






Research Article

Using a Camera System for the In-Situ Assessment of Cordon Dieback due to Grapevine Trunk Diseases

Julie Tang ¹, Olivia Yem,¹ Finn Russell,¹ Cameron A. Stewart ¹, Kangying Lin,¹
Hiranya Jayakody,¹ Matthew R. Ayres,² Mark R. Sosnowski ^{2,3}, Mark Whitty ¹
and Paul R. Petrie ^{1,2,3,4}

¹School of Mechanical and Manufacturing Engineering, University of New South Wales, Kensington, NSW 2052, Australia

²Crop Sciences, South Australian Research and Development Institute, Waite Research Precinct, Urrbrae, SA 5064, Australia

³School of Agriculture, Food and Wine, Waite Research Institute, The University of Adelaide, Glen Osmond, SA 5064, Australia

⁴College of Science and Engineering, Flinders University, Adelaide, SA 5001, Australia

Correspondence should be addressed to Julie Tang; julie.tang92@gmail.com

Received 20 January 2023; Revised 1 August 2023; Accepted 17 August 2023; Published 18 October 2023

Academic Editor: Chris Steel

Copyright © 2023 Julie Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background and Aims. The assessment of grapevine trunk disease symptoms is a labour-intensive process that requires experience and is prone to bias. Methods that support the easy and accurate monitoring of trunk diseases will aid management decisions. **Methods and Results.** An algorithm was developed for the assessment of dieback symptoms due to trunk disease which is applied on a smartphone mounted on a vehicle driven through the vineyard. Vine images and corresponding expert ground truth assessments (of over 13,000 vines) were collected and correlated over two seasons in Shiraz vineyards in the Clare Valley, Barossa, and McLaren Vale, South Australia. This dataset was used to train and verify YOLOv5 models to estimate the percentage dieback of cordons due to trunk diseases. The performance of the models was evaluated on the metrics of highest confidence, highest dieback score, and average dieback score across multiple detections. Eighty-four percent of vines in a test set derived from an unseen vineyard were assigned a score by the model within 10% of the score given by experts in the vineyard. **Conclusions.** The computer vision algorithms were implemented within the phone, allowing real-time assessment and row-level mapping with nothing more than a high-end mobile phone. **Significance of the Study.** The algorithms form the basis of a system that will allow growers to scan their vineyards easily and regularly to monitor dieback due to grapevine trunk disease and will facilitate corrective interventions.

1. Introduction

Grapevine trunk diseases (GTDs), such as Eutypa and Botryosphaeria dieback, are a pervasive and growing issue across the Australian wine industry that gradually reduces vineyard performance. Other trunk diseases, such as esca, Petri disease, Phomopsis dieback, and black foot disease, cause significant issues in other countries but have little impact in Australia [1]. Eutypa dieback causes leaves to become distorted and yellow, shoots to stunt, and cordons to dieback. Botryosphaeria dieback has no distinct foliar symptoms but causes similar cordon dieback. GTDs are detected by the visual assessment of experts, and the control

treatments for GTDs can be labour-intensive and most effective when administered preventively, early in the life of the vineyard [1–4]. Regular vineyard surveys are not feasible for many growers due to the labour resources required.

Methods for estimating GTD dieback from aerial imagery are well-established but are limited by ground vegetation [5]. Recent work by Ouyang et al. [6] used 3D point clouds collected using an unmanned aerial vehicle to detect GTD with an accuracy of 87.4%.

Deep learning techniques are part of a rapidly growing area of machine learning research that is especially effective for image analysis such as the detection of GTD. Deep learning methods typically result in higher classification

accuracy and faster testing times than using traditional machine learning methods, but most critically for this research, they eliminate the need for hand-crafted features [7]. These advantages over traditional machine learning have caused deep learning image analysis to be used in a wide variety of agricultural applications, including disease identification [8–11]. Researchers applied combinations of different networks, both existing and custom architectures, on datasets that they had collected and augmented themselves [7–9, 11, 12].

Mohanty et al. [12] previously achieved an overall accuracy of 99.35% when detecting crop-disease pairs from images of leaves using DL techniques, but there are key differences in the scope of research programs. They identified 26 different diseases in 14 crop species, but the images used were of a single leaf, taken in a controlled environment against a consistent background. Our research aims to detect the presence of a single disease in real time from images taken in the field, which introduces a number of complications. The in-field images introduce uncontrolled backgrounds and conditions, which can reduce the accuracy of the detection.

For in-field images, there has been a wide variety of work in object detection for agriculture, most notably fruit detection. Kuznetsova et al. [13] applied the YOLOv5 algorithms for apple detection with a false positive rate of 3.5% and a false negative rate of 2.8%. For strawberry detection, Chen et al. [14] achieved a false positive rate of 5.7 to 15.4% and a false negative rate between 4.6% and 18.1% on mature fruit. Wang et al. [15] studied various attributes of fruit detection using YOLOv5 and recommended that for single-class object detection, a minimum of 2500 objects should be labelled and used in training.

Beyond object detection, the classification of severity or other fruit attributes has also been studied. In addition to their mature strawberry detection, Chen et al. [14] investigated flower and immature fruit detection, with limited success. Wang et al. [16] adapted a VGG-16 classification model for estimating apple flower distributions, focussing on the maturity stage rather than the frequency of each class. They showed it to be more accurate and slightly faster than YOLOv5 when running on a personal computer.

The aim of this research was to develop an automated edge computing system that would allow growers to quantify the severity of cordon dieback caused by GTDs at a temporal (every season) and spatial (whole vineyard) scale. The system had to use a standard camera mounted on a vineyard vehicle and intelligent algorithms to monitor and map trunk disease and be implemented in such a way that it could be accessed by nontechnical users.

The aim can be split into two components:

- (1) Algorithms for cordon dieback assessment
- (2) System for data collection, processing, and display

This paper presents the first component of the research and evaluates its performance. The algorithm for cordon dieback assessment will be a machine-learning-based image processing algorithm trained using vineyard images collected on a standard camera and will be assessed on the

similarity of the algorithm's results to expert assessment on unseen vines. The scope of this research is limited to vines with bilateral cordons with spurs, trained on a single wire, due to these being more common than quadrilateral cordons in an Australian context.

2. Materials and Methods

2.1. Data. The data used to evaluate the dieback assessment networks were collected in October of 2020 and 2021 in eleven vineyards (cv. Shiraz, *Vitis vinifera* L.) in the McLaren Vale and the Clare and Barossa Valleys, South Australia. Images of the vines were collected using a mobile phone app developed for the purpose and operating on a pair of Samsung Galaxy S21+ phones (model SM-G996B) running Android 11. These phones were mounted on a trailer approximately 300 mm from the ground and the middle of the interrow, with the image sensor facing the vines and the phone orientated so the cordon wire was near the centre of the image. See Figure 1 for the experimental setup. The trailer was driven throughout the vine rows at a speed of approximately 7–9 km/h while imagery was captured and processed by the phone. Images were captured by each phone at a rate of at least 5 frames per second and a resolution of 1280 × 720 pixels. When combined with the wide field of view lens in the phone, this enabled the majority of each vine to be captured, with the trunk at the centre of the image. Further analysis of the achievable framerate is given in Section 4.3.

