ESI Preprints **Not Peer-reviewed**

Clarification of Water Stress in Apple Seedlings Using HSI Texture with Machine Learning Technique

Yanying An
School of Information Technology, Murdoch University, Australia
School of Horticulture, Qingdao Agricultural University, China
Ran Wang
School of Horticulture, Qingdao Agricultural University, China

Doi: 10.19044/esipreprint.1.2024.p518

Approved: 20 January 2024
Posted: 23 January 2024

Copyright 2024 Author(s)
Under Creative Commons CC-BY 4.0
OPEN ACCESS

Cite As:
An Y. & Wang R. (2024). *Clarification of Water Stress in Apple Seedlings Using HSI Texture with Machine Learning Technique.* ESI Preprints.
https://doi.org/10.19044/esipreprint.1.2024.p518

Abstract

Apples are known for their nutrition and economic value. Accurate and rapid diagnosis of water status in apple seedlings on an individual rootstock basis is a prerequisite for precision water management. This study presents a rapid and non-destructive approach for estimating water content in apple seedlings at leaf levels. A PIKA L system collects hyperspectral images(400-1000nm) of apple leaves. To the author's knowledge, no prior work was conducted using the spectral-texture approach in plant water stress. Our research extracts spatial information, gray-level co-occurrence matrix (GLCM), from feature wavelength images of hypercubes. Machine learning methods are applied to these spatial feature matrixs to identify apple leaves under different water stresses. In addition, differences in spectral responses were analysed using machine learning techniques for sorting apple seedlings with varying water treatments (dry, normal, and overwatering). Also, we measure chlorophyll to determine the relationship between hyperspectral characteristics and physiological changes. The achievements of the research indicate that the fusion of texture and hyperspectral imaging coupled with machine learning techniques is promising and presents a powerful potential to determine the water stress in the leaves of apple seedlings.

Keywords: Hyperspectral imaging; machine learning; plant water stress; plant leaf; plant physiology

1. Introduction

Apple (*Malus domestica Borkh*) is one of the world's most widely planted and nutritionally significant fruit crops (Duan et al.,2017). Apples are rich in nutrients vital for good health and disease prevention, making them a valuable choice each day for enhancing the quality of our diet. With increased income and public awareness of balanced nutrition, Global apple consumption is increasing annually (Li et al.,2013). Apple seedlings are tiny and propagated tree stems grafted onto a hardy rootstock (Loucks, 2021). They are hugely profitable. In 2021, according to Willis Orchard Co. (2021), ten seedling trees at the height of 1-2 feet made a profit of $34.95, whereas ten seedling trees at the height of 2-3 feet earned $59.95.

Water management is vital for apple orchards (Apple & Pear Australia Limited,2023). As a result of drought and increasing competition for water, orchardists need to adopt efficient water management strategies (PIRSA,2006). Newly planted apple trees require weekly watering. Ideal apple tree irrigation involves deep root soaking (Ellis,2021). The key is to let the water flow into the soil slowly to allow for deep watering until it is established (The Home Depot,2021). A plant induces leaf senescence prematurely if water shortage exceeds a critical level (Lim and Nam,2007). This process is characterised by loss of chlorophyll and leaf yellowing (Yamaguchi et al.,2010). If wilting leaves are observable, irreversible damage to plants and yield occurs (Behmann et al.,2014). On the other hand, creating standing water and soggy roots can be as damaging as drought conditions for the apple seedlings. Too much water depletes oxygen from the soil, prevents the roots from absorbing necessary minerals, and makes a tree susceptible to rotting and infections (Ellis,2021). The symptoms of overwatering also include wilting, yellowing of leaves, root rot and stunted growth (Canna, 2023). Our research focuses on detecting early drought and overwatering stresses, which are not yet visible to the naked eye. Detecting early symptoms of water stress and distinguishing it from normal variations timely requires further attention.

Water stress can be detected using various methods. Water stress is one of the crucial environmental factors inhibiting photosynthesis (Bradford and Hsiao,1982). Studies have shown that decreased photosynthesis under water stress can be linked to the perturbations of the biochemical process (Graan and Boyer,1990; Lauer and Boyer,1992). In particular, PSII is sensitive to water stress (He et al.,1995). Lu and Zhang (1999) studied the effect of water stress on PSII photochemistry and thermostability of PSII in wheat plants by measurement of chlorophyll

fluorescence. However, chlorophyll fluorescence is destructive and damages the leaf as well. Pressure chambers, measuring plant water potential, are widely used in irrigation management (Waring and Cleary,1967). But they are not well suited to measurements of small plants such as grasses because a petiole must extend through the seal of the pressure chamber. In addition, plants with tender tissues (e.g., new tillers on grasses) are easily damaged by the seal and cannot be used (Water Potential,2022).

Chlorophyll in plant leaves is crucial in plant metabolism and growth. Chlorophyll in apple leaves is essential in photosynthesis (Wang et al.,2016). It constituents a major component of plant leaves and is a useful indicator of the overall health condition of the plant. Determining chlorophyll content in plant leaves can be used to investigate plant physiological and nutritional status and consequently has important implications for plant stress detection (Jin et al.,2020). Generally, as plant stress levels increase, chlorophyll content tends to decrease (Schuerger et al.,2003). Traditional approaches for quantifying chlorophyll content mainly include acetone ethanol extraction, spectrophotometry and high-performance liquid chromatography (Zhang et al.,2022). Such destructive methods based on laboratory procedures are time-consuming and expensive (Gholizadeh et al.,2017). Several spectral indices have been identified using spectroscopy and a portable chlorophyll meter, which can predict chlorophyll content in plant tissues in vivo (Jesperson et al.,2016). However, manually operated portable chlorophyll meters are relatively biased, and spectroscopy techniques cannot measure the spatial distribution of chlorophyll in plant leaves (Gupta et al.,2013). These two methods are too labour-intensive to meet the needs of large-scale screening programs.

Hyperspectral imaging (HSI) is now emerging as a potential tool for rapid, non-destructive, and automated assessment of plant status(Kim et al.,2015). Hyperspectral imaging integrates spectroscopic methods and imaging technology (Williams and Norris,2001). In the pertinent hyperspectral image, each pixel contains a complete spectrum. No physicochemical analyses are needed, making it ideal for agri-food (Ayala,2018). Using hyperspectral technology can identify stress symptoms before a human observer does so, which has obvious benefits (Lowe et al.,2017).

Current HSI in small-scale plant research focuses on chlorophyll and other plant-related endogenous pigments in the visible spectral range(400-800nm) (Bauriegel and Herppich,2014). Chlorophyll could be contrasted with predictions obtained from HSI spectral reflectance (Feng et al.,2019). Based on the characteristics of chlorophyll reflection and absorption of specific wavelength spectra, chlorophyll spectroscopy diagnosis of many plants has been carried out. Wen et al. (2019) visualised chlorophyll content

distribution in apple leaves using hyperspectral imaging technology. Sun et al. (2018) took 382-1019nm bands of hyperspectral imaging data as the research band using the Random Frog Leaping algorithm and partial least squares method, selected the sensitive bands and constructed an estimation model of chlorophyll content in potato leaves. Ding et al. (2015) used hyperspectral imaging technology to predict the chlorophyll content of rape leaves, and the correlation coefficient reached 0.834.

Artificial intelligence technologies are used extensively in image processing and many other fields (Ozdemir and Polat,2020). As a result, classification studies are carried out on hyperspectral images with machine learning methods. Deep Learning, a specialised form of the artificial neural network, is highly preferred due to its unique structure, especially for classifying hyperspectral images containing complex information. In Yin et al's research(2017), four geographical origins of Chinese wolfberries were studied using the HSI technique. The mean spectra of the wolfberry samples were analysed to build a support vector machine (SVM), neural network with radial basis function (NN-RBF) and extreme learning machine (ELM) models. Zhu et al. (2029) identified the freshness of spinach leaves stored at different temperatures using HSI, where Partial least squares discriminant analysis (PLS-DA), support vector machine (SVM), and extreme learning machine (ELM) were used to build models based on full spectra and effective wavelengths. Yamashita et al. (2020) employed five machine learning algorithms, Random Forest (RF), Support Vector Machine (SVM), Cubist, Stochastic Gradient Boosting (SGB), and Kernel-based Extreme Learning Machine (KELM) in the dissection of hyperspectral reflectance to estimate nitrogen and chlorophyll contents in tea leaves. Feng et al. (2019) predicted the chlorophyll contents of crops under salinity stress by combining hyperspectral data with Partial Least Square Regression (PLSR). A high predictive accuracy of SPAD was acquired in this study, as shown by the R-value of 0.838.

Many deep learning-based algorithms focus on deep feature extraction for classification improvement (Li et al.,2018). Features extracted from images are used for image matching and retrieving (Afifi and Ashour, 2012). The texture is one of the essential characteristics of hyperspectral images. Image texture analysis, an important area of research in image processing, aims to provide information about the spatial arrangement of colour or intensities in an image (Xu and Gowen,2019). The texture information in the hyperspectral image is vital for automatic image processing. The texture pattern of different tissues and plants under differing water treatments contains various intensity patterns, which can be exploited for image classification (Bhagat et al.,2019). Besides, much research has demonstrated that adding textural features to characteristic spectral features

could further enhance the performance and accuracy of the model (Jiang et al.,2019; Cheng et al.,2016; Liu et al.,2014). Through band grouping, sample band selection and guided filtering, the texture features of hyperspectral data are improved (Li et al.,2018). Li et al. (2018) proposed a novel hyperspectral classification framework based on optimal Deep Belief Networks (DBNs) and a novel texture feature enhancement (TFE). After TFE, the optimal DBN is employed on the hyperspectral reconstructed data for feature selection and classification. The texture feature extraction methods include many approaches, for example, the Gray level co-occurrence matrix(GLCM). In Lu et al. (2018)'s research using hyperspectral imaging to discriminate yellow leaf curl disease in tomato leaves, 24 texture features were extracted using a gray-level co-occurrent matrix(GLCM). AlSuwaidi et al. (2017) applied hyperspectral imaging on control, stressed Arabidopsis plant leaves, and employed Markov random field model(MRF) to facilitate spectral-texture analysis and enhance prediction performance. Zhang et al. (2022) classified flue-cured tobacco based on the fusion of hyperspectral and texture features; texture features were extracted using GLCM, whereas machine learning techniques KNN, RF, SVM and Naive Bayes Classifier (NB) were employed. GWO-ELM was used in hyperspectral estimation of SPAD value of cotton leaves under Verticillium Wilt stress (Yuan et al., 2023). The smooth and stable sigmoid function, which was easy to derive, was selected as the activation function of ELM, and the multiple trial-and-error method was applied to determine the number of nodes in the hidden layers, the size of the gray wolf population and the maximum number of iterations. Ni et al. (2020) used online sorting of the film on cotton based on GWO-ELM. However, all those research focused on the texture feature of RGB images of hyperspectral imaging cube, our research makes a renovation by combing texture feature of images of main spectra with the spectral features.

