horticulturae

MDPI

Article

The Potential for Hyperspectral Imaging and Machine Learning to Classify Internal Quality Defects in Macadamia Nuts

Michael B. Farrar [1,*], Marcela Martinez [1], Kim Jones [2], Negar Omidvar [1], Helen M. Wallace [3], Thomas Chen [4] and Shahla Hosseini Bai [1]

[1] Centre for Planetary Health and Food Security, School of Environment and Science, Griffith University, Nathan Campus, Nathan, QLD 4111, Australia; marce.martinez12@gmail.com (M.M.); n.omidvar@griffith.edu.au (N.O.); s.hosseini-bai@griffith.edu.au (S.H.B.)
[2] Cropwatch Independent Laboratories, The Centre for Horticulture Research, 494 Bruxner Highway, Alstonville, NSW 2477, Australia; kim.jones@westnet.com.au
[3] School of Biology and Environmental Science, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia; helen.wallace@qut.edu.au
[4] Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia; thomastienhu@student.unimelb.edu.au
* Correspondence: m.farrar@griffith.edu.au

Abstract: Tree nuts are rich in nutrients, and global production and consumption have doubled during the last decade. However, nuts have a range of quality defects that must be detected and removed during post-harvest processing. Tree nuts can develop hidden internal discoloration, and current sorting methods are prone to subjectivity and human error. Therefore, non-destructive, real-time methods to evaluate internal nut quality are needed. This study explored the potential for VNIR (400–1000 nm) hyperspectral imaging to classify brown center disorder in macadamias. This study compared the accuracy of classifiers developed using images of kernels imaged in face-up and face-down orientations. Classification accuracy was excellent using face-up (>97.9%) and face-down (>94%) images using ensemble and linear discriminate models before and after wavelength selection. Combining images to form a pooled dataset also provided high accuracy (>90%) using artificial neural network and support vector machine models. Overall, HSI has great potential for commercial application in nut processing to detect internal brown centers using images of the outside kernel surface in the VNIR range. This technology will allow rapid and non-destructive evaluation of intact nut products that can then be marketed as a high-quality, defect-free product, compared with traditional methods that rely heavily on representative sub-sampling.

Keywords: brown center disorder; internal discoloration; macadamia; machine learning; post-harvest processing; food quality inspection; tree nuts

check for updates

Citation: Farrar, M.B.; Martinez, M.; Jones, K.; Omidvar, N.; Wallace, H.M.; Chen, T.; Hosseini Bai, S. The Potential for Hyperspectral Imaging and Machine Learning to Classify Internal Quality Defects in Macadamia Nuts. *Horticulturae* **2024**, *10*, 1129. https://doi.org/10.3390/horticulturae10111129

Academic Editor: Adriana Guerreiro

Received: 24 September 2024
Revised: 18 October 2024
Accepted: 21 October 2024
Published: 23 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Tree nuts are a healthy and convenient food for human consumption because they are high in monounsaturated fats, vitamins and dietary fiber [1]. Consumption of tree nuts has increased over the last decade and surpassed 5 million metric tons globally in 2021 [2]. Regular consumption of tree nuts can increase nutrient adequacy in adults and help to reduce the risk of heart disease and associated cardiovascular risk factors [1,3,4]. However, tree nuts are susceptible to a range of quality issues that can result in consumer dissatisfaction and economic loss [5,6]. Internal defects include both damage and discoloration and are an important quality issue because they affect a range of nut crops and can develop post-harvest following exposure to impact, heat and humidity [7–9]. Nuts with an internal defect are not easily distinguished from good ones and decrease consumer satisfaction because of a foul and bitter flavor [6,10]. Recently, stricter food safety requirements and higher quality standards are driving the demand for novel processing technologies that can

detect quality defects and improve overall product quality [11]. Therefore, rapid and non-destructive methods are now sought to detect defects for individual nuts and in real-time during post-harvest processing.

Hyperspectral imaging (HSI) has emerged during the last decade as a non-destructive method to investigate the chemical properties of an extremely wide range of objects and materials [12–15]. Hyperspectral imaging collects both spectral and spatial information from target objects, in this case, macadamia nuts [16]. The data can then be used in conjunction with machine learning techniques to develop models to predict chemical properties (chemometrics) or make decisions (classification) regarding the quality class of the nuts [16]. To date, HSI methods have been developed for a wide range of uses, such as food quality applications, because of the ability to capture both spatial and spectral variation from heterogeneous samples found within food, agriculture and soil sciences [17–20]. For each application, a specific method must be developed using particular HSI hardware and machine learning techniques, using a diverse dataset that captures all available variation within the target population [12,21,22]. Most recently, HSI has been used to predict the internal quality of fruit by detecting chemical changes on the skin surface [23,24]. Successful methods have been developed for a range of fruits and quality defects [20,23,25–29]. Concealed damage in almonds has been investigated [8], but a method to detect internal disorders associated with macadamia nuts has not yet been developed.

There is a growing library of machine learning techniques for use in conjunction with HSI to develop models to perform different tasks [20]. Classification is an analytical method that aims to coerce samples into sub-groups based on similarities (or differences) using hyperspectral data [16]. Classification can be conducted in a supervised or unsupervised manner [16]. Supervised classification models create a classification rule based on training with data having known reference classes, and the models can then be used to classify new samples with unknown classes [16]. However, to calibrate efficient classification models, the optimal technique and parameters must be determined first for each specific application [30]. Machine learning models are commonly applied to complex problems using HSI, and each model family has differing underlying assumptions and mathematical theorems [20]. For example, linear discriminant (LD) and support vector machine (SVM) models are suitable for linear class data by reducing dimensionality and identifying separable linear planes between classes, whereas artificial neural network (ANN) and k-nearest neighbor (KNN) models can handle non-linear classes efficiently by identifying clusters or utilizing multiple perceptron layers, respectively [20,30–32]. Therefore, it is imperative that exploratory studies investigate various machine learning techniques to identify the most effective and efficient technique for each specific classification task.

Brown center disorder is a quality defect in macadamia distinguished by internal brown discoloration and an associated foul odor and taste [10]. Various mechanisms can cause brown centers, including the enzymatic browning reaction, the Maillard reaction and/or infection by microorganisms [33,34]. Brown center disorder was first identified within roasted kernels in 1965 [35]; however, brown center disorder is also prevalent in raw kernels and can be induced via incorrect storage and handling during post-harvest processing [7,34,36]. The occurrence of brown center disorder has increased with export volumes, and the issue has cost the world's second-largest exporter, the Australian industry, more than USD 30 million over ten years alone [37–39]. Therefore, novel methods to detect brown center disorder within individual kernels during post-harvest processing will help to ensure adherence with food quality standards and ongoing customer satisfaction.

Currently, modern nut processing facilities use color sorting machines that utilize information from the visible spectral region to identify and sort defects in nuts [40]. However, brown centers are not always visible on the surface of whole kernels and are only detectable after drying [10,41]. Therefore, a skilled and qualified person is required to manually identify brown centers at the final stage, and non-detections remain a problem due to human error and subjectivity. This study explores the potential for laboratory-based HSI and machine learning to classify brown center disorder in macadamia nuts. The aim

was to develop a HSI method for assessment of internal quality that would enable more efficient post-harvest sorting and adoption within modern processing operations. Specifically, this study aimed to (1) identify important spectral regions and appropriate machine learning techniques for brown center disorder classification using HSI, and (2) determine the accuracy of machine learning classifiers using HSI of the inner kernel flesh (face-up images), outer kernel surface (face-down images) or both image orientations combined to form a pooled dataset. These findings will help to maintain food safety requirements while increasing the efficiency of post-harvest sorting by reducing incorrect rejections, leading to higher profitability and consumer confidence.

2. Materials and Methods

2.1. Sample Collection and Preparation

Collecting samples from different orchards and seasons can provide rich variation within a dataset and help to obtain a robust calibration model [22]. Therefore, three batches of macadamia nuts were collected over three seasons from one nut quality laboratory and two commercial processors, where samples from each batch were received from various farms with mixed cultivars to ensure wide representation of different varieties, growing conditions and handling practices (Figure 1). The first batch of samples (n = 85) were selected from 20 consignments of nut-in-shell submitted to a commercial laboratory for independent kernel quality assessment and contained both brown centers (BCs) and premium kernels (PKs) that were unaffected by quality issues. Nut-in-shell samples were dried to 1.5 ± 0.5% moisture concentration according to industry practice and manually cracked [5]. Whole kernels were cut in half using a sharp knife and classified as either brown center or premium using the Australian kernel assessment manual [5]. All samples from the first batch were used during hyperspectral imaging and chemical analysis (Figure 1). Sample preparation and chemical analysis of batch 1 are described in detail in the Supplementary Materials (Supplementary Table S1). The second batch of samples (n = 135) were selected from the processing line of a commercial processor after drying and cracking and classified as brown center or premium by qualified laboratory staff (Figure 1). These samples were used in hyperspectral imaging and classification model development only. The third batch of samples (n = 28) were collected from the processing lines of two commercial processers after drying and cracking and classified as brown center or premium by qualified laboratory staff (Figure 1). These kernels were then used in hyperspectral imaging and classification model development only. In total, 248 macadamia nuts were used in this study (Table 1).

Table 1. Distribution of kernel sample classes of macadamia nuts in the training and independent test datasets that were imaged in face-up and face-down orientations, and for both image orientations combined to form a pooled dataset.

Kernel Orientation	Kernel Class	Training Dataset	Training (%)	Test Dataset	Test (%)	Total
Face-up	BC	96	49.7%	24	50.0%	120
	PK	97	50.3%	24	50.0%	121
	Total	193		48		241
Face-down	BC	88	44.4%	25	50.0%	113
	PK	110	55.6%	25	50.0%	135
	Total	198		50		248
Pooled	BC	181	46.9%	52	53.1%	233
	PK	210	53.7%	46	46.9%	256
	Total	391		98		489

BC: brown center disorder; PK: premium center kernel.

