IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication 2024 – 01.

The Use of Machine-Learning and Data Science for Pest Detection in Viticulture: a Review (January 2024)

A. Barrett-Lennard[1], Masters(IT) Student, CSU
[1] Charles Sturt University, Wagga Wagga, NSW 2650 AU

Corresponding author: A. Barrett-Lennard (e-mail: abl030@gmail.com).

"This work was conducted for ITC571 as part of a Masters of Information Technology – 4913 Words"

ABSTRACT Machine-learning in viticulture has the potential to reduce costs and increase efficiency through the automated detection of pests and diseases in vineyards. However recent studies have mostly been either siloed in laboratory settings or using models that require compute resources inappropriate for farm conditions. This review looks at recent studies in the field of ML for pest and disease detection and determines the technologies used, their effectiveness and their practical use in a real-world viticultural system.

The review finds that current models, especially in computer vision, can be split in two rough groups. With some light weight models, such as YOLOv5, excelling at object detection and bounding in real-time, while other more complex models to classify pathogens require significant compute resources in image pre-processing or classification. We also find a lack of standardised data-sets or testing, with many models being proof of concept only and working on their own datasets.

The review thus presents a framework for future research whereby pest detection could be split into a two tier model. With simpler models running in real-time on edge devices that can detect and create standardised image datasets of vines or vine structures. These datasets being images only could then be either processed on farm or in the cloud by more specialised complex models aimed at detecting specific pathogens. This also solves the problem around model comparison in future research by standardising training and testing datasets.

INDEX TERMS *Agricultural Automation, Agricultural Technology, Classification Models, Cloud Computing, Computer Vision, Disease Detection, Edge Computing, Edge Devices, Image Processing, Machine Learning, Model Comparison, Pest Detection, Precision Agriculture, Real-time Detection, Research Framework, Standardized Datasets, Viticulture*

I. INTRODUCTION

Viticulture is the cultivation and study of grapevines, with the grapes themselves being used primarily for wine, table grapes and raisin production. The estimated total vineyard area of all types was 7.3 mHa in 2022 [1] while the industry had a total global value of 340 billion USD representing 0.4% of global GDP [2] Therefore the industry has a major economic impact all around the world, not only in countries that grow grapes but also industries that consume grape products. Grapevine pests have a significant impact on the viticultural industry with the most recent assessment from Wine Australia from 2010 putting the economic impact at $250 million AUD for Australia alone [3]. The use of machine-learning (ML) and data-science in viticulture represents an evolving tool for increased efficiency and lower costs by automating expensive field work for growth monitoring, pest detection and yield estimation that in the past has been done by expensive manual labour.

Recent reviews have focused on the use of ML in both estimating grapevine vegetative growth and yield estimation [4], [5] with good results in this respect, however, there has been no recent review of the use of ML and data-science to detect pests and diseases in vineyards. Mohimont *et al.* [5] point out in their review the current challenges of moving from lab-based models to the field and importance of future work taking into account the noisy nature of all farm based data. Also these environments are prone to having significant network connectivity challenges, transferring large amounts of data to the cloud is a difficult prospect. It is with this focus and lens that this review will look at the current state of ML and data-science research in pest detection.

Most studies in the field focused to computer-vision to determine pest and disease incidence with only two [6], [7] utilising different sensor data, audio and environmental sensors respectively. Much of the current work centres on lab based imagery with models trained in controlled environments on one specific pathogen. The relevance of this to a more practical farm system is also discussed with many of these models requiring more compute resources than is practical. However a few real-time models based upon edge devices suitable for the field were found with most running YOLO based object bounding neural network models. These are generally more lightweight and suitable to run on devices such as commodity smartphone hardware. However the review found that these models sometimes lack the precision of more complex deep CNN models that are trained on one particular pathogen. Thus the review proposes a framework for future research focusing on a two tier approach, with light weight object bounding models used in the field to create image data-sets of each individual vine or vine structure. These images could then be passed on to a more complex model running in the cloud or farm-based computer which is tuned to specific pathogen types of relevance. This solves several problems around limited compute and noisy field data while also simplifying research by providing standardised data types with which models can be tested and compared.

1) METHODOLOGY

This review will focus on studies published in predominately 2022/2023 with a few highly relevant studies selected from 2021. These studies were classified and compared (see Table 1) for their relevance to an on-farm pest detection system. A framework was then established (see Figure 2) and presented for future research.

II. Literature Review

Grapevines pathogens are varied both in the domain of life they come from and the many symptoms and pathologies that present on the vine. As such the best ML approaches to use in the detection of these pathogens varies considerably based upon the pathogen that is of interest. In the following review we outline the different approaches taken for different pathogen types, their accuracy and suitability for use in real-time applications. Table 1 organises the studies in this review by their technology mix and their practical relevance to on-farm settings, the green highlighted studies are real-time and/or edge device compatible and are therefore the most relevant to this review.

