Hyperspectral Imaging and Machine Learning to Identify Epicuticular Wax Loss in Masena Blueberries for Post-Harvest Freshness

Jonathan Pearse[1*], Yaminn Thawdar[1**], Alicia Sim [1], Melanie Ooi[1], Ben McGuinness[1], Peter Reutemann[2], Dale Fletcher[2], Mike Duke[1]

***Abstract*— Blueberries harvested using mechanical harvesting techniques, like over-the-row (OTR) harvesters, are not suitable for the fresh market due to high damage rates. Assisted harvesting (AH) techniques using hand-held shakers offer a new approach which shows promise in increasing harvesting rates without compromising fruit quality. Epicuticular wax is one of the fruit qualities needing to be preserved during harvesting. This paper investigates the effectiveness of hyperspectral imaging techniques and machine learning to construct a model to identify epicuticular wax rapidly and non-destructively. The best performing model produced was a linear SVM, and had an F1-score of 98.6%. Additionally, this model was used to show that blueberries harvested using hand-held shakers retain more epicuticular wax than traditional hand-harvesting (HH) techniques, a result that shows promise of potentially increased fruit quality using automated methods.**

I. INTRODUCTION

Global demand in agricultural products are continually increasing, and the shift to automation is a step towards meeting these demands [1]. Blueberry harvesting is also shifting from hand-harvesting (HH) towards automation; however, their desired properties are highly susceptible to damage, both in pre- and post-harvest states. Berries harvested using mechanical harvesters are generally lower quality as significant damage occurs during harvesting processes, making them unsuitable for the fresh market [2], [3]. Research has been conducted into both assisted harvesting (AH) and automated harvesting on ways to minimise damage including soft-catch surfaces (SCS) [4], and hand-held shakers [5], [6], [7]. Other properties that contribute to the market price of blueberries include hydration, firmness, size, colour, soluble solids content (SSC) and epicuticular wax (EW), also commonly referred to as bloom [8]. EW is the epidermal wax layer present on many varieties of fruit, appearing as the white, powdery substance on the skin of blueberries. Retaining EW is not only important as it is visually appealing to consumers [9], but also as it contributes significantly to the berry's quality both in pre- and post-harvest states. In the pre-harvest state, it reduces susceptibility to pathogen organisms and reduces the chance of deformations. In the post-harvest state, EW increases moisture retention from non-stomatal areas, prevents microbial infections, and helps resist physical damage caused by external stresses [10], [11], [12], [13]. One of the issues with HH methods is that pickers need to handle the fruit. AH requires less handling of berries, potentially increasing EW present on post-harvest berries. Masena blueberries, a variety of high-bush blueberry (*Vaccinium corymbosum*) selectively bred by Mountain Blue Farms Pty Ltd, have a characteristically thick EW layer and thin skin. This makes Masena an ideal variety for imaging EW. While numerous studies have investigated impact damage on blueberries occurring through automated or AH techniques, accurately assessing EW retention has been impeded by the laborious nature of the process and its subjectivity. To quantify EW, traditional methods include visual inspection, or use of RGB cameras [14], [15]. RGB cameras are well-suited for determining most properties associated with the fruit's grade, however are limited for EW detection as blueberry colour or bruising may affect perception on EW presence. NIR hyperspectral imaging (HSI) captures responses in the 900nm to 1400nm range, which is outside the visible light range. This can provide much higher levels of accuracy by observing greater spectral differences between areas of no EW and areas with EW. HSI is a form of non-destructive analysis that involves capturing multiple images for different spectral groups. Captured images (2-dimensional) are arranged into spatial maps called hypercubes, using the wavelength as a 3rd dimension. Applications of HSI in agriculture have been revolutionary in improving crop quality [16], [17], including blueberries [18], [19], [20], [21], and has only been gaining more ground. This can be seen by the rapid increase of research conducted, with 245 articles published in 2011-2020 in comparison to 97 published in 2001-2010 [22]. As HSI is an expensive and complex imaging process, it is impractical to use it in commercial automated graders. However, it can be used to identify wavelengths characteristic with EW on blueberries, and hence may be the first step towards developing a specialised system utilising wavelength specific imaging for automated grading.