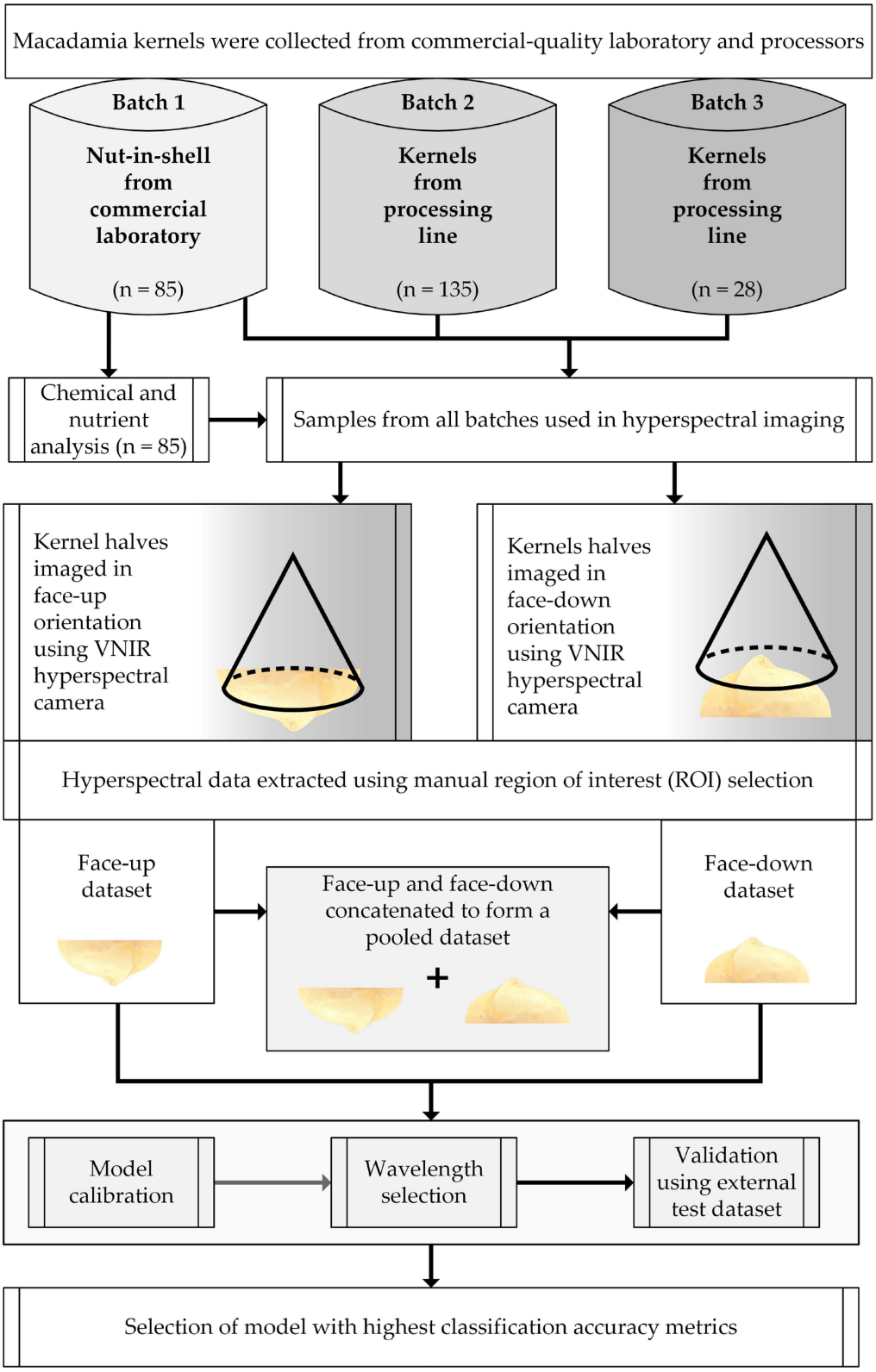


Figure 1. Flowchart showing sample collection, experimental design and workflow employed during this study.

2.2. The Hyperspectral Imaging System, Calibration and Image Acquisition

The benchtop hyperspectral imaging system used to acquire images collected reflectance data in the visible–near-infrared (VNIR) spectral range (400–1000 nm) and in 1.3 nm intervals (Resonon Pika XC2, Bozeman, MT, USA). In total, reflectance data were

collected for each image for 462 spectral bands. The benchtop imaging system was positioned in a darkened room, and kernels were illuminated using four wide-spectrum (400–2500 nm) quartz halogen lights (SoLux 5000 K 35 W, Tailored Lighting Inc., Rochester, NY, USA). The hyperspectral system was physically setup such that the distance between the lights and camera lens and (1) the macadamia nut-in-shell samples was ~185 mm and ~350 mm, respectively, and (2) for the macadamia kernel samples, it was ~200 mm and ~360 mm, respectively. Image calibration was undertaken prior to image acquisition, and this process was facilitated using SpectrononPro (v3.4.3) image acquisition and analysis software (Resonon Inc., Bozeman, MT, USA). Corrected relative reflectance was calculated for each image using Equation (1):

$$R = \frac{(R_0 - D)}{(W - D)} \quad (1)$$

where R_0 is the raw spectral reflectance, D is the reflectance of a dark image captured with the camera lens covered, and W is the reflectance of a Spectralon calibration sheet that reflects approximately 99% of incident light [42]. Camera gain was set to zero to minimize noise, and the following configuration parameters were established during calibration and subsequent imaging: framerate 27.62, integration time 22.33 ms and stage scanning speed 615 pps. All reflectance data were automatically scaled by 10,000 integers by the software.

During sample imaging, raw macadamia kernels were cut in half and placed on black foam in face-up and face-down orientations for imaging (Figure 2). For seven samples, one kernel half was destroyed during cutting and could only be imaged in face-down orientation (Table 1). In total, images of 248 individual macadamia kernels were acquired in face-up orientation and 241 images of kernels in face-down orientation (Table 1).

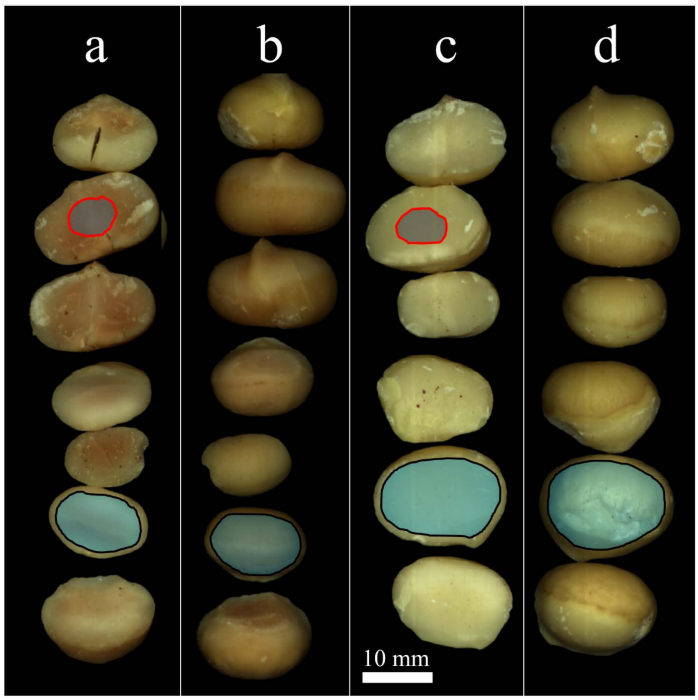


Figure 2. Pseudo-RGB image showing macadamia kernel classes: (**a**) brown center imaged in face-up orientation; (**b**) brown center imaged face-down orientation; (**c**) premium kernel imaged face-up orientation; and (**d**) premium kernel imaged face-down orientation. The spatial region with red outline represents an example of a ≥7 mm diameter sub-sample of kernel flesh that was removed for chemical and nutrient analysis, and the spatial regions with blue shading represent an example of the manually selected regions of interest (ROIs) where mean spectral data were extracted for use in model development.

2.3. Region of Interest Selection, Spectral Data Extraction, Outlier Detection and Dataset Assignment

Each single hyperspectral image contained numerous individual macadamia samples, and therefore, the region of interest for each sample was manually selected to remove background and other samples. The 'lasso tool' native to SpectrononPro software was used to select a region that represented the visible sample surface (Figure 2) (Resonon Inc., Bozeman, MT, USA). Specifically, the mean reflectance for all 462 wavelengths was extracted from all pixels in the manually selected region and for (1) the visible flesh located inside the cut kernel in the face-up orientation, including evidence of brown center disorder where applicable (Figure 2a,c); and (2) the outside kernel surface, visible for kernels positioned in the face-down orientation (Figure 2b,d). In total, mean reflectance was extracted from 241 kernels imaged in the face-up and 248 kernels imaged in the face-down orientation and used in the raw and unprocessed state during model development (Table 1). In this study, each classification sample in both the face-up and face-down image orientations was concatenated to create a 'pooled' dataset (n = 489) that contained all imaged samples (Table 1). The pooled dataset was created to represent real-world conditions during commercial processing where individual kernels lay scattered in a variety of both face-up and face-down orientations on a conveyor surface while transiting a mechanized processing factory.

Outlier detection was performed using principal component analysis (PCA) of all spectral wavelengths for each image orientation dataset separately and by visual inspection of the plotted spectra [11]. No spectral outliers were identified or removed from the dataset. Proportioning data randomly into calibration and independent test datasets and using various validation methods, including cross-validation and independent testing, can help to ensure model validity and robustness [22,43]. In the current study, all samples were randomly partitioned into two separate datasets (80:20, respectively) for use during model development (training dataset) and independent evaluation (test dataset) (Table 1). The distribution of samples following random allocation to training and independent test datasets and their respective true (reference) kernel classes are presented in Table 1.

2.4. Classification Model Development Using Supervised Machine Learning Techniques

MATLAB's Classification Learner was used to investigate a wide variety of algorithms prior to selecting the most promising techniques for greater examination [44]. Initially, all 34 classification models available within the MATLAB Statistics and Machine Learning Toolbox were trained using default parameters and 10-fold cross-validation with unprocessed spectra from all available wavelengths and the pooled training dataset [44,45] (Supplementary Table S2). Models with the highest cross-validation accuracy during training were selected for further investigation using the datasets with different image orientations and selected important wavelengths [45] (Supplementary Table S2).

The five most promising model types selected for further investigation were (1) linear discriminant (LD) analysis, (2) support vector machine (SVM) classification, (3) *k*-nearest neighbors (KNNs), (4) ensemble method using subspace discriminant analysis (ESD) and (5) artificial neural network (ANN) [46]. Models were then developed for the five types using the training data for each image orientation and the pooled dataset, respectively, and with four levels of spectral pre-processing: (1) no pre-processing and using all available (n = 462) wavelengths; (2) latent variables identified by PCA parameterized to explain 99% of the variation in the original wavelengths; (3) using only the ten highest ranking wavelengths identified using the minimum redundancy maximum relevance (MRMR) algorithm; and (4) the same 10 wavelengths identified using MRMR and with decomposition by PCA in latent variables. Using all available hyperspectral data introduces co-linearity and redundancy between wavelengths and can slow computational performance; therefore, spectral processing was used to reduce the dimensionality of hyperspectral data [47]. The four levels were selected to investigate the effects of dimensionality reduction on classifica-

tion accuracy following selection of important wavelengths, reducing the number of latent variables and the combination of these techniques prior to model development [48].