A. FUNGAL CANOPY PATHOGENS

Generally fungal pathogens of grape vines are some of the most detrimental to a vineyard during any one growing season. The main pathogens include Downy Mildew (DM) (Plasmopara viticola), Powdery Mildew (PM) (Erysiphe necator) and Botrytis (Botrytis cinerea). Gutiérrez *et al.* [8] took 841 images of grapevine leaves split into three classifications: infected with DM, infected with Spider Mite and healthy. Spider Mite and DM have remarkably similar pathologies on grapevine leaves and thus their discrimination is hard even for domain experts. Multiple models were then trained on the data as both a multi-class and binary classifier between each possible class. The binary classifiers are of limited value in a wider context, as for every variable introduced one must train a new model which could get computationally expensive. However the multi-class model demonstrated accuracies of 0.94 (F1 score) on the 25% holdout testing data from the dataset. However this study was significantly limited in its practical scope, although the images were taken in the field all computation was done in the lab and the resulting CNN was too computationally expensive for real-time edge applications.

Alternatively both Hernández *et al.* [9] and Zendler *et al.* [10] worked on lab images of leaf discs inoculated with DM spores. In both these works leafs were taken from the field and processed in the lab, they were cut into leaf discs and then inoculated with DM spores and left to sporulate. Zendler *et al.* [10] used RGB images only however Hernández *et al.* [9] utilised both RGB and hyperspectral imagery on the leaf discs. Interestingly the studies differed on their approach to measuring DM severity, with Zendler *et al.* [10] training a SCNN and Hernández *et al.* [9] using the Otsu Method, which is a form of discriminant analysis. Both studies found good correlation with manually annotated expert analysis of the discs, however the Otsu method requires orders of magnitude less compute resources as is thus of interest in edge applications. Hernández *et al.* [9] also trained a CNN on hyperspectral imagery to detect DM infection after three days, before it is visible to the human eye. They had good success at this asymptomatic detection which correlates well with other studies using hyperspectral imagery to detect asymptomatic infections [11], [12], [13].

While the previous studies used variants of computer vision to detect fungal outbreaks Hnatiuc *et al.* [7] used a vast IOT sensor network to monitor and predict when environmental conditions were conducive to fungal growth. An interesting approach that is predictive instead of reactive and novel for all pathogens in the recent literature. Interestingly this approach used traditional ML in the form of a Random Forest classifier and was extremely lightweight and computationally efficient while also having a near 98% accuracy on the obtained dataset. However the study is limited in its wider predictive power due to the limited amount of environmental conditions observed while also only being able to observe an outbreak of PM and no other pathogens. The studies real value, however, is in it's discrimination of variables associated with fungal growth and the computational efficiency of the process.

B. INSECTS

To date there has not been much work done utilising ML to detect insects in vineyards in contrast to the wider agricultural field, which has seen significantly more investment with the creation of standardised data-sets and testing [14]. However there has recently been significant work in Portugal to design a complete date pipeline for an android application to process and identify insects on sticky traps [15], [16], [17]. The pipeline first uses traditional computer vision techniques to focus, correct and standardise smartphones picture of sticky traps taken in the field [15]. This was then accurate enough to utilise these photos to train a ResNet50 model to detect and classify five different insect types relevant to the viticultural area the app was designed for, with accuracies ranging from 82 – 99% depending on class, however the major downside was the models took over 62 seconds to run on a modern smartphone [16] . ResNet50 is a variant of a CNN first made available in 2015 that excels at object and boundary detection, the model is also easy to fine-tune to specific object/s of interest [18]. This pipeline was then used to build an android application 'EyeOnTraps' that using pictures taken on device, counts and catalogues insect types found at each sticky trap [17]. This shows the ability for edge devices to run accurate computer vision models while also stressing how difficult it can be to present the information in a practical way to grape-growers.

C. GRAPEVINE TRUNK DISEASE (GTD)

GTD's are a pernicious pathogen of grapevines in that they kill vines, can be largely asymptomatic and there are limited control measures with the best control being prophylactic, centreing around pruning hygiene and infection prevention [19]. Detection and removal, therefore, of GTD's in the vineyard is vital to viticulturalists, however their stochastic asymptomatic nature makes this difficult. Pérez-Roncal *et al.* [11] and Calamita *et al.* [13] used hyperspectral imaging to detect asymptomatic infections of Esca (a fungal trunk disease) and Armillaria (a fungal root disease) respectively. Hyperspectral images capture a much larger spectrum of wavelengths than just visible light and can thus contain information not visible to the human eye.

The authors of both studies extensively used traditional data-mining techniques to quantise the spectral data and used clustering to find the most predictive and correlative variables. They then trained traditional classifiers with good success at detecting the asymptomatic infections, Naive Bayes had a 75% accuracy for Armillaria [13] and partial least square discriminant analysis with an accuracy of over 95% for Esca [11]. Its important to note that that these techniques were both used because the process yielded information about what in the hyperspectral images was predictive, which was of interest to the authors. If deep learning was used the models are 'black box' and no useful variables would be found. However future work could utilise deep learning models on hyperspectral data which could reduce pre-processing, increase classification speed and make it practical for real-time field applications.

