

Characterization and classification of urban weed species in northeast China using terrestrial hyperspectral images

Research Article

Cite this article: Wang J, Chen G, Ju J, Lin T, Wang R, Wang Z (2023) Characterization and classification of urban weed species in northeast China using terrestrial hyperspectral images. *Weed Sci.* **71**: 353–368. doi: [10.1017/wsc.2023.36](https://doi.org/10.1017/wsc.2023.36)

Received: 1 March 2023

Revised: 8 June 2023

Accepted: 23 June 2023

First published online: 1 August 2023

Associate Editor:

Ramon G. Leon, North Carolina State University

Keywords:


Convolutional neural network; hyperspectral library; physiological activity

Corresponding author:

Jinfeng Wang; Email: jinfeng_w@126.com;

Zhentao Wang; Email: 15770085650@163.com.

*These authors contributed equally to this work.

Jinfeng Wang^{1,*} , Guoqing Chen^{2,*}, Jinyan Ju³, Tenghui Lin², Ruidong Wang² and Zhentao Wang⁴

¹Professor, Vice President, College of Engineering, Northeast Agricultural University, Harbin, China; ²Graduate Student, College of Engineering, Northeast Agricultural University, Harbin, China; ³Research Agronomist, Mechanical Engineering College, Heilongjiang University of Science and Technology, Harbin, China and ⁴Associate Agronomist, College of Engineering, Northeast Agricultural University, Harbin, China

Abstract

Weeds contribute to biodiversity and a wide range of ecosystem functions. It is crucial to map different weed species and analyze their physiological activities. Remote sensing techniques for plant identification, especially hyperspectral imaging, are being developed using spectral response patterns to vegetation for detection and species identification. A library of hyperspectral images of 40 urban weed species in northeast China was established in this study. A terrestrial hyperspectral camera was used to acquire 435 hyperspectral images. The hyperspectral information for each weed species was extracted and analyzed. The spectral characteristics and vegetation indices of different weeds revealed the differences between weed species in the cities of northeast China and indirectly characterized the growth and physiological activity levels of different species, but could not effectively distinguish different species. Five methods—first derivative spectrum (FDS), second derivative spectrum (SDS), standard normal variate (SNV), moving averages (MA), and Savitzky-Golay (SG) smoothing—were used to pretreat the spectral curves to maximize the retention of spectral characteristics while removing the influence of noise. We investigated the application of a convolutional neural network (CNN) with terrestrial hyperspectral remote sensing to identify urban weeds in northeast China. A CNN classification model was established to distinguish weeds from the hyperspectral images and demonstrated a test accuracy of 95.32% to 98.15%. The accuracy of the original spectrum was 97.45%; SNV had the best accuracy (98.15%) and SG was the least accurate (95.32%). This provides a baseline for understanding the hyperspectral characteristics of urban weed species and monitoring their growth. It also contributes to the development of a hyperspectral imaging database with global applicability.

Introduction

Weeds play an important role in enriching biodiversity and ecosystem services, such as by promoting biological evolution, improving the ecological environment, and maintaining climate stability. Evidence for this has been gathered over a long time and is being increasingly reported (Gage and Schwartz-Lazaro 2019; Guo et al. 2018; Korres et al. 2016). The benefits of weeds vary according to their species, growth status, and habitat type. These factors are also the fundamental criteria by which weeds reflect the biodiversity and quality of ecological services (Bretagnolle and Gaba 2015). Therefore, the ability to map weeds and analyze their physiological activity such as photosynthesis and respiration is very significant (Ferreira et al. 2016; Rozenberg et al. 2021). To monitor weeds' physiological activity, data on their biophysical circumstances must be gathered and updated. However, detecting weed physiological activity and species types using traditional visual recognition technology can be challenging (Adegbenjo et al. 2020; Prati et al. 2019). At the same time, many of the costs associated with machine vision are too high, including the construction of complex systems and expensive field surveys (Tzamali et al. 2006), and machine vision provides only limited spectral information as it only records information using three broad bands (625 to 740, 570 to 585, and 492 to 577 nm) (Zhang et al. 2019b).