The MRMR algorithm can identify an optimal set of wavelengths that are mutually and maximally dissimilar and represent the response macadamia nut classes and is well suited to classification problems [46]. The principle of the MRMR algorithm is to minimize redundancy within the available wavelengths and maximize the relevance of the selected wavelengths to the response classes [46,49]. The MRMR algorithm quantifies redundancy and relevance using mutual information contained in the wavelengths and identifies wavelengths that are mutually disparate while maintaining high correlation to the known reference classes [49,50]. The MRMR algorithm was executed in MATLAB and configured to select 10 wavelengths because successful classification using 10 wavelengths could subsequently be used to develop multispectral imaging systems [51]. Wavelength selection using MRMR, PCA and all model development and testing was completed using the Statistics and Machine Learning Toolbox v12.5 in MATLAB (version R2023a, v9.14.0.2206163, The Mathworks Inc., Natick, MA, USA).

2.5. Evaluation of Classification Models

The performance of classification models was evaluated using class predictions for samples during cross-validation and independent test samples during testing, using the following equations:

$$\text{Accuracy } (\%) = \frac{TP + TN}{TP + FN + TN + FN} \times 100 \quad (2)$$

$$\text{Precision } (\%) = \frac{TP}{TP + FP} \times 100 \quad (3)$$

$$\text{Recall } (\%) = \frac{TP}{TP + FN} \times 100 \quad (4)$$

$$\text{F1 score} = \frac{2 \times precision \times recall}{precision \times recall} \times 100 \quad (5)$$

where *TP* are the true positive, *TN* are the true negative, *FP* are the false positive and *FN* are the false negative cases, respectively [52,53]. During evaluation and selection of the best-fit model for each image orientation, priority was given to metrics for the test dataset, and in the event of ties, the F1 score was used to select the highest-performing model.

3. Results

3.1. Spectral Reflectance and Principal Component Analysis of Brown Center and Premium Kernels Imaged in Different Orientations

In general, macadamia kernels with brown center disorder had lower mean reflectance regardless of the image orientation used during imaging or following image concatenation to form a pooled dataset (Figures 3a,b and 4a). Lower reflectance for brown center kernels was especially evident in the face-up images and in the region between 400 and 750 nm (Figure 3a). The spectral signature of brown center and premium kernels was similar above 800 nm, and the spectra extracted from both classes and both image orientations showed a distinct peak at 890 nm followed by a trough between 900 and 950 nm (Figure 3a,b). Lastly, variation in reflectance between the classes in the range 450–650 nm was increased in images taken in the face-up compared with the face-down orientation (Figure 3a,b).

Principal component analyses performed on the spectral data determined patterns in variation for both classes and from images of kernels in the face-up and face-down orientations and their pooled dataset. Principal components (PCs) one and two explained 100% of total variation (87% and 13%, respectively) extracted from images of kernels in the face-up orientation (Figure 3c). The scores plot for PC1 and PC2 showed that premium samples were well grouped along a linear plane, and brown center samples were more varied (Figure 3c). However, the two kernel classes were separated with limited overlap

along a diagonal linear axis (Figure 3c). For spectra extracted from images of kernels in the face-down orientation, PC1 and PC2 explained 99% of total variation (85% and 14%, respectively) (Figure 3d). Samples from the premium class displayed more scattering when imaged in the face-down orientation compared with images of kernels in the face-up orientation, and samples from both classes were more overlapped when using the face-down images (Figure 3c,d). PCA indicated that developing models using linear techniques such as linear discriminant analysis or support vector machines while employing a linear kernel function could be the most appropriate classification approach using the face-up image dataset only.

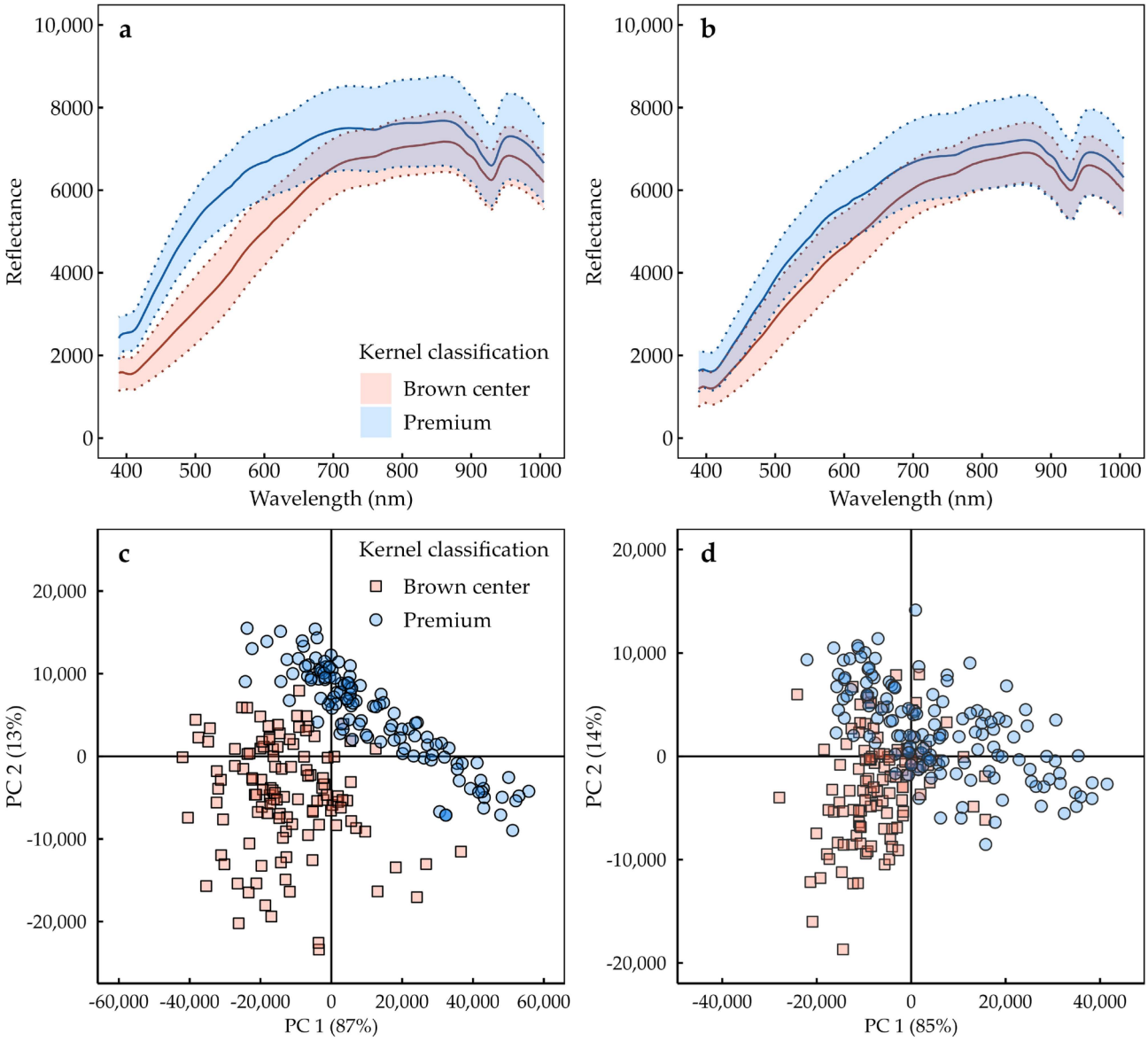


Figure 3. Mean reflectance (solid line) and standard deviation (dotted line) of brown center and premium kernels imaged in (**a**) face-up orientation and (**b**) face-down orientation, and scores plot for principal component analysis (PCA) of reflectance spectra for kernels imaged in (**c**) face-up and (**d**) face-down orientations. Kernel class designated by shape and color: brown center disorder (brown square) and premium kernel (blue circle).

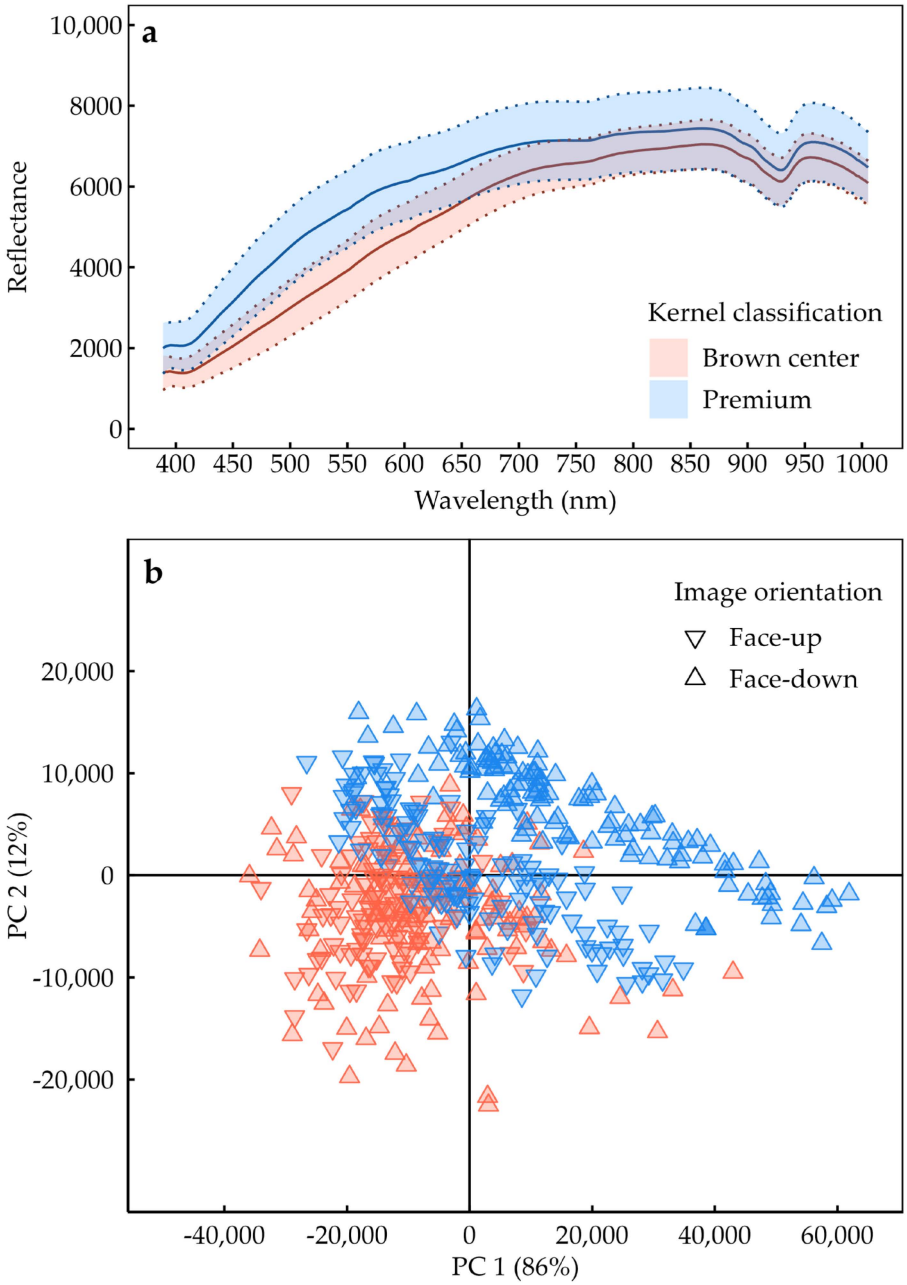


Figure 4. (**a**) Mean reflectance (solid line) and standard deviation (dotted line) of brown center and premium kernels following concatenation to form a pooled dataset, and (**b**) PCA scores plot for brown center (blue) and premium kernel (brown) samples in the pooled dataset. Image orientation and kernel classes are designated by shape and color, respectively: triangle (face-down image), upside-down triangle (face-up image), brown center (brown) and premium kernel (blue).