Our research aims to combine texture and color analysis of hyperspectral imaging with machine learning to identify apple seedling leaves under different water treatments (normalisation, overwatering, drought). The objectives are: (1) investigating and comparing apple leaves' hyperspectral characteristics in fusion with textural and color features under different water treatments using machine learning methods. (2) validating the classification result using chlorophyll content. A spectral-texture analysis framework is proposed for classifying hyperspectral images of apple seedling leaves of different water conditions. Advanced machine learning techniques are used to identify distinctive features in the spectral domain of hyperspectral images. In addition, texture properties are explored in the characteristic wavelength images. The framework integrates these two levels of properties at both feature extraction and classifying decision stages. The main crux of the work lies in the use of significant spectral and texture

features and a decision fusion mechanism to enhance the image properties, thus improving classification accuracy.

2. Material and Methods

2.1 Material

M9T337 apple seedlings with around 20 leaves per plant were grown in the National Agricultural and forestry science and Technology incubator seedling base in Zhucheng, China. Apple seedlings of similar growth were chosen. The height of the rootstock was around 15cm. Apple seedlings were transported to greenhouses at Qingdao Agricultural University. Two weeks after transplanting, all apple seedlings were treated with the complete nutrient solution and supplied with all the essential nutrients. The nutrient solution was provided to each plant using a trickle nozzle.

The temperature of the greenhouse was around 23℃. The humidity of the greenhouse was 95%. Each pot contained one seedling. The substrate was composed of peat and vermiculite (1:1). The experiment occurred between September and November 2021.

2.2 Experiment design

There were three treatment groups, excessive water treatment group, drought group, and control group-normal watering group. For each treatment, 30 apple seedlings were used for hyperspectral imaging collection, and 30 apple seedlings were used for physiological index collection. Hence, there were 180 seedlings used in total. A pre-experiment was conducted to conclude that the daily amount of water a pot of apple seedlings needs was 30mL. The night before the experiment, the three treatments apple seedlings were watered thoroughly. Once the apple seedlings of the drought group were watered, no further watering was required during the experiment. The excessive water treatment group maintained water at a depth of 1.5cm, and the water was uniformly replenished at 6 pm. In addition, 30mL of water was added to apple seedlings in the regular watering group every day at 6 pm.

One leaf in the middle of each plant was employed for hyperspectral data collection. The acquisition occurred every other day after water treatment before apparent symptoms such as wilting and yellowing appeared on apple leaves (day 25). On day 0, the starting point of water treatment, only 30 hyperspectral images were collected (10 for each treatment) due to the similar conditions for all apple seedlings. From then on, 30 images were collected for each treatment. The information and sample collection period ranged from 9:00 am to 11:00 am. Table 1 shows the number of images collected.

Table 1. Images collected

Date	18 Sep	19 Sep	21Sep	23 Sep	25 Sep	27 Sep	29 Sep

No. of images	30	90	90	90	90	90	90

The physiological index group used a random block design, and 30 seedlings constituted a small plot. Samples were measured on September 18 (day 0), September 19 (day 1), till Sep 29 (day 11), respectively. The SPAD of leaves four and seven from the bottom of the apple seedlings were measured using a chlorophyll meter and taken as an average for each pear seedling.

2.3 System Configuration (Imaging System-Hardware)

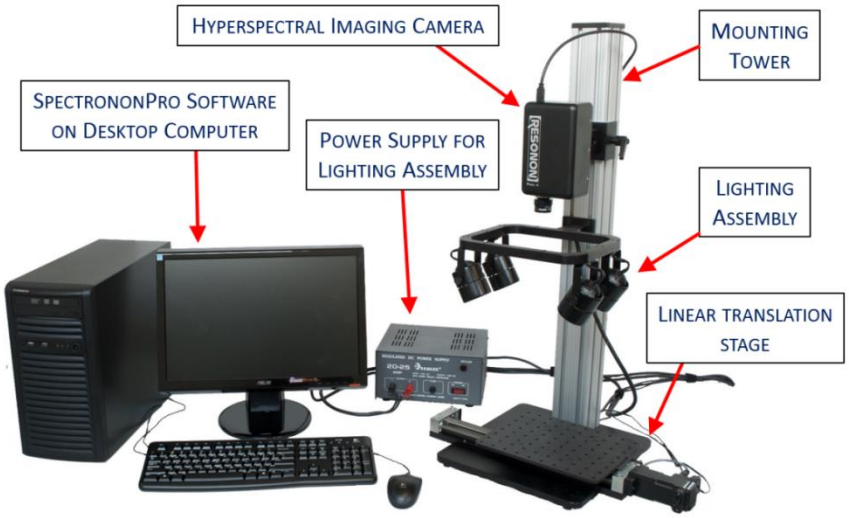


Figure 2. Benchtop hyperspectral imaging system (Resonon Inc.,2019)

Resonon SpectrononPro 5.3 was used in our research(figure 2). It comprised a PIKA hyperspectral imaging camera, linear translation stage, mounting tower, lighting assembly, and software control system (Resonon,2019). The imager and lighting assembly positions were adjusted along the tower's length. Resonon's hyperspectral images are line-scan imagers; two-dimensional images were constructed by translating the sample relative to the camera. A total of 8 Halogen lamps were used.

A Windows® XP operating system with 512MB memory was used. It had a 2.0GHz Pentium 4 compatible processor and 64MB AGP graphics. In addition, a 32-bit standard PCI slot, FireWire 800 port, and OHCI standard IEEE 1394B host controller were used.

2.4 Image Correction

Table 3. Parameter Settings

Camera Settings	
Framerate	30 fps
Integration time	12.904ms
Gain	1.0000dB
Update	
High cut percentage	98%
Stage	
Swath	13.1cm
Scanning speed	0.6993cm/s
Step size	0.0333333333cm
Homing speed	4.000cm/s
Jogging speed	0.07938cm/s

The system collected hyperspectral images in a wavelength range of 400-1000 nm with a spectral resolution of 2.1 nm. Leaves were fixed on black cardboard with negligible reflectance and placed on the mobile platform. During image acquisition, the mobile platform for line scanning had an optimised velocity of 0.6993 cm/s, and the camera integration time was 12.904 ms (Table 3). A corrected hyperspectral image was calculated using the following equation:

R= (image-dark)/(white-dark)

Where: R was the relative reflectance image of the sample, *image* was the raw image of the sample, *white* was the white reference image acquired from a uniform, stable, and high reflectance ceramic tile (reflectance), and *dark* was the current dark image acquired by completely covering the camera lens with its non-reflective opaque black cap.

2.5 Software

Image acquisition was controlled by imaging data acquisition software (SpectrononPro, Resonon Inc., USA). Image correction was done using software ENVI 5.3 (The Environment for Visualizing Images, ITT Visual Information Solutions Corp., USA) and analysed using Matlab R2010a (The Mathworks, Inc., Natick, MA, USA).

2.6 Data preprocessing

Figure 4 demonstrates the workflow of this research. According to Figure 4, the detailed data analysis processes are: The IRIV method was applied to extract spectral fingerprint features. Next, neural networks ELM, and its variations, GWO-ELM, SSA-ELM, were used to evaluate the water stress using the spectral fingerprint features. Then, these three neural networks were employed in the HSI, GLCM feature analysis, in addition to characteristic wavelengths fused with texture feature analysis. Also, the chlorophyll measurement was taken to corroborate the result.

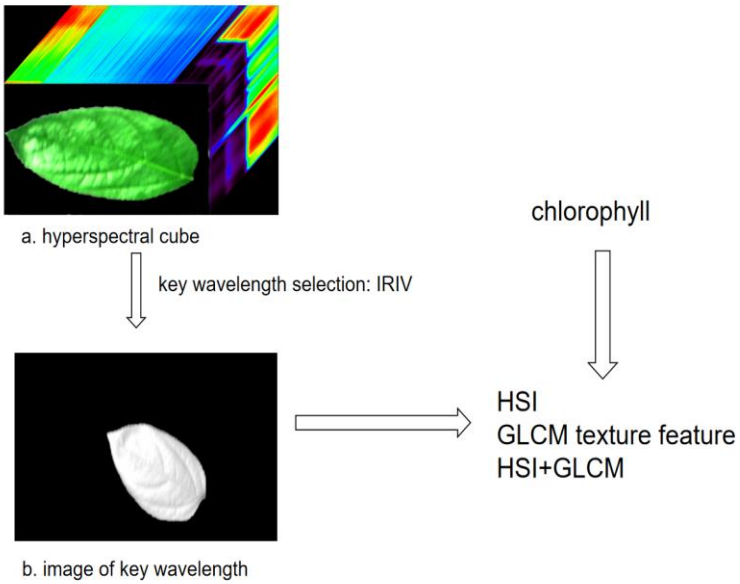


Figure 4. Data processing workflow

In the IRIV selection, Root mean square error of prediction (RMSEP) = 3.8785, RMSEF = 0.5875. 15 bands were selected to represent the 300 bands of hypercubes. These are band185 (766.39nm), band213 (827.34nm), band222 (847.06nm), band223 (849.21nm), band230 (864.65nm) , band234 (873.46nm), band246 (899.97nm), band249 (906.61nm), band257 (924.37nm), band259 (928.82nm),band262 (935.49nm),band276 (966.75nm),band278 (971.22nm), band289 (995.91nm) and band290 (998.16nm).

3. Texture analysis for identifying leaf water stress

Due to significant leaf variations in surface textures, texture analysis was conducted to identify the leaf water stress. Texture analysis characterises regions in an image by their texture content in terms of smoothness, roughness, silkiness, or bumpiness in the context of an image. These characteristics refer to brightness values or grey level variations in this case.

The most commonly used texture measures are derived from the grey-level co-occurrence matrix(GLCM).