This paper aims to improve on existing methods of imaging EW through use of HSI and machine learning techniques, and to identify if there is any significant observable difference in EW retention between AH and HH techniques.

II. MATERIALS AND METHODS

A. Data Collection

Figure 1 provides an overview of the methodologies employed to produce a machine learning model. The experiments presented in this paper were conducted on Masena

[1]Authors are with the School of Engineering, University of Waikato, Hamilton, New Zealand

[2]Authors are with the School of Computing and Mathematical Sciences, University of Waikato, Hamilton, New Zealand

Corresponding authors: jonathan.pearse@waikato.ac.nz*, yt244@students.waikato.ac.nz**

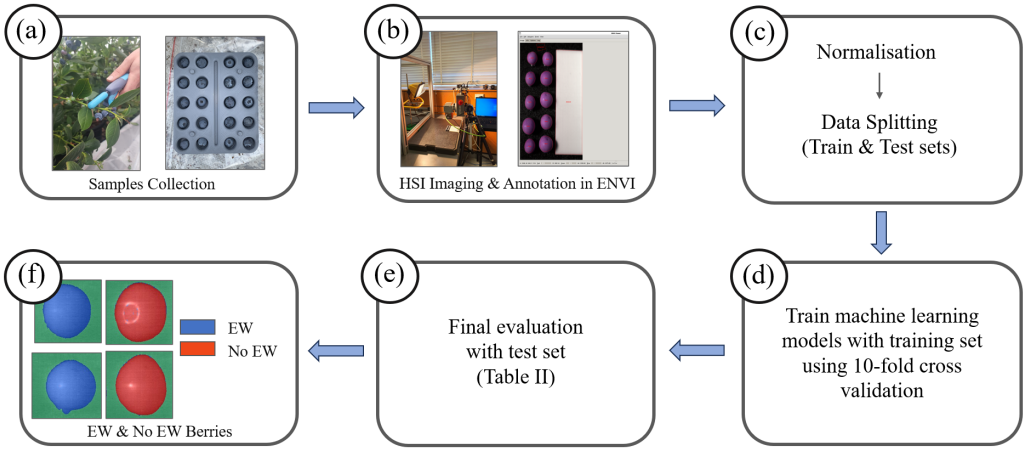

Fig. 1. (a-f) A flowchart representing the procedures conducted for this paper.

blueberries harvested from an orchard located in Pukehina, New Zealand. This orchard had their bushes under tunnels to prevent damage from rain and dew, both of which can remove EW and reduce fruit quality. Data was collected mid-season on November 24, 2023. Blueberries harvested using AH methods used an improved hand-held shaker developed by the University of Waikato [23]. This shaker uses a brushless direct current (BLDC) motor to turn a crank, causing the head of the device to reciprocate, and hence transfer force to a cluster of blueberries.

Using polyvinyl chloride (PVC) gloves, 20 ripe blueberries were carefully harvested by hand, and placed into 3D printed polylactic acid (PLA) carry trays (Figure 4, (a)). These were classified as the 'control' with near perfect EW intact. 10 ripe blueberries were then harvested using traditional HH techniques, without gloves, into a plastic bucket. Using PVC gloves, 10 of these berries were carefully removed from the bucket, ensuring minimal disturbance to remaining EW, and placed into a separate PLA carry tray. Finally, 10 ripe blueberries were harvested into a similar bucket using a hand-held shaker (Figure 2, (a)). These were placed in a separate PLA tray in the same way as the HH berries. Following the harvesting, these berries were placed inside an air-conditioned vehicle and transported within 3 hours to the University of Waikato where they were refrigerated. These blueberries were then imaged using the hyperspectral camera within 9 hours after harvesting.