3.2. Accuracy of Models Developed Using Images of Inner Kernel Flesh in Face-Up Orientation

The best-fit models developed using images of kernels in the face-up orientation were the linear discriminant and ensemble subspace discriminant models; both models classified brown center disorder with 100% accuracy, precision, recall and F1 using the independent test dataset and all 462 wavelengths (Figure 5a and Table 2). All models developed using all available wavelengths had classification accuracy of >95% and >93%, precision of >97% and >92%, recall of >93% and >95% and F1 of >96% and >93% during cross-validation and independent testing, respectively (Figure 5a and Table 2). Following selection of important wavelengths, the best-fit models were also linear discriminant and ensemble subspace discriminant models.

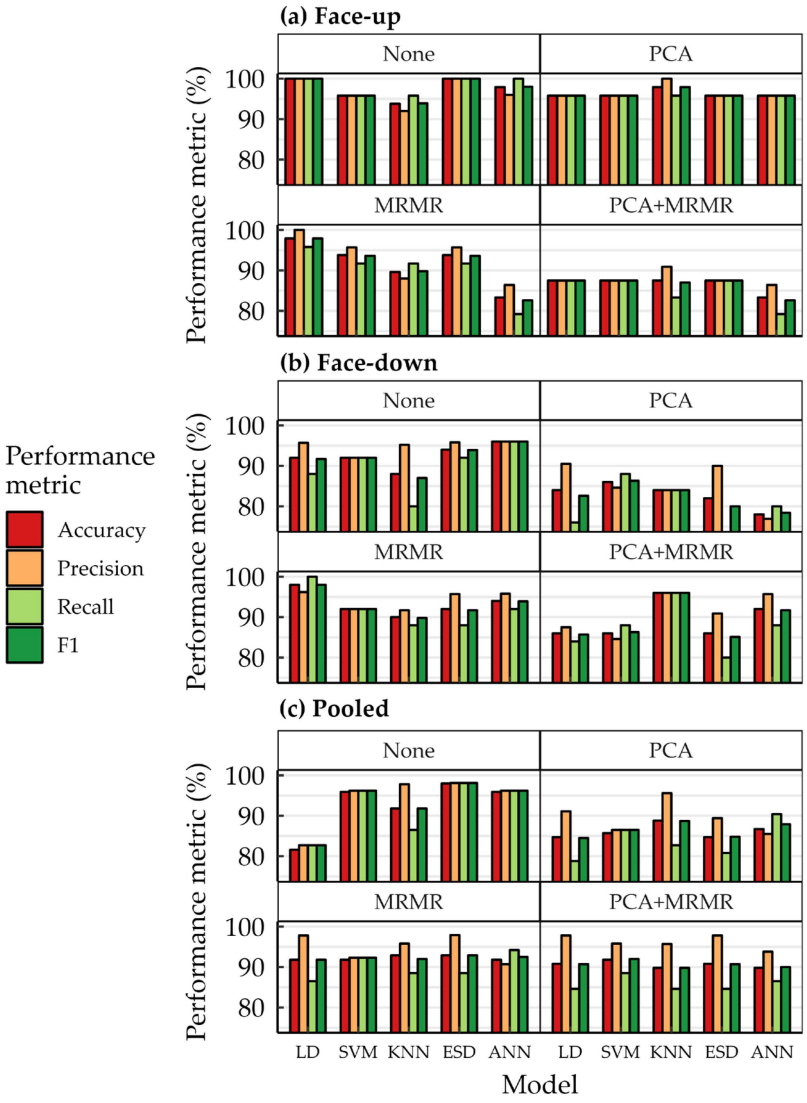


Figure 5. Comparative performance plot for models developed using images of macadamia kernel halves in (**a**) face-up, (**b**) face-down and (**c**) both face-up and face-down images combined to form a pooled dataset. Models developed using all raw spectra (n = 462) with no pre-processing (none) and following pre-processing with principal component analysis (PCA), wavelength selection (n = 10) using the minimum redundancy maximum relevance (MRMR) algorithm, and the combination of PCA and MRMR wavelength selection (PCA + MRMR) are shown.

The most important wavelengths for use with the face-up images identified by MRMR and used in the best-fit linear discriminant model were between 395 and 578 nm and from the visible spectrum (Table 3). The best-fit models using selected wavelengths classified samples during independent testing with accuracy of 97.9% and 93.8%, precision of 100% and 95.7%, recall of 95.8% and 91.7% and F1 of 97.9% and 93.6%, respectively (Table 2 and Figure 5a). In contrast, models developed using SVM, KNN and ANN algorithms all had test accuracy < 94%, precision < 96%, recall < 92% and F1 < 94% (Table 2 and Figure 5a). The ESD model developed using selected wavelengths misclassified brown centers as premium kernels (Type II error 'missed detection') in 8.3% and 4.2% of cases during validation and testing, respectively (Figure 6a,d).

Table 2. Results of classification models developed using hyperspectral images of macadamia kernels in face-up orientation.

Model	Feature Ranking Algorithm	WL	Validation (%)				Test Data (%)			
			Acc	Pr	Rcl	F1	Acc	Pr	Rcl	F1
LD	None	462	96.9	100.0	93.8	96.8	100.0	100.0	100.0	100.0
	PCA	462	94.8	100.0	89.6	94.5	95.8	95.8	95.8	95.8
	MRMR	10	95.3	98.9	91.7	95.1	97.9	100.0	95.8	97.9
	PCA + MRMR	10	92.7	94.6	90.6	92.6	87.5	87.5	87.5	87.5
SVM	None	462	97.4	100.0	94.8	97.3	95.8	95.8	95.8	95.8
	PCA	462	97.4	100.0	94.8	97.3	95.8	95.8	95.8	95.8
	MRMR	10	93.3	93.7	92.7	93.2	93.8	95.7	91.7	93.6
	PCA + MRMR	10	92.7	93.6	91.7	92.6	87.5	87.5	87.5	87.5
KNN	None	462	96.4	97.8	94.8	96.3	93.8	92.0	95.8	93.9
	PCA	462	97.4	100.0	94.8	97.3	97.9	100.0	95.8	97.9
	MRMR	10	92.7	93.6	91.7	92.6	89.6	88.0	91.7	89.8
	PCA + MRMR	10	91.2	92.5	89.6	91.0	87.5	90.9	83.3	87.0
ESD	None	462	98.4	100.0	96.9	98.4	100.0	100.0	100.0	100.0
	PCA	462	94.8	100.0	89.6	94.5	95.8	95.8	95.8	95.8
	MRMR	10	96.9	97.9	95.8	96.8	93.8	95.7	91.7	93.6
	PCA + MRMR	10	92.2	93.5	90.6	92.1	87.5	87.5	87.5	87.5
ANN	None	462	97.9	98.9	96.9	97.9	97.9	96.0	100.0	98.0
	PCA	462	96.4	95.9	96.9	96.4	95.8	95.8	95.8	95.8
	MRMR	10	91.7	91.7	91.7	91.7	83.3	86.4	79.2	82.6
	PCA + MRMR	10	88.6	89.4	87.5	88.4	83.3	86.4	79.2	82.6

LD: linear discriminant; SVM: support vector machine; KNN: k-nearest neighbor; ESD: ensemble subspace discriminant; ANN: artificial neural network; WL: wavelengths used in model; Acc: total accuracy; Pr: precision; Rcl: recall; F1: harmonic mean between precision and recall.

Table 3. Important wavelengths identified using the minimum redundancy maximum relevance (MRMR) algorithm and the training dataset for images of macadamia kernels in face-up, face-down and both orientations pooled.

Kernel Orientation	Wavelengths (nm)
Face-up	395.44, 401.98, 409.82, 436.03, 463.6, 491.24, 511.01, 532.13, 537.41, 578.45
Face-down	390.20, 398.05, 400.66, 408.52, 416.37, 495.19, 514.96, 553.29, 675.58, 789.56
Pooled	391.51, 399.36, 484.65, 493.87, 511.01, 533.45, 558.58, 563.88, 585.08, 941.20

3.3. Accuracy of Models Developed Using Images of Outer Kernel Surface in Face-Down Orientation

The best-fit model developed using images of kernels in the face-down orientation predicted brown center disorder with 98% accuracy, 96.2% precision and 100% recall for the test dataset and was calibrated using a linear discriminant model trained with ten selected wavelengths only (Table 4 and Figure 7). The best-fitting model developed using all wavelengths classified the test dataset with 96% accuracy, precision and recall using an ANN (Table 4). All models developed using all available wavelengths from images of the outer kernel (face-down images) had classification accuracy of >85% and >88%, precision of >84% and >92%, recall of >79% and >80% and F1 of >83% and >87% during cross-validation and independent testing, respectively (Table 4 and Figure 5b). Following

selection of important wavelengths, the best-fit models were developed using LD and ESD and resulted in 90.9% and 98% validation and 98% and 94% test accuracy, respectively (Table 4). The most important wavelengths identified for developing models using the face-down images were between 390 and 675 nm within the visible spectrum, but unlike the face-up dataset, also included 789 nm from the near-infrared range (Table 3). However, redeveloping the ANN model using selected wavelengths only reduced test classification accuracy from 96% to 94% (Table 4). In contrast, the models developed using SVM and KNN algorithms both had test accuracies <94% (Table 4). The best-fit LD model developed using selected wavelengths misclassified premium kernels as brown center kernels (Type I error 'false positive') in 4.5% and 4% of cases during validation and testing, respectively (Figure 7).