The grey-level co-occurrence matrix approach(GLCM) considers the pairs of pixels in certain spatial relations to each other. Co-occurrence matrices are used. They relate the relative frequencies P(i,j|dθ) that two pixels at a constant vector distance (d,θ) from each other have intensity (i,j): in the GLCM P(i,j|dθ), the (i,j)th entry of the matrix, represents the number of occurrences of a pixel having the intensity value i that is separated from another pixel with intensity value j at a distance d in the direction θ. Two forms of co-occurrence matrix exist. In the first case, the matrix is symmetric, where pairs separated by d and –d (for a direction θ) are counted(Haralick et al.,1973). In the second case, the matrix is not symmetric, and only pairs separated by a distance of d are counted(Conners and Harlow,1980). This leads to a square matrix with a dimension equal to the number of intensity levels in the image for each distance d and orientation θ. If pixel pairs in the picture are highly correlated, the entries in the GLCM are gathered along the diagonal of the matrix. There are five features used in our study:

1. **energy**: it provides information on image homogeneity; it has low values when the probabilities of the grey-level pairs are similar and high values otherwise. It is computed as $\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j|d,\theta)^2$
2. **entropy**: it measures the disorder of the GLCM. It is computed as $-\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j|d,\theta) log_2(\mathrm{P(i,j|d,\theta)})$
3. **correlation**: it measures the grey-level linear dependence between pixels (relative to each other) at the specified positions; it has high values when the values are uniformly distributed in the GLCM and low values otherwise.
4. **local homogeneity (also called inverse difference moment)**: it is high when the same pairs of pixels are found (e.g., in the case of a spatial periodicity). It is computed as $\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{P(i,j|d,\theta)}{1+(i-j)^2}$
5. **inertia (also called contrast)**: it quantifies local variations present in the image. It is computed as $\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)^2 P(i,j|d,\theta)$

The GLCM approach consists in considering second-order statistics. The GLCM method studies the grey-level distribution of pairs of pixels. This is why it is also known as the second-order histogram method. In the same way, higher-order statistics analyse the joint distribution of more than two pixels.

4. Machine learning methods

a. ELM

In this study, we compared and validated the ELM model with its variations, GWO-ELM, SSA-ELM and GOA-ELM.

Extreme Learning Machine (ELM) are feedforward neural networks, invented by Huang (2006). According to Huang (2006), this algorithm tends to provide good generalization performance at extremely fast learning speed. ELM doesn't require gradient-based backpropagation to work. It used Moore-Penrose generalized inverse to set its weights.

$f_L(x)=\sum_{i=1}^{L}\beta_i g_i(x)=\sum_{i=1}^{L}\beta_i g(\omega_i * x_j + b_i), j = 1,\ldots,N$ (4)

Where: L is a number of hidden units, N is a number of training samples, β is weight vector between the hidden layer and output, ω is a weight vector between input and hidden layer, g is an activation function, b is a bias vector, x is an input vector.

b. GWO-ELM

The gray wolf optimisation algorithm, proposed in 2014, is a new swarm intelligence optimisation algorithm (Mirjalili et al.,2014). The GWO algorithm mainly maps the group predation behaviour of gray wolves to the machine learning field to realise the optimisation of parameters and algorithms. Its main features include fewer adjustment parameters, a simple structure, high efficiency and convenience, etc. The algorithm has a convergence factor adaptive mechanism and an information feedback mechanism, which can effectively avoid various defects of the traditional algorithm. It has advantages in the convergence speed and solution accuracy of the problem(Wang,2021).

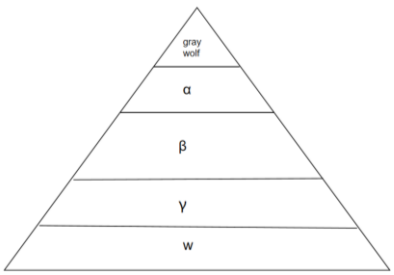


Figure 5. Gray wolf population classification

The wolf group in the grey wolf optimisation algorithm is divided into four levels **α**, **β**, **δ**, **ω** in the shape of a pyramid (figure 5). Among which **α** is the head wolf, whose main task is to coordinate the group's decision-making and conduct internal management, **β**' s main task is to assist **α** in decision-making, and **δ** is mainly responsible for observation, hunting and nursing. **ω** is responsible for coordinating the relationship within the group. The wolves will constantly update their location according to their level of fitness.

c. Sparrow Search Algorithm (Sparrow Search Algorithm, SSA) optimised ELM.

In SSA, the first sparrow to obtain food is often the discoverer with the highest fitness value, leading the entire population to determine the location and direction of food. Therefore, the finder will search for food in a broader range than other sparrows and constantly explore the fitness values in different directions and positions. The position of the population needs to be updated iteratively to find out the best foraging position.

Randomly initialise the population size of sparrows *n* and the maximum number of iterations *max iter*, and determine its objective function; the sparrow group position is initialised as:

$$X=\begin{bmatrix} \boldsymbol{f}([\boldsymbol{x_{1,1}} & \cdots & \boldsymbol{x_{1,d}}]) \\ \vdots & \vdots & \vdots \\ \boldsymbol{f}([\boldsymbol{x_{n,1}} & \cdots & \boldsymbol{x_{n,d}}]) \end{bmatrix} \quad (5)$$

In the formula, the value of each row in Fx represents the fitness value of the individual (Liu et al., 2022).

5. Neural network training

The training target accuracy (goal) is set to 0.01, and the number of training is set to 100 times. When the training curve's fitting accuracy(best) (train) reaches the target accuracy, that is, when the end point of the training curve, the fitting accuracy, and the target accuracy coincide, the neural network automatically stops training. Suppose the fitting accuracy never reaches the target accuracy. In that case, the neural network will repeat the training process until the number of training times is completed. 75% of the data were randomly selected as the training sample set, 15% as the validation data, and 15% as the test data. In table 6, the numbers of HSI data cube is the exact amount of data cubes taken during the experiment session (11 days collectively). The number of GLCM input data is the number of key wavelengths (15) multiplied by the number of data cubes collected, it is the same with the input data HSI fused with GLCM. Table 7 demonstrates the parameters of different learning algorithms. In both ELM variations GWO-ELM and SSA-ELM, the number of hidden layers is 50, and the max iteration is 100.

Table 6. Numbers of input data

Input data	number
HSI	30+30×3×6
GLCM	15×(30+30×6)
HSI+GLCM	15×(30+30×6)

Table 7. Neural network parameters

Machine learning networks	parameters
ELM	hiddennum = 50; activate_model = 'sig';
GWO-ELM	hiddennum = 50; SearchAgents_no = 5; Max_iteration = 100;
SSA-ELM	hiddennum = 50; Pop= 5; activate_model = 'sig'; Max_iteration = 100;

5.2.1 Evaluation metrics

A. Confusion matrix

Performance measurement is vital in defining the effectiveness of a program. Confusion matrixes are a common evaluation tool used in machine learning (An, 2020). Generally, they consist of a n× n table plotting actual class against predicted class (n denoting the number of classes, so a binary classifier would utilize a 2× 2 table), to which the true and false (determined by the actual classes) positives and negatives (determined by the predicted classes) fit within (Visa et al., 2011).

B. Accuracy

In practical applications, we should take the accuracy of the classifier into consideration. Because scientists and farmers are more concerned with the situations where the classifier sorts the drought pear seedling leaves as sound ones if the classifier makes the wrong decision, which will hinder the timely watering of the plants, leading to more significant potential economic losses than discarding the plants.

$$\text{Accuracy}=\frac{correctly\ classified\ samples}{total\ number\ of\ samples} \tag{18}$$

Accuracy (total correct divided by the total number of assessments), however, does not consider the significance of misidentified class (Halimu et al., 2019) and tends to be an overaly optimistic performance indicator.

C. Evaluation metrics for classification algorithms

Commonly used evaluation metrics for classification algorithms include recall, precision, F1 score and mse-loss. Recall measures the ability

to identify positive samples, precision measures the accuracy of positive sample predictions, and F1 score is a metric that combines recall and precision. MSE-loss is a criterion that measures the mean squared error between each element in the input and target.These metrics can be selected and weighted based on specific requirements. The formulas for these metrics are provided accordingly.

$$\text{Recall}=\frac{TP}{TP+FN} \quad (19)$$

$$\text{Precision}=\frac{TP}{TP+FP} \quad (20)$$

$$\text{F1 score}=\frac{2\times P\times R}{P+R} \quad (21)$$

Where TP is true positive, TN is true negative, FP is false positive, FN is false negative, P is precision, and R is recall (Shu et al., 2023).

6. Physiological Data

6.1 Chlorophyll

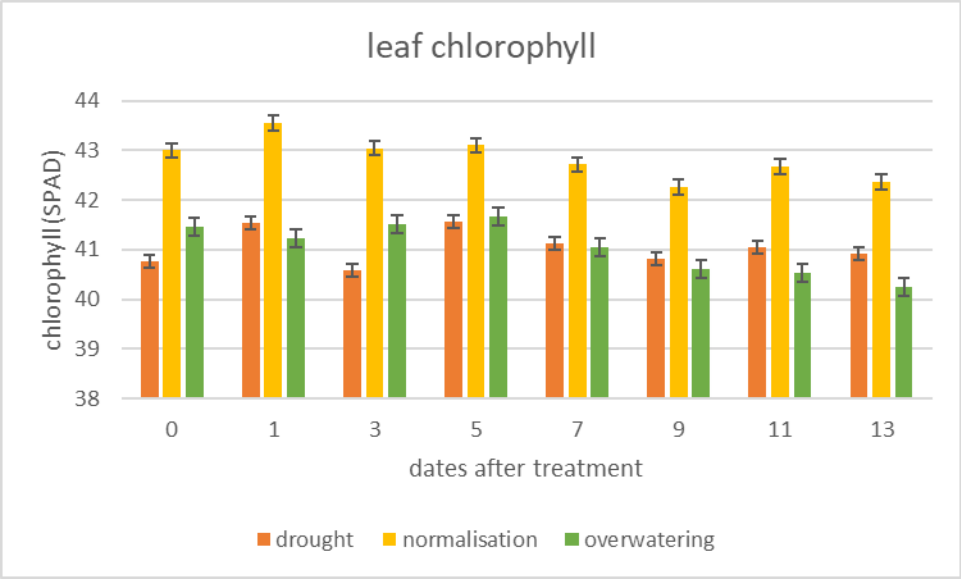


Figure 8. SPAD of apple leaves under different water treatments. The small vertical line on the graph represents the standard deviation of the mean(LSD=0.05).

The leaf chlorophyll content is an essential indicator of the growth and photosynthesis of apples under water stress (Sun et al.,2021). Figure 8 shows the amount of chlorophyll produced by photosynthesis in leaves under different water treatments. There was a significant difference in chlorophyll among other water treatment groups. The chlorophyll of apple leaves in the normal treatment group increased, reached its maximum on October 5th, and then slowly decreased. This can be explained by the fact that the leaves in the normal water treatment group are sufficient in water, and the

photosynthesis capability is vital. By comparison, the chlorophyll of apple leaves in the drought treatment group slowly decreased as plants did not have enough water. As Pour-Aboughadareh et al. (2019) stated, water deficit stress reduced the relative chlorophyll content (SPAD). Chlorophyll is the main component of photosynthesis-one of the physiological processes most sensitive to environmental stress (Hussain et al.,2019). Like high temperature, water deficiency may induce lipid peroxidase and electrolytic leakage from chloroplast and thylakoid membranes, leading to a loss of chlorophyll content (Ristic et al.,2007). The chlorophyll of leaves in the overwatering group increased slightly but soon decreased as the water provided exceeded what the plant needed. In waterlogged soil, there is not enough oxygen for the roots. The roots start to rot fast, thus leading to pigmentation loss (Green,2022).