Symptomatic GTD infection such as Eutypa Dieback show predictable foliar symptoms especially in spur pruned vines and are thus easier to detect. Tang *et al.* [20] utilised smartphones mounted to a trailer and towed through a vineyard to detect Eutypa dieback severity in a Shiraz vineyard. They trained a YOLOv5 model on over 15000 annotated images that ran at over 5 frames a second on a low-end modern smartphone with an accuracy of 85% in a multi-class severity estimate. YOLO models, first introduced in 2016, are shallower CNN's than ResNet and therefore more computationally efficient but still perform well at object bounding and detection - [21]. Figure 1. [20] shows the resulting imagery and bounding generated by the smartphone in real time. This work shows the promise of real time pathogen detection and mapping in a vineyard under field conditions using commodity hardware.

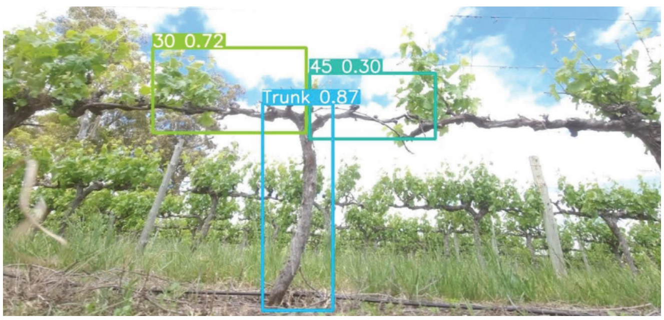

FIGURE 1. Smartphone imagery with YOLOv5 detection bounding and severity estimation [1]

Finally work done by Rudenko *et al.* [22] is of significant interest as it is also one of only four real-time models in this review. Working with a semi-automated grapevine grafting assembly line the authors created a fuzzy classifier to look at lesions within cutting scions. The classifier look for trunk diseases or lesions within the cuttings to assess their suitability for grafting. Being a compute efficient fuzzy classifier the model was able to work in real-time with the assembly line drastically reducing grafting time. However this work was done in a factory based setting with automated standardised imagery capture meaning the actual implementation does not translate well to on-farm conditions.

D. VIRUSES

Similar to GTD infection grapevine viruses can have both asymptomatic and symptomatic pathologies while also having limited control measure apart from detection, removal and hygiene [23]. Sawyer *et al.* [12] used hyperspectral imagery, again in a lab environment, to train both a CNN model and a Random Forest model to classify leaves infected asymptomatically with grapevine leafroll-associated viruses and grapevine red blotch virus. As in [11], [13] Sawyer *et al.* [12] had to utilise significant pre-processing for the random forest model to quantise the spectral data and find the most highly associated variables while the CNN model was just trained on standardised images. The random forest model saw an accuracy of 82.8% while the CNN saw an accuracy approaching 87%. This again shows the power of deep learning to utilise extra information in hyperspectral images and classify asymptomatic grapevine etiologies. However this study was again limited in that it was in a lab setting only, the practicality of such a system needs to be proven in the field in future work.

Wang *et al.* [24] had a novel strategy for identifying grapevine virus A infection in Shiraz vines by mapping canopy density over time throughout the growing season. A drone was flown over the vineyard and various times and the data processed back in the lab by an extremely simply Random Forest Classifier. The classifier identified pixels as either grapevine, soil, shadow or weeds with each class in the training set only containing 5-7 pixels. The model was then accurate enough to identify the canopy spread in pixels over time between flights. The delayed bud burst caused by grapevine virus A could then be easily identified in the data. The model was extremely lightweight, simple and usable on field data, showing that creativity and domain knowledge can lead to significant results.

E. BIRDS

Unique in this review is the work by Cinkler *et al.* [6] in that they used audio to train their model to detect birds in the vineyard. Birds are significant pests in vineyards with netting and other associated preventative measure costing wineries significantly [25]. Cinkler *et al.* [6] utilised a two tier approach to their system with a simpler SVM model running on the sensors within the vineyard, once this model hit a certain threshold of certainty about a particular bird call it then sent the information up the cloud. There a second CNN model was then utilised to gain further confidence on the species of bird call detected. This two tier system was shown to be computationally and network efficient for the limited environment on-farm as well as being accurate enough to detect starlings in near real-time. The two-tier idea is conducive to further study in it's applicability to other edge applications, it may translate well to visual data but more work is needed in this regard.

F. LESIONS AND ANOMALOUS BERRIES

Another way of thinking about pathogens is that they create a disease 'state' within the grapevine, this state varies in appearance from a 'healthy' or 'normal' state of a grapevine. This difference can then be used to train a computer vision model to detect these diseased or different states. This method does not detect a specific pathogen as in the previous studies but merely a state that differs from normal and therefore could potentially detect any pathogen whether known or unknown while also limiting re-training every time a new pathogen is added to detection software. This method has been used to detect changes both at a bunch [26] and grape [27] level.