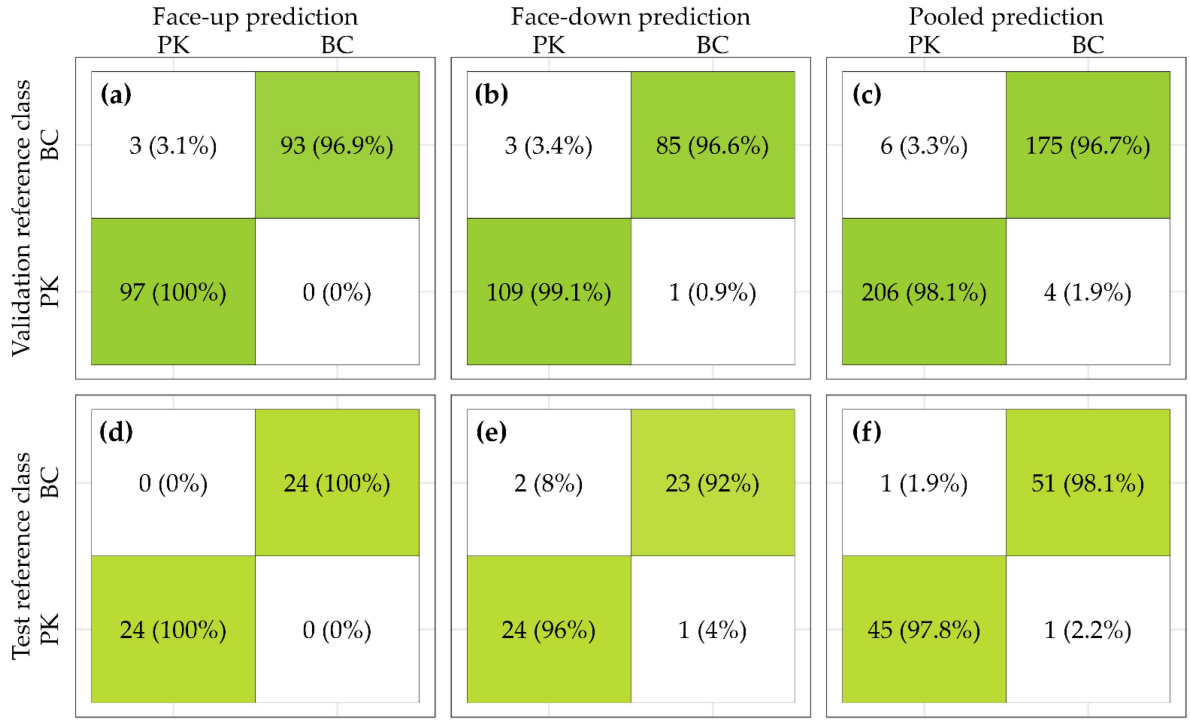


Figure 6. Confusion matrix showing cross-validation (**a**–**c**) in dark green and independent test (**d**–**f**) in light green classification results (recall) for the best-fit models developed using the ensemble subspace discriminant (ESD) model and all (n = 462) VNIR wavelengths extracted from images of macadamia kernels in the (**a**,**d**) face-up orientation, (**b**,**e**) face-down orientation and (**c**,**f**) pooled dataset. Classes within each matrix cell are denoted as BC: brown center and PK: premium kernel. Values represent the number of samples and (recall%).

Table 4. Results of classification models developed using hyperspectral images of macadamia kernels in face-down orientation.

Model	Feature Ranking Algorithm	WL	Validation (%)				Test Data (%)			
			Acc	Pr	Rcl	F1	Acc	Pr	Rcl	F1
LD	None	462	98.5	97.8	98.9	98.3	92.0	95.7	88.0	91.7
	PCA	462	82.3	86.3	71.6	78.3	84.0	90.5	76.0	82.6
	MRMR	10	90.9	93.8	85.2	89.3	98.0	96.2	100.0	98.0
	PCA + MRMR	10	84.3	88.0	75.0	81.0	86.0	87.5	84.0	85.7

Table 4. *Cont.*

Model	Feature Ranking Algorithm	WL	Validation (%)				Test Data (%)			
			Acc	Pr	Rcl	F1	Acc	Pr	Rcl	F1
SVM	None	462	86.9	84.4	86.4	85.4	92.0	92.0	92.0	92.0
	PCA	462	85.4	85.5	80.7	83.0	86.0	84.6	88.0	86.3
	MRMR	10	88.9	83.7	93.2	88.2	92.0	92.0	92.0	92.0
	PCA + MRMR	10	91.9	90.0	92.0	91.0	86.0	84.6	88.0	86.3
KNN	None	462	85.9	87.5	79.5	83.3	88.0	95.2	80.0	87.0
	PCA	462	85.9	84.9	83.0	83.9	84.0	84.0	84.0	84.0
	MRMR	10	90.9	90.7	88.6	89.7	90.0	91.7	88.0	89.8
	PCA + MRMR	10	91.4	91.8	88.6	90.2	96.0	96.0	96.0	96.0
ESD	None	462	98.0	98.8	96.6	97.7	94.0	95.8	92.0	93.9
	PCA	462	82.8	88.6	70.5	78.5	82.0	90.0	72.0	80.0
	MRMR	10	88.9	94.6	79.5	86.4	92.0	95.7	88.0	91.7
	PCA + MRMR	10	83.3	87.7	72.7	79.5	86.0	90.9	80.0	85.1
ANN	None	462	88.9	91.3	83.0	86.9	96.0	96.0	96.0	96.0
	PCA	462	80.8	76.6	81.8	79.1	78.0	76.9	80.0	78.4
	MRMR	10	91.4	89.9	90.9	90.4	94.0	95.8	92.0	93.9
	PCA + MRMR	10	89.9	89.5	87.5	88.5	92.0	95.7	88.0	91.7

LD: linear discriminant; SVM: support vector machine; KNN: k-nearest neighbor; ESD: ensemble subspace discriminant; ANN: artificial neural network; WL: wavelengths used in model; Acc: total accuracy; Pr: precision; Rcl: recall; F1: harmonic mean between precision and recall.

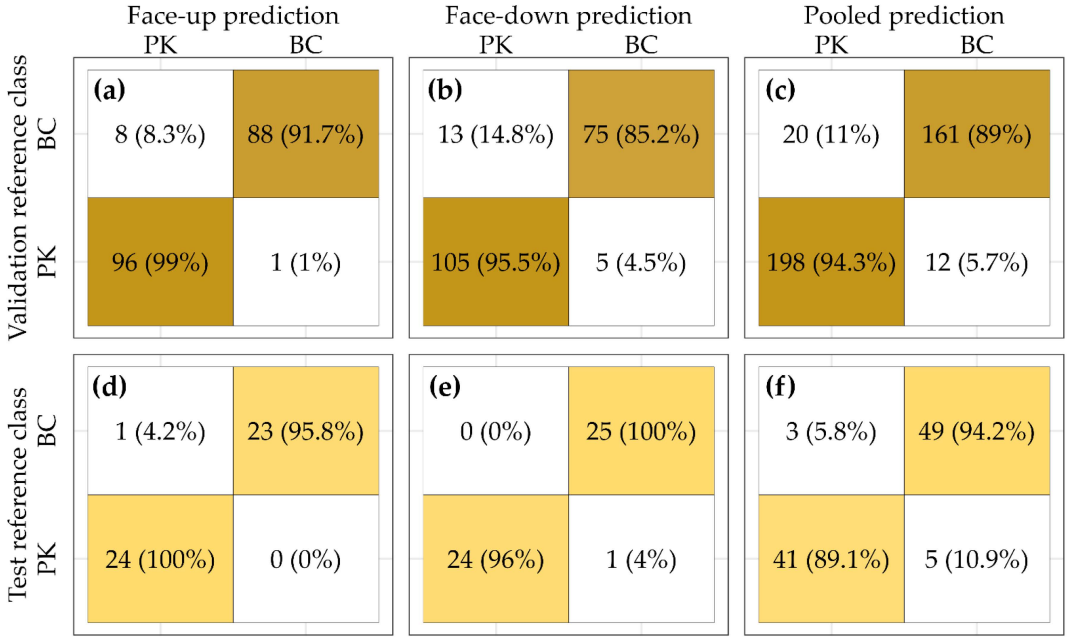


Figure 7. Confusion matrix showing cross-validation (**a**–**c**) in dark yellow and independent test (**d**–**f**) in light yellow classification results for the best-fit models developed using ten wavelengths selected by the MRMR algorithm and extracted from images of macadamia kernels in (**a**,**d**) face-up orientation using the linear discriminant model, (**b**,**e**) face-down orientation using the LD model and (**c**,**f**) pooled dataset using the artificial neural network (ANN) model. Classes within each matrix cell are denoted as BC: brown center and PK: premium kernel. Values represent sample number of samples and (recall%) in parentheses.

3.4. Accuracy of Models Developed Using All Images Combined to Form a Pooled Dataset

The best-fit model developed using all wavelengths from images of kernels in the pooled dataset were ESD and ANN and predicted brown center disorder with 98% and 95.9% accuracy in the test dataset, respectively (Table 5 and Figure 6). All models developed using all wavelengths from images in the pooled dataset had classification accuracy of >87% and >81%, precision of >89% and >82%, recall of >80% and >82% and F1 of >85% and >82% during cross-validation and independent testing, respectively (Table 5 and Figure 5b). Following redevelopment of models using selected wavelengths only, the best-fit models were SVM and ANN, and both models classified the test dataset with 91.8% accuracy (Table 5). The ANN model developed using selected wavelengths misclassified brown center kernels as PK (Type II error 'missed detection') in 11% and 5.8% of cases during validation and testing, respectively (Figure 7). However, using all wavelengths, the best-fit ESD model misclassified brown center kernels as PK (Type II error 'missed detection') in 3.3% and 1.9% of cases during validation and testing, respectively (Figure 6). The most important wavelengths selected for use with images from the pooled dataset were between 391 and 585 nm and from the visible spectrum, but unlike the face-up dataset, also included 941 nm from the near-infrared range (Table 3).

Table 5. Results of classification models developed using hyperspectral images of macadamia kernels in face-up and face-down orientations concatenated to form a pooled dataset.