B. Hyperspectral Imaging Methods Used

HSI involves capturing an image in many spectral wavelengths. A typical camera captures an image in 3 spectral ranges (red, green and blue), while an HSI camera captures it in the hundreds of spectral wavelengths. Electromagnetic radiation, typically given by an illuminant, either from natural or artificial sources, is reflected off the surface of the sample being imaged. Depending on the chemical and physical makeup of the sample, this will be reflected differently [24]. Thus, spectral reflectance can be observed using HSI, typically in the visible and near-infrared (NIR) wavelengths,

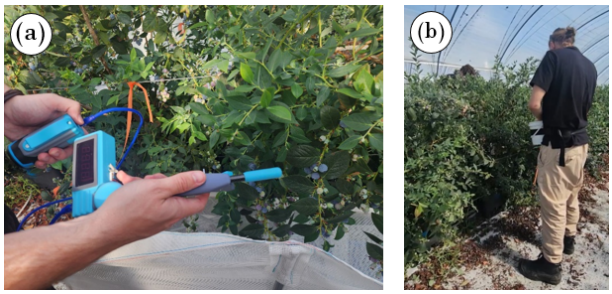

Fig. 2. (a) Hand-held shaker developed by the University of Waikato shaking a cluster of Masena blueberries; (b) Traditional hand-harvesting method where workers place blueberries into a waist-strapped bucket.

between 400nm and 2500nm. The imaging conducted in this paper was between 900nm 1700nm. For reference, the visible light spectrum is between 380nm and 700nm, hence all HSI conducted was in the NIR region.

The configuration of a hyperspectral camera consists of several fundamental elements: a lens, an image-capturing spectrograph and an area-covering image sensor. For this paper, a Specim FX17e camera captured NIR spectral bands ranging from 900nm to 1700nm, divided into 224 spectral bands with the resolution of 640x640 pixels. This camera utilizes a line scanning configuration referred to as push-broom technique. This means the camera scans the spectral data of the target one line at a time [25]. This linear scanning camera is also facilitated with a rotary stage. Using the Specim Lumo Scanner software, the speed and angle of scanning are effectively managed.

1) Physical setup: In the physical setup for the camera, the HSI camera is fixed to a tripod stand, which is positioned on an even surface. It was crucial to ensure that the tripod's legs remained stationary once the optimal setup had been achieved. The link between the computer and the camera was established by connecting an Ethernet cable to the computer and attaching the other end to the camera using a standard GigE M12 X-coded Ethernet connector. The camera was rotated at a speed of 5.32 degrees per second while capturing

TABLE I
DATA SPLITTING

1. Total Fruits Imaged	**39 images**
a. Fruit with EW	30 images
b. Fruit with no EW	9 images
2. Total Pixels of Fruit with EW	**316,474**
a. In Training set with 10-fold cross-validation	182,513
c. In Test set	133,961
d. Background (excluded from training data)	15,378
e. White Reference (excluded from training data)	992,785
3. Total Pixels of Fruit with no EW	**148,547**
a. In Training set with 10-fold cross-validation	99,200
c. In Test set	49,347
d. Background (excluded from training data)	10,112
e. White Reference (excluded from training data)	488,466

the images of blueberries. From the control berries with fully intact EW, 9 berries were carefully wiped using cotton cloths to remove the wax before imaging. This allowed for a definitive comparison between 'perfect EW' and 'no EW'. In total, there were 5 sets of spectral hypercubes collected by the HSI camera. These were 'No EW', 'Perfect EW', 'No EW vs Perfect EW', 'Assisted Harvested Blueberries (AHB)', and 'Hand Harvested Blueberries (HHB)'. Table I shows the composition of captured images.

2) Illumination: In HSI imaging, the illumination system is the primary source of light that interacts with the object under observation. This interaction is vital for achieving optimal image quality and precise spectral data [26]. For the imaging conducted in this paper, the Philips Plusline Halogen 500W R7s 240V was utilized. It was positioned facing the subject to ensure consistent and evenly distributed illumination. The halogen lamp was placed at approximately 320 mm from the HSI camera.

C. Image Processing and Machine Learning

1) Normalization of Data: In the analysis of raw spectral data, normalization is required [27]. When using dark and white references in normalization, the spectral data is more reliable and dependable as noise is removed. It also helps with uniformity and consistency across multiple imaging sessions and different environmental conditions. Background noise and interference can be eliminated by subtracting the dark reference, and the adjustment for illumination variations is accomplished by dividing a white reference.

