Computers and Electronics in Agriculture 235 (2025) 110341

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Synthetic hyperspectral reflectance data augmentation by generative adversarial network to enhance grape maturity determination

Hongyi Lyu [a,*], Miles Grafton [a,*], Thiagarajah Ramilan [a], Matthew Irwin [a], Eduardo Sandoval [b]

[a] School of Agriculture and Environment, Massey University, Palmerston North 4410, New Zealand
[b] Massey Agri-Food (MAF) Digital Lab., Massey University, Palmerston North 4410, New Zealand

ARTICLE INFO

Keywords:
Hyperspectral imaging system
Grape maturity
Generative adversarial network
Deep learning

ABSTRACT

Non-destructive and rapid grape maturity detection is important for the wine industry. The ongoing development of hyperspectral imaging techniques and deep learning methods has greatly helped in non-destructive assessing of grape quality and maturity, but the performance of deep learning methods depends on the volume and the quality of labeled data for training. Building non-destructive grape quality or maturity testing datasets requires damaging grapes for chemical analysis to produce labels which are time consuming and resource intensive. To solve this problem, this study proposed a conditional Wasserstain Generative Adversarial Network (WGAN) with the gradient penalty data augmentation technique to generate synthetic hyperspectral reflectance data of two grape maturity categories (ripe and unripe) and different Total Soluble Solids (TSS) values. The conditional WGAN with the gradient penalty was trained for a range of epochs: 500, 1000, 2000, 8000, 10,000, and 20,000. After training of 10,000 epochs, synthetic hyperspectral reflectance data were very similar to real spectra for each maturity category and different TSS values. Thereafter, contextual deep three-dimensional CNN (3D-CNN), Spatial Residual Network (SSRN) and Support Vector Machine (SVM) are trained on original training and synthetic + original training datasets to classify grape maturity. The synthetic hyperspectral reflectance data, incrementally added to the original training set in steps of 250, 500, 1000, 1500, and 2000 samples, consistently resulted in higher model performance compared to training solely on the original dataset. The best results were achieved by augmenting the training dataset with 2000 synthetic samples and training with a 3D-CNN, yielding a classification accuracy of 91 % on the testing set. To better assess the effectiveness of GAN-based data augmentation methods, two widely used regression models: Partial Least Squares Regression (PLSR) and one-dimensional CNN (1D-CNN) were used based on same data augmentation method. The best result was achieved by adding 250 synthetic samples to the original training set when training 1D-CNN model, yielding an R^2 of 0.78, RMSE of 0.63 °Brix, and RPIQ of 3.36 on the testing set. This study indicated that deep learning models combined with conditional WGAN with the gradient penalty data augmentation technique had a good application prospect in the grape maturity assessment.

1. Introduction

New Zealand wineries producing high-quality wines that are exported globally. One critical factor in the production of high-quality wines is the maturity of the grapes used to make the wine. Bramley et al., (2003) showed that it's more profitable for wineries to classify the maturity of grapes to produce wines of different qualities than to mix grapes of different maturity to produce a single quality wine. Traditional methods for measuring grape maturity involves measuring grape Total Soluble Solids (TSS) based on refractometry or enzymatic tests (Rolle et al., 2022). These methods can provide accurate measurements, but are often labor-intensive, time-consuming, and require destructive sampling. However, the grape TSS value showed marked spatial variability within a block or in the same vine (Baluja et al., 2013). Using traditional destructive techniques, we only analyzed a small subset of TSS values and use this value to represent the grape maturity of an entire region. This results in uneven maturity of the grapes used for wine-making during the post-harvest stage. Therefore, it is desirable to develop a non-destructive method to effectively sort grape maturity during the post-harvest stage.

* Corresponding authors.
E-mail addresses: hlyu@massey.ac.nz (H. Lyu), m.grafton@massey.ac.nz (M. Grafton).

https://doi.org/10.1016/j.compag.2025.110341
Received 8 December 2024; Received in revised form 19 February 2025; Accepted 23 March 2025
Available online 3 April 2025

The development of image technologies, e.g., monochromatic, red–green–blue (RGB), multispectral and hyperspectral imaging for non-destructive assessing of fruit quality or maturity has advanced rapidly over the last decades (Li et al., 2019; Lu et al., 2020, 2022). Among these, hyperspectral imaging technology can acquire spatial and spectral information simultaneously (Sun, 2010). Compared with other imaging technologies, hyperspectral imaging can detect subtle changes in the spectral reflectance of grape berries, which are closely related to their TSS values. As grapes mature, their chemical composition, such as sugar content, water content, and anthocyanin concentration, changes, leading to variations in spectral reflectance, especially in the visible and near-infrared (VNIR) regions (Tsakiridis et al., 2023; Lyu et al., 2024). Compared to traditional RGB or multispectral imaging technologies, HSI captures subtle spectral variations, making it a powerful tool for precise quality and maturity assessment. Currently, hyperspectral bands in the range of 400–1000 nm are relatively easy to obtain due to advancements in sensor technology, and the reduced cost of equipment has facilitated its deployment in industrial applications. However, for other spectral ranges, such as near-infrared (1000–2500 nm) or far-infrared bands, the cost of acquisition remains high, posing a significant barrier to widespread adoption (Gomes et al., 2017). Additionally, hyperspectral imaging generates high-dimensional data, requiring significant storage, transmission, and computational resources, as well as advanced data processing algorithms.

Integrated with image technologies, deep learning improved efficacy and efficiency in grading and sorting grape berries. Deep learning, especially convolutional neural networks (CNNs) are widely being used in data-driven visual recognition (e.g., image classification, semantic segmentation and object detection) tasks in agricultural production (Koirala et al., 2019a; Wang et al., 2021). Various architectures of CNNs such as AlexNet, VGG, GoogLeNet, DenseNet, and Inceptionv3, are being used to classify fruit maturity status based on RGB images (Zhao et al., 2023). Das and Yadav, (2020) proposed a revised version of AlexNet to classify tomato maturity. Ashtiani et al., (2021) compared the accuracy of five different CNNs models to classify the maturity of white and black mulberries. The result showed that AlexNet and ResNet-18 models exhibited the best accuracy score for white and black mulberries, respectively. Additionally, Ramos et al., (2021) used VGG-19 to classify the maturity status of Syrah and Cabernet Sauvignon grapes. The result showed that VGG-19 achieved maturity status classification accuracy of 91.30 % and 80.97 % for Syrah and Cabernet Sauvignon, respectively. Wei et al., (2022) used back-propagation neural network (BPNN) and RGB images to predict greenhouse grape maturity in solar greenhouse, achieving up to 88 % accuracy by combing color and physical–chemical indicators. Tsakiridis et al., (2023) proposed two different autoencoder architectures to transform raw spectra into standardized reflectance spectra, overcoming issues with varying illumination, when using hyperspectral imaging predict grape quality. The result showed deep convolutional autoencoder outperformed deep fully connected autoencoder, with a mean R^2 of 0.70. Benelli et al., (2021) used an on-the-go hyperspectral imaging system and Partial Least Squares Regression (PLSR) model to predict grape maturity, achieving good prediction accuracy, with R^2 values of 0.77 (RMSE = 0.79 °Brix) for TSS.

One of the challenges using deep learning for visual recognition tasks is the need for extensive and diverse datasets to train CNNs models effectively (Abbas et al., 2021). Deep learning models usually need to train several million parameters to capture sufficient variations needed for vision recognition tasks (Miranda et al., 2023). Models like You Only Look Once (YOLO) come pretrained on large, open datasets such as COCO (Common Objects in Context) or ImageNet to improve their effectiveness for general object detection tasks involving the types of objects they were trained on (Koirala et al., 2019a; Wang et al., 2019). However, when the application involves detecting objects or patterns that are not well-represented in these standard datasets, using a pretrained YOLO model may not yield optimal results. In this case, deep learning models only have high performance and generalizability on sufficiently large and diverse datasets. For example, (Koirala et al., 2019b) used 1515 images of trees from different orchards to create the MangoYOLO models to detect the orchard fruit. Additionally, Zhao et al., (2023) used 3849 images from a "Nine Peach" dataset to train different segmentation models to determine peach maturity. However, creating extensive and diverse datasets needs tremendous effort to ground truth data collection and labelling which is time consuming and resource intensive (Lu and Young, 2020). This is especially relevant for non-destructive fruit quality or maturity testing datasets. Labeling often necessitates damaging the fruit to measure internal qualities. To mitigate the scarcity of datasets, one common approach is data augmentation (Gour et al., 2020; Khalifa et al., 2022).