Model	Feature Ranking Algorithm	WL	Validation (%)				Test Data (%)			
			Acc	Pr	Rcl	F1	Acc	Pr	Rcl	F1
LD	None	462	90.5	89.6	90.1	89.8	81.6	82.7	82.7	82.7
	PCA	462	82.9	83.5	78.5	80.9	84.7	91.1	78.8	84.5
	MRMR	10	91.3	95.1	85.6	90.1	91.8	97.8	86.5	91.8
	PCA + MRMR	10	85.4	83.0	86.2	84.6	90.8	97.8	84.6	90.7
SVM	None	462	94.9	96.0	92.8	94.4	95.9	96.2	96.2	96.2
	PCA	462	84.4	85.3	80.1	82.6	85.7	86.5	86.5	86.5
	MRMR	10	92.3	92.2	91.2	91.7	91.8	92.3	92.3	92.3
	PCA + MRMR	10	83.1	85.3	76.8	80.8	91.8	95.8	88.5	92.0
KNN	None	462	87.0	90.6	80.1	85.0	91.8	97.8	86.5	91.8
	PCA	462	84.4	90.0	74.6	81.6	88.8	95.6	82.7	88.7
	MRMR	10	89.8	89.8	82.9	86.2	92.9	95.8	88.5	92.0
	PCA + MRMR	10	87.0	87.5	85.1	86.3	89.8	95.7	84.6	89.8
ESD	None	462	97.4	97.8	96.7	97.2	98.0	98.1	98.1	98.1
	PCA	462	82.1	83.2	76.8	79.9	84.7	89.4	80.8	84.8
	MRMR	10	90.8	95.0	84.5	89.5	92.9	97.9	88.5	92.9
	PCA + MRMR	10	84.9	83.5	84.0	83.7	90.8	97.8	84.6	90.7
ANN	None	462	93.6	93.8	92.3	93.0	95.9	96.2	96.2	96.2
	PCA	462	83.6	83.1	81.2	82.1	86.7	85.5	90.4	87.9
	MRMR	10	91.8	93.1	89.0	91.0	91.8	90.7	94.2	92.5
	PCA + MRMR	10	82.1	79.4	82.9	81.1	89.8	93.8	86.5	90.0

LD: linear discriminant; SVM: support vector machine; KNN: k-nearest neighbor; ESD: ensemble subspace discriminant; ANN: artificial neural network; WL: wavelengths used in model; Acc: total accuracy; Pr: precision; Rcl: recall; F1: harmonic mean between precision and recall.

4. Discussion

4.1. Classification of Brown Center Disorder Using Chemometric Models

Brown center disorder was accurately classified using images of kernels in all orientations and model types investigated. Successful classification using chemometrics was possible in our study because numerous chemical and nutritional attributes of brown center kernels were significantly higher than those of premium kernels. Sucrose is a disaccharide carbohydrate that contains a fructose and a glucose molecule that can be reduced via sucrose hydrolysis in the presence of water [34]. In this study, fructose and glucose were significantly higher in brown centers compared with premium kernels. This result is consistent with our complimentary study that identified that brown center kernels contain elevated levels of the reducing sugars fructose and glucose, indicating that sucrose hydrolysis had occurred in our brown center samples [54]. Previously, NIR spectroscopy has shown potential to accurately predict fructose and glucose in fruit juice and pear fruits [55,56]. This is possible because sugars are organic molecules and are mainly composed of C–C, C–H, C–O and O–H bonds [57]. The hydrogen-containing bonds O–H and C–H omit sharp overtones and combination signatures in the NIR region [56,58]. Specifically, our samples showed reflectance troughs (increased absorbance) around 720–780 nm and 940 nm that may be associated with the third overtone of C–H_2 stretching and the third overtone of O–H stretching, respectively [56]. Additionally, fructose was only present within brown center samples and glucose in very small quantities. Therefore, we suggest classification of brown center disorder may be possible using the prediction of reducing sugars as a proxy for brown center disorder classification.

Successful classification of brown center disorder from the outside kernel surface can also be explained by detectable and significant differences in mineral nutrient concentrations between the kernel classes. In this study, kernel nutrients Ca, Fe, K, Mg, P, S and Zn were significantly higher in the flesh of brown center kernels compared with premium kernels. In a complementary study where samples were collected from one variety, the difference among mineral nutrients was not as pronounced, but the sugar differences remained the same [34,54]. Chemical differences are not visible to the naked eye, and there was overlap in the reflectance signature for some samples between 750 and 1000 nm, indicating that some brown centers may be very difficult to identify. The important wavelengths we identified are similar to wavelengths identified in another study that predicted nutrients in the *Canarium indicum* nut using the VNIR spectrum [59]. Specifically, wavelengths at 580, 590 and 790 nm for Fe; 440, 490, 510, 540, 670, 680, 790 and 940 nm for K; 400, 580 and 940 nm for Mg; 400 nm for S; and 400, 580 and 790 nm for Zn have been identified as important to predict both the nutrients in *Canarium indicum* [59] and the brown center kernels imaged in this study. Additionally, using HSI within the VNIR spectrum has previously predicted Ca, Mg, Fe and Zn in avocado foliar tissue [13], and Ca in avocado fruit flesh [23]. We suggest that successful brown center disorder classification was partly supported by the detection of chemical differences between brown center and premium classes for Fe, K, Mg, S and Zn, and the differences were identifiable within the VNIR spectral region.

4.2. The Effect of Kernel Orientation and Model Type on Brown Center Disorder Classification Accuracy

Best-fit classification models (100%) were developed using all wavelengths and spectra from images of face-up kernels and were marginally more accurate than models developed using images of kernels in face-down (98%), or pooled image (98%) orientations. The best-fit model using VNIR face-down images was more accurate than another study using the NIR region (980–1680 nm) that reported 88.2% accuracy [60]. This may be attributed to the hyperspectral images in this study collecting both spectral and spatial data and, therefore, allowing inspection of a greater kernel surface area in comparison with the NIR point method. High test accuracy (98%) was achieved using images of the outside kernel surface (face-down images) when combined with linear discriminant (LD) analysis and

following wavelength selection. In another study, a chemometric method using LD analysis successfully classified immaturity, mold, insect damage and discoloration in macadamia using the shortwave infrared region (SWIR) (1000–2500 nm) [61]. Our results suggest brown center disorder classification is also possible using LD models with images from the VNIR spectral range. The best-fit model we developed using the pooled dataset had 98% accuracy and was developed using an ensemble subspace discriminant (ESD) model without wavelength selection. An ensemble classifier is a predictive model that comprises a weighted combination of multiple classification models, and in general, combining multiple classification models increases predictive performance [46]. Therefore, ESD models using a linear subspace were more accurate, or similar, compared with the respective LD models and for all image orientations without wavelength selection. Finally, 100% classification accuracy was possible using both ESD and LD models with face-up images.

All the different types of classification algorithms developed using images of kernels in the face-up orientation had high (>95%) validation and test accuracy. This was unsurprising because brown center disorder normally presents as brown discoloration within the internal kernel flesh and with limited visible presence on the outside kernel surface [10]. The kernels in this study were deliberately cut in half and placed face-up and perpendicular to the camera lens to present a clear view of the inside flesh. Therefore, our face-up imaging method maximized the potential for brown center disorder to be in direct line of sight with the camera and with minimal scattering effects, i.e., a perfect scenario. It should be cautioned that commercial macadamia processing does not allow for perfect scenario imaging of face-up kernels because markets place higher value on whole kernels, and therefore, processors aim to minimize mechanical damage and ensure high proportions of whole kernels [62]. Additionally, where whole kernels are accidentally broken into smaller sections or perfect halves, the likelihood of perfect perpendicular exposure to a hyperspectral camera remains unlikely. Therefore, the results for models developed using the face-down and pooled image datasets more closely resemble the reality of commercial nut processing.

4.3. Important Wavelengths Selected for Brown Center Disorder Classification

Best-fit models developed with only ten selected wavelengths maintained high accuracy using both face-up (97.9%) and face-down (98%) images, and 91.8% accuracy was also possible using all images combined into a pooled dataset. The important wavelengths identified by MRMR for the face-up dataset were all located between 395 and 578 nm and restricted to within the visible region of the cameras' spectral range and not within the NIR region. This suggests that for models developed using face-up images, successful classification of brown center disorder is based on the brown color the disorder manifests rather than changes in chemical and nutrient properties. On the contrary, and for models developed using both the face-up and pooled datasets, MRMR identified two important wavelengths at 790 nm and 940 nm, respectively. Both wavelengths are within the NIR spectral region that is extremely sensitive to organic matter concentrations because of molecules containing hydrogen, especially C–H, O–H and N–H [57,58]. The hyperspectral system used in this study can record spectral data for 462 individual wavelengths, and therefore identifying a single wavelength (out of ten) for each dataset in this region was of particular interest. Previous studies to predict sugar concentration have identified useful spectral information between 930 and 950 nm and related to C–H bonds [56,63].

In this study, models were developed using 10 important wavelengths in order to evaluate the potential for acceptable classification accuracy using more cost-effective multispectral systems [43,51]. Best-fit models developed with 10 wavelengths and the pooled dataset decreased classification accuracy, recall and F1 by >5%; however, precision was not substantially reduced. This suggests that higher rates of missed brown center detection are a considerable issue should our method be adopted in multispectral systems; however, incorrect detection of premium nut product as a brown center reject was not exacerbated by using fewer wavelengths. This result supports the potential adaptation to multispectral

systems that can be fitted to existing processing infrastructure. Considering the pooled dataset only, MRMR identified a single important wavelength at 941.2 nm, and we also observed a distinct reflectance trough (absorption peak) between 910 and 950 nm. Additional wavelengths within this region and previously related to bonds in sugars are the third overtones of C–H stretching at 910 nm and CH_2 stretching at 930 nm [63]. Additionally, an earlier successful method that detected internal damage in almond nuts using transmittance and ratios of wavelength pairs identified absorbance peaks around 710 and 930 nm that were attributed to the oxidation of oil during exposure to moisture [64].

4.4. Implications and Limitations of This Scoping Study

Commercial application of classifiers developed using VNIR hyperspectral images will require further investigation and redevelopment with more samples while considering the effects of local biases, the physical environment and other macadamia quality disorders. Disorders such as mold and basal discoloration manifest symptoms within the visible spectrum. For example, basal discoloration may be incorrectly rejected as a brown center because basal discoloration manifests within the visible spectral range on the kernel surface and can occur in 2.6–4.2% of an entire crop [65]. Basal discoloration is staining of the kernel surface that can occur over the entire lower hemisphere (one half), may present as blotches or patches, and can range from brown to grey, green, orange, salmon, yellow and purple [5]. Therefore, in some cases, basal discoloration can appear like brown center and be incorrectly rejected because, under the Australian classification system, premium-grade kernels may contain a limited amount of basal discoloration [5]. Using HSI may help to avoid confusion between premium kernels with basal discoloration and reject kernels containing brown center disorder.