Remote sensing techniques can map the distribution of weeds and effectively monitor their changes in complex urban environments (Segarra et al. 2020). Hyperspectral imaging techniques, in particular, are evolving from spectral response models to allow species identification and vegetation monitoring. Many studies have used hyperspectral data obtained by satellite (Khaliq et al. 2019), airborne (Maes and Steppe 2019), and ground-based (Behmann et al. 2018) cameras for species identification, vegetation monitoring, and crop classification (Mariotto et al. 2014). However, many factors affect satellite-based detection accuracy, including

altitude, atmospheric conditions, and the systematic errors related to orbit angle and solar radiation pressure known to affect the Chinese BeiDou Navigation Satellite System, geostationary orbit satellites, and inclined geostationary orbit satellites (Guo et al. 2016). Although a variety of sensors can be used to improve accuracy, space missions have limited availability, making it difficult to obtain specific hyperspectral data (Zhong et al. 2019). Airborne acquisition techniques, such as unmanned aerial vehicles (UAVs), have the advantage of lower flight altitudes; however, they do not produce accurate correction signals (Rossini et al. 2015). Therefore, a ground-based hyperspectral technique has been adopted by researchers to acquire high-spatial-resolution images with minimal atmospheric effects (Katkovsky et al. 2018).

Databases of plant-specific information have been established, providing important resources for various fields of research. The International Water Management Institute provides a wetland vegetation hyperspectral database that includes a broad-spectrum database for coastal wetland vegetation communities under different bioclimatic, soil, and disturbance conditions. This facilitates the monitoring of detailed changes in wetland vegetation structure and species composition (Zomer et al. 2009). Manjunath et al. (2014) established a spectral database of Himalayan vegetation species, including the spectra of leaves and branching canopies of various vegetation types, and calculated 22 vegetation indices such as the normalized difference vegetation index (NDVI), simple ratio, and soil-adjusted vegetation index (SAVI). At the same time, the biochemical parameters such as chlorophyll *a*, chlorophyll *b*, and protein were measured using standard methods, and the close correlation between them was studied. The information provided in spectral libraries can be used to explore the chemical composition, growth behavior, and ecological environment of Himalayan vegetation (Yang et al. 2021). Khder and Yones (2021) established an innovative spectral library of common wild plants on the northwest coast of Egypt. It analyzed the characteristics of 27 wild vegetation species, such as saltbush (*Atriplex halimus* L.) and thyme [*Thymus capitatus* (L.) Hoffsgg. & Link], and strengthened the observation of wild vegetation through remote sensing used to identify common wild plant species. Although existing spectral libraries cover a wide range of spectral information, there is limited information on the spectral features of weed species in different cities. Many species native to the middle temperate zones, such as musk thistle (*Carduus nutans* L.) and sand cinquefoil (*Potentilla supina* L.), are unlikely to be found in these spectral libraries. Northeast China is renowned for its distinctive landscape features and has a mid-temperate continental monsoon climate with a great diversity of weed species. Most weeds grow in an intense urban environment (Brunzel et al. 2009; Von der Lippe and Kowarik 2007) with long winters and have short growth cycles. Therefore, this study analyses the abundant weed species of northeast China to contribute to the development of a hyperspectral imaging database. This enriches studies on weed species in urban areas of northeast China with hyperspectral data.

In light of this, the process of this study was: (1) a hyperspectral library of urban weeds in northeast China was established, including 435 hyperspectral images of 40 species and 23 families; (2) the spectral profiles and vegetation indices were used to indirectly characterize the growth and physiological activity of weed species; (3) five different pretreatments (first derivative spectrum [FDS], second derivative spectrum [SDS], standard normal variate [SNV], moving averages [MA], and Savitzky-Golay [SG] smoothing) were used to maximize the retention of spectral

features, combined with a convolutional neural network (CNN) built to identify different weed species.

Materials and Methods

Study Area and Target Weed Selection

This study was conducted in Heilongjiang Province, northeast China. The selected weed species (Figure 1) were located at seven different sites: (1) Harbin City (44.067°N to 46.667°N, 125.7°E to 130.167°E), (2) Qiqihar (45°N to 48°N, 122°E to 126°E), (3) Hegang City (47.067°N to 48.35°N, 129.65°E to 132.517°E), (4) Shuangyashan City (46.333°N to 47.9°N, 130.9°E to 131.783°E), (5) Daqing City (45.767°N to 46.917°N, 124.317°E to 125.2°E), (6) Jiamusi City (45.933°N to 48.467°N, 129.483°E to 135.083°E), and (7) Suihua City (45.05°N to 48.033°N, 124.217°E to 128.5°E). To construct representative data sets and establish robust models to characterize and identify common weeds in northeastern cities, weed selection criteria included but were not limited to weed species, size, similarity of appearance, proximity to roads, and traffic flow on nearby roads.