There are two categories of data augmentation approaches, including traditional image augmentation and augmentation based on deep learning methods. The traditional image augmentation include geometric (e.g., flipping and cropping) and color processing approaches (Shorten and Khoshgoftaar, 2019). However, methods such as cropping remove critical context from an image, making it harder for the model to learn the relationships between different parts of the image. Basic augmentation techniques often follow predictable patterns (e.g., fixed rotation angles, specific crops). This can lead to a lack of diversity in the training data, limiting the model's exposure to truly novel examples. Additionally, traditional image augmentation can only create limited variations of existing data and is unable to learn the variations or invariant features across the samples (Chen et al., 2024). Recently, Generative Adversarial Networks (GANs) have been used for data augmentation to solve the drawback of traditional approaches in the agricultural community. The Generative Adversarial Network (GAN), originally proposed by Goodfellow et al., (2014) was used to synthesize new data with the same characteristics of training instances. GANs generate high-quality, realistic synthetic data that retain the statistical properties of the original dataset, enhancing the diversity without compromising the data quality (Shorten and Khoshgoftaar, 2019). Compared with traditional data augmentation techniques, GANs offer several additional advantages. They can model complex data distributions, generating novel samples that reflect subtle variations in the data, which are often beyond the reach of basic augmentation methods (Shorten and Khoshgoftaar, 2019). This is particularly valuable in situations where data is scarce or imbalanced, as GANs can produce synthetic samples for underrepresented classes, improving model performance. Furthermore, GANs can generate data that mimics domain-specific features, enabling more effective domain adaptation and style transfer. These strengths make GANs a powerful tool for augmenting data in deep learning tasks, especially in fields like medical imaging, remote sensing, and object recognition, where data acquisition can be costly or limited (Chen et al., 2024).

Several studies have used GANs as a data augmentation technique to synthesize fruit images in post-harvest quality assessment tasks (Guo et al., 2021; Bird et al., 2022; Miranda et al., 2023; Tan et al., 2024). For example, Bird et al., (2022) used Conditional GAN to produce synthetic lemon images for fruit quality image classification. When using synthetic fruit images in VGG-16 models, the classification performance is higher than that without data augmentation. However, most of the previous studies applied GANs to generate synthetic RGB or monochromatic fruit images. Currently, the potential of GANs in fruit hyperspectral image generation remains largely unexplored. Compared with RGB images, hyperspectral imaging can provide detailed spectral information for accurate substance identification, classification, and non-destructive internal quality assessment (Sun, 2010).

This study aims to develop a hyperspectral reflectance data augmentation technique based on GAN network to improve the deep learning model performance both grape maturity classification (e.g., unripe, ripe) and TSS content estimation as a continuous variable. This dual evaluation aims to support industrial producers in choosing the appropriate approach based on their specific operational requirements, whether it is the simplicity of a classification system or the precision of a

regression model. The specific objectives of this research are:

- Using different TSS values or maturity labels to generate corresponding synthetic grape hyperspectral reflectance data based on GAN network.
- Combining the generated synthetic data with the original dataset to evaluate whether GAN-based data augmentation can improve the performance of different models

2. Methodology

2.1. Data collection

The samples used in this study were collected from Palliser Estate Winery located in Martinborough, New Zealand (Lyu et al., 2023). The winegrape variety used in this study was 'Pinot Noir', and compared to other varieties, the maturity status during the harvest stage greatly affects the value of subsequent wines. The 'Pinot Noir' were planted in an organic vineyard named Hua Nui. The vines were planted in 1998, with vine spacing of 1.7 m and row spacing of 2.2 m. A total of 231 samples were collected from a 2.6 ha block in Hua Nui vineyard (Fig. 1). A total of 77 samples were collected, weekly from March 01 to March 14, 2024, during the harvest stage. To acquire representative samples, the sampling locations were collected from a predetermined nested grid. Each grid had locations taken at intervals of 4 m, 8 m, 16 m and 32 m. During each measurement, a healthy vine closest to the sampling location was selected for sampling. In each sampling vine, three bunches were selected for grape berry collection. The selected bunches were divided into three positions: top, middle, and bottom. From each position, one berry was handpicked resulting in a total of nine berries per sampling vine. It is worth noting that nine grape berries per vine met the minimum criteria for using the digital refractometer (PAL-BX/ACID F5 Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan) to measure TSS. The nine grape berries were placed in plastic bags with labels, and spectrum measurements were immediately taken after sampling.

2.2. Hyperspectral image acquisition

The reflectance data of nine grapes were acquired from a push-broom VNIR hyperspectral imaging system which utilize 1800 pixels to collect 186 channels of spectral data from 406.8 to 995.8 nm. The VNIR hyperspectral imaging system is composed of the following constituents: two 150 W customized direct current linear light source (HySpex, Norsk Elektro Optikk, Norway), a spectra camera (HySpex VNIR-1800, Norsk Elektro Optikk, Norway), laboratory rack, a white reference tile made of aluminum oxide (SphereOptics GmbH, Germany), and a computer (Fig. 2). The spectra sampling interval of HySpex VNIR-1800 is 3.26 nm with a 17° field of view. The hyperspectral imaging system was turned on for 30 min before the image acquisition to increase the spectral data quality. The height between the camera and samples was 20.4 cm. HySpex Ground v4 12.7 software was used to process image acquisition within a dark room at room temperature. The frame period was 11,200 us, integration time was 11,085 us, and the moving speed of the conveyor belt was 5 cm/s. According to the measurement protocol, the white reference was measured prior to each data acquisition session to ensure reliable calibration and consistency of hyperspectral measurements throughout the experiment. During each measurement, nine grape berries were squared on the conveyor belt, pedicel facing the white reference. After acquiring hyperspectral images, the software HySpex Rad v3.4 was used to obtain the reflectance data of the image. Subsequently, the region of interest (ROI) tool was used in the Environment for Visualizing Images (ENVI) V 5.6 software (Research Systems Inc., Boulder, Co, USA) to manually separate the grape berries from the background. After extracting the berry spectra, Savitzky-Golay smoothing (window size = 10, polynomial = 2) was applied to reduce noise and enhance spectral features.

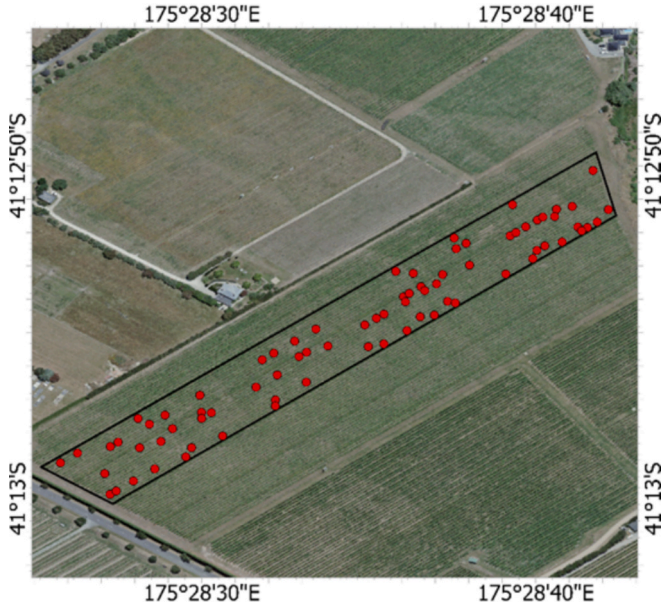

Fig. 1. Sampling location in Hua Nui vineyard (red points represent each sampling point).

2.3. Determination of total soluble solids

After hyperspectral image acquisition, the samples were stored in the fridge at 4 °C until their analysis for grape total soluble solids. Before each measurement, the samples were stabilized to the room temperature at around 25 °C. During each measurement, nine grape berries were squeezed together into juice Then, 0.2 ml of the juice was applied onto a portable digital refractometer (PAL-BX/ACID F5 Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan). The TSS value (°Brix) was a direct reading from the refractometer with a measuring range between 0 and 60 % (0.1 % resolution). For each sample, the TSS value was measured twice. The average of these two measurements was used to represent the TSS value of the nine grapes.

2.4. Synthetic hyperspectral reflectance data generation via GAN model

In order to improve the classification model performance, conditional Wasserstein GAN (WGAN) with the gradient penalty was used as the hyperspectral imaging data augmentation technique to increase the

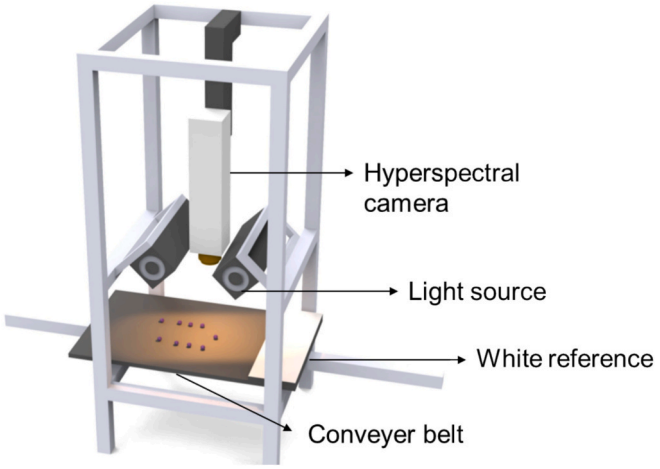

Fig. 2. The hyperspectral imaging system used in this study.

size of the dataset. Traditional GAN consists of a generator and discriminator model which are trained in opposition to each other (Goodfellow et al., 2014). The work of the generator is to take random noise data as inputs and produce fake data. The work of discriminator is to receive the fake image from the generator and real data and distinguish whether the data are synthetic or real. The generator and discriminator train simultaneously and try to outdo each other. The discriminator makes sure that the data generated by the generator are as close to the real images as possible. The generator aims to produce data that the discriminator cannot distinguish from real data, while the discriminator aims to improve its accuracy in distinguishing real from synthetic data. This adversarial process ensures that the data generated by the generator becomes increasingly similar to the real images.