In this study, VNIR hyperspectral images successfully classified brown center disorder (>90%) using the pooled image dataset that included images of kernels in both different orientations. Out of the three imaging methods investigated during this study, using the pooled image dataset most closely resembled the reality of commercial nut processing, where falling nuts land in random orientations on a conveyor system during transit around the factory, increasing variability within images. In this study, each classification model was trained at one time only using the training dataset and was then evaluated using its respective test dataset. However, this approach is limited to the knowledge gained from the initial training dataset, and the accuracy of further classifications cannot be guaranteed when new samples become available [66]. In contrast, incremental learning or implementing online updates of the original model can allow learning from a small amount of new training samples as they become available and help a model to adapt to variability in new samples and changing environmental conditions and improve generalizability [66,67]. Additionally, some macadamia kernels break into perfect halves during cracking because of an epicuticular wax coating on the surface of each cotyledon, and some varieties are more prone to halving than others [68]. Finally, this study assumes respective limitations for each model family, such as linearity and other physical constraints that may not represent industrial settings. For example, variability in factory lighting, humidity and air quality may affect hyperspectral image quality and reliability. Macadamia nuts have high oil content, and the presence of oily macadamia dust in the air can cause visually obstructive issues by settling on surfaces including camera lenses and screens [69]. Therefore, our study suggests future work needs to consider large and diverse pooled image datasets to mimic real-word scenarios during model training.

5. Conclusions

Combining spectral data from the VNIR range with various chemometric and machine learning tools demonstrated high accuracy classification of brown center disorder in macadamia kernels. Best-fit classification models developed using all wavelengths and all different image orientations provided high test (>98%) accuracy. Successful brown center disorder classification was possible using various chemometric methods because of

significant differences between brown center and premium kernels for the reducing sugars fructose and glucose. Selecting the important wavelengths correlated with the bonds in reducing sugar molecules and mineral nutrients reduces classifier model complexity while maintaining high accuracy and therefore allows for potential application to commercial in-line processing systems. Importantly, using a pooled library containing images of kernels in both face-up and face-down orientations resulted in similar accuracy because this method is the most representative of commercial macadamia processing. Our study provided strong evidence that HSI combined with machine learning can be utilized in nut processing lines to detect and reject brown center kernels while reducing human subjectivity errors and preventing economic loss by rejecting nuts that are not actual rejects.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae10111129/s1. Supplementary Table S1: Chemical and nutrient properties of the macadamia nut samples in batch one collected during this study; Supplementary Table S2: List of all available classifiers in the MATLAB machine learning toolbox. Refs. [70–73] are referred to in this part.

Author Contributions: Conceptualization, S.H.B. and H.M.W.; methodology, S.H.B., H.M.W., M.M., K.J. and M.B.F.; validation, S.H.B., M.B.F. and T.C.; formal analysis, M.B.F.; investigation, M.B.F., M.M. and S.H.B.; resources, S.H.B., H.M.W. and K.J.; data curation, M.B.F., M.M. and N.O.; writing—original draft preparation, M.B.F., M.M. and S.H.B.; writing—review and editing, H.M.W., T.C., K.J., M.M. and N.O.; visualization, M.B.F.; supervision, S.H.B.; project administration, M.M. and M.B.F.; funding acquisition, S.H.B. and H.M.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Queensland Government, Department of Tourism, Innovation and Sport (Advance QLD Industry Fellowship No: AQIRF092-2021RD4), the Australian Macadamia Society, CL Macs, Cropwatch Independent Laboratory, Macadamias Direct, Marquis Macadamias and Suncoast Gold Macadamias. The authors would like to thank Helen Wallace, Stephen Trueman, Steve Lee, Joel Michael, Peter Zummo, Jeff Clements, Mariana Proksch, Jon Perrin, Michael Green, Lucy Andrews, Julian Lancaster-Smith and Royce Alcorn for sharing expert knowledge, helping with the guidance of the project and providing technical support.

Data Availability Statement: Data are unavailable due to IP arrangements.

Acknowledgments: The authors thank the staff at Cropwatch Independent Laboratory, CL Macs, Macadamias Direct, Marquis Macadamias, Suncoast Coast Gold Macadamias, Hinkler Park Plantations, Macadamias Australia, RFM Macadamias and TQ Holdings Pty Ltd. for supplying samples and financial support. The authors would also like to thank Stephen Trueman, Steve Lee, Joel Michael, Peter Zummo, Jeff Clements, Mariana Proksch, Jon Perrin, Michael Green, Lucy Andrews, Julian Lancaster-Smith, Royce Alcorn and Leoni Kojetin for sharing expert knowledge, helping with the guidance of the project and providing technical support.

Conflicts of Interest: The authors declare there are no conflicts of interest.

References

1. Ros, E. Health benefits of nut consumption. *Nutrients* **2010**, *2*, 652–682. [CrossRef] [PubMed]
2. INC. *Nuts & Dried Fruits Statistical Yearbook 2022/2023*; The International Nut and Dried Fruit Council Foundation: Reus, Spain, 2023.
3. Mereles, L.; Ferro, E.; Alvarenga, N.; Caballero, S.; Wiszovaty, L.; Piris, P.; Michajluk, B. Chemical composition of Macadamia integrifolia (Maiden and Betche) nuts from Paraguay. *Int. Food Res. J.* **2017**, *24*, 2599–2608.
4. O'Neil, C.E.; Nicklas, T.A. Tree nut consumption is associated with better nutrient adequacy and diet quality in adults: National Health and Nutrition Examination Survey 2005–2010. *Nutrients* **2015**, *7*, 595–607. [CrossRef]
5. AMS. *Kernel Assessment Manual*; Australian Macadamia Society: Lismore, Australia, 2021; Volume 7, pp. 1–51.
6. Gama, T.; Wallace, H.M.; Trueman, S.J.; Hosseini-Bai, S. Quality and shelf life of tree nuts: A review. *Sci. Hortic.* **2018**, *242*, 116–126. [CrossRef]
7. Walton, D.A.; Randall, B.W.; Le Lagadec, M.D.; Wallace, H.M. Maintaining high moisture content of macadamia nuts-in-shell during storage induces brown centres in raw kernels. *J. Sci. Food Agric.* **2013**, *93*, 2953–2958. [CrossRef] [PubMed]

8. Rogel-Castillo, C.; Boulton, R.; Opastpongkarn, A.; Huang, G.; Mitchell, A.E. Use of near-infrared spectroscopy and chemometrics for the nondestructive identification of concealed damage in raw almonds (*Prunus dulcis*). *J. Agric. Food Chem.* **2016**, *64*, 5958–5962. [CrossRef]
9. Kader, A. Impact of nut postharvest handling, de-shelling, drying and storage on quality. In *Improving the Safety and Quality of Nuts*; Harris, L.J., Ed.; Woodhead Publishing Limited: Cambridge, UK, 2013; pp. 22–34. [CrossRef]
10. Le Lagadec, M. Kernel brown centres in macadamia: A review. *Crop Pasture Sci.* **2009**, *60*, 1117–1123. [CrossRef]
11. Chen, C.; Pan, Z. Postharvest processing of tree nuts: Current status and future prospects—A comprehensive review. *Compr. Rev. Food Sci. Food Saf.* **2022**, *21*, 1702–1731. [CrossRef]
12. Manley, M. Near-infrared spectroscopy and hyperspectral imaging: Non-destructive analysis of biological materials. *Chem. Soc. Rev.* **2014**, *43*, 8200–8214. [CrossRef]
13. Hapuarachchi, N.S.; Trueman, S.J.; Kämper, W.; Farrar, M.B.; Wallace, H.M.; Nichols, J.; Bai, S.H. Hyperspectral imaging of adaxial and abaxial leaf surfaces for rapid assessment of foliar nutrient concentrations in Hass avocado. *Remote Sens.* **2023**, *15*, 3100. [CrossRef]
14. Farrar, M.B.; Wallace, H.M.; Tahmasbian, I.; Yule, C.M.; Dunn, P.K.; Hosseini Bai, S. Rapid assessment of soil carbon and nutrients following application of organic amendments. *Catena* **2023**, *223*, 106928. [CrossRef]
15. Dung, C.D.; Trueman, S.J.; Wallace, H.M.; Farrar, M.B.; Gama, T.; Tahmasbian, I.; Bai, S.H. Hyperspectral imaging for estimating leaf, flower, and fruit macronutrient concentrations and predicting strawberry yields. *Environ. Sci. Pollut. Res.* **2023**, *30*, 114166–114182. [CrossRef] [PubMed]
16. Park, B.; Lu, R. *Hyperspectral Imaging Technology in Food and Agriculture*; Springer: New York, NY, USA, 2015. [CrossRef]
17. Malmir, M.; Tahmasbian, I.; Xu, Z.; Farrar, M.B.; Bai, S.H. Prediction of macronutrients in plant leaves using chemometric analysis and wavelength selection. *J. Soils Sediments* **2020**, *20*, 249–259. [CrossRef]
18. Su, W.H.; Sun, D.W. Fourier transform infrared and Raman and hyperspectral imaging techniques for quality determinations of powdery foods: A review. *Compr. Rev. Food Sci. Food Saf.* **2018**, *17*, 104–122. [CrossRef]
19. Siche, R.; Vejarano, R.; Aredo, V.; Velasquez, L.; Saldaña, E.; Quevedo, R. Evaluation of food quality and safety with hyperspectral Imaging (HSI). *Food. Eng. Rev.* **2016**, *8*, 306–322. [CrossRef]
20. Saha, D.; Manickavasagan, A. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review. *Curr. Res. Food Sci.* **2021**, *4*, 28–44. [CrossRef]
21. Gama, T.; Farrar, M.B.; Tootoonchy, M.; Wallace, H.M.; Trueman, S.J.; Tahmasbian, I.; Bai, S.H. Hyperspectral imaging predicts free fatty acid levels, peroxide values, and linoleic acid and oleic acid concentrations in tree nut kernels. *LWT* **2024**, *199*, 116068. [CrossRef]
22. Nicolaï, B.M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K.I.; Lammertyn, J. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* **2007**, *46*, 99–118. [CrossRef]
23. Kämper, W.; Trueman, S.J.; Tahmasbian, I.; Bai, S.H. Rapid determination of nutrient concentrations in Hass avocado fruit by Vis/NIR hyperspectral imaging of flesh or skin. *Remote Sens.* **2020**, *12*, 3409. [CrossRef]
24. Arendse, E.; Fawole, O.A.; Magwaza, L.S.; Opara, U.L. Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review. *J. Food Eng.* **2018**, *217*, 11–23. [CrossRef]
25. Farrar, M.B.; Wallace, H.M.; Brooks, P.; Yule, C.M.; Tahmasbian, I.; Dunn, P.K.; Hosseini Bai, S. A performance evaluation of Vis/NIR hyperspectral imaging to predict curcumin concentration in fresh turmeric rhizomes. *Remote Sens.* **2021**, *13*, 1807. [CrossRef]
26. Davur, Y.J.; Kämper, W.; Khoshelham, K.; Trueman, S.J.; Bai, S.H. Estimating the ripeness of Hass avocado fruit using deep learning with hyperspectral imaging. *Horticulturae* **2023**, *9*, 599. [CrossRef]
27. Tziotzios, G.; Pantazi, X.E.; Paraskevas, C.; Tsitsopoulos, C.; Valasiadis, D.; Nasiopoulou, E.; Michailidis, M.; Molassiotis, A. Non-Destructive Quality Estimation Using a Machine Learning-Based Spectroscopic Approach in Kiwifruits. *Horticulturae* **2024**, *10*, 251. [CrossRef]
28. Workhwa, S.; Khanthong, T.; Manmak, N.; Thompson, A.K.; Teerachaichayut, S. Detection of Hardening in Mangosteens Using near-Infrared Hyperspectral Imaging. *Horticulturae* **2024**, *10*, 345. [CrossRef]
29. Wang, B.; Yang, H.; Li, L.; Zhang, S. Non-Destructive Detection of Cerasus Humilis Fruit Quality by Hyperspectral Imaging Combined with Chemometric Method. *Horticulturae* **2024**, *10*, 519. [CrossRef]
30. Devos, O.; Ruckebusch, C.; Durand, A.; Duponchel, L.; Huvenne, J.-P. Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. *Chemom. Intellig. Lab. Syst.* **2009**, *96*, 27–33. [CrossRef]
31. Ma, J.; Sun, D.W.; Pu, H.; Cheng, J.H.; Wei, Q. Advanced techniques for hyperspectral imaging in the food industry: Principles and recent applications. *Annu. Rev. Food Sci. Technol.* **2019**, *10*, 197–220. [CrossRef] [PubMed]
32. Zareef, M.; Chen, Q.; Hassan, M.M.; Arslan, M.; Hashim, M.M.; Ahmad, W.; Kutsanedzie, F.Y.; Agyekum, A.A. An overview on the applications of typical non-linear algorithms coupled with NIR spectroscopy in food analysis. *Food. Eng. Rev.* **2020**, *12*, 173–190. [CrossRef]
33. Srichamnong, W.; Srzednicki, G. Internal discoloration of various varieties of Macadamia nuts as influenced by enzymatic browning and Maillard reaction. *Sci. Hortic.* **2015**, *192*, 180–186. [CrossRef]
34. Martinez, M.; Wallace, H.M.; Searle, C.; Elliott, B.; Bai, S.H. Understanding chemical pathways of brown centre formation in laboratory induced and conventionally dried nut-in-shell macadamia kernels. *Heliyon* **2024**, *10*, e25221. [CrossRef]