Image Acquisition Plan

To obtain high-spatial-resolution data with minimal atmospheric disturbance, a field acquisition method was adopted. A new, portable, handheld hyperspectral camera, SPECIM IQ (model SN: 190-1100381, SPECIM, Spectral Imaging, Oulu, Finland), was used to obtain hyperspectral images of weeds. This camera integrates hyperspectral data acquisition, analysis, processing, and visualization of the results. It weighs 1.3 kg and measures 207 × 91 × 74 mm (lens: 125.5 mm). The camera is supported by Specim IQ Studio software and can be remotely connected via universal serial bus or wireless fidelity for the remote control of all camera functions except focusing. SPECIM IQ takes full hyperspectral images without the need for external movement. In the optical and near-infrared wavelengths of the electromagnetic spectrum (approximately 397 to 1,003 nm), the camera obtains hyperspectral images containing a 204 spectral narrow band with a spectral resolution of about 7 nm. The camera provides 512 × 512 pixels of spatial sampling, covering a field of view of 31° × 31°. When the camera samples 1 m away from the target, it captures an area of 0.55 × 0.55 m, and the peak signal-to-noise ratio is >400:1.

To maintain the accuracy of weed data sampling under different time of day and weather conditions, a white reference board was used for calibration (Wang et al. 2023). Irradiance fluctuations brought on by atmospheric factors and changes in solar illumination are the most important factor in field reflectance measurements. Most reflectance measurements are still single field-of-view measurements, so the time between the target and reference measurements should be as short as possible to avoid any potential atmospheric effects. This improvement is also necessary because all weed samples and white reference frame measurements must have the same lighting characteristics. Any change in distance between the camera and weed samples and in-field light intensity will affect the reflectance measurements. To ensure the accuracy of the reflectance measurements of all samples, quick atmospheric correction (QUAC) must be considered (Zhang et al. 2021). QUAC aims to eliminate atmospheric, light, radiation, and surface temperature effects on the spectral reflectance of samples to obtain more accurate reflectance data. In addition, it does not need to provide the specific location and time of sample measurement, and it automatically collects spectral information on the samples and

Species, abbreviation, and family	Photos	No. of images	Species, abbreviation, and family	Photos	No. of images
<i>Artemisia annua</i> L. Asteraceae		12	<i>Agrostis canina</i> L. Poaceae		12
<i>Artemisia capillaris</i> Thunb. Asteraceae		10	<i>Arctium lappa</i> L. Asteraceae		12
<i>Agastache rugosa</i> (Fisch. & C.A.Mey.) Kuntze Labiatae		11	<i>Chenopodium album</i> L. Chenopodiaceae		10
<i>Cirsium arvense</i> (L.) Scop. var. <i>integrifolium</i> Wimm. & Grab. Asteraceae		10	<i>Capsella bursa-pastoris</i> (L.) Medik. Brassicaceae		11
<i>Coryza canadensis</i> (L.) Cronquist. Asteraceae		12	<i>Chelidonium majus</i> L. Papaveraceae		10
<i>Carduus nutans</i> L. Asteraceae		11	<i>Catolobus pendulus</i> (L.) Al-Shehbaz Brassicaceae		10
<i>Crepis tectorum</i> L. Asteraceae		10	<i>Duchesnea indica</i> (Andrews) Teschem. Rosaceae		13
<i>Draba nemorosa</i> L. Brassicaceae		9	<i>Fallopia multiflora</i> (Thunb.) Haraldson Polygonaceae		10
<i>Geranium carolinianum</i> L. Geraniaceae		10	<i>Glechoma hederacea</i> L. Lamiaceae		11
<i>Galium spurium</i> L. Rubiaceae		10	<i>Hosta plantaginea</i> (Lam.) Asch. Asparagaceae		11

Figure 1. Descriptions and photos of selected weed species collected in Heilongjiang Province, northeast China.





