Compared with traditional GAN, WGAN use Wasserstein distance metric as its loss function. In WGAN, the discriminator is replaced by 'critic' science, it evaluates the quality of generated data by assigning them a score rather than classifying them as real or fake (Gulrajani et al., 2017). Conditional WGAN extends WGAN by incorporating additional information (e.g., class labels) into the network. Both the generator and critic are conditioned on this additional information, allowing the generator to produce class-specific data. This study introduces an additional classifier model in the network (Fig. 3). The classifier ensures that the generated data not only appears realistic but also adheres to the given conditional labels (Audebert et al., 2018). This setup helps in aligning the generated data distribution more closely with the conditional labels provided. The objective function for the generator and critic, including the classifier, is:

$$\min_G \max_D \Big(E_{x\ P_{data}}[D(x,c)] - E_{z\ P_z, c\ P_c}[D(G(z,c),c)] + \lambda E_{\hat{x}\ P_{\hat{x}}}\left[\left(\|\nabla_{\hat{x}} D(\hat{x},c)\|_2 - 1 \right)^2 \right] \Big)$$

where x is real data, P_{data} is the data distribution, $D(x,c)$ is the critic's score for the real data x conditioned on c. z is noise sample, P_z is the noise distribution, c is conditional label, P_c is label distribution, $G(z,c)$ is fake data from the generator using noise z and condition c. $D(G(z,c),c)$ is critic's score for the fake data $G(z,c)$ conditioned on c. λ is coefficient for the gradient penalty term, $\hat{x}$ are samples interpolated between real data x and fake data $G(z,c)$, P_x is the distribution of these interpolated samples, $D(\hat{x},c)$ is critic's score for interpolated samples $\hat{x}$ conditioned on c.

The conditional WGAN with the gradient penalty generator model consists of several dense layers followed by Leaky rectified linear unit (ReLU) activation functions and a final Sigmoid activation function. The model begins with an input layer that takes in the noise sample and condition vectors. This is followed by a sequence of three dense layers, each followed by a Leaky ReLU activation layer. The final dense layer is followed by a Sigmoid activation layer to produce smooth outputs in the range [0,1]. The critic and classifier model in the conditional WGAN with the gradient penalty have similar architecture, including an input layer, three dense layers followed by Leaky ReLU activation functions. The final dense layer in critic model outputs a single value, which indicates the authenticity of the input data. The final dense layer in the classifier model outputs a vector with a size equal to the number of classes, which represents the classification scores for each class. The model applies a weight initialization function to all layers, following the He et al., (2016) policy, which initializes weights using the Kaiming normal distribution. In addition, we also tested Random Initialization methods, to compare their effects on the model's performance. However, the results showed that the Kaiming initialization method outperformed the others in terms of model performance He et al., (2016).

The code and hyperparameter setting of synthetic hyperspectral reflectance data generation via GAN model are available in Table 1 and https://github.com/hlyu821/GANs. In this study, the generated hyperspectral reflectance data was evaluated based on Mahalanobis Distance (MD) and visual analysis.

Table 1
The hyperparameter setting of synthetic hyperspectral reflectance data generation based on WGAN model.

Hyperparameter	Value
Batch Size	32
Learning rate	0.00005
Alpha of Leaky ReLU	0.01
Noise dimension	30
Noise Type	Random
Number of features	186
Hidden layer size	512
Discriminator training steps per generator training step	5
Auxiliary classifier weight	0.2
Mixup	0.1

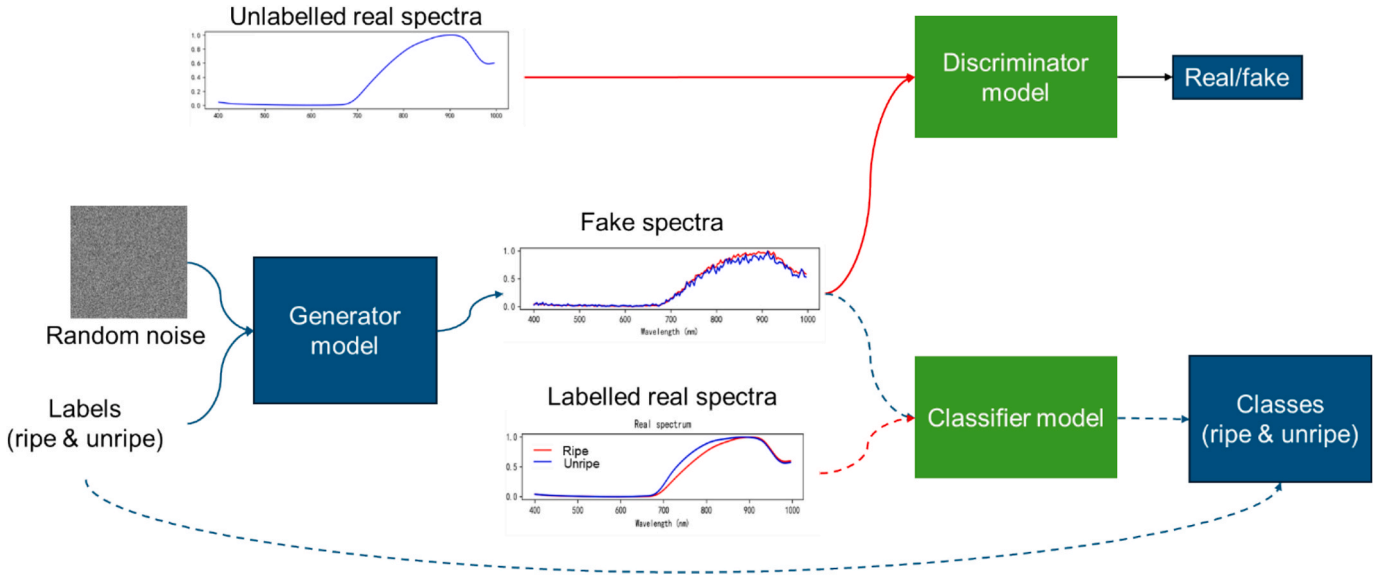

Fig. 3. The proposed conditional WGAN with the gradient penalty. Red lines represent the training of the discriminator; blue lines represent the training of the generator; red and blue dashed lines represent the training of the classifier.

2.5. Grape maturity classification with different CNNs models

As this study aimed to assess the effectiveness of GAN-based data augmentation methods in enhancing grape maturity determination performance, two widely used hyperspectral imaging classification CNN models: contextual deep three-dimensional CNN (3D-CNN) and Spatial Residual Network (SSRN) were used. 3D-CNN is an end-to-end fully convolutional network proposed by Lee and Kwon, (2017). The 3D-CNN uses a multi-scale convolutional filter bank. This multi-scale approach captures features at various spatial resolutions to analysis both spectral and spatial information in the hyperspectral data (Fig. 4). The SSRN uses 3-D convolutional layers to process the hyperspectral data cubes, incorporating both spectral and spatial information effectively (Zhong et al., 2017). SSRN incorporates spectral and spatial residual blocks to improve training efficiency and performance (Fig. 5). Spectral residual blocks use $1 \times 1 \times m$ convolutions, focusing on spectral features, while spatial residual blocks use convolutions that maintain spatial dimensions to extract spatial features. To regularize the training process and reduce overfitting, batch normalization and dropout techniques are applied within the network. Tables S1 and S2 show the hyperparameter setting of 3D-CNN and SSRN. In addition, one traditional machine learning model Support Vector Machine (SVM) with radial kernel were used to assess the effectiveness of GAN-based data augmentation methods.

The samples were divided into two groups (unripe and ripe) separated by threshold limiting values: 21 °Brix for TSS, which is based on a local viticulturist suggestion. In this study, two sets of experiments have been performed. The original samples (n = 231) were randomly split into training and testing sets using an 80/20 ratio. The training set (n = 184) was used for generating synthetic datasets and optimize classification models' hyperparameters. First of all, the conditional WGAN with the gradient penalty has been trained for 500, 1000, 2000, 8000, 10,000, and 20,000 epochs to generate synthetic hyperspectral grape reflectance data of two categories (unripe and ripe). After successful completion of the training of conditional WGAN with the gradient penalty, 1000 synthetic hyperspectral data of grape berries were generated for each category.

In the second step, the hyperspectral reflectance data were transformed to a 3D cubic data and input in the CNNs model. Then the classification model (3D-CNN, SSRN and SVM) has been trained on the training set and synthetic dataset. To systematically evaluate the impact of data augmentation on classifier performance, the original training set (n = 184) is used as a reference baseline to train classifiers. This ensures that any observed improvements can be attributed to the addition of synthetic data rather than other confounding factors. Initially, classifiers are trained exclusively on synthetic data to assess their standalone effectiveness. Subsequently, synthetic samples are incrementally incorporated into the original training dataset (n = 184) in steps of, 250, 500, 1000, 1500, and 2000. This stepwise approach allows for a controlled analysis of how different quantities of synthetic data influence model performance, enabling the identification of an optimal augmentation level. A 5-fold cross-validation was applied on the training set to tune model hyperparameters, ensuring that the evaluation process was robust. The testing set (n = 47) was used to evaluate the performance based on the Kappa coefficient, accuracy (Acc), recall, and precision which were calculated from the confusion matrices of each model. The calculation of Acc, recall, and precision was defined bellow:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

where *TP*, *FN*, *FP*, and *TN* are the true positive, false negative, false positive, true negative.

The model in this work were implemented with PyTorch (version 1.8) and CUDA 11.1 and trained on an RTX 2080Ti GPU (4352 CUDA cores).