The spectral reflectance of leaves is closely related to the leaf surface characteristics, leaf thickness, water content, and the contents of chlorophyll and other pigments (Feng et al., 2004). The wavelength region (500 to 900 nm) contains wavelengths with pigment absorption features (Merzlyak et al., 2003) as well as the red edge (700 to 750 nm) (Mutanga and Skidmore, 2007). The results of this study indicated that changes in chlorophyll content could be used to validate hyperspectral imaging changes in apple seedlings exposed to water stress

7. Classification results

Table 9 shows the classification results from different machine learning methods using different input data. According to the table, when fused with GLCM texture features, the performance of these three models significantly increased. GWO-ELM and SSA-ELM performed better than ELM; where the combination of HSI and GLCM can reach a test accuracy of 65.73% and 63.64% respectively. In GWO-ELM, when HSI is fused with GLCM features, the precision rate is 0.6657, recall rate is 0.6561, and F1 value is 0.6518.

Table 9. Performance metrics

Machine learning algorithms	Different data input	Test accuracy	Precision	Recall	Loss	F1
ELM	HSI	64.94%	0.6957	0.5714	77.92%	0.6275
	GLCM	61.04%	0.6857	0.7059	74.03%	0.6957
	HSI+GLCM	64.94%	0.7826	0.6000	70.13%	0.6792
GWO-ELM	HSI	62.69%	0.6258	0.6369	82.09%	0.6270
	GLCM	53.85%	0.5363	0.5358	90.21%	0.5342
	HSI+GLCM	65.73%	0.6657	0.6561	61.54%	0.6518
SSA-ELM	HSI	58.21%	1	1	91.04%	1
	GLCM	47.56%	1	1	83.92%	1
	HSI+GLCM	63.64%	1	1	84.62%	1

7.1 Confusion matrix

A. ELM

Figure 10 shows the confusion matrix of ELM with three different input data. In the training data of HSI, 120 out of 160 normal samples were correctly identified (75%). But 26 overwatering samples were wrongly treated as normal, and 7 overwatering samples were wrongly treated as drought (26.6%). In the test data of HSI, 17 out of 25 samples were correctly identified as overwater (68%), but 4 normal samples were recognized as overwatered and 8 normal samples were treated as drought (42.9%).

In the training data of GLCM, 106 out of 154 normal samples were correctly recognized (68.8%), but 19 overwater samples were treated as normal, 41 drought samples were recognized as overwatered. In the test data of GLCM, 24 out of 35 samples were correctly identified as normal (68.6%), while 8 normal samples were treated as overwatered.

In the training data of HSI fused with GLCM, 113 out of 158 normal samples were correctly identified (71.5%), while 17 overwater samples were treated as normal, and 16 overwater samples were treated as drought (25.6%). In the test data of HSI fused with GLCM, 18 normal samples and 18 overwater samples were correctly identified.

(1) HSI

(2) GLCM

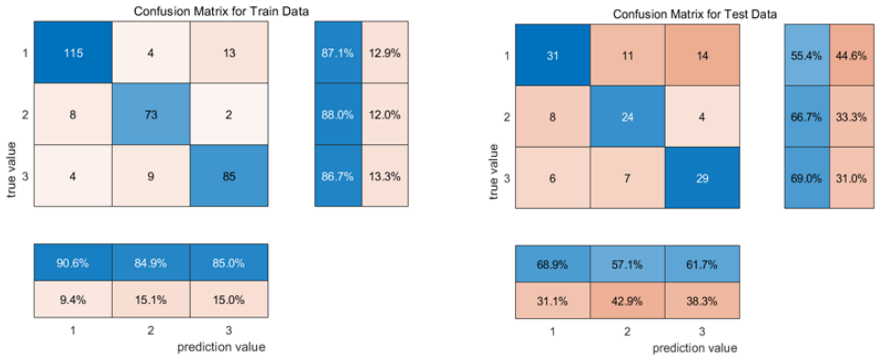
(3) HSI fused with GLCM

Figure 10. Confusion matrix for ELM using HSI (1), GLCM (2) and HSI fused with GLCM (3)

B. GWO-ELM

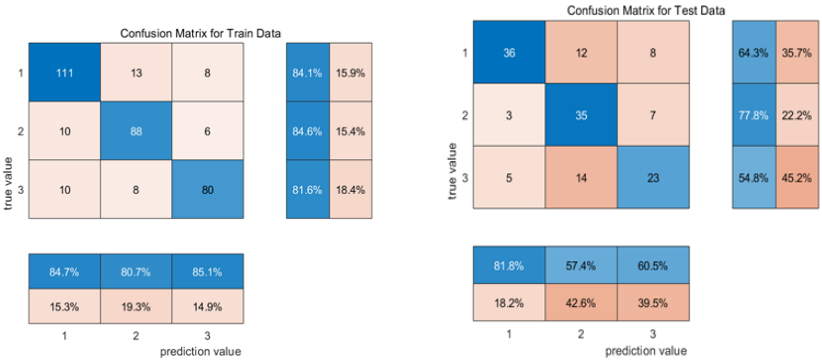
Figure 11 shows the confusion matrix for GWO-ELM using three inputs. In the training data of HSI, 115 out of 132 normal samples were correctly identified (87.1%), but 13 normal samples and 2 overwater samples were treated as drought (15.5). In the test data of HSI, 24 out of 42 samples were correctly identified as overwatered (57.1%), but 8 and 4 overwater samples were mis-treated as normal and drought respectively (33.3%).

In the training data of GLCM, 101 out of 126 samples were correctly labeled as normal (80.2%), but 9 overwater samples were wrongly taken as normal and 13 overwater samples as drought (21.2%). In the test data of GLCM, 26 out of 52 samples were correctly identified as overwater (50.0%), but 9 overwater and 9 drought samples were wrongly classified as normal (36.7%), In the training data of HSI fused with GLCM, 80 out of 98 drought samples were correctly identified (81.6%), but 8 normal samples and 6 overwater samples were wrongly labeled as drought (14.9%). In the test data of HSI fused with GLCM, 36 out of 44 samples were correctly identified as normal (81.8%), while 14 drought samples were wrongly taken as overwatered.



(a) HSI

(b) GLCM



(c) HSI fused with GLCM

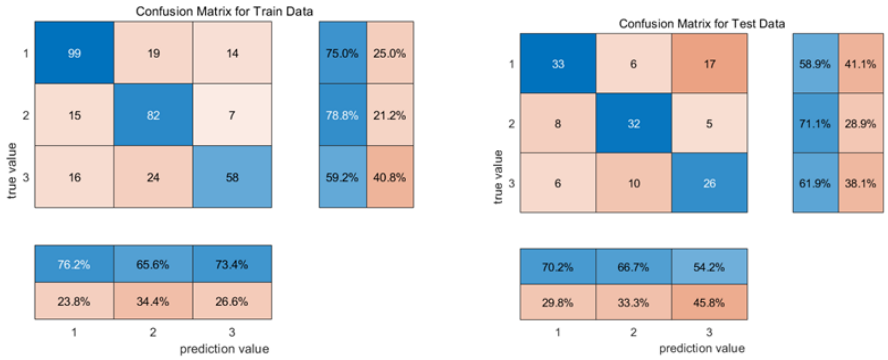
Figure 11. Confusion matrix for GWO-ELM using HSI (1), GLCM (2) and HSI fused with GLCM (3)

Figure 12 demonstrates the confusion matrix for SSA-ELM using HSI, GLCM and HSI fused with GLCM. Figure 12 demonstrates the confusion matrix for SSA-ELM using HSI, GLCM and HSI fused with GLCM. In the training data of HSI, 107 out of 135 samples were correctly identified as normal (79.3%), but 15 normal samples and 8 overwater samples were wrongly taken as drought (22.5%). In the test data of HSI, 20 out of 36 overwater samples were correctly identified (55.6%), but 10 drought samples and 7 drought samples were taken as normal and overwater respectively (40.5%).

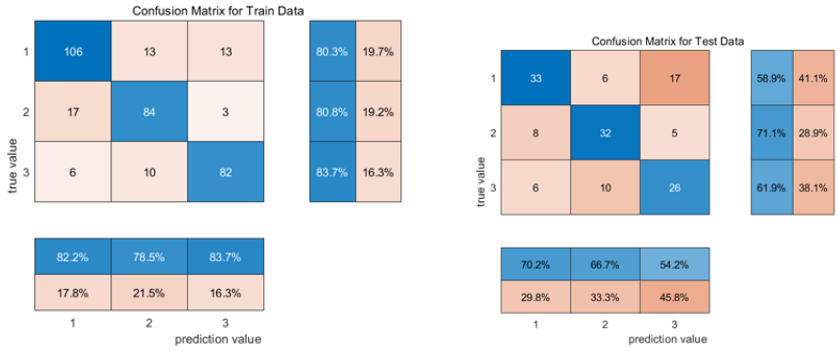
In the training data of GLCM, 99 out of 132 normal samples were correctly identified (75%), but 19 normal and 14 normal samples were wrongly taken as overwater and drought respectively (25%). In the test data of GLCM, 33 out of 56 normal samples were correctly identified (58.9%), but 4 drought samples were wrongly taken as normal, 13 overwater samples were wrongly taken as drought.

In the training data of HSI fused with GLCM, 106 out of 132 samples were correctly identified as normal (80.3%), but 17 overwater samples were wrongly taken as normal, 3 overwater samples were wrongly taken as drought (19.2%). In the test data of HSI fused with GLCM, 17 normal samples were wrongly taken as overwatered. Only 26 out of 42 drought samples were correctly identified (61.9%).

(a) HSI



(b) GLCM



(c) HSI fused with GLCM

Figure 12. Confusion matrix for SSA-ELM using HSI (1), GLCM (2) and HSI fused with GLCM (3)

8. Conclusion and discussion

This study explored the feasibility of using hyperspectral imaging techniques and texture features combined with machine learning to detect and identify water stress on apple leaves. The results proved that texture features fused with hyperspectral imaging performed better than merely using hyperspectral imaging for machine learning classification. The spectral behaviour was closely related to the changes in the leaf's physiological status caused by the water stress. Apple leaves under water stress produced a different specific and characteristic reflectance spectrum from regular leaves, which was difficult to observe by eyes in the early stage.