<i>Humulus scandens</i> auct. non (Lour.) Merr. Cannabaceae		10	<i>Hylotelephium spectabile</i> (Boreau) H. Ohba Crassulaceae		11
<i>Iris lactea</i> Pall. Iridaceae		12	<i>Lepidium apetalum</i> Willd. Brassicaceae		13
<i>Lagopsis supina</i> (Steph. ex Willd.) Ikonn.-Gal. Lamiaceae		10	<i>Oenanthe javanica</i> (Blume) DC. Apiaceae		8
<i>Polygonum aviculare</i> L. Polygonaceae		12	<i>Plantago depressa</i> Willd. Plantaginaceae		11
<i>Platycodon grandiflorus</i> (Jacq.) A. DC. Campanulaceae		10	<i>Persicaria lapathifolia</i> (L.) Gray Polygonaceae		13
<i>Potentilla supina</i> L. Rosaceae		13	<i>Rumex acetosa</i> L. Polygonaceae		11
<i>Stellaria media</i> (L.) Vill. Caryophyllaceae		11	<i>Salvia × sylvestris</i> L. (pro sp.) [<i>nemorosa</i> × <i>pratensis</i>] Labiatae		11
<i>Taraxacum mongolicum</i> Hand.-Mazz Asteraceae		11	<i>Trigonotis peduncularis</i> (Trevir.) Benth. ex Baker & S. Moore Boraginaceae		11
<i>Trifolium repens</i> L. Fabaceae		11	<i>Urtica angustifolia</i> Fisch. ex Hornem. Urticaceae		11
<i>Urtica laetevirens</i> Maxim. Urticaceae		9	<i>Viola philippica</i> Cav. Violaceae		12
Total		435			

Figure 1. (continued).

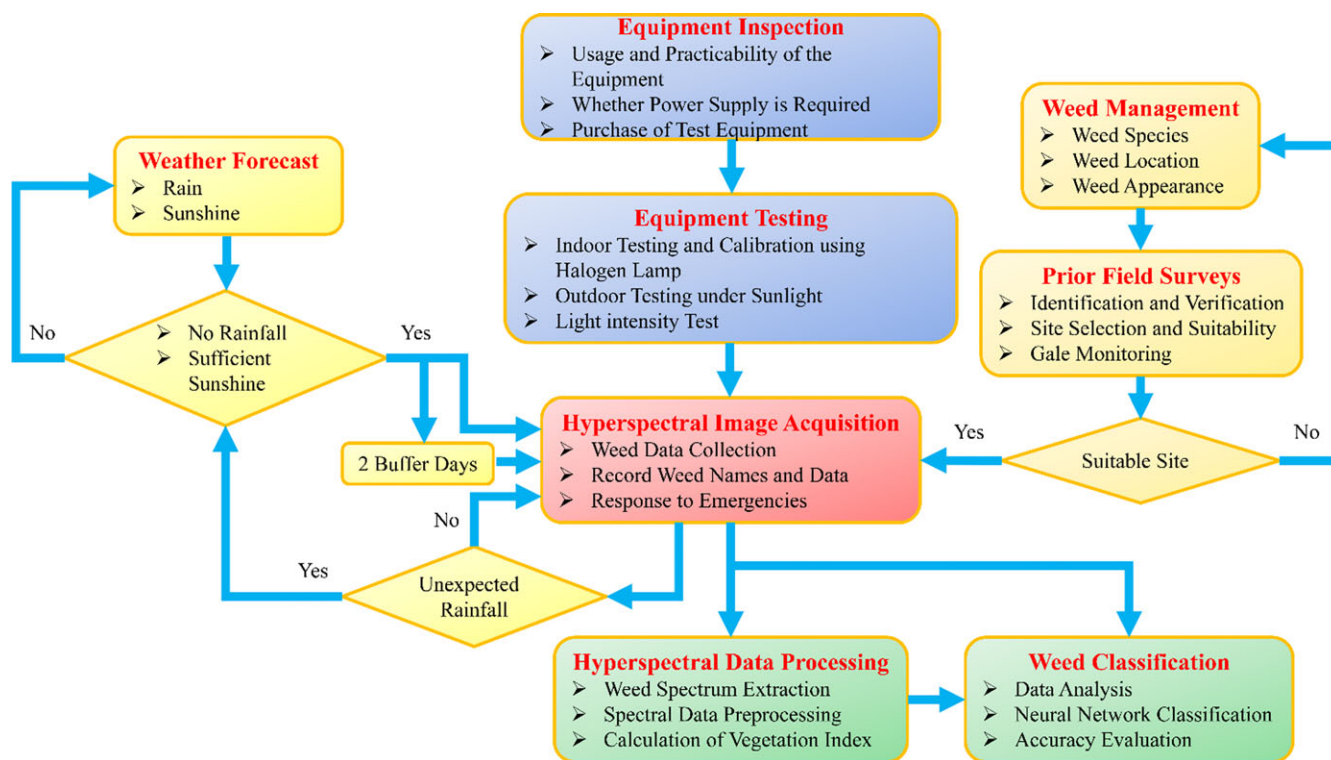


Figure 2. Flowchart of the hyperspectral detection process used in this study.