2) Software Utilised: Lumo Scanner is a software for recording data from Specim's HSI camera and scanner systems. It provides a user interface for real-time monitoring and adjustment of camera settings during imaging processes [28]. It identifies over-exposed areas for correction, ensuring the best setup. An in-house developed software called 'ENVI Viewer' [29] was used to annotate the images, as seen in Figure 3. The samples were categorised into four groups: 'EW', 'No EW', 'Background', and 'White reference'. The images were then saved in PNG format, and the annotations were exported as JSON files using the Object-Predictions-Exchange (OPEX) Format [30]. All the annotated data were converted into .mat files to process in MATLAB.

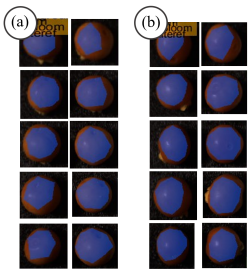

Fig. 3. (a) AH Berry Annotations in ENVI Software; (b) HH Berry Annotations in ENVI Software.

3) Data Splitting: In this experiment, 39 images of blueberries were collected, whereby 30 were of blueberries with EW and 9 without EW. This data was then divided into training and test sets, 70% and 30% each, respectively. The validation set was obtained from the training using 10-fold cross-validation. The distribution of pixels in the training set, testing set, and validation set of blueberries with EW and without EW are shown in Table I.

4) Machine Learning Approaches Used: Deep learning models can be applied in this paper; however, classical machine learning models provided acceptable outcomes. Data availability is another reason not choosing deep learning models [33] as deep learning requires a large number of datasets which was not available for this project. Compared to classical machine learning, deep learning requires a huge amount of hyperparameters to tune which was beyond the scope in this particular study. In this study, supervised learning models were used to train on datasets which have been labelled. There are several algorithms in supervised learning; however only three models such as decision trees, discriminant analysis and Support Vector Machine (SVM) were chosen. The types of algorithms and their characteristics are described in Table II.

5) Validation Method: There are three common methods to evaluate the performance of produced models in machine learning such as holdout method, k-fold cross-validation and leave-one-out cross-validation [34]. However, k-fold cross-validation was selected to assess the accuracy of the produced models. In this method, 'k' numbers of subsets (folds) are used to split the datasets. Typical values of k are 5, 10 and 20 [34], and are dependent on the size of the datasets used for training. For this sample, as the dataset was small, a k value of 10 was used. The training dataset is split into 10 folds and the model is trained 10 times. For each iteration, 9 of them are used for the training data and the remaining 1-fold is for the validation set. This action is repeatedly processed for each of the 10 folds.

6) Evaluation Metrics: In this paper, the 7 machine learning models described in Table II were trained using the training set and applied to the test set for predicting

TABLE II
MACHINE LEARNING MODELS

Algorithm	Characteristics
1. Decision Tree	Decision tree is a graphical representation in a tree structure which consists of nodes and branches [29]. It is easily interpretable, and is efficient and fast in training and making predictions although it has low predictive accuracy
1.1 Fine Tree	A fine tree can perform well to capture complex patterns and it has a high accuracy on training data. However, it has a high chance of overfitting [30].
1.2 Medium Tree	A medium tree has fewer split nodes than fine tree; however, it has more splits than a coarse tree.
1.3 Coarse Tree	A coarse tree with limited split nodes leads to faster training time than a fine tree [30]. Although its flexibility and training accuracy are likely to be lower compared to fine tree, it has less risk of overfitting. This means it is able to provide more reliable predictions on unseen data.
2. Support Vector Machine (SVM)	SVM is one of the most widely used supervised learning models for analysing data which is used for classification and regression analysis [29]. SVM is also popular for performing linear and nonlinear classifications.
2.1 Linear SVM	Linear SVM has a straight-line linear boundary and it is applied to the data which is linearly separable [31].
2.2 Cubic SVM	Cubic SVM has non-linear boundaries and can capture more complex pattern in the data [30]. In contrast, it is likely to have the risk of overfitting when the model is too complex.
3. Discriminant Analysis	This is one of the supervised machine learning algorithms and is used for classification of multiple classes and dimensionality in data analysis. It is popular for it precision, fast performance and easy interpretability [32].
3.1 Linear Discriminant Analysis (LDA)	Linear Discriminant Analysis creates linear separation between classes
3.2 Quadratic Discriminant Analysis (QDA)	Quadratic Discriminant Analysis creates nonlinear quadratic separation between classes

the EW on Masena blueberries. To evaluate these models' performance, precision, recall and F1-score play a crucial role.