To identify the categories of water effectively, standard normal variate was used to preprocess the data. Subsequently, the multivariate stepwise linear regression method was applied to extract spectral fingerprint features, simplifying the training model and accelerating the calculation speed. These characteristic wavelengths correlated to spectral behaviours and cellular structure were selected in association with chlorophyll. The number of input spectral fingerprint features was only five, which are Band 85(564.26nm), Band 174(753.41nm), Band 156(714.65nm), Band 116(629.43nm), Band 194(796.79nm), Band 144(688.95nm) and Band 285(998.16nm), which decreased by more than 97% compared with the original spectral data. The absorbances are associated with the combinations of fundamental vibrations of C-H and O-H functional groups (Kamruzzaman et al.,2016). The 710–760 nm (red-edge) band and band around 700 nm are related to chlorophyll (ElMasry et al., 2007; Pacumbaba and Beyl, 2011). Also, chlorophyll absorptions are found in the visible region of the spectrum near 430, 460, 640, and 660 nm (Curran, 1989). Subtle absorption at 780nm and 980nm may be associated with the third and second overtones of O-H stretching, which may be relevant to water (ElMasry et al.,2012). Absorption at 940 nm is related to C-H third overtone(Kamruzzaman et al.,2016). According to Lu et al. (2022), a small reflectance peak appeared in the band around 550 nm, which is caused by the reflection of chlorophyll; a trough appeared in the red band around 680 nm, which is caused by the strong absorption of chlorophyll. Finally, diverse classifiers (ELM, GWO-ELM, SSA-ELM) were developed and evaluated on hyperspectral images and spectral-textural hyperspectral imaging analysis. For water stress detection, all classification models obtain relatively good accuracy (>90%) and robust performance, recognising the water stress before the signs of wilting or yellowing appear notably in the apple leaf. Also, the neural networks regarding textural analysis in fusion with hyperspectral imaging reached a satisfactory result. It should be mentioned that the classification result of the selected texture parameters fused with hyperspectral imaging outperformed those using only hyperspectral imaging data.

Recent research has made efforts to integrate spectral and image characteristics in order to obtain sufficient information and enhance the practical outcomes of hyperspectral imaging (HSI) applications (Wang et al., 2015; Ru et al., 2019). When compared to the accuracy achieved in identifying yellow rust in wheat leaves using spectral data alone, the accuracy of identification increases by 7.3% when utilizing both spectral and texture features (Guo et al., 2020). Furthermore, the incorporation of spectral, texture, and morphological features can lead to a 2% improvement in accuracy for the germ side and a 1.3% improvement for the endosperm side (Yang et al., 2015).

Previous studies have employed statistical analysis methods, such as the gray level co-occurrence matrix (GLCM) and morphological parameter calculations, to extract texture and morphological features from images. However, these methods are known for their complexity, time-intensive nature, reliance on spatial scale considerations, and the need for prior information (Sachar and Kumar, 2021). In recent years, deep learning has emerged as a powerful technique for feature extraction and has found widespread application, particularly in the field of imaging (Yu et al., 2020).

Our study confirmed the results of other research. In Zhang et al. (2022)'s analysis of distinguishing flue-cured tobacco samples, the texture features of flue-cured tobacco were selected by gray-level co-occurrence matrix (GLCM). Then the texture features and spectral features were normalised and fused. SPA chose twenty-five feature bands, and the classification accuracy using feature bands was decreased. The classification effect was improved significantly after the fusion of hyperspectral and texture characteristics. Zhu et al. (2017) investigated the feasibility and potentiality of presymptomatic detection of tobacco disease using hyperspectral imaging, combined with the variable selection method and machine learning classifiers. Four texture features, including contrast, correlation, entropy, and homogeneity were extracted according to grey-level co-occurrence matrix (GLCM). Additionally, different machine-learning algorithms were developed and compared to detect and classify disease stages with effective wavelengths, texture features and data fusion respectively. Yu et al. (2018) researched the prediction of polysaccharides and total sugar in wolfberry based on spectral and texture information from hyperspectral images. According to the absolute value of the average weight coefficient of each wavelength under the first three principal component images of 300 wolfberry samples, five characteristic wavelengths were selected (578. 42, 622. 97, 699. 61,809. 39, 865. 40 nm). *Graycomatrix* was used to extract the texture features of hyperspectral images corresponding to five characteristic wavelengths. After the screening, partial Least Squares

Regression was used to examine the correlation between polysaccharide content and spectra data. The correlation result was satisfactory.

Based on the spectral fingerprint features combined with the supervised classification algorithm developed in this study, a hyperspectral imaging detection system was potentially designed for identifying the categories and water stress levels in apple leaves. In future research, we will focus on (1) contrasting the spectral fingerprint features extracted in the field conditions (uncontrolled lighting) with those in the laboratory, (2) increasing more samples with different stress levels to generate and collect more image profiles of apple seedling leaves, therefore, making the algorithm more accurate and robust, (3) building a detailed time course to record the physiological and hyperspectral characteristics of leaves under different water treatments in differential time frames and (4) expanding our finding to other plant species, and applying the methodology in a real-world sorting environment for online and reliable plant or leaf water stress detection, to achieve excellent plant water and nutrition management.

This research introduces a new method, texture analysis of hyperspectral imaging and machine learning tools to identify apple seedling leaves under different water treatments. This otherwise cannot be identified by human eyes, except at the late stages, when the leaves are turning yellow or have wilted. The proposed method offers a new approach and the possibility of effective differentiation of leaves under different water treatments. Furthermore, leaves' chlorophyll corroborates our findings using HSI and machine learning.

Author Contributions: Conceptualization, W.R..; methodology, A.A.; software, A.A.; validation, A.A.; formal analysis, A.A.; investigation, A.A.; resources, A.A.; data curation, A.A.; writing—original draft preparation, A.A.; writing—review and editing, A.A.; visualisation, A.A.; supervision, W.R.; project administration, A.A.; funding acquisition, W.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: The data presented in this study is available on request from the corresponding author. The data is not publicly available due to district restrictions on data collection.

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

References:

1. Abdulridha, J., Ehsani, R., & De Castro, A. (2016). Detection and differentiation between laurel wilt disease, phytophthora disease, and salinity damage using a hyperspectral sensing technique. *Agriculture*, 6(4), 56.
2. Afifi, A, & Ashour, W. (2012). Image retrieval based on content using colour feature. *International scholarly research notices*.https://doi.org/10.5402/2012/248285.
3. Agjee, N. E., Mutanga, O., & Ismail, R. (2015). Remote sensing bio-control damage on aquatic invasive alien plant species. *South African Journal of Geomatics*, 4(4), 464-485.
4. Alexandridis, A., Chondrodima, E., & Sarimveis, H. (2016). Cooperative learning for radial basis function networks using particle swarm optimization. *Applied Soft Computing*, 49, 485-497.
5. AlSuwaidi A, Grieve B, and Yin H. (2017). Spectral-texture approach to hyperspectral image analysis for plant classification with SVMs. *2017 IEEE International Conference on Imaging Systems and Techniques* (IST), 1-6, doi: 10.1109/IST.2017.8261496.
6. Apple & Pear Australia Limited. (2023). Guidelines for irrigation management for apple and pear growers. https://apal.org.au/wp-content/uploads/2019/09/fo-ow-handout-09-sept-irrigation-guidelines.pdf
7. Ayala Martini D. (2018). Bachelor's Thesis. Universidad Pública de Navarra; Navarra, Spain: Automatización del Análisis de Imágenes Hiperespectrales para Identificación de Aptitud de Patatas.
8. Backhaus, A., Bollenbeck, F., & Seiffert, U. (2011, June). Robust classification of the nutrition state in crop plants by hyperspectral imaging and artificial neural networks. In 2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) (pp. 1-4). IEEE.
9. Bajji, M, Lutts, S, and Kinet, J. (2000). Physiologi cal changes after exposure to and recovery from PEG (Polyethylene Gilcol) -induced water deficit in callus cultures issued from durum wheat (*Triticum durum Desf*.)cultivars differing in drought resistance. *J.Plant Physiol*, 156, 75-83.
10. Bauriegel, E. & Herppich, W. (2014). Hyperspectral and chlorophyll fluorescence imaging for early detection of plant diseases, with special reference to *Fusarium spec.* Infections on wheat. *Agriculture*, 4(1), 32-57.
11. Behmann, J., Steinrücken, J., & Plümer, L. (2014). Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 98-111.

12. Bhagat, P., Choudhary, P., Manglem Singh, K. (2019). A comparative study for brain tumour detection in MRI images using texture features.Chapter 13 - A comparative study for brain tumour detection in MRI images using texture features. In *Advances in ubiquitous sensing applications for healthcare, Sensors for Health Monitoring*, 5, 259-287.
13. Bhugra, S., Chaudhury, S., & Lall, B. (2015). Use of leaf colour for drought stress analysis in rice. In *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics* (NCVPRIPG) (pp. 1-4). IEEE.
14. Bradford, K & Hsiao, T.（1982). Physiological response to moderate stress. In: Lange OL, Nobel PS, Osmond CB, Ziegler H,ets. *Encyclopedia of plant physiology. Physiological plant ecology. II. Water relations and carbon assimilation*. New York, Berlin, Heidelberg: Springer, 263-324.
15. Broomhead, D., & Lowe, D. (1988). Multivariable functional interpolation and adaptive networks, complex systems, vol. 2.
16. Canna. (2023). The most occurring forms of plant stress-part 2. https://www.cannagardening.com/most_occurring_forms_plant_stress_part_2
17. Chen, X., Pang, T., Tao, H., Lin, M., & Yang, H. (2017). Prediction of apple internal qualities using hyperspectral imaging techniques. In 2017 International Conference on Advanced Mechatronic Systems (ICAMechS) (pp. 450-455). IEEE.
18. Cheng L, Zhu X, Gao W, Wang L. and Zhao G. (2016). Hyperspectral Estimation of Phosphorus Content in Apple Leaves Based on Random Forest Model. *Journal of Fruit Science*, 33, 1219-1229.
19. Cheng, W., Sun, D., Pu, H., and Liu, Y. (2016). *LWT-Food Sci. Technol.* 72, 322-329.
20. Conners, R. & Harlow, C. (1980). A theoretical comparison of texture algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2,3: 204-222.
21. Curran, P., Dungan, J., Gholz, H. (1990). Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. *Tree Physiol*, 7, 33–48. doi: 10.1093/treephys/7.1-2-3-4.33.
22. Ding, X., Liu, F., Zhang, C., et al. (2015). Detection of SPAD value of rape leaves based on hyperspectral imaging technology. *Spectroscopy and Spectral Analysis*, 35(2), 486-491.
23. Duan, N., Bai, Y., Sun, H., Wang, N., Ma, Y., Li, M., ... & Chen, X. (2017). Genome re-sequencing reveals the history of apples and

