seedling-level analysis for phenotyping activities. Future work should revisit new systems that could provide imagery covering multiple spectral channels with simultaneous height estimation using emerging ranging sensors.

Future steps

Industry partners are being sought to test these methods in a commercial setting. The most likely avenues for this will be through submission as a possible Pre-Seed Accelerator Fund candidate project. If successful, the following steps will be required to validate the methods in a commercial setting.

Stage 1: Stakeholder engagement
During this stage, the exact requirements of commercial nurseries will be identified along with the key factors determining the cost/benefit trade-off. Interviews with potential end-users are typically the best path to quickly identify end-user requirements for a viable commercial solution, existing approaches and likely roadblocks. Partners for the implementation pathway will also need to be identified, with existing commercial service providers to commercial nurseries being the most likely choice.

Stage 2: Method validation
In this stage, it will be crucial to validate the methods tested in this study for use in commercial nurseries. A repeatable process that works in a range of settings must be developed. Methods such as the use of fiducial markers and streamlined image processing routines will need to be tested for rapid, large-scale imaging of commercial seedling beds. The baseline models will also need to be enhanced with substantially more training data covering a much wider range of conditions. The image processing and model prediction pipeline will also need to be refined to validate the scalability of the models. This may require larger GPUs, simpler models, testing lower resolution input imagery or some combination of these elements. This process will provide the information necessary to inform the business case.

Stage 3: Deployment
In this stage, the model will be deployed in parallel alongside existing methods with a large amount of human oversight and checking. This will represent a stop/go process before considering a full roll-out as a commercial service. If the model achieves acceptable accuracies, a process of continuous improvement can be used to further refine the model and gradually replace manual inventories with automated ones.

Addendum

After the completion of this project. Scion tested an experimental combination of deep learning models to enable real-time tracking of seedlings in videos. This system was adapted from models designed to track people and vehicles and repurposed to perform dense-object tracking of topped seedlings. Although the model lost some precision due to the small dataset available being used to parameterise two deep nueral networks, the overall approach worked well. Importantly, this approach removes the requirements for image processing completely. All that is required is a UAV equipped with a nadir-facing video camera. The system identifies and tracks individual seedlings and adds them to a counter once the seedlings have been tracked across a region of interest defined in the video frame. The output records the final seedling count as well as metadata on detection confidence and tracking errors that can be applied to the final count if needed. A demonstration of the system can be found on Scion's YouTube channel at https://www.youtube.com/watch?v=KUhk2wULENM.

Acknowledgements

This work was funded by the Forest Growers Levy Trust with additional funding contributed by Scion's Strategic Science Investment Funding (SSIF) administered by the Ministry for Business Innovation and Employment. Honey Estarija provided many hours of careful annotation. Peter Massam and Honey Estarija were responsible for the trial layout, field measurements and UAV data acquisition. John Lee and David Pont designed and fabricated the Phenocart and intern Niva Sharma made valuable contributions to the code. We would also like to thank the Scion nursery staff for their valuable assistance with establishing and maintaining the nursery trial.

References

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, *19*(1), 52–61. https://doi.org/10.1016/j.tplants.2013.09.008

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *ArXiv:2004.10934 [Cs, Eess]*. http://arxiv.org/abs/2004.10934

Champ, J., Mora-Fallas, A., Goëau, H., Mata-Montero, E., Bonnet, P., & Joly, A. (2020). Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots. *Applications in Plant Sciences*, *8*(7). https://doi.org/10.1002/aps3.11373

Crain, J. L., Wei, Y., Barker, J., Thompson, S. M., Alderman, P. D., Reynolds, M., Zhang, N., & Poland, J. (2016). Development and Deployment of a Portable Field Phenotyping Platform. *Crop Science*, *56*(3), 965–975. https://doi.org/10.2135/cropsci2015.05.0290

David, E., Madec, S., Sadeghi-Tehran, P., Aasen, H., Zheng, B., Liu, S., Kirchgessner, N., Ishikawa, G., Nagasawa, K., Badhon, M. A., Pozniak, C., de Solan, B., Hund, A., Chapman, S. C., Baret, F., Stavness, I., & Guo, W. (2020, August 20). *Global Wheat Head Detection (GWHD) Dataset: A Large and Diverse Dataset of High-Resolution RGB-Labelled Images to Develop and Benchmark Wheat Head Detection Methods* [Research Article]. Plant Phenomics; Science Partner Journal. https://doi.org/10.34133/2020/3521852