2.6. Grape maturity prediction with different regression models

In order to better assess the effectiveness of GAN-based data augmentation methods in enhancing grape maturity determination performance, two widely used regression models: Partial Least Squares Regression (PLSR) and one-dimensional CNN (1D-CNN) were used. PLSR is the dominant chemometric regression approach for non-invasive detection of fruit attribute level (Walsh et al., 2020a). PLSR projects the spectra input into a few latent variables that capture maximum covariance. PLSR handles highly collinear and noisy data by projecting both predictors and responses into a new space. 1D-CNN is a widely used deep learning architecture to process hyperspectral data in a regression task. 1D-CNN automatically extracts relevant features from raw spectral data by applying convolutional filters along the wavelength axis (Pullanagari et al., 2021). The 1D-CNN typically includes an input layer that processes the spectra data, followed by hidden layers comprising convolutional layers that extract features, pooling layers that reduce dimensionality, and fully connected layers that integrate the extracted features. The final output layer produces predictions based on the learned features. In this study a custom 1D-CNN architecture was proposed. 1D-CNN features a flexible architecture with multiple convolutional layers, each consisting of a 1D convolution, ReLU activation, and max-pooling. These layers extract and downsample features from the input data. The output is flattened and passed through a series of fully connected layers with ReLU activation and dropout for regularization. Finally, a linear layer produces output. This study used the cosine annealing learning rate scheduler to adjust the learning rate during training. Additionally, an open-source hyperparameter optimization framework based on Bayesian optimization (using the Optuna library) was employed to automate the search for optimal hyperparameters in the 1D-CNN model. This framework adaptively samples the hyperparameter space by balancing exploration and exploitation, aiming to minimize the model's validation loss. The code and hyperparameter setting of 1D-CNN are available in Table S3 and https://github.com/hlyu821/GANs.

In this study, two sets of experiments have been performed to assess the effectiveness of GAN-based data augmentation methods in regression models. The original samples (n = 231) were randomly split into training and testing sets using an 80/20 ratio. The training set (n = 184) was used for generating synthetic datasets and optimizing regression models' hyperparameters. First of all, the conditional WGAN with the

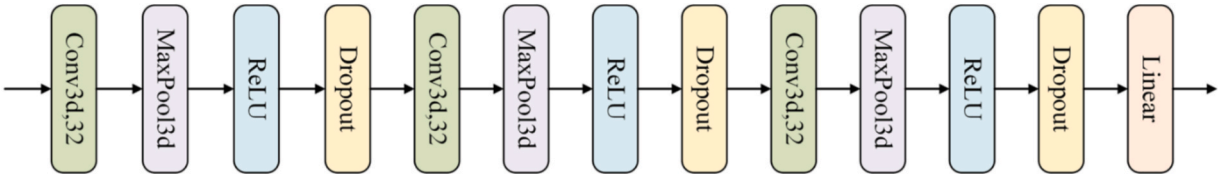

Fig. 4. The network architecture of 3D-CNN.

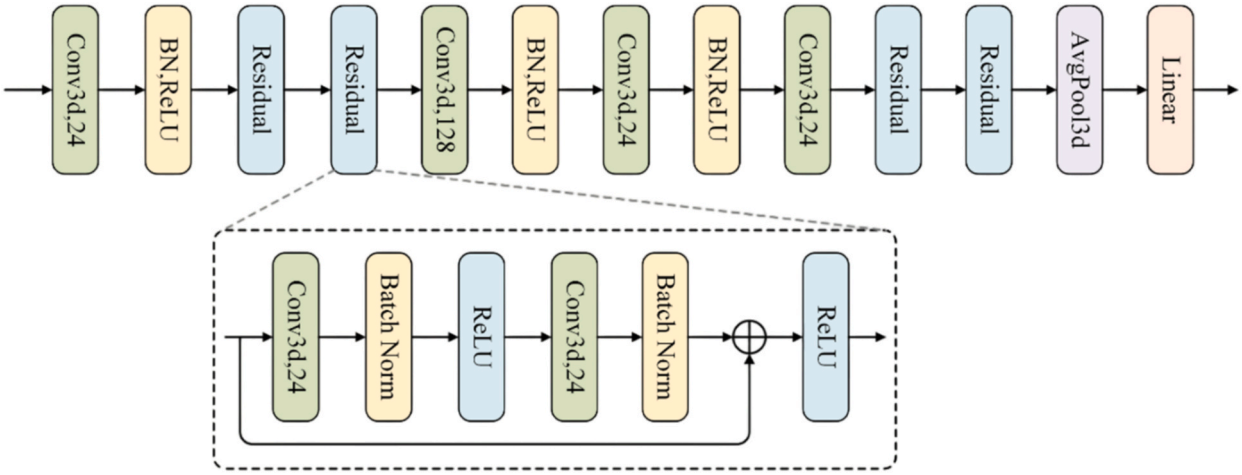

Fig. 5. The network architecture of SSRN.

gradient penalty has been trained for 500, 1000, 2000, 8000, 10,000, and 20,000 epochs to generate synthetic hyperspectral grape reflectance data of different grape TSS values. After successful completion of the training of conditional WGAN with the gradient penalty, 40 synthetic hyperspectral data of grape berries were generated for each grape TSS values.

In the second step, the regression model (PLSR and 1D-CNN) has been trained on the training set and synthetic dataset. In order to compare the performance of each model before and after data augmentation, the original training set (n = 184) was defined as a benchmark. Initially, regression models are trained exclusively in synthetic data to assess their standalone effectiveness. Subsequently, synthetic samples are incrementally incorporated into the original training dataset (n = 184) in steps of 250, 500, 1000, 1500, and 2000. A 5-fold cross-validation was applied on the training set to tune model hyperparameters, ensuring that the evaluation process was robust. The testing set (n = 47) was used to evaluate the performance based on the values of RMSE, ratio of performance to interquartile range (RPIQ) and R^2. The calculation of RMSE, RPIQ and R^2:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

$$RPIQ = \frac{RMSE}{Q_3 - Q_1}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where n is the number of samples used to fit the model, y_i is the ground truth value of the i th sample, $\bar{y}$ is the mean response value, $\hat{y}_i$ is the model estimated value of the i th sample, Q_3 is the upper quartile, Q_1 is the lower quartile.

3. Result

3.1. Basic statistical analysis of measured TSS and hyperspectral reflectance data

The basic statistical description (i.e., sampling date, mean, maximum, minimum, stand deviation, coefficient of variation and variance) of the TSS value is shown in Table 2. The TSS value of grapes increases closer to the harvest date. The reflectance curves for 'Pinot Noir' samples across the wavelength range of 406.8–995.8 nm were shown in Fig. 6. The spectrum exhibits lower reflectance in visible regions (400–700 nm), which is primarily influenced by pigments such as chlorophylls, carotenoids and anthocyanins (Nagy et al., 2016). There is a sharp increase in reflectance in the red edge region (700–750 nm) which is related to chlorophyll absorption and internal cellular structure. High reflectance in NIR (750–950 nm) is influenced by berry structure, water content, sugars and other organic compounds (e.g., TSS), The slight dips near 970 nm is related O—H bond vibration (Walsh et al., 2020).

3.2. Performance of synthetic hyperspectral reflectance data

After training from the conditional WGAN with the gradient penalty, this study generates 1000 hyperspectral reflectance data for each category (unripe and ripe) and 40 hyperspectral reflectance data for each TSS value. To evaluate the quality of synthetic data, Figs. 7 and 8 shows comparison results of real spectrum and synthetic data spectrum for different maturity classes and four random grape TSS values, respectively. Fig. 7 shows that when the epoch = 500, the synthetic data spectrum was random noise. As the number of epochs grows to 2000 epoch, the synthetic data spectrum gradually smoothed, while the synthetic data spectrum cannot show the characteristics of the different maturity classes. Finally, it could be seen that after epoch = 10,000 the synthetic data spectrum of different maturity classes show difference in the Red Edge and NIR region, consistent with the real spectrum. Table 3 shows the MD values corresponding to different categories (unripe and ripe) across different epochs. MD decreases significantly from 500 to 10,000 epochs, indicating improved similarity between generated and

Table 2
Descriptive statistics for TSS value of Pinot Nor during different sampling date (where SD = stand deviation, CV = coefficient of variation).