supports a two-stage model for fruit enlargement. *Nature communications*, 8(1), 249.
24. Ellis, M. (2021). Apple tree irrigation- how to water an apple tree in the landscape. *Gardening*. https://www.gardeningknowhow.com/edible/fruits/apples/apple-tree-watering.htm.
25. ElMasry, G., Kamruzzaman, M., Sun, D.-W., Allen, P. (2012). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: A review. *Crit. Rev. Food Sci. Nutr*, 52, 999–1023.
26. ElMasry, G., Wang, N., ElSayed, A., Ngadi, M. (2007). Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *J. Food Eng*, 81, 98–107. doi: 10.1016/j.jfoodeng.2006.10.016
27. Escobar -Gutierrez, A J, Zipperlin , B, Carbonne, F , Moing , A ,and Gaudillere , JP. (1998). Photosynthesis, carbon partitioning and metabolite content during drought stress in the peach seedling.*Aust. J Plant Physiol*, 25, 197-205.
28. Feng, X., Zhan, Y., Wang, Q., Yang, X., Yu, C., Wang, H., ... & He, Y. (2020). Hyperspectral imaging combined with machine learning as a tool to obtain high‐throughput plant salt‐stress phenotyping. *The Plant Journal*, 101(6), 1448-1461.
29. Feng, X., Chen, X., Bao, A. M., Sun, L., Wang, D., & Ma, Y. (2004). Analysis of the cotton physiological change and its hyperspectral response under the water stress condition. *Arid Land Geography*, 2, 250-255.
30. Grann, T. & Boyer, J. (1990). Very high CO2 partially restores photosynthesis in sunflowers at low water potentials. *Planta*, 181,378-384.
31. Green, D. (2022). Leaves turning transparent or translucent: The causes explained. https://thegardenhows.com/leaves-turning-transparent-or-translucent-the-causes-explained/.
32. Gupta, S., Ibaraki, Y., & Pattanayak, A. (2013). Development of a digital image analysis method for real-time estimation of chlorophyll content in micro propagated potato plants. *Plant Biotechnol Rep*, 7, 91–97.
33. Guo, A., Huang, W., Ye, H., Dong, Y., Ma, H., Ren, Y., & Ruan, C. (2020). Identification of wheat yellow rust using spectral and texture features of hyperspectral images. *Remote Sensing*, 12(9), 1419.
34. Jiang, B., Wang, P., Zhuang, S., Li, M., Li, Z., & Gong, Z. (2018). Detection of maize drought based on texture and morphological features. *Computers and electronics in agriculture*, 151, 50-60.

35. Haralick, R., Shanmugam, K. & Dinstein, I. (1973). Textural feature for image classification. *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, 6, 610-621.
36. He, K., Zhang, X., Ren, S., Sun, J. (2016). Identify mappings in deep residual networks. *Computer Vision and Pattern Recognition*. arXiv:1603.05027.
37. He, J., Wang, J., & Liang, H. (1995). Effects of water stress on photochemical function and protein metabolism of photosystem II in wheat leaves. *Physiologia Plantarum*, 93, 771-777.
38. Hintzsche, H., & Stopper, H. (2012). Effects of terahertz radiation on biological systems. *Critical Reviews in Environmental Science and Technology*, 42(22), 2408-2434.
39. Holizadeh, A., Saberioon, M., BorVka, L., Wayayok, A., & Soom, M. (2017). Leaf chlorophyll and nitrogen dynamics and their relationship to lowland rice yield
40. for site-specific paddy management. *Inf Process Agric*, https://doi.org/10.1016/j.inpa.2017.08.002.
41. Huang, Y. (2017). Monitoring of nitrogen nutrition in wheat based on hyperspectral imaging.Nanjing Agricultural University.
42. Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, *70*(1-3), 489-501.
43. Humeau-Heurtier. (2019). Texture feature extraction methods: A survey. *IEEE Access*, 7, 8975-9000.
44. Hussain, H. A., Men, S., Hussain, S., Chen, Y., Ali, S., Zhang, S., ... & Wang, L. (2019). Interactive effects of drought and heat stresses on morpho-physiological attributes, yield, nutrient uptake and oxidative status in maize hybrids. *Scientific reports*, 9(1), 3890.
45. Jespersen, D., Zhang, J., & Huang, B. (2016). Chlorophyll loss associated with heat-induced senescence in bentgrass. *Plant Sci*, 249, 1–1.
46. Jia, Z., Ou, C., Sun, S., Wang, J., Liu, J., Li, M., Jia, S. & Mao, P. (2023). A novel approach using multispectral imaging for rapid development of seed pellet formulations to mitigate drought stress in alfalfa. *Computers and Electronics in Agriculture*, 212, 108136.
47. Jiang, H, Yoon, S., Zhuang, H., Wang, W., Li, Y. & Yang, Y. (2019). Spectrochim. *Acta*, Part A, 213,118-216.
48. Jin, X., Zarco-Tejada, P., Schmidhalter, U., Reynolds, M., Hawkesford, M., Varshney, R., Yang, T., Nie, C., Li, Z., Ming, B., Xiao, Y., Xie, Y., & Li, S. (2020). High-throughput estimation of crop traits: a review of ground and aerial phenotyping platforms.

IEEE Geosci Remote Sens Mag. https://doi.org/10.1109/MGRS.2020.2998816.
49. Jones, H. (2004). Application of thermal imaging and infrared sensing in plant physiology and ecophysiology. *Advances in Botanical Research*, 41, 107–163.
50. Kamruzzaman, M., Makino, Y., Oshita, S. (2016). Hyperspectral imaging for real-time monitoring of water holding capacity in red meat. *LWT-Food Sci. Technol*, 66, 685–691.
51. Katsoulas, N., C. Kittas, G. Dimokas, and Ch. Lykas. (2006). Effect of irrigation frequency on rose flower production and quality. *Bios. Eng*, 93, 237–244.
52. Kim, D., Zhang, H., Zhou, H., Du, T., Wu, Q., Mockler, T. & Berezin, M. (2015). Highly sensitive image-derived indices of water-stressed plants using hyperspectral imaging in SWIR and histogram analysis. *Sci Rep*, 5, 15919. https://doi.org/10.1038/srep15919.
53. Lakso, A., Geyer, A. & Carpenter, S. (1984). Seasonal osmotic relations in apple leaves of different ages. *J.Amer.Soc.Hort.Sci*, 109, 544-547.
54. Lauer, M. & Boyer, J. (1992). Internal CO2 measures directly in leaves: abscisic acid and low leaf water potential cause opposing effects. *Plant Physiology*, 98, 1010-1016.
55. Li, Z., Liu, G., Zhang, X. (2013). Review of researches on the Valsa canker of apple trees. *North. Fruits*, 4, 13.
56. Lim, P., & Nam, H. (2007). Aging and senescence of the leaf organ. *Journal of Plant Biology*, 50, 291-300.
57. Lippmalm, R. (2002). Neural Nets for Computing.Acoustics, speech, and signal processing.*IEEE International Conference on ICASSP*, 88, 1-6.
58. Liu, D., Pu, H., Sun, D., Wang, L. & Zeng, X. (2014). *Food Chem.*160, 330-337.
59. Liu, G., Wei, X., Zhang, S., et al. (2019). Analysis of epileptic seizure detection method based on improved genetic algorithm optimisation back propagation neural network [J]. *Journal of Biomedical Engineering*, 36(1), 24-32.
60. Li, B., Wang, R., Ma, J., & Xu, W. (2020). Research on crop water status monitoring and diagnosis by terahertz imaging. *Frontiers in Physics*, 8, 571628.
61. Li, R., Lu, Y., Peters, J.M.R. et al. (2020). Non-invasive measurement of leaf water content and pressure–volume curves using terahertz radiation. *Sci Rep*, 10, 21028. https://doi.org/10.1038/s41598-020-78154-z

62. Li, J., Xi, B., Li, Y., Du, Q., & Wang, K. (2018). Hyperspectral Classification Based on Texture Feature Enhancement and Deep Belief Networks. *Remote Sensing*, 10(3), 396. MDPI AG. Retrieved from http://dx.doi.org/10.3390/rs10030396
63. Liu, X., Zhang, Z., Meng, F., & Zhang, Y. (2022). Fault diagnosis of wind turbine bearings Based on CNN and SSA–ELM. Journal of Vibration Engineering & Technologies, 1-17.
64. Loucks J. (2021). How to grow apple tree seedlings. *SFGATE*. https://homeguides.sfgate.com/grow-apple-tree-seedlings-59390.html.
65. Lowe, A., Harrison, N., & French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant methods*, 13(1), 80.
66. Lu, Z., Zhang, Z., Zhang, X., Yang, W., Zhu, L., & Huang, C. (2022). Intelligent identification of cotton verticillium wilt based on spectral and image feature fusion.*Research Square*, 8(2). https://doi.org/10.21203/rs.3.rs-1903428/v1
67. Lu J, Zhou M, Gao Y, and Jiang H. (2018). Using hyperspectral imaging to discriminate yellow leaf curl disease in tomato leaves. *Precision Agriculture*, 19, 379-394. https://doi.org/10.1007/s11119-017-9524-7
68. Lu, C., & Zhang, J. (1999). Effects of water stress on photosystem II photochemistry and its thermostability in wheat plants. *Journal of Experimental Botany*, 50(336), 1199-1206.
69. Mack, J., Schindler, F., Rist, F., Herzog, K., Töpfer, R., & Steinhage, V. (2018). Semantic labeling and reconstruction of grape bunches from 3D range data using a new RGB-D feature descriptor. *Computers and Electronics in Agriculture*, 155, 96-102.
70. Mirjalili, S., Mirjalili, S., & Lewis, A. (2014). Grey wolf optimizer[J]. Advances in engineering software. 69, 46-61.
71. Levenberg-Marquardtbackpropagation.https://www.mathworks.com/help/deeplearning/ref/trainlm.html;jsessionid=39efd735b2ae310e3e285a17b507.
72. Ma, C., Zhang, H. H., & Wang, X. (2014). Machine learning for big data analytics in plants. *Trends in plant science*, 19(12), 798-808.
73. Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., Solovchenko, A. E., & Pogosyan, S. I. (2003). Application of reflectance spectroscopy for analysis of higher plant pigments. *Russian Journal of plant physiology*, 50, 704-710.
74. Mutanga, O., & Skidmore, A. K. (2007). Red edge shift and biochemical content in grass canopies. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(1), 34-42.