The proportion of cordon dieback on each vine was also visually assessed by two experts in the vineyard, and the score was recorded for each of the assessed vines [4, 17]. Cordon dieback in these vineyards is predominantly caused by GTDs, as evidenced by the presence of *Eutypa* dieback foliar symptoms, but it should be acknowledged that other factors such as nematodes, viruses, and other vineyard management practices may have contributed to the cordon dieback [18]. Each cordon was assigned a score in the range of (0, 50) in increments of 5, representing the percentage of dieback on the cordon as a total of the vine. Class 0 represents a complete and healthy canopy, and class 50 represents a cordon with no shoots or leaves. The assessment of dieback can vary between experts, and there is a particular difficulty in differentiating between the lower classes of 0, 5, and 10. These scores were matched with the images of the vines, and the images were labelled with bounding boxes around the trunk and around each cordon with the dieback score. During the growing season for the 2021 vintage, 12,642 bilateral cordon vines were scored and imaged, with 5,570 in the McLaren Vale and 7,072 in the Clare Valley. In the 2022 vintage growing season, an additional 1,149 bilateral cordon vines were scored and imaged, 568 from the McLaren Vale and 581 from the Clare Valley. The vines imaged in the 2022 vintage growing season were also imaged in the previous growing season. Overall, 13,791 vines were imaged and scored.

2.2. Algorithm Development. The model chosen for the dieback assessment network was YOLOv5s, as it is small enough to deploy on edge computers while maintaining



FIGURE 1: Two pairs of smartphones mounted facing the opposite rows for continuous data collection. This prototype setup also includes two action cameras and two vertically orientated phones; note that only images collected from the horizontally orientated phones were included here. Other than a companion phone to control these phones from the driver’s seat, no additional infrastructure was needed. See Figure 2 for an example image.

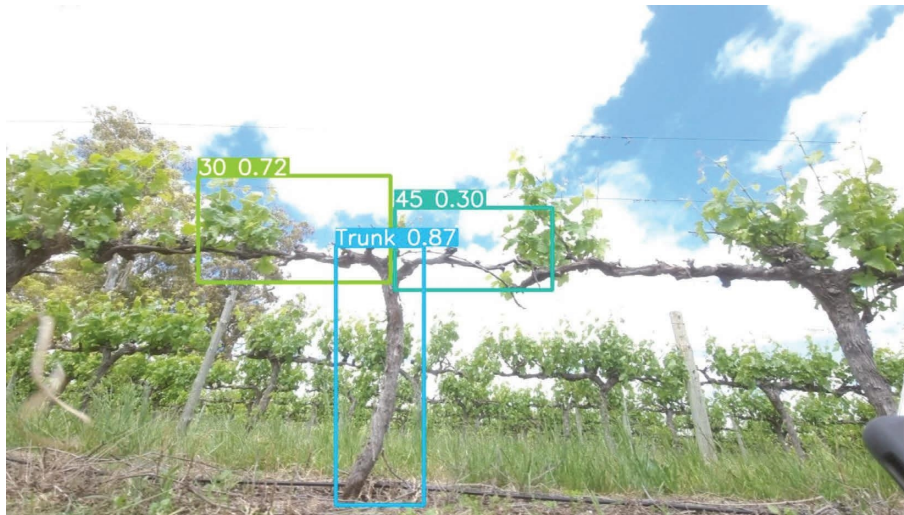


FIGURE 2: Detection from the unseen test set. The dieback on the right cordon is underestimated as the shoots from the adjacent cordon extend into the bounding box.

good detection results. The total dataset of all scored vine image across each vineyard resulted in an unequal number of instances between cordon classes (Figure 3(a)). This class imbalance can result in the assessment network overfitting to certain classes, artificially increasing the probability of assessing certain classes. This is particularly detrimental to the classes with very few training examples, such as classes 45 and 50 (Figure 3(a)). A subset of the total dataset with a balanced class distribution was created and used as the balanced training dataset for the network (Figure 3(b)). The number of training examples was greatly reduced when a balanced training dataset was created, as the number of training examples in each class was reduced to approximately the number of instances in the smallest class (class 50), and the majority of training examples consisted of classes 5, 10, and 15. Experimentation was used to explore the effects of various combinations of training sets across years.

Data augmentation techniques were used to increase the number of training examples, so that the network would be more robust to changes in orientation and variable environmental conditions. Each training image was flipped horizontally with a probability of 85%, which would simulate driving the vehicle carrying the camera in each direction along the row of vines, capturing images of both sides of the vine. A Gaussian blur was applied to the training images to increase the number of training instances and to increase the robustness of the algorithm to lower-quality images which may occur when capturing images from a moving platform. The weather conditions greatly affected the brightness of the grapevine images, so each of the images had its brightness both increased and decreased using a gamma correction function to simulate a range of weather conditions. Gamma correction applies a mathematical function to each pixel that either lightens or darkens the image overall, depending on the parameters used. The augmentations applied increased

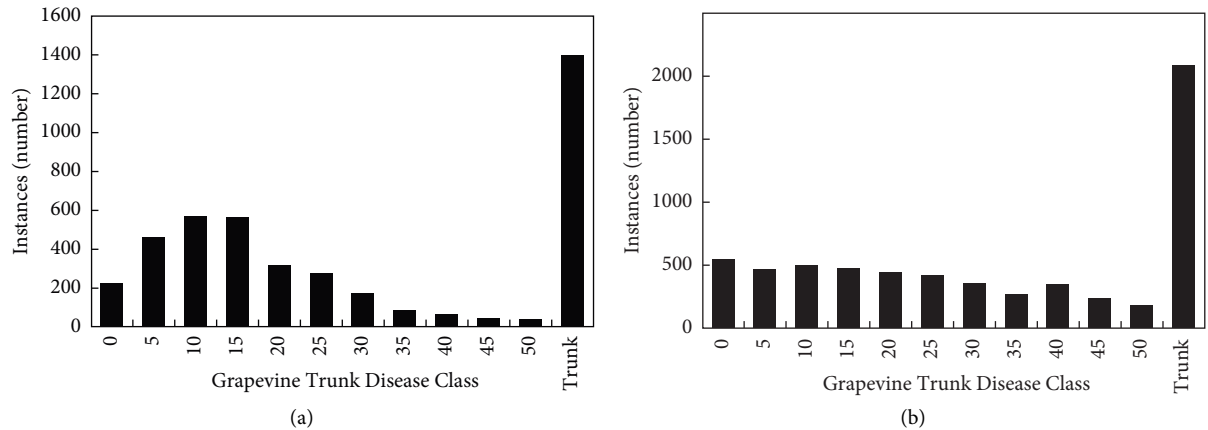


FIGURE 3: (a) The distribution of vine cordon GTD labels for imagery collected in the 2022 vintage. (b) The more balanced distribution of vine cordon GTD labels used for training the models, using data collected from both 2021 and 2022 vintages. Zero indicates no GTD symptoms observed, and 50 indicates complete dieback for that cordon. The total number of cordons is approximately twice the number of vines as indicated by the number of trunks identified.

the number of training images from 2084 images to 13076 images. The validation and test set images were not augmented in any way (Table 1).