Date	Mean (°Brix)	Maximum(°Brix)	Minimum(°Brix)	SD	CV	Variance
01 March	19.48	21.3	18	0.71	4 %	0.5
06 March	20.67	23.1	19	0.83	4 %	0.67
14 March	22.06	23.9	20	0.76	3 %	0.57

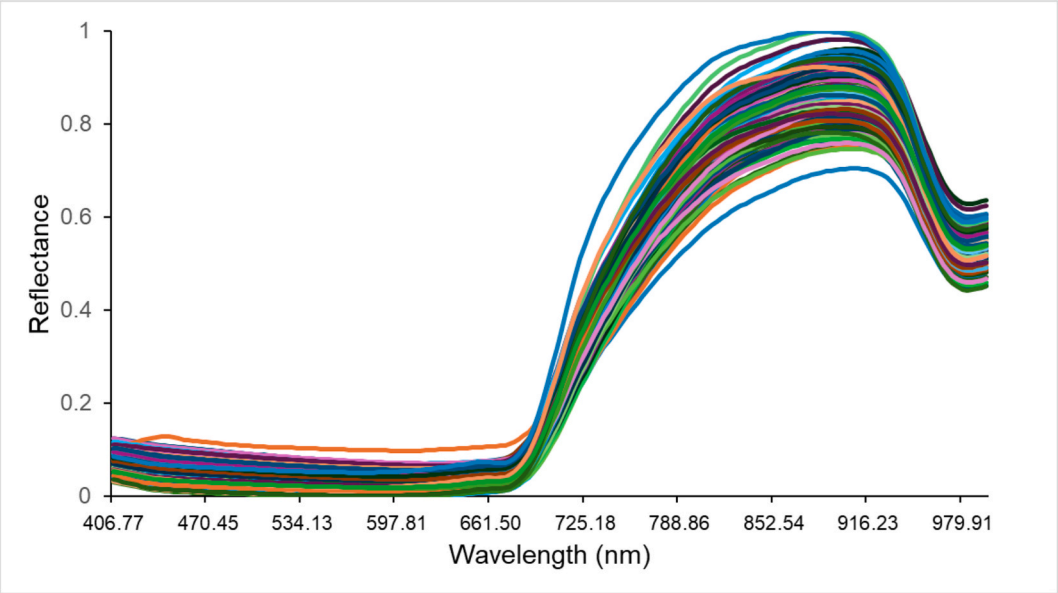

Fig. 6. Reflectance spectra of the 231 'Pinot Noir' samples collected based on HySpex VNIR hyperspectral imaging system. Color of each spectrum corresponding to TSS value.

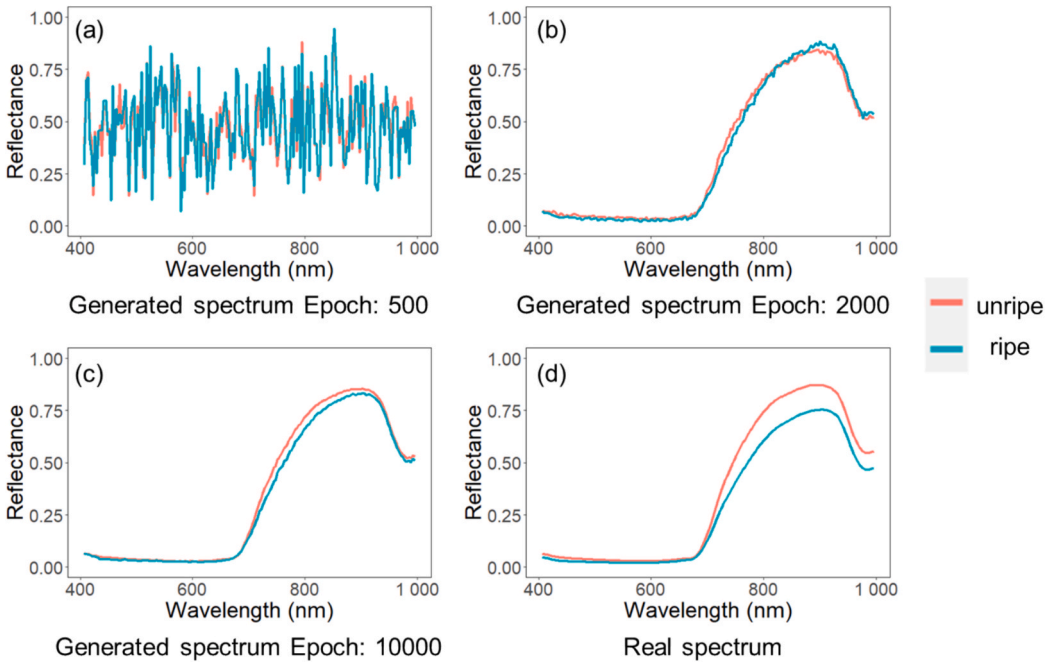

Fig. 7. Spectrum of different grape maturity generated by the conditional WGAN with the gradient penalty when epoch was 1000 (a); 2000 (b); 10,000 (c) and the real spectrum (d).

real data. After 10,000 epochs, MD starts increasing, suggesting potential overfitting or instability. Therefore, in subsequent classification model training, this study used hyperspectral reflectance data generated at epoch = 10,000 to augment the original data.

When using the conditional WGAN with the gradient penalty generate different TSS values, the synesthetic data spectrum was random noise when epoch = 1000 (Fig. 8a). As the number of epochs grows to 8000, the synthetic data spectrum produces a spectral curve for grape berries with large noise values and cannot show the characteristics of the different TSS values (Fig. 8c). Finally, it can be observed that at epoch 10,000, the synthetic data exhibit spectral characteristics corresponding to different TSS values in the NIR region, which are consistent with those of the real spectra (Fig. 8d, f). However, when epoch = 20,000, the synthetic data spectrum cannot show the characteristics of the different between 19.4 and 19.7 °Birx TSS value (Fig. 8e). One possible reason is that, after 10,000 epochs, the WGAN model might have overfitted to the spectral patterns associated with different TSS values. Fig. 9 presents the MD values corresponding to different TSS values across various epochs. The MD values were high for all TSS values at epochs 500 and 1000, suggesting that the model initially struggled to generate realistic data. After 1000 epochs, the MD values began to decrease across all TSS values, indicating an improvement in the model's ability to generate realistic data. However, after 10,000 epochs, the MD values started increasing, suggesting potential overfitting or

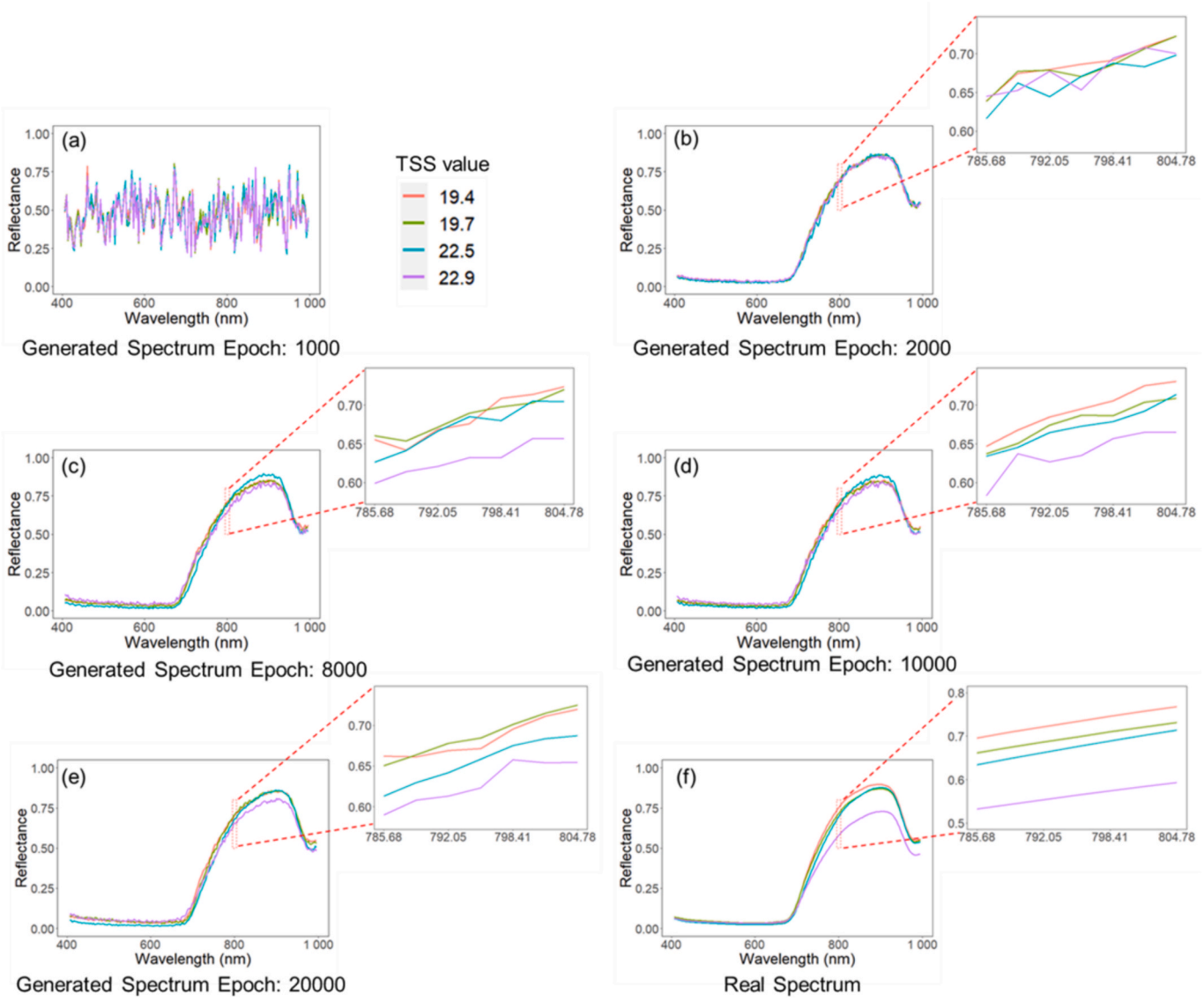

Fig. 8. Spectrum of four random grape TSS values generated by the conditional WGAN with the gradient penalty when epoch was 1000 (a); 2000 (b); 8000 (c); 10,000 (d); 20,000 (e) and real spectrum (f).

Table 3
The Mahalanobis Distance (MD) results for the generated hyperspectral reflectance data across different epochs.