75. Ni, C., Li, Z., Zhang, X., Sun, X., Huang, Y., Zhao, L., ... & Wang, D. (2020). Online sorting of the film on cotton based on deep learning and hyperspectral imaging. Ieee Access, 8, 93028-93038.
76. Ozdemir A & Polat K. (2020). Deep Learning Applications for Hyperspectral Imaging: A Systematic Review. *Journal of the Institute of Electronics and Computer*, 2, 39-56. https://doi.org/10.33969/JIEC.2020.21004.
77. Pacumbaba R. O., Jr. & Beyl C. A. (2011). Changes in hyperspectral reflectance signatures of lettuce leaves in response to macronutrient deficiencies. *Adv. Space Res*, 48, 32–42. doi: 10.1016/j.asr.2011.02.020
78. Plaza, A., Martínez, P., Plaza, J., Pérez R. (2003). Spatial/spectral analysis of hyperspectral image data. In Advances in Techniques for Analysis of Remotely Sensed Data, 2003 IEEE Workshop on. IEEE.
79. Pearcy, M. & R, Hindle. (1989). A new method for the non-invasive three-dimensional measurement of human back movement. *Clinical Biomechanics*, 4(2), 73-39.ISSN 0268-0033.https://doi.org/10.1016/0268-0033(89)90042-9.
80. PIRSA. (2006). Irrigating Horticulture Crops with Reduced Water Supplies. Fact Sheet No. 18/06, www.pir.sa.gov.au/factsheets.
81. Ranney, T., Bassuk, N., & Whitlow, T. (1991). Osmotic adjustment and solute constituents in leaves and roots of water-stressed cherry (prunus)trees. *J Amer.Soc.Hort.Sci*,116 (4), 684-688.
82. Resonon Inc, Resonon. (2019). *SpectrononPro Manual*. Release 5.3.
83. Ristic, Z., Bukovnik, U., & Prasad, P. V. (2007). Correlation between heat stability of thylakoid membranes and loss of chlorophyll in winter wheat under heat stress. *Crop Science*, 47(5), 2067-2073.
84. Ru, C., Li, Z., & Tang, R. (2019). A hyperspectral imaging approach for classifying geographical origins of rhizoma atractylodis macrocephalae using the fusion of spectrum-image in VNIR and SWIR ranges (VNIR-SWIR-FuSI). *Sensors*, 19(9), 2045.
85. Sabzi, S., Pourdarbani, R., Rohban, M. H., Fuentes-Penna, A., Hernández-Hernández, J. L., & Hernández-Hernández, M. (2021). Classification of Cucumber Leaves Based on Nitrogen Content Using the Hyperspectral Imaging Technique and Majority Voting. *Plants*, 10(5), 898. https://doi.org/10.3390/plants10050898
86. Sachar, S., & Kumar, A. (2021). Survey of feature extraction and classification techniques to identify plant through leaves. *Expert Systems with Applications*, 167, 114181.
87. Schuerger, A., Capelle, G., Benedetto, J., Mao, C., Chi, N., & Evans, M. (2003). Comparison of two hyperspectral imaging and two laser-induced fluorescence instruments for the detection of zinc stress and

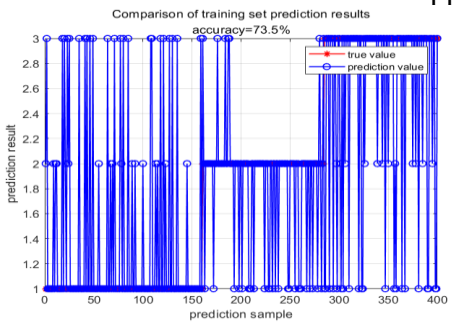
chlorophyll concentration in Bahia grass (Paspalum notatum Flugge). *Remote Sens Environ*, 2003, 84(4), 572–88.
88. Shao C. (2020). Data classification by quantum radical-basis-function networks. *Physics Review*, 102, 042418.
89. Singh, S. & Hemachandran, K. (2012). Content-Based Image Retrieval using Color Moment and Gabor Based Image Retrieval using Color Moment and Gabor Texture Feature.*IJCSI International Journal of Computer Science Issues*.ISSN (Online): 1694-0814. www.IJCSI.org. 9(5), 1.
90. Smart, R. & Bingham, G. (1974). Rapid estimates of relative water content. *Plant Physiology*, 53(2), 258-260.
91. Sun, J., Yang, L., Yang, X., Wei, J., Li, L., Guo, E. & Kong, Y. (2021). Using Spectral Reflectance to Estimate the Leaf Chlorophyll Content of Maize Inoculated With Arbuscular Mycorrhizal Fungi Under Water Stress. *Front. Plant Sci*, 12, 646173. doi: 10.3389/fpls.2021.646173.
92. Sun H, Zheng T, Liu N, et al. (2018). Hyperspectral image detection of the vertical distribution of chlorophyll content in potato leaves. *Transactions of the Chinese Society of Agricultural Engineering*, 34(1), 149-156.
93. Tan, K., Wang, R., Li, M., & Gong, Z. (2019). Discriminating soybean seed varieties using hyperspectral imaging and machine learning. Journal of Computational Methods in Sciences and Engineering, 19(4), 1001-1015.
94. The Home Depot. (2021). How to grow trees. https://www.homedepot.com/c/ah/how-to-grow-apples/9ba683603be9fa5395fab901f33a6977.
95. Tian, H.,Wang, C., Hao, M.,et al. (2013). Rapid positive potato dry matter content detection based on spectral differential filtering and multivariate calibration[J].*Inner Mongolia Agricultural University Journal* (Natural Science Edition), 34(5), 93-97.
96. Tumer, A. & Akkus, A. (2019). Application of Radial Basis Function Networks with Feature Selection for GDP per Capita Estimation Based on Academic Parameters. *International Journal of Computer Systems Science & Engineering*, 3, 145-150.
97. Wang, Y., Li, L., Shen, S., Liu, Y., Ning, J. & Zhang, Z., 2020. Rapid detection of the quality index of postharvest fresh tea leaves using hyperspectral imaging. Journal of the Science of Food and Agriculture. 100, 3803-3811.
98. Wang J. (2015). Research on nondestructive testing of Tan mutton tenderness based on spectral image information fusion technology[D].Yinchuan: Ningxia University.

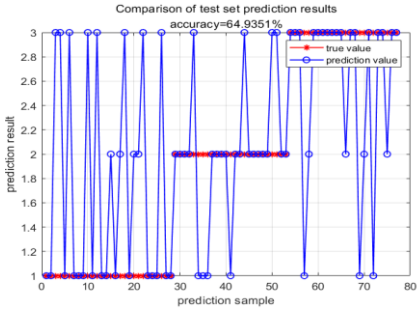
99. Wang, L., Liu, D., Pu, H., Sun, D. W., Gao, W., & Xiong, Z. (2015). Use of hyperspectral imaging to discriminate the variety and quality of rice. *Food analytical methods*, 8, 515-523.
100. Wang, Z., Liu, H., & Duan, H. (2006). Study on Hyperspectral Inversion Model of Soybean Chlorophyll Content. *Agriculture Engineering News*, 8, 16-21.
101. Water Potential. (2022). Advantages and disadvantages of the pressure chamber. *European Medical*. https://www.europeanmedical.info/water-potential/ii-advantages-and-disadvantages-of-the-pressure-chamber.html
102. Wang, Z., Quebedeaux, B. & Stutte, G. (1995). Osmoti c adjustment: Effect of water stress on carbohydrates in apple under water stress. *Aust. J. Plant Physiol*, 22, 747- 754(1995).
103. Wen, P., Li, H., Lei, H., & Zhang, F. (2021). High spectral data acquisition and pretreatment method of water-injection mutton. *Journal of Inner Mongolia Agricultural University* (Natural Science Edition).doi:10.16853 /j.cnki.1009-3575.2021.02.015.42(2).
104. Wen, X., Zhu, X., Yu, R., Xiong, J., Gao, D., Jiang, Y., Yang, G. (2019). Visualisation of chlorophyll content distribution in apple leaves based on hyperspectral imaging technology. *Agricultural Sciences*, 10, 6.
105. Williams, P. & Norris, K. (2001). Near-infrared technology in the agricultural and food industries. Vol. 2, edited by St. Paul, Minn. *American Assn. of Cereal Chemists.*
106. Willis Orchard Company. (2021). Apple seedlings. https://www.willisorchards.com/product/apple-seedlings#.YZjPaJfYtPY.
107. Wu, X., Wu, X., Huang, H., Zhang, F., & Wen, Y. (2023). Characterization of Pepper Ripeness in the Field Using Hyperspectral Imaging (HSI) with Back Propagation (BP) Neural Network and Kernel Based Extreme Learning Machine (KELM) Models. Analytical Letters, 1-16.
108. Xu, J., & Gowen, A. (2020). Spatial‐spectral analysis method using texture features combined with PCA for information extraction in hyperspectral images. *Journal of Chemometrics*, 34(2), e3132.
109. Yamaguchi, I., Cohen, J. D., Culler, A. H., Quint, M., Slovin, J. P., Nakajima, M., ... & Sakagami, Y. (2010). 4.02–plant hormones. *Comprehensive Natural Products II, Chemistry and Biology* (ed. Liu HW & Mander L.), 9-125.
110. Yamashita, H., Sonobe, R., Hirono, Y. et al. (2020). Dissection of hyperspectral reflectance to estimate nitrogen and

chlorophyll contents in tea leaves based on machine learning algorithms. *Sci Rep*, 10, 17360. https://doi.org/10.1038/s41598-020-73745-2