Dutta, A., & Zisserman, A. (2019). The VIA Annotation Software for Images, Audio and Video. *Proceedings of the 27th ACM International Conference on Multimedia*. https://doi.org/10.1145/3343031.3350535

Dyrmann, M., Skovsen, S., Laursen, M. S., & Jørgensen, R. N. (2018). Using a fully convolutional neural network for detecting locations of weeds in images from cereal fields. *International Conference on Precision Agriculture*.

Guo, J., Li, X., Li, Z., Hu, L., Yang, G., Zhao, C., Fairbairn, D., Watson, D., & Ge, M. (2018). Multi-GNSS precise point positioning for precision agriculture. *Precision Agriculture*, *19*(5), 895–911. https://doi.org/10.1007/s11119-018-9563-8

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Recent Developments on Deep Big Vision*, *187*, 27–48. https://doi.org/10.1016/j.neucom.2015.09.116

Khan, Z., & Miklavcic, S. J. (2019). An Automatic Field Plot Extraction Method From Aerial Orthomosaic Images. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00683

Milioto, A., Lottes, P., & Stachniss, C. (2018). Real-Time Semantic Segmentation of Crop and Weed for Precision Agriculture Robots Leveraging Background Knowledge in CNNs. *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2229–2235. https://doi.org/10.1109/ICRA.2018.8460962

MPI. (2019). *2019 Forestry Labour Requirements Survey*. Ministry for Primary Industries. https://www.mpi.govt.nz/dmsdocument/33484-2019-Forestry-Labour-Requirements-Survey

Parraga-Alava, J., Cusme, K., Loor, A., & Santander, E. (2019). RoCoLe: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. *Data in Brief*, *25*, 104414. https://doi.org/10.1016/j.dib.2019.104414

Pearse, G. D. (2017). *Nursery Seedling Counts* [Scion Internal Report]. Scion.

Rawat, W., & Wang, Z. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, *29*(9), 2352–2449. https://doi.org/10.1162/neco_a_00990

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection. *ArXiv:1506.02640 [Cs]*. http://arxiv.org/abs/1506.02640

Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7263–7271.

Roth, L., Aasen, H., Walter, A., & Liebisch, F. (2018). Extracting leaf area index using viewing geometry effects—A new perspective on high-resolution unmanned aerial system photography. *ISPRS Journal*

of Photogrammetry and Remote Sensing, *141*, 161–175. https://doi.org/10.1016/j.isprsjprs.2018.04.012

Shannon, D. K., Clay, D. E., & Kitchen, N. R. (2020). *Precision Agriculture Basics* (1st Edition). ACSESS.

Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S., & Batra, N. (2020). PlantDoc: A Dataset for Visual Plant Disease Detection. *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, 249–253. https://doi.org/10.1145/3371158.3371196

Skovsen, S., Dyrmann, M., Mortensen, A. K., Laursen, M. S., Gislum, R., Eriksen, J., Farkhani, S., Karstoft, H., & Jorgensen, R. N. (2019). *The GrassClover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture*. 0–0. https://openaccess.thecvf.com/content_CVPRW_2019/html/CVPPP/Skovsen_The_GrassClover_Image_Dataset_for_Semantic_and_Hierarchical_Species_Understanding_CVPRW_2019_paper.html

Thorp, K. R., & Tian, L. F. (2004). A Review on Remote Sensing of Weeds in Agriculture. *Precision Agriculture*, *5*(5), 477–508. https://doi.org/10.1007/s11119-004-5321-1

Tresch, L., Mu, Y., Itoh, A., Kaga, A., Taguchi, K., Hirafuji, M., Ninomiya, S., & Guo, W. (2019). Easy MPE: Extraction of Quality Microplot Images for UAV-Based High-Throughput Field Phenotyping. *Plant Phenomics*, *2019*, 2591849. https://doi.org/10.34133/2019/2591849

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

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., & Yan, J. (2020). Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. *Molecular Plant*, *13*(2), 187–214. https://doi.org/10.1016/j.molp.2020.01.008