35. Prichavudhi, K.; Yamamoto, H.Y. Effect of drying temperature on chemical composition and quality of macadamia nuts. *Food Technol.* **1965**, *19*, 1153–1156.
36. Kowitz, T.; Mason, R. Poor control over drying macadamia nut-in-shell on-farm causes abnormal kernel browning during roasting. In Proceedings of the 2nd International Macadamia Symposium, Tweed Heads, Australia, 29 September–4 October 2003; pp. 55–57.
37. DAF. *Macadamia Industry Benchmark Report. 2009 to 2021 Seasons*; Project MC18002; The State of Queensland: Brisbane, Australia, 2022.
38. Jones, K. *Brown Centres Ain'T Brown Centres*; Industry Information Bulletin; Cropwatch Independent Laboratories: Wardell, Australia, 2020.
39. Jones, K. *2022 Season Report*; Industry Information Bulletin; Cropwatch Independent Laboratories: Wardell, Australia, 2022.
40. End-To-End Processing. Available online: https://marquis.com/processing/our-process/ (accessed on 5 January 2024).
41. McConachie, I. *Investigations into Problems of Brown Centres in 1992*; Industry Information Bulletin; Australian Macadamia Society: Lismore, Australia, 1992.
42. Sanz, J.M.; Extremiana, C.; Saiz, J.M. Comprehensive polarimetric analysis of Spectralon white reflectance standard in a wide visible range. *Appl. Opt.* **2013**, *52*, 6051–6062. [CrossRef] [PubMed]
43. Wu, D.; Sun, D.W. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review—Part I: Fundamentals. *Innov. Food Sci. Emerg. Technol.* **2013**, *19*, 1–14. [CrossRef]
44. Appakaya, S.B.; Sankar, R.; Ra, I.-H. Classifier comparison for two distinct applications using same data. In Proceedings of the 9th International Conference on Smart Media and Applications, Jeju, Republic of Korea, 17–19 September 2020; pp. 429–432.
45. Yaman, O.; Yetiş, H.; Karaköse, M. Image processing and machine learning-based classification method for hyperspectral images. *J. Eng.* **2021**, *2021*, 85–96. [CrossRef]
46. The MathWorks Inc. *Statistics and Machine Learning Toolbox Documentation*; MathWorks: Natick, MA, USA, 2022.
47. ElMasry, G.; Kamruzzaman, M.; Sun, D.W.; Allen, P. Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: A review. *Crit. Rev. Food Sci. Nutr.* **2012**, *52*, 999–1023. [CrossRef]
48. Zhang, C.; Guo, C.; Liu, F.; Kong, W.; He, Y.; Lou, B. Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine. *J. Food Eng.* **2016**, *179*, 11–18. [CrossRef]
49. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [CrossRef] [PubMed]
50. Park, K.; Hong, Y.K.; Kim, G.H.; Lee, J. Classification of apple leaf conditions in hyper-spectral images for diagnosis of Marssonina blotch using mRMR and deep neural network. *Comput. Electron. Agric.* **2018**, *148*, 179–187. [CrossRef]
51. Bendel, N.; Backhaus, A.; Kicherer, A.; Köckerling, J.; Maixner, M.; Jarausch, B.; Biancu, S.; Klück, H.-C.; Seiffert, U.; Voegele, R.T. Detection of two different grapevine yellows in *Vitis vinifera* using hyperspectral imaging. *Remote Sens.* **2020**, *12*, 4151. [CrossRef]
52. Vujović, Ž. Classification model evaluation metrics. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 599–606. [CrossRef]
53. Ebrahimi, S.; Pourdarbani, R.; Sabzi, S.; Rohban, M.H.; Arribas, J.I. From harvest to market: Non-destructive bruise detection in kiwifruit using convolutional neural networks and hyperspectral imaging. *Horticulturae* **2023**, *9*, 936. [CrossRef]
54. Martinez, M.; Wallace, H.M.; Searle, C.; Elliott, B.; Hosseini Bai, S. Chemical differences between brown centre and white macadamia kernels. *J. Agric. Food Res.* **2023**, *14*, 100878. [CrossRef]
55. Chen, W.; Li, H.; Zhang, F.; Xiao, W.; Zhang, R.; Chen, Z.; Du, Y. Handheld short-wavelength NIR spectroscopy for rapid determination of sugars and carbohydrate in fresh juice with Sampling Error Profile Analysis. *Infrared Phys. Technol.* **2021**, *115*, 103732. [CrossRef]
56. Li, J.; Wang, Q.; Xu, L.; Tian, X.; Xia, Y.; Fan, S. Comparison and optimization of models for determination of sugar content in pear by portable Vis-NIR spectroscopy coupled with wavelength selection algorithm. *Food Anal. Methods* **2019**, *12*, 12–22. [CrossRef]
57. Zhang, D.; Xu, L.; Liang, D.; Xu, C.; Jin, X.; Weng, S. Fast prediction of sugar content in dangshan pear (*Pyrus* spp.) using hyperspectral imagery data. *Food Anal. Methods* **2018**, *11*, 2336–2345. [CrossRef]
58. Ghosh, P.K.; Jayas, D.S. Use of spectroscopic data for automation in food processing industry. *Sens. Instrum. Food Qual. Saf.* **2009**, *3*, 3–11. [CrossRef]
59. Bai, S.H.; Tahmasbian, I.; Zhou, J.; Nevenimo, T.; Hannet, G.; Walton, D.; Randall, B.; Gama, T.; Wallace, H.M. A non-destructive determination of peroxide values, total nitrogen and mineral nutrients in an edible tree nut using hyperspectral imaging. *Comput. Electron. Agric.* **2018**, *151*, 492–500. [CrossRef]
60. Rahman, A.; Wu, Q.; Chang, H.; Wang, S.; Yan, J.; Xu, H. Assessment of intact macadamia nut internal defects using near-infrared spectroscopy. In Proceedings of the 2020 ASABE Annual International Virtual Meeting, Virtual, 13–15 July 2020. [CrossRef]
61. De Carvalho, L.C.; Pereira, F.M.V.; de Morais, C.d.L.M.; de Lima, K.M.G.; de Almeida Teixeira, G.H. Assessment of macadamia kernel quality defects by means of near infrared spectroscopy (NIRS) and nuclear magnetic resonance (NMR). *Food Control* **2019**, *106*, 106695. [CrossRef]
62. AMS. *Kernel Quality Standard for Processors*; Version AMKQS–V1; Australian Macadamia Society: Lismore, Australia, 2018; pp. 1–6.
63. Golic, M.; Walsh, K.; Lawson, P. Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. *Appl. Spectrosc.* **2003**, *57*, 139–145. [CrossRef]
64. Nakariyakul, S.; Casasent, D.P. Classification of internally damaged almond nuts using hyperspectral imagery. *J. Food Eng.* **2011**, *103*, 62–67. [CrossRef]

65. Penter, M.; Nkwana, E.; Nxundu, Y.; Kruger, F. An investigation into the occurrence of kernel dicolouration in the 'Beaumont'cultivar. *S. Afr. Macadamia Grow. Assoc. Yearb.* **2007**, *15*, 9–12.
66. Huang, M.; Tang, J.; Yang, B.; Zhu, Q. Classification of maize seeds of different years based on hyperspectral imaging and model updating. *Comput. Electron. Agric.* **2016**, *122*, 139–145. [CrossRef]
67. Jiang, H.; Ye, L.; Li, X.; Shi, M. Variety identification of Chinese Walnuts using hyperspectral imaging combined with chemometrics. *Appl. Sci.* **2021**, *11*, 9124. [CrossRef]
68. Walton, D.A.; Wallace, H.M. Ultrastructure of Macadamia (*Proteaceae*) embryos: Implications for their breakage properties. *Ann. Bot.* **2005**, *96*, 981–988. [CrossRef]
69. Maestri, D. Groundnut and tree nuts: A comprehensive review on their lipid components, phytochemicals, and nutraceutical properties. *Crit. Rev. Food Sci. Nutr.* **2023**, *64*, 7426–7450. [CrossRef]
70. AMS. *Kernel Assessment Workshops, Lismore and Bundaberg, June 2022*; Australian Macadamia Society: Lismore, Australia, 2021.
71. Richards, T.E.; Kämper, W.; Trueman, S.J.; Wallace, H.M.; Ogbourne, S.M.; Brooks, P.R.; Nichols, J.; Hosseini Bai, S. Relationships between nut size, kernel quality, nutritional composition and levels of outcrossing in three macadamia cultivars. *Plants* **2020**, *9*, 228. [CrossRef]
72. The Jamovi Project. *Jamovi*, version 2.5; The Jamovi Provect: Sydney, Australia, 2024.
73. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.