Epoch Class	500	1000	2000	5000	8000	10,000	20,000
Ripe	369484.1	428858.6	12547.94	4402.58	3341.48	2456.98	3545.77
Unripe	612687.9	936421.2	17635.98	4764.43	4021.2	3061.07	7974.58

instability. Therefore, in subsequent regression model training, this study used hyperspectral reflectance data generated at epoch = 10,000 to augment the original data. Among all TSS values, those between 20 and 21 consistently exhibited higher MD values, indicating greater difficulty in accurate generation.

3.3. Performance of classification model

This study evaluated the performance of a proposed model for binary classification (unripe and ripe) task. To assess the impact of data augmentation on three common hyperspectral imaging classification models' (SSRN, 3D-CNN and SVM) performance, classifiers are first trained on the original training set as a baseline. When trained on the original training set, the 3D-CNN, SSRN, and SVM achieved classification accuracies of 79 %, 77 %, and 74 %, respectively, on the testing set (Tables 4-7). In addition, classification models are trained exclusively on synthetic data to assess their standalone effectiveness. However, classification models trained exclusively on synthetic data showed poor performance on the testing set (Tables 4-7), failing to effectively distinguish between the maturity levels of grapes. Subsequently synthetic samples are incrementally added to the original training set in steps of 250, 500, 1000, 1500, and 2000 to measure their contribution. The results from different classifiers along with augmentation approaches can be observed in Tables 4-6. All methods of adding synthetic samples resulted in higher accuracy compared to training solely on original training set, demonstrating that WGAN augmentation has had a positive effect on the learning process. The best results were achieved by augmenting the dataset with 2000 synthetic samples and

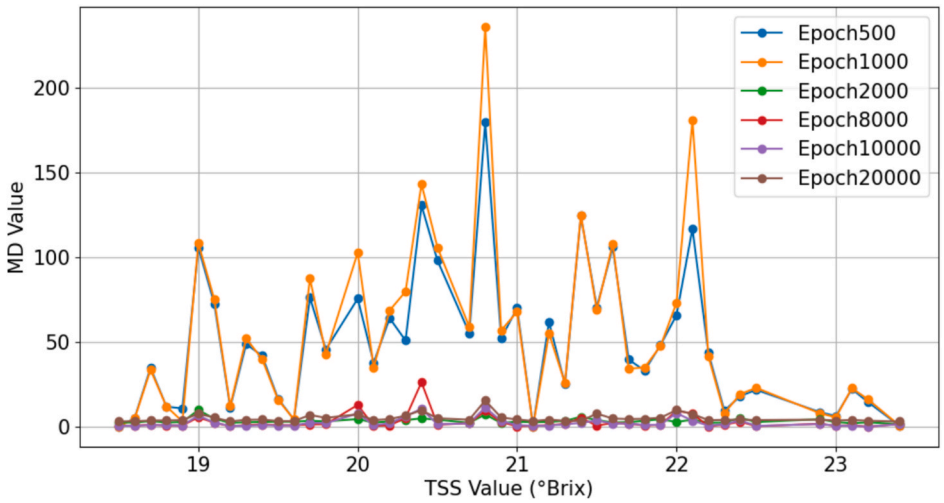

Fig. 9. The Mahalanobis Distance (MD) results of the generated hyperspectral reflectance data corresponding to different TSS values across different epochs.

Table 4
Classification performance of SSRN on test set.

Original Data	Synthetic Data	Accuracy	Precision	Recall	Kappa Coefficient
184	0	0.77	0.77	0.75	0.51
184	250	0.85	0.87	0.83	0.69
184	500	0.83	0.84	0.81	0.64
184	1000	0.81	0.82	0.79	0.6
184	1500	0.89	0.89	0.89	0.78
184	2000	0.81	0.84	0.78	0.59
0	2000	0.74	0.76	0.72	0.46

Table 5
Classification performance of 3D-CNN on test set.

Size of Original Data	Size of synthetic Data	Accuracy	Precision	Recall	Kappa Coefficient
184	0	0.79	0.83	0.76	0.54
184	250	0.85	0.87	0.87	0.71
184	500	0.81	0.84	0.78	0.59
184	1000	0.87	0.88	0.89	0.75
184	1500	0.85	0.87	0.87	0.71
184	2000	0.91	0.92	0.93	0.83
0	2000	0.74	0.75	0.75	0.49

Table 6
Classification performance of SVM on test set.

Size of Original Data	Size of synthetic Data	Accuracy	Precision	Recall	Kappa Coefficient
184	0	0.74	0.75	0.75	0.49
184	250	0.79	0.78	0.78	0.56
184	500	0.74	0.75	0.75	0.49
184	1000	0.74	0.75	0.75	0.49
184	1500	0.81	0.84	0.78	0.59
184	2000	0.79	0.78	0.78	0.56
0	2000	0.72	0.72	0.73	0.45

training with a 3D-CNN, yielding a classification accuracy of 91 % on the testing set. In deep learning models (SSRN and 3D-CNN), although adding different numbers of synthetic samples resulted in varying classification accuracies, it is noteworthy that even the weakest augmentation approach (when 500 synthetic samples were added to the 3D-CNN) increased the classification accuracy ability from 79 % to 81 %.

Table 7
Regression performance of PLSR on test set.

Size of Original Data	Size of synthetic Data	R^2	RMSE	RPIQ
184	0	0.61	0.84	2.49
184	250	0.61	0.84	2.5
184	500	0.66	0.79	2.67
184	1000	0.66	0.79	2.67
184	1500	0.67	0.77	2.72
184	2000	0.68	0.76	2.76
0	2000	0.24	1.18	1.79

The result show that SSRN, 3D-CNN and SVM have better classification performance after adding synthetic hyperspectral reflectance data than using the original dataset only.

3.4. *Performance of regression model*

This study evaluated the performance the effectiveness of GAN-based data augmentation methods for regression tasks. Tables 7 and 8 represent the regression performance of the proposed method on the original dataset and augmented dataset with different regression models. The regression performance on the original dataset was defined as the benchmark to compare the impact of GAN data augmentation techniques on model performance. In this study, two common regression models (PLSR and 1D-CNN) were used to assess the GAN data augmentation techniques. When trained on the original training set, PLSR and 1D-CNN achieved moderate regression performance, with R^2 values of 0.61 and 0.73, RMSEs of 0.84 and 0.70 °Brix, and RPIQs of 2.49 and 2.99, respectively, on the testing set (Tables 7 and 8) Additionally, regression models were trained exclusively on synthetic data to evaluate their standalone effectiveness. However, like the classification tasks, training solely on synthetic data resulted in poor performance on the testing set. Subsequently synthetic samples are incrementally added to

Table 8
Regression performance of 1D-CNN on test set.

Size of Original Data	Size of synthetic Data	R^2	RMSE	RPIQ
184	0	0.73	0.7	2.99
184	250	0.78	0.63	3.36
184	500	0.78	0.63	3.31
184	1000	0.75	0.67	3.14
184	1500	0.74	0.69	3.05
184	2000	0.71	0.72	2.9
0	2000	0.34	1.09	1.96

the original training set in steps of 250, 500, 1000, 1500, and 2000 to measure their contribution. The best result was achieved by adding 250 synthetic samples to the original training set when training 1D-CNN model, yielding an R^2 of 0.78, RMSE of 0.63 °Brix, and RPIQ of 3.36 on the testing set. Additionally, adding 2000 synthetic samples improved the PLSR prediction performance, increasing the R^2 from 0.61 to 0.68. Fig. 10 shows the 1:1 line relationship between predicted and measured TSS values on the test set for PLSR and 1D-CNN models trained on the original dataset and the original dataset augmented with synthetic samples. The result shows that both PLSR and 1D-CNN have better regression performance after adding synthetic hyperspectral reflectance data than using the original dataset only.

4. Discussion

Non-destructive assessing of grape maturity during the post-harvest stage is important to wineries to classify the grapes and produce wines of different quality. This study aims to use hyperspectral imaging system (400–1000 nm) and deep learning techniques to determine grape maturity in a non-destructive and rapid way. The TSS classification accuracy of using 3D-CNN and SSRN outperformed predictions using SVM (Tables 4-6). This demonstrates that deep learning models outperform traditional machine learning approaches in handling hyperspectral data, likely due to their ability to capture complex spatial-spectral features. The classification result of this study support the finding of Ramos et al., (2021), who used VGG-19 To classify the maturity status of Syrah and Cabernet Sauvignon. The ability of using CNNs models to process big data is no longer the bottleneck, instead these methods require larger datasets than are usually available or easily obtained. In addition, this study used regression models (PLSR and 1D-CNN) to predict grape TSS values. For the model trained on the original training data, the best performing is 1D-CNN, with a result of R^2 0.73, RMSE of 0.7 °Brix and RPIQ of 2.99 in the test set. In recent study have demonstrated the potential of more advanced deep learning models, such as InceptionTime, Residual Networks, and the lightweight MiniRocket model, for similar tasks (Silva et al., 2024). These models have been found to outperform traditional 1D-CNN approaches, showing superior generalization capacity in the prediction of sugar content from hyperspectral data. Future research could explore the integration of these novel models to further improve the accuracy and robustness of grape maturity and quality predictions.