Precision is defined as the ratio of the instances of true positive predictions to the sum of true and false positives predictions. It detects the accuracy of positive prediction of the model [35].

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

Recall can be defined as the ratio of positive predictions to all positive predictions in the datasets [35].

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

F1-score is a statistical calculation which is the combination of precision and recall into a single metric [35].

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

III. RESULTS AND DISCUSSION

A. Trained Model Results

After training various models, the results, as seen in Table III, show that the linear SVM model is the best model due to high precision (99% in EW and 94% in No EW), recall (97.8% in EW and 99.8% in No EW) and F1-score (98.6%) in all classes. Some models, such as Cubic SVM and Medium Tree, appear promising for some classes but are inconsistent in comparison to the linear SVM model.

There are some limitations in the trained model. The main limitation of SVM models is the kernel selection [36]. This means that finding the most appropriate kernel of SVM is challenging for a specific situation. Additionally, computational limitation, which in this case results in increased training time, is another drawback of linear SVM [37].

Despite its limitations, the trained model had certain strengths. Linear SVM models have better generalization abilities [38] than other models in machine learning. This means it can perform more accurate predictions and classification on testing data and not on part of the training data. Therefore, it is an effective model for classification of agricultural data [39]. Moreover, it predicts the outcomes well and shows high precision and recall across all classes compared to other types of machine learning models.

Decision trees have different categories based on some criteria. The benefits of decision trees are interpretability and feature selection. In terms of interpretability, they are easy to classify and interpret due to the tree structure which make predictions by following the branches and nodes. When looking at feature selection, they always select the most important feature to separate the data into different categories based on some criteria [40]. Overfitting is one of the limitations of decision trees. Which means that they can perform high accuracy on training sets; however, it leads to poor performance on testing data.

The advantage of discriminant analysis is dimension reduction technique [41]. It means that the discriminant analysis can change the original data into simpler feature space by determining a linear combination. By doing this, it results in easy classification in new data based on the new dimensions. However, it can have difficulty when the number of sample size is smaller than the number of features [41]. It causes unreliable performance of models, overfitting and limited generalization which means the performance of the model is inaccurate on testing data although it can perform well on training data.

B. Application of HSI -SVM on Epicuticular Wax Retention

In this section, the best performing model (Linear SVM) for the training set of hyperspectral images is referred to as HSI-SVM. This model is applied on test images AH and HH berries shown in Figure 4 and Figure 5 respectively.

Figure 4 (a) indicates that there is no EW lost from the harvesting process, which seems inaccurate as Figure 4 (b) shows several scrapes and areas that are darker, indicating EW loss. Figure 5 (a) indicates that there are areas with no

TABLE III

PRECISION, RECALL AND F1 SCORE FOR VARIOUS TRAINED MODELS ON TEST SET

		Decision Trees			Support Vector Machine (SVM)		Discriminant Analysis	
		Coarse Decision Tree	Fine Decision Tree	Medium Decision Tree	Cubic SVM	Linear SVM	Linear Discriminant Analysis (LDA)	Quadratic Discriminant Analysis (QDA)
Background	Precision	0.999	0.999	0.999	0.994	0.999	0.988	1.000
	Recall	0.910	0.997	0.997	1.000	1.000	1.000	0.998
EW	Precision	0.879	0.926	0.901	0.964	0.999	0.999	0.741
	Recall	0.877	0.956	0.959	0.807	0.978	0.939	0.902
No EW	Precision	0.659	0.869	0.866	0.636	0.945	0.858	0.351
	Recall	0.672	0.792	0.715	0.916	0.998	0.994	0.144
F1-score		0.832	0.923	0.904	0.875	0.986	0.961	0.672

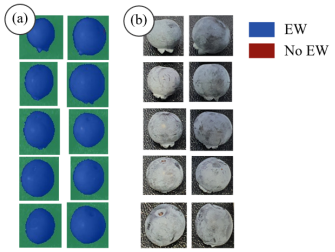

Fig. 4. (a) HSI-SVM results for AH berries; (b) RGB images for AH berries.