Interestingly both these studies differed in their fundamental approach with Pinheiro *et al.* [26] hand annotating two bunch classes, healthy and damaged and training 3 different YOLO models to multi-classify the images. The classification accuracy was high at 97% however the mAP was low, meaning the model sometimes failed to accurately detect the bunches themselves. In contrast Miranda *et al.* [27] simplified the training schema in that they trained a convolutional variational autoencoder on only healthy berries and thus the model became a binary classifier with deviations from this being labelled as damaged berries. Accuracy for this approach was much higher with 92% of damaged berries labelled as such. This is an interesting approach that potentially could generalise to many other pathogens without re-training the model, though more work is needed in this regard.

III. Discussion

Machine-learning and data science in precision viticulture is a nascent field still in its early stages of development, many of the current studies focus on lab-based imagery [9], [10], [11], [12], [13], [22] (see Table 1.). Which, while useful as proof of concepts, have limited applications to the field in their current form. Accuracy for these lab based approaches, however, is generally excellent with most lab based approaches reaching over 90% accuracy for the chosen pathogen [10], [11], [22] except for some asymptomatic infections [9], [12], [13]. The controlled environment of these approaches leads to successful knowledge finding, especially in the field of hyperspectral imagery where spectral quantisation and variable clustering enables researchers to find specific disease indicators not visible to the human eye [9], [11], [12], [13]. The approaches use traditional ML techniques and deep learning all with good results, however all the approaches except Rudenko *et al.* [22] are significantly

Paper	Imagery Type	Technique	Algorithm	Training Dataset	Dataset: Lab/Field	Dataset Details	Pest Detected	Accuracy	Real-Time	Cloud or Edge
(Cinkler et al., 2022) [6]	N/A – Audio	Deep Learning	SVM & CNN	Generated for the Paper	Field	95 hours of various sounds, including 84 hours of bird song. Split into 4 second classified segments	Starlings	96.00%	Yes	Edge & Cloud
(Tang et al., 2023) [20]	RGB	Deep Learning	YOLOv5	Generated for the Paper	Field	15000 images of vines	Eutypa Dieback	84.00%	Yes	Edge
(Rudenko et al., 2023) [22]	RGB	Traditional	Fuzzy Classifier	Generated for the Paper	Lab – Factory	1000 images of scions	Various GTD's. Anomalous stem appearance	90.00%	Yes	Cloud
(Hnatiuc et al., 2023) [7}	N/A – Multi-class Sensor Data	Traditional	Random Forest	Generated for the Paper	Field	IOT Vineyard Sensor Data – Annotated as Plant state 'Diseased' 'Healthy'	Powdery Mildew, Botrytis Cinerea	98.00%	Yes	Cloud
(Gonçalves et al., 2022) [16]	RGB	Deep Learning	SSD ResNet50	Generated for the Paper	Field	168 smartphone images of sticky traps. Annotated with bound boxes of target insects totalling 8966 annotations.	Grapevine Moth, Green Leafhopper, "Flavescence Doree" Leafhopper, Tomato Moth, Morphotype C (Moth)	82 -99% Depending on insect class	No	Edge
(Calamita et al., 2021) [13]	Hyperspectral	Traditional	Naive Bayes	Generated for the Paper	Lab	105 Leaf Images	Armillaria	75% Asymptomatic	No	Cloud
(Hernández et al., 2021) [9]	RGB & Hyperspectral	Traditional & Deep Learning	Otsu Method & CNN	Generated for the Paper	Lab	Unknown number of images.	Downy Mildew (on leaf discs)	82% Early Detection – Hyperspectral.	No	Cloud
(Pérez-Roncal et al., 2022) [11]	Hyperspectral	Traditional	PLS-DA	Generated for the Paper	Lab	54 images of leaves.	Esca	95.26% Asymptomatic	No	Cloud
(Sawyer et al., 2023) [12]	Hyperspectral	Traditional & Deep Learning	Random Forest and CNN	Generated for the Paper	Lab	500 images of leaves	Leafroll Virus	RF: 82.8% CNN: 87%	No	Cloud
(Zendler et al., 2021) [10]	RGB	Deep Learning	SCNN	Generated for the Paper	Lab	~950 images per class	Downy Mildew (on leaf discs)	95.00%	No	Cloud
(Gutiérrez et al., 2021) [8]	RGB	Deep Learning	CNN	Generated for the Paper	Field	250 images per class	Downy Mildew & Rust Mite	94.00%	No	Cloud
(Miranda et al., 2022) [27]	RGB	Deep Learning	Variational Autoencoder	Generated for the Paper	Field	662 images	"Damaged" Berries	92.30%	No	Cloud
(Pinheiro et al., 2023) [26]	RGB	Deep Learning	YOLO (v5, v7 and R)	Generated for the Paper and made publicly available.	Field	968 original images, permuted into 10010 images for the dataset.	Berry Lesions	97% accurate for detections, though a low mAP shows many bunches not detected	No	Cloud
(Wang et al., 2023) [24]	RGB	Traditional	Random Forest	Generated for the Paper	Field	Pixel based polygons, only 5-7 per class.	Shiraz Disease	Near 100%, but only classifying pixels as a measure of canopy density.	No	Cloud