In real practical conditions, building CNNs models need thousands of

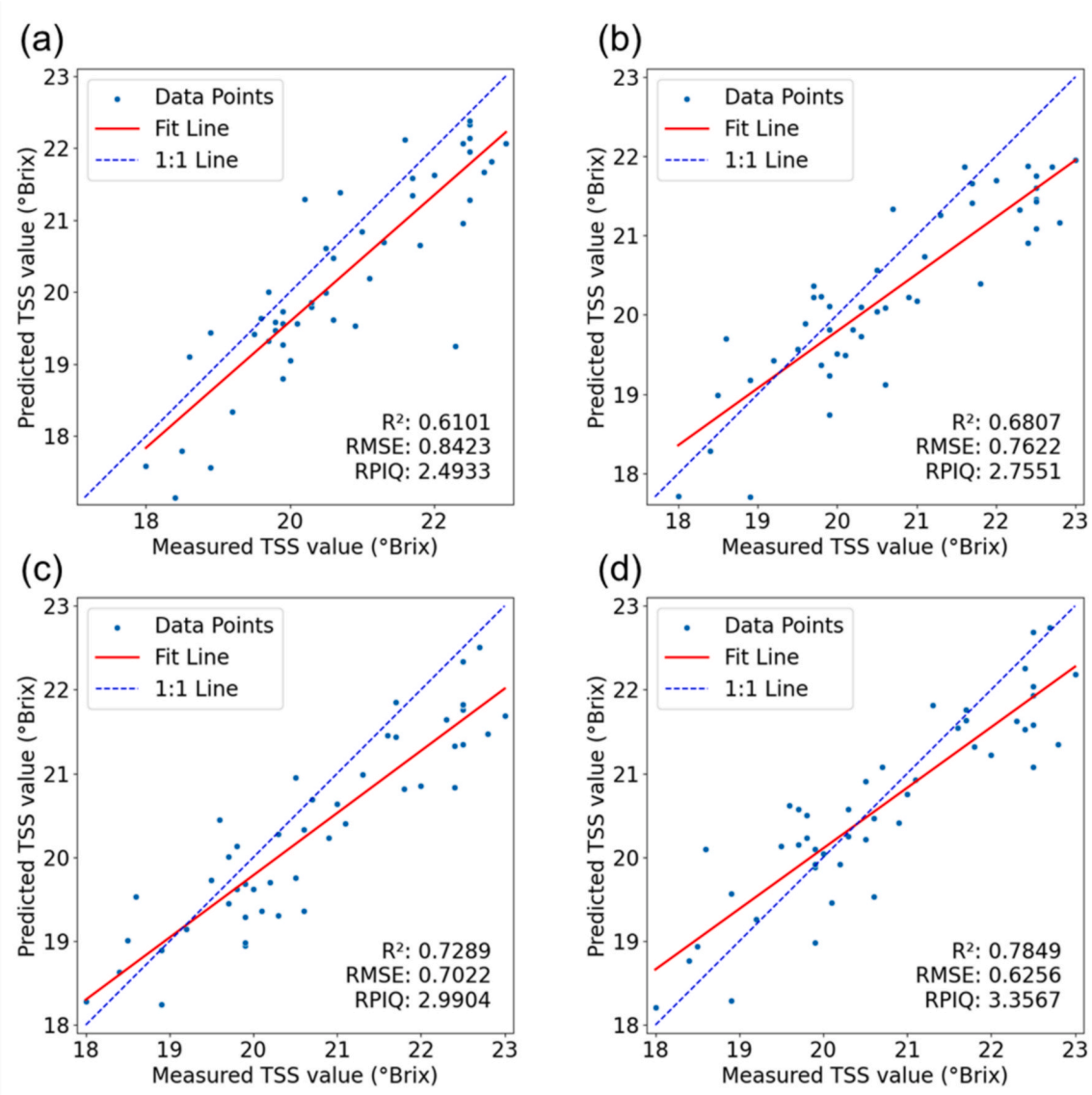

Fig. 10. The 1:1 line relationship between predicted and measured TSS values for the test set by PLSR on original dataset (a), PLSR on original + 2000 synthetic data (b), 1D-CNN on original dataset (c), 1D-CNN on original + 250 synthetic data (d).

images or data to train the model to prevent the model from overfitting. However, collecting thousands of images and data for non-destructive grape maturity needs large resources to ground truth data collection and labelling which is time consuming (Lu and Young, 2020). To mitigate the scarcity of datasets, this study proposed a data augmentation method based on deep learning architecture. Conditional WGAN with the gradient penalty was used to produce synthetic hyperspectral reflectance data. Compared with traditional GANs architecture, this study introduces an additional classifier model in the network. The classifier improves the quality and diversity of the generated samples by conditioning them on specific class labels. To balance the trade-off between data quality and computational efficiency, the conditional WGAN with the gradient penalty was trained for a range of epochs: 500, 1000, 2000, 8000, 10,000, and 20,000. Visual inspection and Mahalanobis Distance (MD) analysis reveal that the synthetic data generated at 10,000 epochs exhibit the highest similarity to the real data (Figs. 7-9, Table 3). It is noticeable that increasing the number of epochs to 20,000 led to a decrease in the similarity of the generated data to the real data, suggesting that beyond a certain point, further training does not significantly improve the quality of the generated hyperspectral reflectance data and may even introduce undesirable artifacts. Most previous studies applied GANs to generate RGB images in the agricultural and food production field (Abbas et al., 2021; Bird et al., 2022; Lu et al., 2022). This study demonstrates the potential of using GANs to generate hyperspectral reflectance data. This has helped by generating a large and diverse hyperspectral reflectance dataset.

The results using different classification models (3D-CNN, SSRN, and SVM) and regression models (PLSR and 1D-CNN) with synthetic hyperspectral reflectance data can be observed from Tables 4-6. The synthetic hyperspectral reflectance data, incrementally added to the original training set in steps of 250, 500, 1000, 1500, and 2000 samples, consistently resulted in higher model performance compared to training solely on the original dataset. This demonstrates that data augmentation positively impacts the classification and regression models' performance. Although performance varied depending on the model type and the number of synthetic samples added, the models trained with augmented datasets consistently outperformed those trained exclusively on the original training set in the test set, regardless of the model or task. This is similar to the work of Abbas et al., (2021) who used conditional GAN to generate tomato plant disease images. They showed the deep learning models have higher classification accuracy after adding synthetic image data during the training process. Future studies should validate the effectiveness of GAN-generated synthetic data on a broader range of machine learning and deep learning models to further assess its applicability and robustness in grape maturity estimation. Our findings highlight that GAN-augmented hyperspectral reflectance data can enhance the performance of both classification and regression models for grape maturity assessment. This result provides practical guidance for viticulture practitioners to select the task type (classification or regression) that aligns best with their production demands, such as rapid maturity screening using classification or detailed quality evaluation through TSS prediction.

Although the synthetic hyperspectral data generated after 10,000 epochs yielded the best performance in both grape maturity categories and continuous TSS values, compared to the real spectral data, the generated spectra still exhibited noticeable noise (Fig. 7 and 8). This noise can primarily be attributed to the complexity and high dimensionality of hyperspectral data. Unlike RGB images, hyperspectral reflectance spectra consist of hundreds of narrow, continuous bands, with subtle correlations between neighboring wavelengths. Capturing these fine spectral patterns poses a considerable challenge for GANs, often resulting in spectral fluctuations that appear as noise (Hennessy et al., 2021). Furthermore, although the conditional WGAN with gradient penalty (WGAN-GP) was employed to improve training stability, GAN models inherently involve a delicate balance between the generator and discriminator. This adversarial process can still introduce small irregularities, especially when generating high-dimensional spectral features. Future work could focus on developing more advanced GAN architectures tailored to hyperspectral data, such as spectral attention mechanisms or hybrid models, to reduce noise and further improve the realism of synthetic spectra (Qi et al., 2024). In addition, the industrial use of cameras only focuses on a specific spectrum region, exploring the ability of using GANs in specific regions that can be better applied to industrial data generation. While this study successfully demonstrates the potential of conditional WGAN with gradient penalty in generating synthetic hyperspectral reflectance data for grape maturity assessment, the development of generative models for hyperspectral data synthesis remains an emerging area. Future research could explore alternative and potentially more advanced generative approaches, such as diffusion models, transformer-based generative architectures, or hybrid methods that combine GANs with variational autoencoders (VAEs). These methods may further improve the quality and diversity of synthetic hyperspectral data, ultimately enhancing model robustness and generalizability in practical applications (ur Rahman et al., 2024).

5. Conclusion

Deep learning models have shown promising results in fruit quality or maturity detection during the post-harvest stage. However, training a deep learning model needs extensive and diverse datasets to avoid overfitting. In addition, labelling the grape maturity status also needs destructive analysis, usually from a limited dataset. This study proposes conditional WGAN with the gradient penalty to generate synthetic hyperspectral reflectance data to provide an extensive and diverse grape maturity dataset. The performance of proposed data augmentation techniques was evaluated on three classification models (SSRN, 3D-CNN and SVM) and two regression models (PLSR and 1D-CNN). Synthetic hyperspectral reflectance data plus original dataset scored higher classification and regression performance than training only on the original dataset, showing that augmentation has had a positive effect on the classification and regression model performance. The best classification results were achieved by augmenting the training dataset with 2000 synthetic samples and training with a 3D-CNN, yielding a classification accuracy of 91 % on the testing set. The best regression result was achieved by adding 250 synthetic samples to the original training set when training 1D-CNN model, yielding an R^2 of 0.78, RMSE of 0.63 °Brix, and RPIQ of 3.36 on the testing set. This study demonstrates the potential of using GANs to generate hyperspectral reflectance data and improve the deep learning model performance. In the future, GANs are expected to be applied on more grape varieties and specific spectrum regions.