Experiments were carried out to evaluate the suitability of the proposed algorithm by varying the hyperparameters and the data used for training each model (Supplementary Table 1 to Supplementary Table 6). The Ultralytics YOLOv5 version 6.1 Python library was used to implement the algorithm [19]. Training was carried out on a personal computer with 16 Intel® Core™ i9-9900KF CPUs using Python 3.7.3 and Ubuntu 18.04.6 LTS. The data used to train a deep learning image processing network is crucial and one of the defining factors in the results.

All models were evaluated on an unseen test set consisting of all the assessed vines in one block. This was to ensure that there was no overlap between the training and test data and that the results of each experiment could be directly comparable. The primary variables that were investigated were the data used to train the network and the training hyperparameters. All experiments were trained to completion, with completion being defined as the trend of the accuracy on the validation set across training epochs appearing to stabilise, with training lasting at least 300 epochs.

The most accurate model was evaluated not only on the unseen test set (Block 4) but also on a much larger set of images from the remaining blocks, again ensuring that these images were not included in either the training or validation sets.

2.2.1. Algorithm Evaluation Metrics. The success of the dieback assessment algorithm was measured using the following criteria:

- (i) Percentage of trunks detected
- (ii) Percentage of cordons detected
- (iii) Percentage of cordons with dieback scores identified correctly (class accuracy)
- (iv) Percentage of cordons with dieback scores identified within 5% of correct score (variation accuracy $\pm 5\%$)

- (v) Percentage of cordons with dieback scores identified with 10% of correct score (variation accuracy $\pm 10\%$)

The percentage of trunks detected should be as high as possible, as the system used to analyse the images relies on the detection of a trunk or half cordon to denote the results of the dieback assessment algorithm. By detecting the trunk and using images only where the trunk appeared close to the centre of the image, double-counting of successive half cordons was avoided. The algorithm must be able to detect the grape vine cordons in order to identify the extent of dieback, so the successful detection of cordons must occur for the algorithm to be effective. The assessment of the extent of dieback is subjective and can vary between experts. Therefore, the identification of the dieback score for each cordon will be assessed on an exact match to the in-field scoring as well as with a margin of 5% or 10% error.

2.3. System Overview. To manage, control, and observe the scanning process with ease, a smartphone-based two-application system was designed with a “controller” and a “scanner” application (Figures 4 and 5). The system only needs to connect to external devices on two occasions: for the initial fast localisation of the GNSS system or when downloading the map data for display on a computer. The system is able to process the images and automatically generate a map of the GTD in real time using only the “scanner” phone, the results of which are displayed on the “controller” phone or a computer. Further details of the system are outside the scope of this paper and available on request.

3. Results

3.1. Dieback Assessment Algorithm. Following the experiments used for training the dieback assessment algorithms, model 6 gave the best overall performance (Table 2). The trunk class was excluded from the confusion matrix (Figure 6) for the best performing model (model six) as all

TABLE 1: Number of images in each data subset.

Data subset	Number of images
All data	13,791
Balanced training data	2,084
Augmented balanced training data	13,076
Balanced validation data	821
Test data from a single unseen vineyard with no augmentations or balancing	172

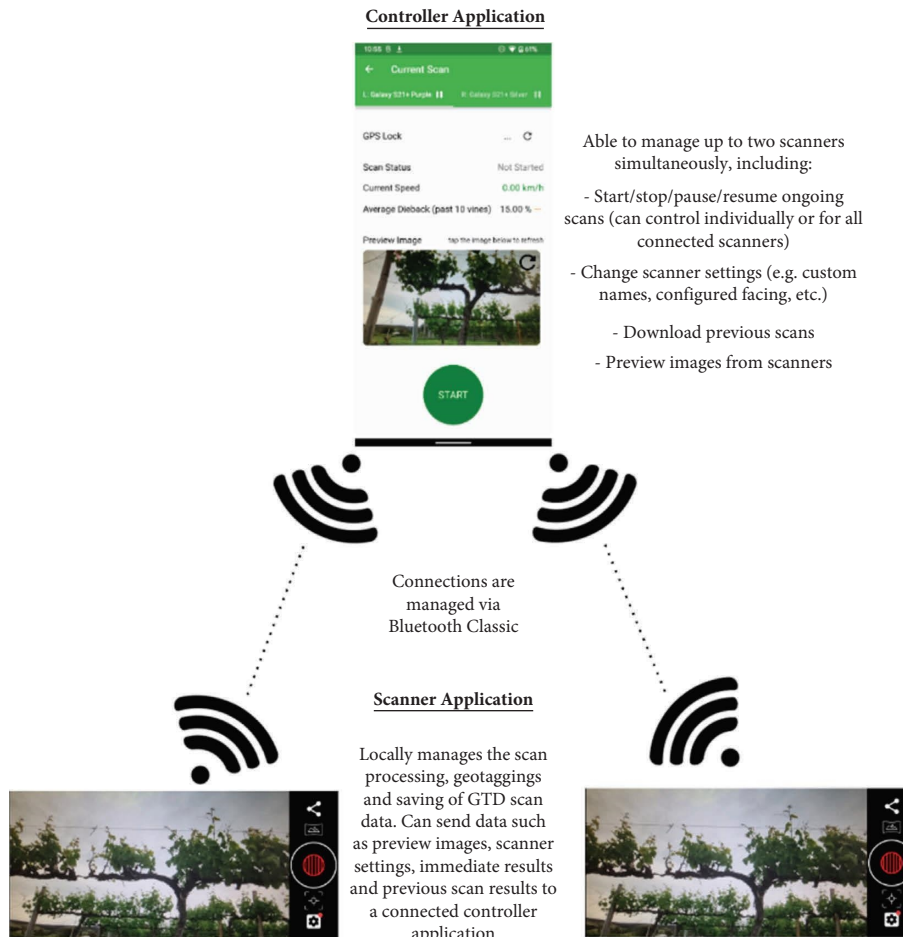


FIGURE 4: Application ecosystem overview.

trunks were correctly detected in the test set. Missing cordons, which are cordons that were labelled, but not detected by the algorithm, were designated a separate class (“M”) in the confusion matrix.

Model 6 was applied to images collected in the same blocks used for training. Even though these vines and images were not seen by the model during training or validation, excellent correlation with ground truth is seen, with over 99% of vines having an estimated GTD dieback severity within 10% of the manual ground truth (Table 2 and Figure 6(b)).