Table 1: Comparison of models, training techniques and data-sets. Green denotes real-time and/or edge based compute.

compute intensive and not-suitable for real time applications, either in the pre-processing or classification model. As such this review cannot make any specific recommendations for these approaches to on-farm practical uses, except to say that further research is needed to take these lab based approaches into the field. Hyperspectral imagery is also an interesting area of research that needs to be further studied, the suitability of current hyperspectral imaging devices to the field needs to be established.

The remainder of the studies in this review used images or sensor data gathered in the field and they can further be split up into two relevant sections. Real-time detection models and models that were either not tested in a real time capacity or the computer requirements were too high. Table 1 highlights the real-time and edge compatible models in green, we can see that of the recent studies published in the field they represent less than a third.

Starting with the non real-time models there was a ResNet CNN model trained on static images of insects [16], this model ran on modern smartphones (an edge device) however the compute complexity of the model meant that classifying one image took over a minute, well below the five frames a second needed for real time. There were also two CNN based models, one a basic CNN [8] and the other a Variational Autoencoder [27], detecting different disease states of the grapevine , both these models were not tested in a real time capacity. However the complexity of CNN models hints that they may not be suitable for computer limited environments. A YOLOv5 was also the best at detecting diseased bunches but this was only used on static images of grape bunches [26], as previously discussed YOLO models are remarkably lightweight for their accuracy and thus this model does look promising in a real-time application but it was never tested. Lastly a simple Random Forest model simply classify pixels of an aerial image of a vineyard [24], due to the low complexity of this model it would be easily run on edge compatible devices in a real time manner but the real problem with this approach is vine tracking. The model needs to keep track of canopy sizes of each individual vine, with a static aerial image it is quite easy by simply taking the same image from the same spot every time but in moving video this becomes a significant hurdle.

For real-time field based models we are left with only three studies in the last three years, interestingly they all use different types of data to train their models. Firstly a hybrid approach to bird song detection whereby a local SVM and cloud-based CNN were trained on four second snippets to identify starlings [6]. This approach was successful in that the SVM model was lightweight enough to operate on IOT devices in real-time and the hybrid approach limited network bandwidth. Secondly a simple Random Forest algorithm was trained on environmental IOT sensor data to detect periods of time conducive to fungal pathogen growth [7], interestingly this study did not detect the pathogen itself but predicted periods of fungal spread. The algorithm was lightweight enough to run continuously in real-time though it was run in the cloud with streamed data as opposed to on the IOT devices themselves, limiting its usage to farms with good connectivity. Lastly the most applicable study to real-time pathogen detection was work done by Tang *et al.* [20] to train a YOLOv5 model on video taken from commodity smartphones mounted to a trailer and towed through a vineyard. Initially trained on 15000 grapevine images the model was able to run at at least 5 frames a second for the entire run through the vineyard block while maintaining a high accuracy when detecting Eutypa Dieback severity. Representing true real-time detection and mapping this is a landmark study, however it is also limited in its applicability to other pest and diseases. This is because it exploits the ability of YOLO models to easily detect object boundaries, from the size of the cordon ‘objects’ (see Figure 1) canopy size is inferred and thus dieback severity. The model does not ‘understand’ the images, merely measures their size, it is this simplicity that lets the model run in real time on edge devices.

If we focus on computer vision the current state of the art in ML for grapevine pest detection therefore can be split into two camps. Firstly there are highly specialised models that can in some ways ‘understand’ the images that are portrayed. These models have high accuracy and can discriminate different pathogens to a high degree however they do not generalise well and need images to be a certain form before they can be used see [10], [13], [16] for examples. They are not necessarily good at object detection or bounding in a real-time situation, this is for more specialised object bounding networks such as YOLO models [21] which are the second group. These models are more computationally efficient and have been shown to have good success at detecting cordon size [20] and detecting bunches [26]. It is this differentiation that is the fundamental finding of this review, to create future models that work on compute-limited edge devices on field data we must embrace this difference. The framework proposed in this review for future research centres around utilising lightweight object bounding neural networks, such as YOLO models, to detect objects of interest in the vineyard such as leaves or bunches. This model could be trained to store static images of these objects which could be utilised in a data pipeline similar to [15] whereby the images are stabilised, corrected and standardised. With these images in hand they can then be fed into more specialised models which can detect specific pathogens, this could be done by multiple different models or similar to [27] detect any variations from a ‘healthy’ plant state. This framework is outlined descriptively in Figure 2.