CRediT authorship contribution statement

Hongyi Lyu: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Miles Grafton:** Writing – review & editing, Funding acquisition, Data curation, Conceptualization. **Thiagarajah Ramilan:** Writing – review & editing, Supervision, Conceptualization. **Matthew Irwin:** Writing – review & editing, Software, Resources, Formal analysis. **Eduardo Sandoval:** Resources, Formal analysis, Data curation.

Funding

This research received no external funding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the assistance of Mr. Guy McMaster, manager and winemaker of Palliser Estate Ltd.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2025.110341.

Data availability

Data will be made available on request.

References

Abbas, A., Jain, S., Gour, M., Vankudothu, S., 2021. Tomato plant disease detection using transfer learning with C-GAN synthetic images. Comput. Electron. Agric. 187, 106279.

Ashtiani, S.-H.-M., Javanmardi, S., Jahanbanifard, M., Martynenko, A., Verbeek, F.J., 2021. Detection of mulberry ripeness stages using deep learning models. IEEE Access 9, 100380–100394.

Audebert, N., Le Saux, B., Lefèvre, S., 2018. Generative adversarial networks for realistic synthesis of hyperspectral samples. In: In: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 4359–4362.

Baluja, J., Tardaguila, J., Ayestaran, B., Diago, M.P., 2013. Spatial variability of grape composition in a Tempranillo (Vitis vinifera L.) vineyard over a 3-year survey. Precis. Agric. 14, 40–58.

Benelli, A., Cevoli, C., Ragni, L., Fabbri, A., 2021. In-field and non-destructive monitoring of grapes maturity by hyperspectral imaging. Biosyst. Eng. 207, 59–67.

Bird, J.J., Barnes, C.M., Manso, L.J., Ekárt, A., Faria, D.R., 2022. Fruit quality and defect image classification with conditional GAN data augmentation. Sci. Hortic. 293, 110684.

Bramley, R., Pearse, B., Chamberlain, P., 2003. Being profitable precisely - a case study of precision viticulture from Margaret River. Aust. N. Z. Grapegrow. Winemak. Annu. Tech. Issue 473a, 84–87.

Chen, D., Qi, X., Zheng, Y., Lu, Y., Huang, Y., Li, Z., 2024. Synthetic data augmentation by diffusion probabilistic models to enhance weed recognition. Comput. Electron. Agric. 216, 108517.

Das, P., Yadav, J.P.S., 2020. Transfer learning based tomato ripeness classification. In: 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). IEEE, pp. 423–428.

Gomes, V., Fernandes, A., Faia, A., Melo-Pinto, P., 2017. Comparison of different approaches for the prediction of sugar content in new vintages of whole Port wine grape berries using hyperspectral imaging. Comput. Electron. Agric. 140, 244–254.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. Adv. Neural Inf. Process Syst. 27.

Gour, M., Jain, S., Sunil Kumar, T., 2020. Residual learning based CNN for breast cancer histopathological image classification. Int. J. Imaging Syst. Technol. 30, 621–635.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C., 2017. Improved training of wasserstein gans. Adv. Neural Inf. Process Syst 30.

Guo, Z., Zheng, H., Xu, X., Ju, J., Zheng, Z., You, C., Gu, Y., 2021. Quality grading of jujubes using composite convolutional neural networks in combination with RGB color space segmentation and deep convolutional generative adversarial networks. J. Food Process Eng. 44, e13620.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

Hennessy, A., Clarke, K., Lewis, M., 2021. Generative adversarial network synthesis of hyperspectral vegetation data. Remote Sens. 13, 2243.

Khalifa, N.E., Loey, M., Mirjalili, S., 2022. A comprehensive survey of recent trends in deep learning for digital images augmentation. Artif. Intell. Rev. 55, 2351–2377.

Koirala, A., Walsh, K.B., Wang, Z., McCarthy, C., 2019a. Deep learning–Method overview and review of use for fruit detection and yield estimation. Comput. Electron. Agric. 162, 219–234.

Koirala, A., Walsh, K.B., Wang, Z., McCarthy, C., 2019b. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO.'. Precis. Agric. 20, 1107–1135.

Lee, H., Kwon, H., 2017. Going deeper with contextual CNN for hyperspectral image classification. IEEE Trans. Image Process. 26, 4843–4855.

Li, S., Luo, H., Hu, M., Zhang, M., Feng, J., Liu, Y., Dong, Q., Liu, B., 2019. Optical non-destructive techniques for small berry fruits: a review. Artif. Intell. Agric. 2, 85–98.

Lu, Y., Chen, D., Olaniyi, E., Huang, Y., 2022. Generative adversarial networks (GANs) for image augmentation in agriculture: a systematic review. Comput. Electron. Agric. 200, 107208.

Lu, Y., Saeys, W., Kim, M., Peng, Y., Lu, R., 2020. Hyperspectral imaging technology for quality and safety evaluation of horticultural products: a review and celebration of the past 20-year progress. Postharvest Biol. Technol. 170, 111318.

Lu, Y., Young, S., 2020. A survey of public datasets for computer vision tasks in precision agriculture. Comput. Electron. Agric. 178, 105760.

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., Sandoval, E., 2024. Hyperspectral imaging spectroscopy for non-destructive determination of grape berry total soluble solids and titratable acidity. Remote Sens. 16, 1655.

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., Wei, H.-E., Sandoval, E., 2023. Using remote and proximal sensing data and vine vigor parameters for non-destructive and rapid prediction of grape quality. Remote Sens. 15, 5412.

Miranda, J.C., Gené-Mola, J., Zude-Sasse, M., Tsoulias, N., Escolà, A., Arnó, J., Rosell-Polo, J.R., Sanz-Cortiella, R., Martínez-Casasnovas, J.A., Gregorio, E., 2023. Fruit sizing using AI: a review of methods and challenges. Postharvest Biol. Technol. 206, 112587.

Nagy, A., Riczu, P., Tamás, J., 2016. Spectral evaluation of apple fruit ripening and pigment content alteration. Sci. Hortic. 201, 256–264.

Pullanagari, R.R., Dehghan-Shoar, M., Yule, I.J., Bhatia, N., 2021. Field spectroscopy of canopy nitrogen concentration in temperate grasslands using a convolutional neural network. Remote Sens. Environ. 257, 112353.

Qi, H., Huang, Z., Jin, B., Tang, Q., Jia, L., Zhao, G., Cao, D., Sun, Z., Zhang, C., 2024. SAM-GAN: an improved DCGAN for rice seed viability determination using near-infrared hyperspectral imaging. Comput. Electron. Agric. 216, 108473.

Ramos, R.P., Gomes, J.S., Prates, R.M., Simas Filho, E.F., Teruel, B.J., dos Santos Costa, D., 2021. Non-invasive setup for grape maturation classification using deep learning. J. Sci. Food Agric. 101, 2042–2051.

Rolle, L., Segade, S.R., Paissoni, M.A., Giacosa, S., Gerbi, V., 2022. Assessment and control of grape maturity and quality. White Wine Technology. Elsevier 1–16.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6, 1–48.

Silva, R., Freitas, O., Melo-Pinto, P., 2024. Evaluating the generalization ability of deep learning models: an application on sugar content estimation from hyperspectral images of wine grape berries. Expert Syst. Appl. 250, 123891.

Sun, D.-W., 2010. Hyperspectral imaging for food quality analysis and control. Elsevier.

Tan, H., Hu, Y., Ma, B., Yu, G., Li, Y., 2024. An improved DCGAN model: data augmentation of hyperspectral image for identification pesticide residues of Hami melon. Food Control 157, 110168.

Tsakiridis, N.L., Samarinas, N., Kokkas, S., Kalopesa, E., Tziolas, N.V., Zalidis, G.C., 2023. In situ grape ripeness estimation via hyperspectral imaging and deep autoencoders. Comput. Electron. Agric. 212, 108098.

ur Rahman, Z., Asaari, M.S.M., Ibrahim, H., Abidin, I.S.Z., Ishak, M.K., 2024. Generative Adversarial Networks (GANs) for image augmentation in farming: a review. IEEE Access.

Walsh, K.B., Blasco, J., Zude-Sasse, M., Sun, X., 2020. Visible-NIR 'point'spectroscopy in postharvest fruit and vegetable assessment: the science behind three decades of commercial use. Postharvest Biol. Technol. 168, 111246.

Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P., Li, X., 2021. A review of deep learning used in the hyperspectral image analysis for agriculture. Artif. Intell. Rev. 54, 5205–5253.

Wang, Z., Walsh, K., Koirala, A., 2019. Mango fruit load estimation using a video based MangoYOLO—Kalman filter—hungarian algorithm method. Sensors 19, 2742.

Wei, X., Wu, L., Ge, D., Yao, M., Bai, Y., 2022. Prediction of the maturity of greenhouse grapes based on imaging technology. Plant Phenomics.

Zhao, Z., Hicks, Y., Sun, X., Luo, C., 2023. Peach ripeness classification based on a new one-stage instance segmentation model. Comput. Electron. Agric. 214, 108369.

Zhong, Z., Li, J., Luo, Z., Chapman, M., 2017. Spectral–spatial residual network for hyperspectral image classification: a 3-D deep learning framework. IEEE Trans. Geosci. Remote Sens. 56, 847–858.