When the most successful model (model 6) was applied to the unseen test set; that is, with vines from a block completely unseen in the training or validation, the shape of the distribution is well matched against ground truth data (Figure 7(a)). Similar patterns were seen for the blocks used

as part of the validation (within the training process) (Supplementary Figure 1). Examples of detections in images are shown in Figures 8 and 9.

3.2. Evaluation of the Selected Model across Eleven Test Sites.

Data from the eleven sites used for training and validating the algorithm were processed with model 6 using the smartphone—with an additional block in the Barossa Valley also mapped (Block 1). Histograms were used to display the distribution of GTD severity across the block (Supplementary Figure 1). The vines and severity of GTD were georeferenced and plotted on aerial images (Figures 10–12).

In Block 1 (Figure 10), the mapped data displayed a high degree of average severity uniformly distributed across all of the surveyed vines. Whilst there are pockets of higher-

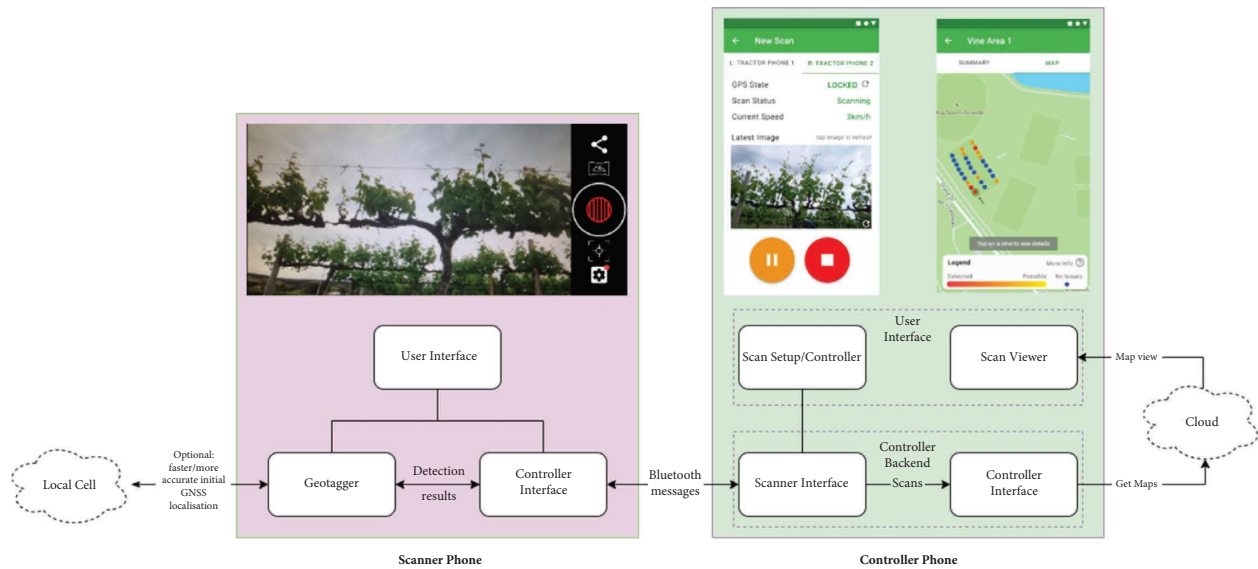


FIGURE 5: System overview diagram. The internal phone GNSS receiver is used to geotag each image. As such, the positional accuracy is dependent on having mobile phone reception on initial setup.

severity vines (such as in the centre of the top row), most vines exhibit symptom severities in the 40–60 percent range. This could indicate an older block, where the disease has had time to spread throughout most vines and less attention has been placed on remedial treatment.

In Block 5 (Figure 11), a high concentration of vines exhibiting severe symptoms were located at the northern end of the rows. Grapevine trunk disease does not normally follow a spatial pattern—so the grouping of the affected vines in one section of the vineyard was surprising. On further investigation, it was identified that the northern end of the block had reduced vigour as it is prone to frost, and a frost event had occurred several weeks before the assessment. Regardless of the cause, this gives growers an indicator that this is an area where the vines are performing poorly. A further manual inspection would often be made of the worst-affected areas to confirm the cause of an unusually concentrated area of increased dieback.

The mapping of Block 8 (Figure 12) exhibits less severe symptom severity. The high-symptom severity vines are clustered into small groups and distributed across the eastern portions of the block.

Vine symptom severity was usually normally distributed across the respective block, with a skew towards lower levels of severity (Figure 13). The results across blocks were typically clustered to a 10–20% range with some outliers. Blocks 3 and 5 exhibit results with a wider spread, with lower peaks, and a flatter distribution. In Block 5, this was a cause of the severe concentration of vine symptom severity in a small section of the block (see Figure 11).

3.3. Application Performance and Optimisation. The target framerate (5 FPS) was achieved consistently as a result of optimisation of the phone application. Images were captured at 1280×720 pixels and processed at 640×360 pixels using model 6. The two test phones used (128 GB and 256 GB

models of the SM-G996B Samsung Galaxy S21+ 5G) were both able to maintain a throughput of at least 5 FPS, shown as the ability to process individual images consistently in less than 200 ms over 110 minutes (Figure 14). The increase in processing time observed in the 128 GB model at the 60 minute mark is likely due to processor throttling as the phone heated up over time; however, the 200 ms threshold was not exceeded.

4. Discussion

4.1. GTD Detection Algorithm

4.1.1. Trunk and Cordon Detection. Trunk detection was high across all the experiments, with at least 97% of trunks being detected in each experiment and trunk detection as high as 100% in two of the trained models. Trunk detection was consistently high because of the number of instances in the training data and the appearance of the trunks. The trunks are visually distinct from the cordons, most notably due to their orientation. For every grapevine, there is a single trunk and two bilateral cordons with spurs; quadrilateral-cordon vines were considered out of scope for this research due to their distinctly different appearance. Given that the cordons are broken down into 10 classes based on the extent of dieback, the number of instances of trunks is much higher than any other class. Deep learning object detection algorithms require many examples to accurately detect objects in images; therefore, the high number of training instances for trunks ensures that the trunk detection was successful.

The algorithm must detect the grapevine cordons in order to classify them based on the extent of dieback, which makes the percentage of cordons detected critical to the overall performance of the algorithm. The percentage of detected cordons rose with the number of training examples and the increase in the left-right flip during training. The increase of the left-right flip hyperparameter also effectively

TABLE 2: Results of training GTD detection algorithm when applied to a completely unseen test set (Block 4).