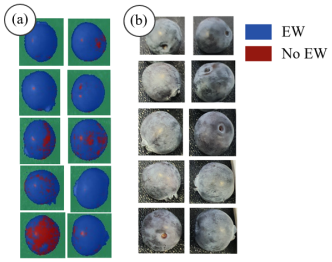

Fig. 5. (a) HSI-SVM results for HH berries; (b) RGB images for HH berries.

EW, and this is confirmed with the RGB image showing where EW seems to have been wiped off during the harvesting process. As this produced model assumes a single fault, i.e. loss of EW, the linear SVM is a good fit. If more faults were to be investigated, including but not limited to bruising, mould or ripeness, models will need to be retrained. One of the advantages of using the HSI-SVM model for determining EW loss is that it reduces the misclassifications compared to using traditional RGB imaging methods. As can be seen for the upper 3 berries in the second column of both Figure 4 (b) and Figure 5 (b), these berries appear darker, as if they have significantly less EW than the rest of the berries, and would likely be rejected. However this darker appearance is only due to inconsistent lighting. This inconsistent lighting does not affect the HSI-SVM results. The generated model can also be used to calculate a percentage of EW lost, by comparing the number of pixels of No EW and EW.

C. Limitations and Recommendations

Although the results have indicated that there is a difference in EW retention between HH and AH techniques, these results have been derived from a limited data set. There is also significant variability in the HH results seen in Figure 5 (a), which may be due to variation in EW quality of the berries prior to harvesting, rather than the harvesting method itself. It is recommended that further testing be done with significantly larger data sets to confirm the trend associated with non HH techniques. The HSI imaging has also assumed absence of bruising and fungal infections, along with uniform firmness, all of which can influence the NIR reflectance. If more complexities like these were to be addressed, deep learning methodologies should be considered for future models. Additionally, as it has been found that EW can be identified in the NIR region, further identification of these wavelength signatures could be used for developing cheaper imaging systems which capture reflectance only at required wavelengths. Automated techniques are the next step after AH, which can be further driven by the ability to retain more EW, a property that has clearly been demonstrated to drastically increase fruit quality.

IV. CONCLUSIONS

This paper presented a new method of quantifying EW loss on Masena blueberries using HSI techniques with machine learning models. 7 machine learning models were evaluated on HSI images captured for berries harvested using either HH or AH techniques. From the trained models produced, Linear SVM performed the best, and attained an F1-score of 98.6%. Using the Linear SVM model, it was found that HH techniques removed more EW than AH techniques. This indicates that automated techniques using end effectors similar to the hand-held shakers could increase post-harvest quality in blueberries by preserving more EW.

ACKNOWLEDGMENT

This work was supported by MBIE contract UOAX1810. The authors of this paper would like to extend their gratitude to Janelle Taft of Nanric Blue for providing access to her orchard, and sharing invaluable insight into blueberry harvesting processes.

REFERENCES

[1] R. K. Gallardo and D. Zilberman, "The Economic Feasibility of Adopting Mechanical Harvesters by the Highbush Blueberry Industry," HortTechnology, vol. 26, no. 3, pp. 299–308, Jun. 2016.

[2] B. Casamali, J. G. Williamson, A. P. Kovaleski, S. A. Sargent, and R. L. Darnell, "Mechanical Harvesting and Postharvest Storage of Two Southern Highbush Blueberry Cultivars Grafted onto Vaccinium arboreum Rootstocks," Hortscience, vol. 51, no. 12, pp. 1503–1510, Dec. 2016

[3] Aitazaz Ahsan Farooque et al., "Effect of ground speed and header revolutions on the picking efficiency of wild Blueberry Harvester," 2013 Kansas City, Missouri, July 21 - July 24, 2013, Jan. 2013.

[4] S. A. Sargent, F. Takeda, J. G. Williamson, and A. D. Berry, "Harvest of Southern Highbush Blueberry with a Modified, Over-The-Row Mechanical Harvester: Use of Handheld Shakers and Soft Catch Surfaces," Agriculture, vol. 10, no. 1, pp. 4–4, Dec. 2019.