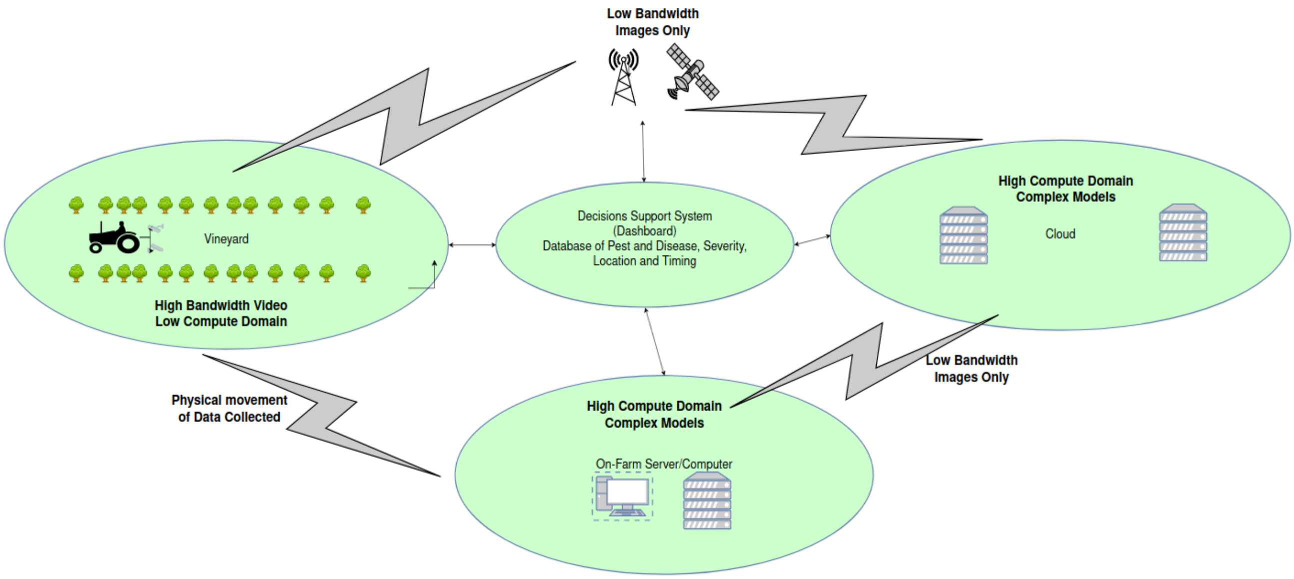

FIGURE 2. Demonstrates the two-tier framework. Lightweight models on-farm generate standard image datasets of the vineyard which are either transferred to the computing infrastructure on-farm or on the cloud. There more complex models inference on the data and provide information for a decision support system.

The proposed framework for researching computer vision in viticulture has many benefits, the biggest hurdle to the application of ML in pest detection is getting the data out to the grower in a useful way. It is only by finding practical, on-farm and easy to use solutions that provide accurate results will the technology then be adopted. By creating this two-model approach it would allow simple models to run on edge hardware, for example a smartphone mounted on a spray rig behind a tractor. This initial model could use object detection to determine the appropriate objects to the specific task or could even be a generic model that simply catalogues vines. This model could automate the collection of photographs of vines and standardise them, thus creating vast datasets of standard images that other more complex models could then classify. This hybrid model enables the use of cheap commodity hardware on the farm to easily generate standard data from the noise of a farming system.

The non-vision related ML systems employed in [6], [7] are similar in scope to the proposed visual framework above, Cinkler *et al.* [6] employed a two tier system with the smaller model efficient enough to run on IOT devices on the field. Larger, more accurate models were employed in the cloud and only run when the smaller model reached a certain confidence level. Hnatiuc *et al.* [7] did not use a two tier model system however they did use an IOT sensor network to generate a data-set of environmental conditions. This dataset could then be fed into any number of specialised models in the cloud, similar to a two tier system. The first model creates a generic data-set automatically with which further inferences can then be ascertained through specialised models.

IV. Future Work

As described state of ML and data-science in viticultural pest management is nascent and evolving, with much work to be done until it can generate actionable information that growers can use in their decision making. This work provides a framework with which more research can be conducted, there is a need to generate a real-time lightweight model that can automate the process of generating images of grapevines and their relevant structures.

Much work today has been on specialised models using data-sets generated only for that body of work. The ability for these models to generalise across different data-sets and/or real world collected images needs to be addressed. Just as in field crops [14] there is a need for the creation of published test data-sets that can compare models objectively, the siloed nature of current research makes independent comparison difficult and meaningless. This would allow for the creation and comparison of the specialised models needed to work on the data-sets created by smaller object detecting field models. This work shows the ability for both deep learning and traditional ML models to be effective in the field, both should be included in further research.