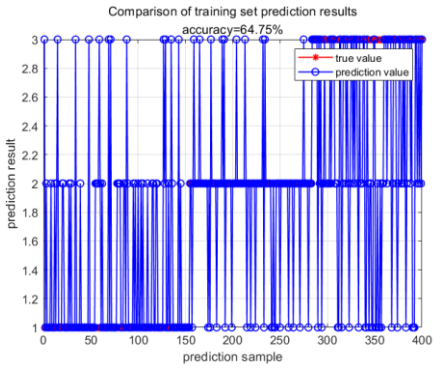
111. Yang, X., Hong, H., You, Z., & Cheng, F. (2015). Spectral and image integrated analysis of hyperspectral data for waxy corn seed variety classification. *Sensors*, 15(7), 15578-15594.
112. Yin, W., Chu, Z., Zhu, H., Zhao, Y.,& Yong, H. (2017). Application of near-infrared hyperspectral imaging to discriminate different geographical origins of Chinese wolfberries. *Plos One*.12(7), e0180534.
113. Yu, C., Li, F., Chang, C. I., Cen, K., & Zhao, M. (2020). Deep 2D convolutional neural network with deconvolution layer for hyperspectral image classification. In *Communications, Signal Processing, and Systems: Proceedings of the 2018 CSPS* Volume II: Signal Processing 7th (pp. 149-156). Springer Singapore.
114. Yu H, Wang R, Yin Y, Liu Y. (2018). Prediction of polysaccharide and total sugar in wolfberry based on spectral and texture information from the hyperspectral image. *Journal of Nuclear Agricultural Sciences*, 32(3), 0523-0531.
115. Yuan, X., Zhang, X., Zhang, N., Ma, R., He, D., Bao, H., & Sun, W. (2023). Hyperspectral Estimation of SPAD Value of Cotton Leaves under Verticillium Wilt Stress Based on GWO–ELM. *Agriculture*, *13*(9), 1779.
116. Zhang, H., Zhang, W., Zhang, Y. (2022). Research on classification method of flue-cured tobacco based on the fusion of hyperspectral and texture features[J]. *Acta Tabacaria Sinica*, 28(3). doi:10.16472/j.chinatobacco.2021.T0139.
117. Zhang, H., Ge, Y., Xie, X., Atefi, A., Wijewardane, N. K., & Thapa, S. (2022). High throughput analysis of leaf chlorophyll content in sorghum using RGB, hyperspectral, and fluorescence imaging and sensor fusion. *Plant Methods*, 18(1), 1-17.
118. Zhang, J., Zhang, W., Xiong, S., Song, Z., Tian, W., Shi, L. & Ma, X. (2021). Comparison of the new hyperspectral index and machine learning models for winter wheat leaf water content prediction. *Plant Methods*, 17, 34. https://doi.org/10.1186/s13007-021-00737-2.
119. Zhu, H., Chu, B., Zhang, C., Liu, F., Jiang, L., & He, Y. (2017). Hyperspectral imaging for presymptomatic detection of tobacco disease with successive projections algorithm and machine-learning classifiers. *Scientific reports*, 7(1), 4125.

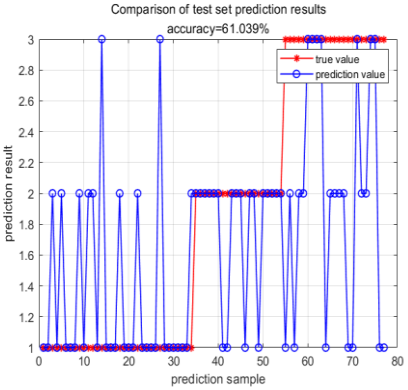
Appendix



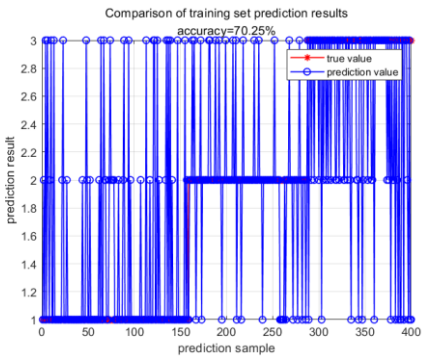


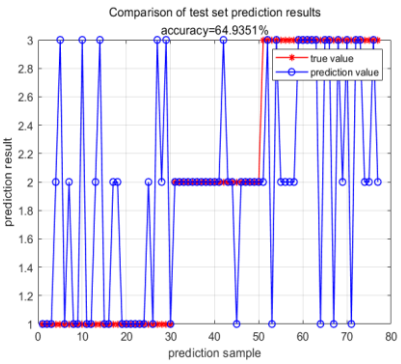
a. HSI





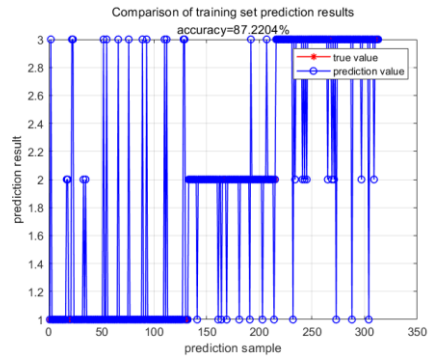
b. GLCM

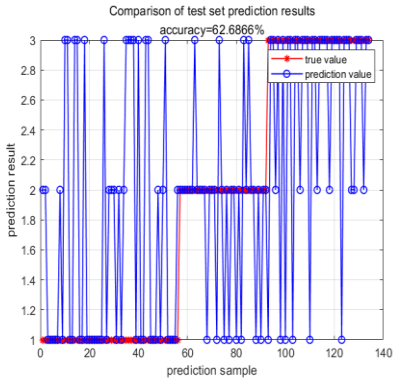




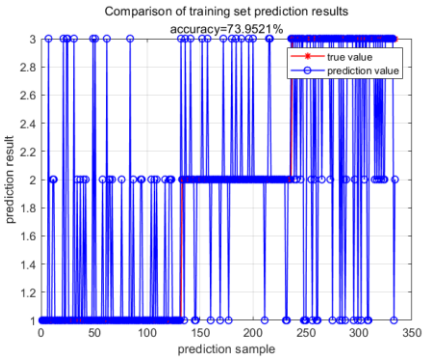
c. HSI fused with GLCM

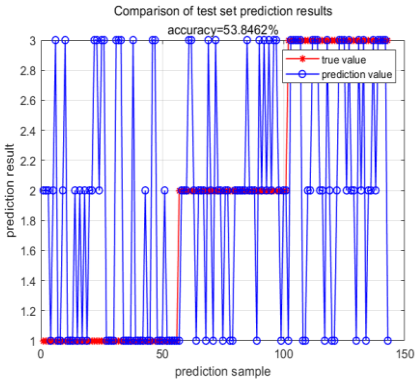
(1) ELM



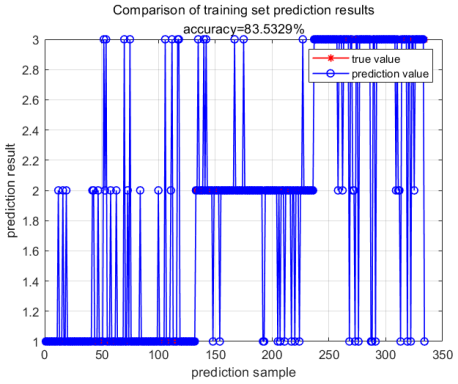


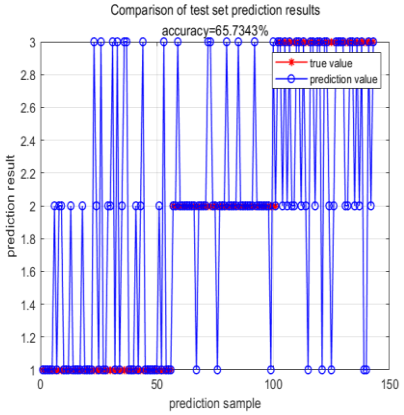
a.HSI





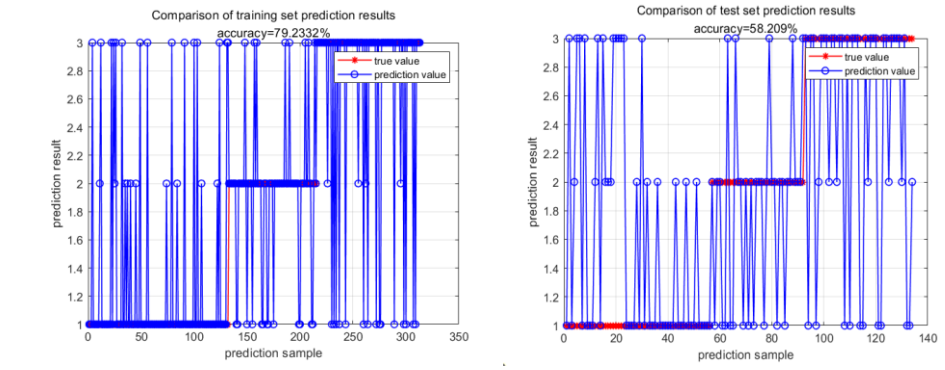
b.GLCM



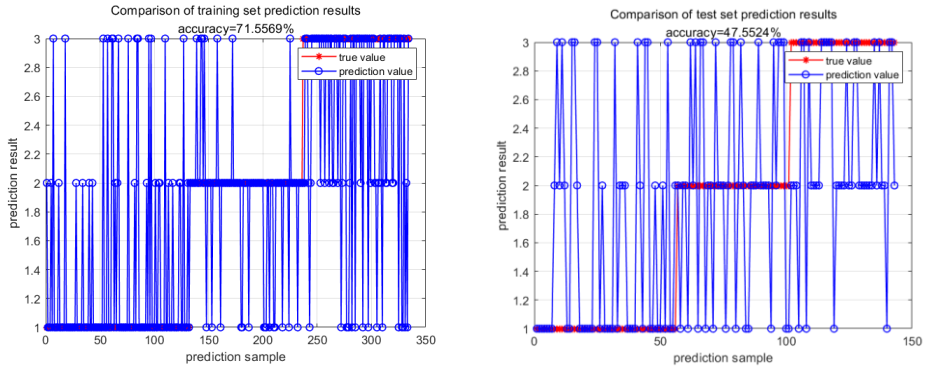


c.HSI+GLCM

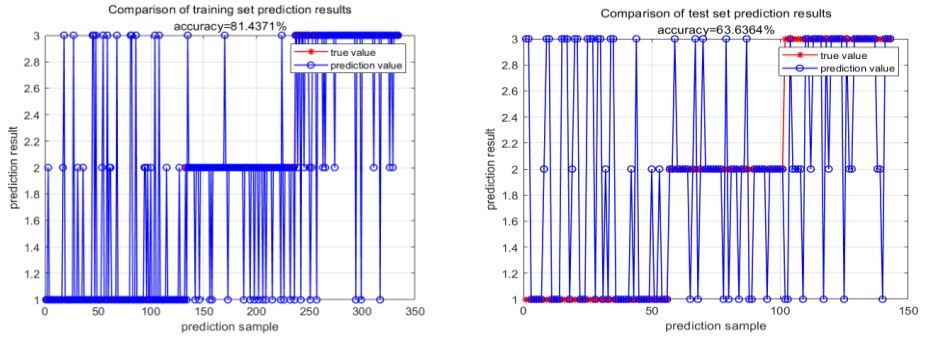
(2) GWO-ELM



a. HSI



b. GLCM



c.HSI+GLCM

(3) SSA-ELM

Figure 12. Classification result of ELM (1), GWO-ELM (2) and SSA-ELM (3) with an input of HSI (a), GLCM (b) and HSI+GLCM (c)