[5] F. Takeda et al., "Applying New Technologies to Transform Blueberry Harvesting," Agronomy, vol. 7, no. 2, p. 33, May 2017.

[6] A. Malladi, Tripti Vashisth, and S. NeSmith, "Development and Evaluation of a Portable, Handheld Mechanical Shaker to Study Fruit Detachment in Blueberry," Hortscience, vol. 48, no. 3, pp. 394–397, Mar. 2013.

[7] E. Kim, A. Freivalds, F. Takeda, and C. Li, "Ergonomic Evaluation of Current Advancements in Blueberry Harvesting," Agronomy, vol. 8, no. 11, p. 266, Nov. 2018.

[8] R. K. Gallardo et al., "Breeding Trait Priorities of the Blueberry Industry in the United States and Canada," HortScience, vol. 53, no. 7, pp. 1021–1028, Jul. 2018.

[9] R. Saftner, J. Polashock, M. Ehlenfeldt, and B. Vinyard, "Instrumental and sensory quality characteristics of blueberry fruit from twelve cultivars," Postharvest Biology and Technology, vol. 49, no. 1, pp. 19–26, Jul. 2008.

[10] I. Lara, B. Belge, and L. F. Goulao, "The fruit cuticle as a modulator of postharvest quality," Postharvest Biology and Technology, vol. 87, pp. 103–112, Jan. 2014.

[11] Patiwit Loypimai, Sudpiti Paewboonsom, L. Damerow, and M. M. Blanke, "The wax bloom on blueberry: Application of luster sensor technology to assess glossiness and the effect of polishing as a fruit quality parameter," Journal of applied botany and food quality, vol. 90, pp. 154–158, May 2017.

[12] W. Chu, H. Gao, H. Chen, X. Fang, and Y. Zheng, "Effects of cuticular wax on the postharvest quality of blueberry fruit," Food Chemistry, vol. 239, pp. 68–74, Jan. 2018.

[13] M. C. N. Nunes, J.-P. Emond, and J. K. Brecht, "Quality Curves for Highbush Blueberries as a Function of the Storage Temperature," Small Fruits Review, vol. 3, no. 3–4, pp. 423–440, Jul. 2004.

[14] S. Matiacevich, P. Silva, J. Enrione, and F. Osorio, "Quality assessment of blueberries by computer vision," Procedia Food Science, vol. 1, pp. 421–425, 2011.

[15] G. A. Leiva-Valenzuela and J. M. Aguilera, "Automatic detection of orientation and diseases in blueberries using image analysis to improve their postharvest storage quality," Food Control, vol. 33, no. 1, pp. 166–173, Sep. 2013.

[16] S. Pascucci, S. Pignatti, R. Casa, R. Darvishzadeh, and W. Huang, "Special Issue 'Hyperspectral Remote Sensing of Agriculture and Vegetation,'" Remote Sensing, vol. 12, no. 21, p. 3665, Nov. 2020.

[17] M. Davies, M. B. Stuart, M. J. Hobbs, Andrew, and J. R. Willmott, "Image Correction and In Situ Spectral Calibration for Low-Cost, Smartphone Hyperspectral Imaging," Remote Sensing, vol. 14, no. 5, pp. 1152–1152, Feb. 2022.

[18] Y. Jiang, C. Li, and F. Takeda, "Nondestructive Detection and Quantification of Blueberry Bruising using Near-infrared (NIR) Hyperspectral Reflectance Imaging," Scientific Reports, vol. 6, no. 1, Oct. 2016

[19] Z. Zheng, Z. An, X. Liu, J. Chen, and Y. Wang, "Finite Element Analysis and Near-Infrared Hyperspectral Reflectance Imaging for the Determination of Blueberry Bruise Grading," Foods, vol. 11, no. 13, pp. 1899–1899, Jun. 2022.

[20] J.-Y. Choi et al., "Hyperspectral imaging technigue for monitoring moisture content of blueberry during the drying process," Korean Journal of Food Preservation, vol. 28, no. 4, pp. 445–455, Jul. 2021.