Identifier	Training data	Training hyperparameters	Trunks detected (%)	Cordons detected (%)	Class accuracy (%)	Variation accuracy ($\pm 5\%$)	Variation accuracy ($\pm 10\%$)
Model 1	V2022 training and validation data. No augmentations applied	See Supplementary Table 1	99.42	15.79	6.43	11.68	14.02
Model 2	V2022 training and validation data. No augmentations. 85% left-right flip applied	See Supplementary Table 2	99.42	82.56	24.12	56.09	71.49
Model 3	V2022 training and validation data. Exposure and blur augmentations applied	See Supplementary Table 3	100.00	88.66	24.99	63.35	77.00
Model 4	V2021 and V2022 training and validation data. No augmentations. 85% left-right flip applied	See Supplementary Table 4	97.09	85.17	26.74	60.74	74.96
Model 5	V2021 and V2022 balanced training and validation data. No augmentations	See Supplementary Table 5	98.84	97.1	25.72	64.39	81.52
Model 6	V2021 and V2022 balanced training and validation data. Exposure and blur augmentations applied. 85% left-right flip applied	See Supplementary Table 6	100.00	99.42	26.16	63.56	84.18



FIGURE 6: Confusion matrices for cordon class detections for model 6 (as described in Table 2). (a) Images for the unseen test set were collected from vineyards where other images from those vineyards were used in training. (b) Images for the unseen vines comprised of vines from Block 4 which had not been used for any model training or validation the yellow off-diagonal terms have been highlighted as ± 10 threshold has been applied in the evaluation of results in this paper. The general diagonal form of the results is evident, with some outliers (discussed below).

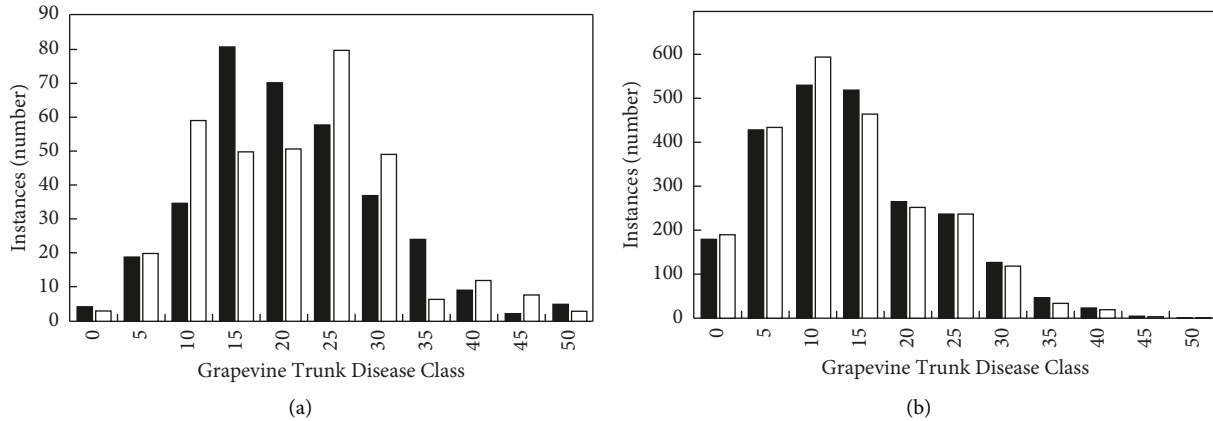


FIGURE 7: A comparison of manually assessed and detected severity of grapevine trunk disease in test sets when using model 6 described in Table 2 for (a) images collected from Block 4 and (b) unseen vines in vineyards used for training the algorithm. Black bars indicate manually assessed severity and white bars show detected severity. A score of 0 represents no cordon dieback and a score of 50 represents no foliage present.

increased the number of training images as the images are reversed horizontally, as most clearly seen in the 67% increase in the detected cordons between experiments 1 and 2 when a 0.85 left-right flip was applied. The effects of increasing the training examples diminished as more training examples were used, but the network achieved the correct detection of 99% of cordons in the unseen test set which underpins the rest of the analysis.

Two cordons were not detected in the test set. In the first example of a missing cordon, the cordon was not detected as

there was a tree in the background with the foliage extending above and below the cordon, so the cordon was not distinguished from the background (Figure 15). For the second missed cordon detection, the photo is blurred and the leaves are pale in the image, but the cordon is not unrecognisable to a human observer (Figure 16). There are other considerations for the algorithm in this example. First, the right cordon was detected with a confidence score of 0.41, low compared to the majority of cordon confidence scores, which suggests that the light conditions and the blur (more

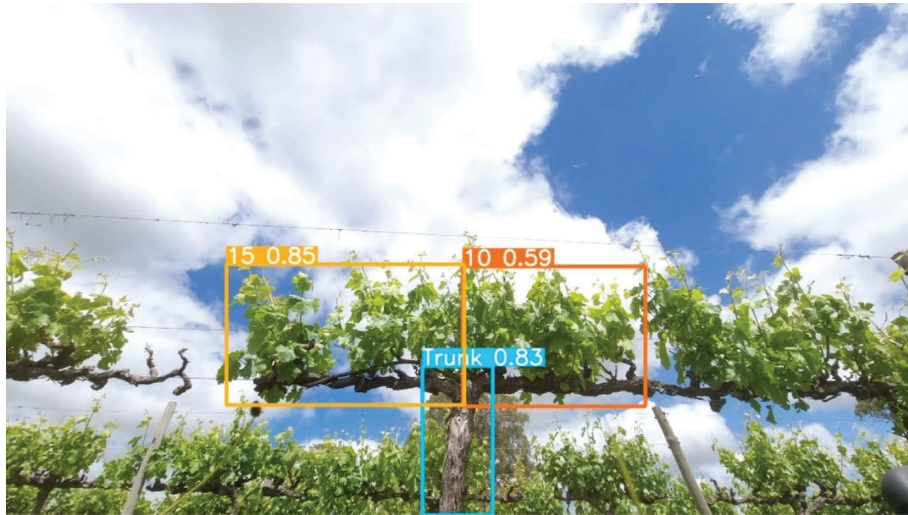


FIGURE 8: Detections from model 6. The labelled classes were 15 on the left and 15 on the right. The detected classes were 15 for the left cordon and 10 for the right, as the detected classes are given first in the bounding box labels. The decimal in the bounding box is the confidence score, an expression of how sure the algorithm is of what is detected.

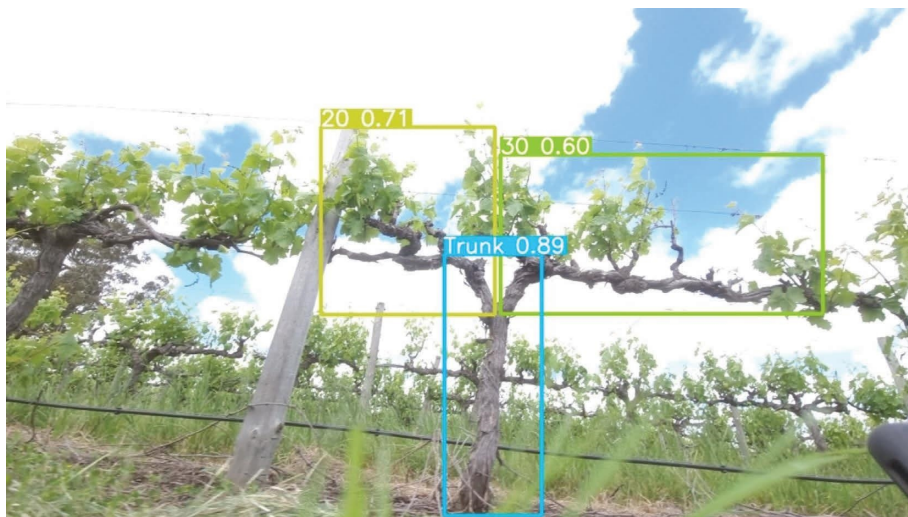


FIGURE 9: Detections from model 6. The labelled classes were 25 on the left and 30 on the right. The detected classes were 20 for the left cordon and 30 for the right, as the detected classes are given first in the bounding box labels. The decimal in the bounding box is the confidence score, an expression of how sure the algorithm is of what is detected.

significant on the left cordon) were a major factor in any detections of the left cordon falling under the confidence threshold of 0.25. The bounding box for the trunk is also much wider than the typical trunk bounding boxes as the trunk itself is slanted (Figure 16). The cordons do not originate from the centre of the trunk bounding box, as is the norm, and combined with the thinness of the left cordon, this creates additional difficulties for the detection of this cordon.