Finally more work is needed to create the smaller edge models that will run in the field, promising areas in this respect are the lightweight YOLO models that have been shown to run on modern smartphones with good results [20]. These models can detect objects in real-time video footage captured by the phones, however the technology for the central thesis of this review, the creation of standardised image data-sets from YOLO models, needs to be created and is a significant area of potential future study.

V. Conclusion

The potential of ML in viticulture for pest and disease management remains strong, with this review outlining the many accurate models classifying images, video, audio and environmental sensor data. Currently most work is siloed without standardisation of data-sets or comparison testing, it is mostly in the proof of work concept stage. Much of the work occurs in a lab setting or uses hardware that is inappropriate for on-farm field work such as GPU compute heavy neural networks. As such their remains a gap for on-farm edge compatible models that can create real value for grape growers.

This review proposes a framework for future research, a two tier system whereby a at the lower tier a simpler object detection deep learning framework is used to generate standardised image datasets of vineyard blocks. These datasets could then be either moved to the cloud or on-farm computer infrastructure upper tier where more specialised models could be used to classify specific pathogens and their severity. These models could be either deep learning or traditional data-mining as both have been shown to be effective in difference scenarios. This solves the problem of limited compute and network connectivity on farm while also creating a pathway for standardised model testing and comparison. It also simplifies future model research in that when training a new model on a novel pathogen existing data-sets and data-collection methods can be utilised in a plug and play manner.

APPENDIX

Abbreviations

- *CNN – Convolutional Neural Network*
- *RGB – Red/Green/Blue (optical wavelengths)*
- *mAP – mean Average Precision*
- *YOLO – You Only Look Once neural network model (outlined in [21])*
- *SVM – Support Vector Machine*
- *IOT – Internet of Things*

References

[1] OIV, "State of the World Vitivinicultural Sector in 2022." OIV, 2023. Accessed: Dec. 15, 2023. [Online]. Available: https://www.oiv.int/sites/default/files/documents/OIV_State_of_the_world_Vine_and_Wine_sector_in_2022_2.pdf

[2] Statista, "Topic: Global wine market," Statista. Accessed: Dec. 15, 2023. [Online]. Available: https://www.statista.com/topics/7802/global-wine/

[3] Wine Australia, "Assessment of economic cost of endemic pests and diseases on the Australian grape and wine industry." Accessed: Jan. 28, 2024. [Online]. Available: https://www.wineaustralia.com/research_and_innovation/projects/assessment-of-economic-cost-of-endemic-p

[4] M. V. Ferro and P. Catania, "Technologies and Innovative Methods for Precision Viticulture: A Comprehensive Review," *Horticulturae*, vol. 9, no. 3, Art. no. 3, Mar. 2023, doi: 10.3390/horticulturae9030399.

[5] L. Mohimont, F. Alin, M. Rondeau, N. Gaveau, and L. A. Steffenel, "Computer Vision and Deep Learning for Precision Viticulture," *Agronomy*, vol. 12, no. 10, Art. no. 10, Oct. 2022, doi: 10.3390/agronomy12102463.

[6] T. Cinkler, K. Nagy, C. Simon, R. Vida, and H. Rajab, "Two-Phase Sensor Decision: Machine-Learning for Bird Sound Recognition and Vineyard Protection," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11393–11404, Jun. 2022, doi: 10.1109/JSEN.2021.3134817.

[7] M. Hnatiuc *et al.*, "Intelligent Grapevine Disease Detection Using IoT Sensor Network," *Bioengineering*, vol. 10, no. 9, Art. no. 9, Sep. 2023, doi: 10.3390/bioengineering10091021.

[8] S. Gutiérrez, I. Hernández, S. Ceballos, I. Barrio, A. M. Díez-Navajas, and J. Tardaguila, "Deep learning for the differentiation of downy mildew and spider mite in grapevine under field conditions," *Computers and Electronics in Agriculture*, vol. 182, p. 105991, Mar. 2021, doi: 10.1016/j.compag.2021.105991.

[9] I. Hernández, S. Gutiérrez, S. Ceballos, R. Iñíguez, I. Barrio, and J. Tardaguila, "Artificial Intelligence and Novel Sensing Technologies for Assessing Downy Mildew in Grapevine," *Horticulturae*, vol. 7, no. 5, p. 103, 2021, doi: 10.3390/horticulturae7050103.

[10] D. Zendler, N. Malagol, A. Schwandner, R. Töpfer, L. Hausmann, and E. Zyprian, "High-Throughput Phenotyping of Leaf Discs Infected with Grapevine Downy Mildew Using Shallow Convolutional Neural Networks," *Agronomy*, vol. 11, no. 9, Art. no. 9, Sep. 2021, doi: 10.3390/agronomy11091768.

[11] C. Pérez-Roncal, S. Arazuri, C. Lopez-Molina, C. Jarén, L. G. Santesteban, and A. López-Maestresalas, "Exploring the potential of hyperspectral imaging to detect Esca disease complex in asymptomatic grapevine leaves," *Computers and Electronics in Agriculture*, vol. 196, p. 106863, May 2022, doi: 10.1016/j.compag.2022.106863.