[21] Y. Huang, D. Wang, Y. Liu, H. Zhou, and Y. Sun, "Measurement of Early Disease Blueberries Based on Vis/NIR Hyperspectral Imaging System," Sensors, vol. 20, no. 20, p. 5783, Oct. 2020.

[22] B. Lu, P. Dao, J. Liu, Y. He, and J. Shang, "Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture," Remote Sensing, vol. 12, no. 16, p. 2659, Aug. 2020.

[23] A. Sim, B. McGuinness, S. H. Lim, H. Williams, and M. Duke, "Redefining Selective Blueberry Harvesting - Localised Shaking with Minimal Damage," ACRA, 2023.

[24] H. El-Mesery, H. Mao, and A. Abomohra, "Applications of Non-destructive Technologies for Agricultural and Food Products Quality Inspection," Sensors, vol. 19, no. 4, p. 846, Feb. 2019.

[25] Specim FX17 – User Manual 2.2, SPECIM Spectral Imaging Ltd., Finland, 2020.

[26] J. Katrašnik, F. Pernuš, and B. Likar, "A method for characterizing illumination systems for hyperspectral imaging," Opt. Express, vol. 21, no. 4, p. 4841, 2013.

[27] F. Cao, Z. Yang, J. Ren, M. Jiang, and W. Ling, "Does normalization methods play a role for Hyperspectral image classification?," ArXiv, vol. abs/1710.02939, 2017.

[28] Lumo Scanner – User Guide 1.3, SPECIM Spectral Imaging Ltd., Finland, 2015.

[29] "Hyperspectral Imaging Group," hsi.eng.waikato.ac.nz. https://hsi.eng.waikato.ac.nz/ (accessed Feb. 21, 2024).

[30] "WaikatoLink2020/objdet-predictions-exchange-format," GitHub, Feb. 17, 2023. https://github.com/WaikatoLink2020/objdet-predictions-exchange-format (accessed Feb. 27, 2024).

[31] B. Mahesh, "Machine Learning Algorithms – A Review," ResearchGate, vol. 9, no. 1, pp. 381, Jan, 2020.

[32] "Choose classifier options - MATLAB & Simulink - MathWorks Australia," Mathworks.com. [Online]. Available: https://au.mathworks.com/help/stats/choose-a-classifier.html. [Accessed: 21-Feb-2024].

[33] A. Ozdemir and K. Polat, "Deep Learning Applications for Hyperspectral Imaging: A Systematic Review," Journal of the Institute of Electronics and Computer, vol. 2, no. 1, pp. 39–56, 2020

[34] Z. Xiong, Y. Cui, Z. Liu, Y. Zhao, M. Hu, and J. Hu, "Evaluating explorative prediction power of machine learning algorithms for materials discovery usingk-fold forward cross-validation," Comput. Mater. Sci., vol. 171, no. 109203, p. 109203, 2020.

[35] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," Lecture Notes in Computer Science, pp. 345–359, 2005.

[36] H. Bhavsar and M. Panchal, "A Review on Support Vector Machine for Data Classification," International Journal of Advanced R esearch in Computer Engineering & Technology (IJARCET), vol. 1, no. 10, pp. 2278–1323, 2012.

[37] I. W. Tsang, J. T. Kwok, and P.-M. Cheung, "Core Vector Machines: Fast SVM Training on Very Large Data Sets," Journal of Machine Learning Research, vol. 6, no. 13, pp. 363–392, Dec. 2005.

[38] V. Vapnik, The Nature of Statistical Learning Theory. Springer Science & Business Media, 2013. Accessed: Mar. 13, 2024.

[39] L. Shi, Q. Duan, X. Ma, and M. Weng, "The Research of Support Vector Machine in Agricultural Data Classification," Computer and Computing Technologies in Agriculture V, pp. 265–269, 2012.

[40] H.H. Patel and P. Prajapati, "Study and Analysis of Decision Tree Based Classification Algorithm", IJCSE. vol. 6, no. 10, pp. 74-78, Oct, 2018.

[41] C. H. Park and H. Park, "A comparison of generalized linear discriminant analysis algorithms," Pattern Recognition, vol. 41, no. 3, pp. 1083–1097, Mar. 2008.