In terms of pure object detection accuracy using YOLOv5, Kuznetsova et al. [13] obtained an accuracy of 97.1% in counting apples in general images. It is not surprising that the trunk detection results in this work are slightly higher in accuracy given the size and uniqueness of the shape compared with apples.

4.1.2. GTD Dieback Detection. Detecting trunks and cordons in vineyard images allows the algorithm to fulfill the aim of detecting the extent of dieback. Model 4 had the highest class accuracy, with 27% of cordons classified by the algorithm matching the labels given in the vineyard (Table 2). As previously stated (see Section 3.1), the dieback scoring is subjective and can vary between different experts, and there is particular difficulty in differentiating between classes 0, 5, and 10. When the variation accuracy within 5% and 10% was considered, model 6 had the highest $\pm 10\%$ variation accuracy (84%), as well as a higher $\pm 5\%$ variation accuracy and more cordons detected than model 4. Models 4 and 6 had slightly different training hyperparameters, but the main difference between these models was the training data



FIGURE 10: Block 1 vine symptom mapping—each point represents the overall vine assessment combining both cordons.



FIGURE 11: Block 5 vine symptom mapping—each point represents the overall vine assessment combining both cordons.



FIGURE 12: Block 8 vine symptom mapping—each point represents the overall vine assessment combining both cordons.

used. Model 6 used the V2021 and V2022 augmented data, resulting in many more training examples which improved the performance similarly to the cordon detection. The augmented data adjusted the exposure and blur of the training images, which theoretically would make the algorithm more robust to changing light conditions and changes in photo quality, as shown by Wang et al. [15]. However, the test set images were taken on a single day using one model of phone. Therefore, the effects of changing illumination due to weather conditions and changing the photo quality due to the camera used were not directly assessed.

The $\pm 10\%$ variation accuracy is quite reasonable given that some classes included as little as 50 training images (prior to augmentation), and Wang et al. [15] recommended 2500 training images for a single class. The detection algorithm is more likely to overestimate the extent of the dieback rather than underestimate it, with 19 cordons underestimated by at least 15% and 28 cordons overestimated by at least 15%, the manual scoring factors in all the shoots extending from each cordon, including when a shoot extends over the adjacent cordon, although this is not common. The detection algorithm estimates the extent of dieback based on the volume of leaves around the cordon, as the training images were labelled with a bounding box around the cordon, and the volume of leaves in the bounding box is largely consistent with the

amount of dieback. If a shoot extends to another cordon, the algorithm will estimate the extent of dieback incorrectly (see Figure 2). The right cordon in this example was given a manual score of 50, but the detection algorithm assigned a score of 30 due to the shoots from the adjacent cordon extending into the bounding box. The accuracy of 84% on an individual vine level compares well with that of Ouyang et al. [6], who achieved 87% accuracy on an aggregated row level. The number of classes of severity used by Ouyang et al. [6] was slightly smaller, which would also lead to improved results.

When the frequencies of each class in the manual scoring and detections were compared between a set of vines from a vineyard that was not used in training (Figure 7(a)) and a larger test set consisting of the unseen vines in vineyards that were used in training the algorithm (Figure 7(b)), the algorithm performed better on the unseen vines in vineyards used in training. The images of the unseen vineyard are often overexposed, although some of the unseen vines in the training blocks are overexposed as well, these are a higher proportion in the unseen vineyard. The images in the unseen vineyard are blurred in addition to the light conditions that would cause more difficulties in accurately estimating the dieback.

There are two possible courses of action to potentially improve the results. The training images could be given new scores based on the volume of leaves around each cordon.