[12] E. Sawyer *et al.*, "Phenotyping grapevine red blotch virus and grapevine leafroll-associated viruses before and after symptom expression through machine-learning analysis of hyperspectral images," *Frontiers in Plant Science*, vol. 14, 2023, Accessed: Nov. 23, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpls.2023.1117869

[13] F. Calamita, H. A. Imran, L. Vescovo, M. L. Mekhalfi, and N. L. Porta, “Early Identification of Root Rot Disease by Using Hyperspectral Reflectance: The Case of Pathosystem Grapevine/Armillaria,” *Remote Sensing*, vol. 13, no. 13, p. 2436, 2021, doi: 10.3390/rs13132436.
[14] T. Kasinathan, D. Singaraju, and S. R. Uyyala, “Insect classification and detection in field crops using modern machine learning techniques,” *Information Processing in Agriculture*, vol. 8, no. 3, pp. 446–457, Sep. 2021, doi: 10.1016/j.inpa.2020.09.006.
[15] P. Faria, T. Nogueira, A. Ferreira, C. Carlos, and L. Rosado, “AI-Powered Mobile Image Acquisition of Vineyard Insect Traps with Automatic Quality and Adequacy Assessment,” *Agronomy*, vol. 11, no. 4, p. 731, 2021, doi: 10.3390/agronomy11040731.
[16] J. Gonçalves *et al.*, “Edge-Compatible Deep Learning Models for Detection of Pest Outbreaks in Viticulture,” *Agronomy*, vol. 12, no. 12, Art. no. 12, Dec. 2022, doi: 10.3390/agronomy12123052.
[17] L. Rosado *et al.*, “EyesOnTraps: AI-Powered Mobile-Based Solution for Pest Monitoring in Viticulture,” *Sustainability*, vol. 14, no. 15, p. 9729, 2022, doi: 10.3390/su14159729.
[18] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
[19] J. Kenfaoui *et al.*, “A Panoramic View on Grapevine Trunk Diseases Threats: Case of Eutypa Dieback, Botryosphaeria Dieback, and Esca Disease,” *Journal of Fungi*, vol. 8, no. 6, p. 595, Jun. 2022, doi: 10.3390/jof8060595.
[20] J. Tang *et al.*, “Using a Camera System for the In-Situ Assessment of Cordon Dieback due to Grapevine Trunk Diseases,” *Australian Journal of Grape and Wine Research*, vol. 2023, p. e8634742, Oct. 2023, doi: 10.1155/2023/8634742.
[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jun. 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.
[22] M. Rudenko, Y. Plugatar, V. Korzin, A. Kazak, N. Gallini, and N. Gorbunova, “The Use of Computer Vision to Improve the Affinity of Rootstock-Graft Combinations and Identify Diseases of Grape Seedlings,” *Inventions*, vol. 8, no. 4, Art. no. 4, Aug. 2023, doi: 10.3390/inventions8040092.
[23] M. Fuchs, “Grapevine viruses: a multitude of diverse species with simple but overall poorly adopted management solutions in the vineyard,” *J Plant Pathol*, vol. 102, no. 3, pp. 643–653, Aug. 2020, doi: 10.1007/s42161-020-00579-2.
[24] Y. M. Wang, B. Ostendorf, and V. Pagay, “Evaluating the Potential of High-Resolution Visible Remote Sensing to Detect Shiraz Disease in Grapevines,” *Australian Journal of Grape and Wine Research*, vol. 2023, p. e7376153, May 2023, doi: 10.1155/2023/7376153.
[25] Z. Wang, D. Fahey, A. Lucas, A. S. Griffin, G. Chamitoff, and K. C. Wong, “Bird damage management in vineyards: Comparing efficacy of a bird psychology-incorporated unmanned aerial vehicle system with netting and visual scaring,” *Crop Protection*, vol. 137, p. 105260, Nov. 2020, doi: 10.1016/j.cropro.2020.105260.
[26] I. Pinheiro, Moreira, G., Queirós da Silva, D., Magalhães, S., Valente, A., Moura Oliveira, P., Cunha, M., and Santos, F., “Deep Learning YOLO-Based Solution for Grape Bunch Detection and Assessment of Biophysical Lesions,” *Agronomy*, vol. 13, no. 4, Art. no. 4, Apr. 2023, doi: 10.3390/agronomy13041120.
[27] M. Miranda, L. Zabawa, A. Kicherer, L. Strothmann, U. Rascher, and R. Roscher, “Detection of Anomalous Grapevine Berries Using Variational Autoencoders,” *Frontiers in Plant Science*, vol. 13, 2022, Accessed: Nov. 23, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpls.2022.729097

A Barrett-Lennard (CSU 24) Currently employed in Wine, Andy is undertaking a masters degree in information technology specialising in data science and software development. This is his final report for ITC571 – Emerging Technologies and Innovation.