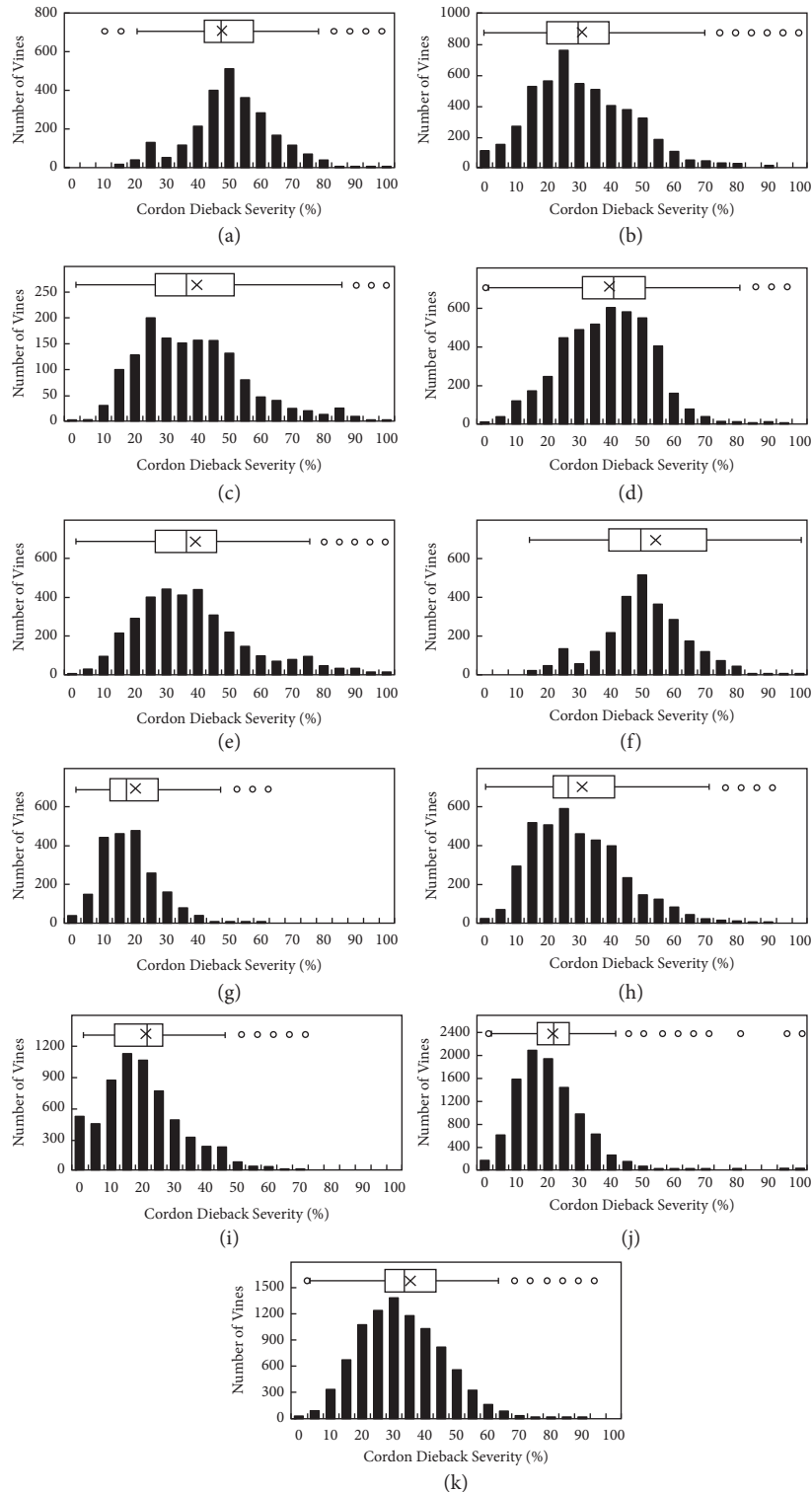


FIGURE 13: Cordon dieback severity distribution across blocks, relative to 25% of total number of vines scanned in that block. Subfigures (a–k) correspond to Blocks 1 to 11, respectively.

This would not reflect the manual scoring system as closely but may align more with the needs of the growers. The presence of growth in the area is more important than exactly which vine it extends from. Alternatively, semantic segmentation could be used to identify exactly which shoots

extend from each cordon to align with the manual scoring more closely. Semantic segmentation would not appear to currently be a feasible technique for real-time processing on a mobile phone due to the need to classify each pixel rather than identify three bounding boxes. The training data would

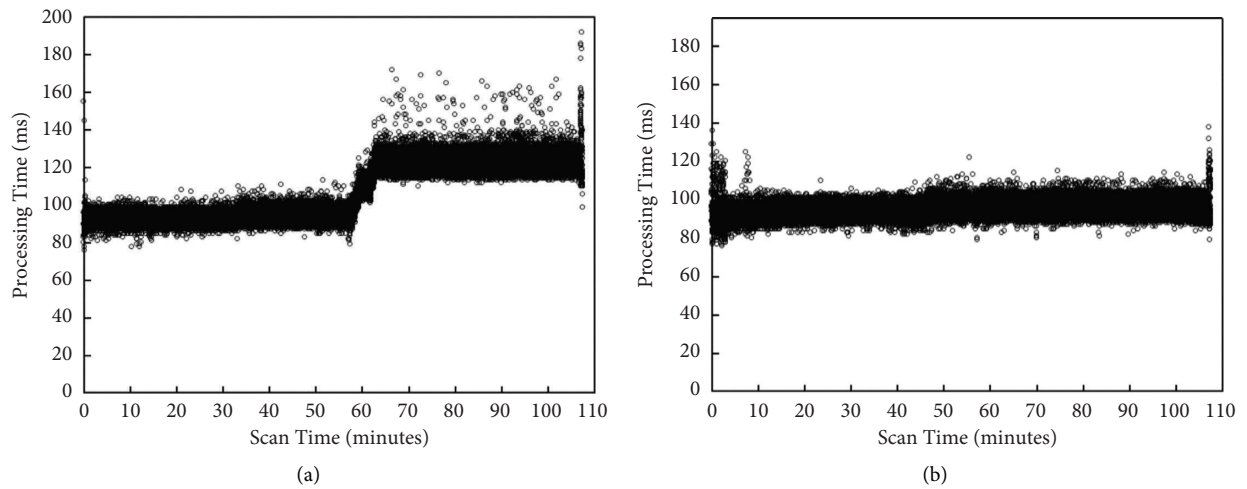


FIGURE 14: Sustained processing time of GTD dieback scanning across different S21+ models—128 GB (a) and 256 GB (b).



FIGURE 15: A missing cordon detection from the test set.



FIGURE 16: Another missing cordon detection from the test set.

also need to be labelled by experts in manually scoring dieback as correctly allocating each shoot to the correct cordon is very important.

4.2. Evaluation of the Selected Model across Eleven Test Sites. By comparing the results across all the test sites, significant variation in the spatial pattern, incidence, and severity of GTD dieback symptoms was observed. This may be due to different ages of vines, different GTDs involved, or local climatic conditions. The incidence and severity of GTDs increases with age of vines [3, 4, 20]. The distribution of pathogens that cause *Eutypa* and *Botryosphaeria* dieback varies between Australian regions [21, 22] which may also explain some of the variability observed between regions in the current study. Rainfall is required for infection, and certain temperature and humidity conditions favour the different causal pathogen species [23, 24]. Nonetheless, this work lays the foundation for the analysis of data over multiple seasons or before and after remediation activities to monitor changes in GTD symptoms.

Very few blocks have a low average severity (Supplementary Figure 1), partially due to grapevines being a natural system and not growing uniformly despite the best endeavours of growers. It also highlights the potential for growers to tend to underestimate the severity throughout their blocks, as once the canopy is more fully grown, shoots will tend to spread out and disguise diseased sections of the cordon.

4.3. Evaluation of Smartphone Application. After running for many hours in-field conditions, the smartphone “scanner” application was able to successfully collect, process, and geo-reference all the images across the eleven test sites. Despite the hard requirement of 5 FPS processing, the phone was able to sustain this performance consistently in tests lasting more than an hour. Compared with the aerial method of Ouyang et al. [6], the ability to undertake the survey using only a mobile phone mounted on a vehicle is somewhat simpler yet of comparable accuracy, giving greater opportunity for industry adoption.

5. Conclusions

This paper presents and evaluates an algorithm to detect and map grapevine trunk disease dieback using only a smartphone. The YOLOv5-based algorithm was successfully applied in a smartphone app to collect and process data from more than 13,000 vines in the McLaren Vale, Clare Valley, and Barossa Valley regions of South Australia across two growing seasons and ten vineyards.

The algorithm was effective, as it was able to classify 99% of cordons within 10% of expert visual dieback assessment on unseen vines from the same blocks as used in the training and validating the model. When tested on vines from a different block, again unseen by the model, a classification accuracy of 84% was achieved and 99.5% of cordons were detected.

Furthermore, the algorithm reliably operated at a frame rate of 5 FPS on a commercially available smartphone,

including capturing, processing, and mapping the data with GNSS.

Further research into the robustness of the algorithm under different weather conditions and image quality is recommended to ensure that the system remains effective for many models of phone used and that the system is not reliant on good weather conditions. A variation of the algorithm that can be used in vineyards with different training systems (e.g., multiple cordons) would also be a recommended area of further research. A reliance on existing deep learning algorithms mean the GTD level had to be discretised; further work could examine methods for providing a continuous numerical output.

Being able to transform a deep learning model trained on a server to run in real-time on a smartphone has provided a powerful tool for growers to attach to a vehicle and obtain maps of GTD dieback symptoms. This opens the potential for rapid assessment of GTD more widely across the industry on bilateral cordon-trained vines. It also highlights the potential for deep learning models to be trained to detect visual symptoms of other diseases and to be applied in the field with just a smartphone.

Data Availability

The underlying data for this study are stored in the UNSW Data Archive. The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by the South Australian Wine Industry Development Scheme (2020) and was led by the South Australian Research and Development Institute (SARDI). Fieldwork, ground truth assessment, system testing, and grower liaison were undertaken by SARDI, and the algorithm and system were developed by UNSW Sydney (UNSW). This research would not have been completed without the assistance of DJs Growers, especially Mr. Joe Siebert, and the Clare Valley Wine and Grape Association, especially Ms. Anna Baum, who provided advice and access to a number of growers who in turn made their blocks available for the study. We appreciate the support provided by the growers in the McLaren Vale and the Clare Valley in allowing us to access their vineyards.

Supplementary Materials

Supplementary Tables 1 to 6 include the training hyperparameters used to train the YOLOv5s object detection network to identify grapevine trunks and assess the dieback, resulting in the results given as experiments 1 to 6. Supplementary Figure 1 shows validation results on blocks used during training, albeit with images not contributing to the training process. (*Supplementary Materials*)

References

- [1] D. Gramaje, J. R. Úrbez-Torres, and M. R. Sosnowski, "Managing grapevine trunk diseases with respect to etiology and epidemiology: current strategies and future prospects," *Plant Disease*, vol. 102, no. 1, pp. 12–39, 2018.
- [2] M. R. Sosnowski, T. J. Wicks, and E. S. Scott, "Control of Eutypa dieback in grapevines using remedial surgery," *Phytopathologia Mediterranea*, vol. 50, pp. S277–S284, 2011.
- [3] M. Sosnowski and D. Mundy, "Prevalence of grapevine trunk disease in New Zealand – observations from vineyard surveys," *New Zealand Winegrower*, vol. 117, pp. 116–121, 2019.
- [4] M. Sosnowski, "Grapevine trunk disease—best practice management guide," 2021, <https://www.wineaustralia.com/growing-making/pest-and-disease-management/eutypa-dieback>.
- [5] C. Poblete-Echeverria, S. Fuentes, S. Ortega-Farias, J. Gonzalez-Tallice, and J. A. Yuri, "Digital cover photography for estimating leaf area index (LAI) in apple trees using a variable light extinction coefficient," *Sensors*, vol. 15, no. 2, pp. 2860–2872, 2015.
- [6] J. Ouyang, R. De Bei, and C. Collins, "Assessment of canopy size using UAV-based point cloud analysis to detect the severity and spatial distribution of canopy decline," *OENO One*, vol. 55, no. 1, pp. 253–266, 2021.
- [7] N. Zhu, X. Liu, Z. Liu et al., "Deep learning for smart agriculture: concepts, tools, applications, and opportunities," *International Journal of Agricultural and Biological Engineering*, vol. 11, no. 4, pp. 21–28, 2018.
- [8] Q. Zhang, Y. Liu, C. Gong, Y. Chen, and H. Yu, "Applications of deep learning for dense scenes analysis in agriculture: a review," *Sensors*, vol. 20, no. 5, p. 1520, 2020.
- [9] L. Santos, F. N. Santos, P. M. Oliveira, and P. Shinde, "Deep learning applications in agriculture: a short review," in *Iberian Robotics Conference*, pp. 139–151, Springer, Cham, Switzerland, 2019.
- [10] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: a review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.
- [11] A. Kamilaris and F. X. Prenafeta-Boldú, "A review of the use of convolutional neural networks in agriculture," *The Journal of Agricultural Science*, vol. 156, no. 3, pp. 312–322, 2018.
- [12] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.
- [13] A. Kuznetsova, T. Maleva, and V. Soloviev, "Detecting apples in orchards using YOLOv3 and YOLOv5 in general and close-up images," in *Advances in Neural Networks—ISNN 2020*, M. Han, S. Qin, and N. Zhang, Eds., vol. 12557, Cham, Switzerland, Springer, 2020.
- [14] Y. Chen, W. S. Lee, H. Gan et al., "Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages," *Remote Sensing*, vol. 11, no. 13, p. 1584, 2019.
- [15] X. Wang, J. Tang, and M. Whitty, "Data-centric analysis of on-tree fruit detection: experiments with deep learning," *Computers and Electronics in Agriculture*, vol. 194, Article ID 106748, 2022.
- [16] X. Wang, J. Tang, and M. Whitty, "DeepPhenology: estimation of apple flower phenology distributions based on deep learning," *Computers and Electronics in Agriculture*, vol. 185, Article ID 106123, 2021.
- [17] B. Henderson, M. R. Sosnowski, M. G. McCarthy, and E. S. Scott, "Incidence and severity of Eutypa dieback in grapevines are related to total surface area of pruning wounds," *Australian Journal of Grape and Wine Research*, vol. 27, no. 1, pp. 87–93, 2021.
- [18] P. O'Brien, R. De Bei, M. Sosnowski, and C. Collins, "A review of factors to consider for permanent cordon establishment and maintenance," *Agronomy*, vol. 11, no. 9, p. 1811, 2021.
- [19] Ultralytics, "YOLOv5," 2022, <https://github.com/ultralytics/yolov5>.
- [20] J. A. Duthie, G. P. Munkvold, and J. J. Marois, "Relationship between age of vineyard and incidence of Eutypa dieback," *Phytopathology*, vol. 81, p. 1183, 1991.
- [21] W. Pitt and R. Huang, "Identification, distribution and current taxonomy of Botryosphaeriaceae species associated with grapevine decline in New South Wales and South Australia," *Australian Journal of Grape and Wine Research*, vol. 16, no. 1, pp. 258–271, 2010.
- [22] F. P. Trouillas, W. M. Pitt, M. R. Sosnowski et al., "Taxonomy and DNA phylogeny of Diatrypaceae associated with *Vitis vinifera* and other woody plants in Australia," *Fungal Diversity*, vol. 49, no. 1, pp. 203–223, 2011.
- [23] J. R. Úrbez-Torres, "The status of Botryosphaeriaceae species infecting grapevines," *Phytopathologia Mediterranea*, vol. 50, pp. 5–45, 2011.
- [24] J. S. Mayorquin, D. H. Wang, M. Twizeyimana, and A. Eskalen, "Identification, distribution, and pathogenicity of Diatrypaceae and Botryosphaeriaceae associated with citrus branch canker in the southern California desert," *Plant Disease*, vol. 100, no. 12, pp. 2402–2413, 2016.