other background materials in the image, so as to invert the true reflectance of the samples. At present, QUAC supports a hyperspectral range of 400 to 2,500 nm, which is suitable for hyperspectral imaging in various complex environments and bands.

We collected 435 hyperspectral images of 40 weed species. The specific information on the collected weed samples and corresponding pictures are shown in Figure 1. A detailed flowchart of the hyperspectral image acquisition and data analysis processes is shown in Figure 2.

Methods

For spectral processing, ENVI v. 5.3 software (Exelis Visual Information Solutions, Boulder, CO, USA) was used to extract spectral data from the hyperspectral images obtained by SPECIM IQ. To improve the prediction accuracy and deal with the inevitable negative impacts of environmental and system noise, the 981 to 1,003 nm band was eliminated, as it was heavily affected by noise. Therefore, the band for our actual pretreatment and spectral analysis is 397 to 981 nm. In addition, before the neural network classification model was established, five methods (FDS, SDS, SNV, MA, and SG) were used to preprocess the hyperspectral data. The five pretreatment methods, kernel principal component analysis (KPCA), and classification models were run in MATLAB R2022a software (MathWorks, Natick, MA, USA); principal component analysis (PCA) was run in IBM SPSS Statistics 27 software (International Business Machines, Armonk, NY, USA).

Smoothing

The SG smoothing and MA (smoothing window of five points) techniques were adopted to reduce the influence of high-frequency noise and baseline translation noise while retaining the unique

characteristics of the samples (Chen et al. 2012; Zhang et al. 2020). SG is the most commonly used smoothing algorithm. It obtains a best estimate of the spectral smoothing points by weighted filtering and polynomial fitting of the data within a window of a certain width. By changing the size of the window and the order of the polynomial function, one can change the degree of SG smoothness (Chen et al. 2011). On the other hand, the MA approach smooths the window with an odd width and averages the data within the window to remove noise. MA smooths single-sample data rather than moving and scaling all sample data and can accurately identify the noise in each segment of the spectrum and the unique features of the spectral curve. In addition, to avoid the smoothed spectral data being lower than the original data, the spectral data were zeroed before data processing and then fit.

Differentiation

The core of this study is to distinguish the characteristics of various weeds and improve the accuracy of the neural network, which is often affected by baseline translation, smoothing background interference, and spectral mixing and overlap (Cheoi et al. 2020; Hu et al. 2019). The FDS and SDS of spectral reflectance data are commonly used to avoid the influences of these factors (Yeow and Leong 2005) and are widely used in vegetation detection. Previous studies have reported that differential spectroscopy can further improve the ability to identify vegetation using spectral data (Qian et al. 2013), reflect the waveform changes caused by light absorption by chlorophyll and other substances in plants, and reveal the characteristics of the peak in the spectrum (Becker et al. 2005). Therefore, to highlight the subtle absorption features in weed data, this study used FDS and SDS to preprocess the original spectral data, so as to improve the classification accuracy of the neural network.

Standard Normal Variate Transformation

In this study, weed hyperspectral data were collected in a field environment, which was easily affected by surface scattering and optical path changes. To eliminate the influence of these factors and reflect the spectral characteristics of weeds, SNV transformation was used to pretreat the spectral data of weeds before using the neural network classification model to identify them. The difference between this algorithm and others is that SNV processes one spectrum and can better extract the spectral characteristics of weeds (Kachrimanis et al. 2007). The formula for SNV is as follows:

$$R_{i,SNV} = \frac{R_{i,k} - \bar{R}_i}{\sqrt{\frac{\sum_{k=1}^m (R_{i,k} - \bar{R}_i)^2}{m-1}}} \quad [1]$$

where $R_{i,SNV}$ is the transformed spectrum, $\bar{R}_i = \frac{\sum_{k=1}^m R_{i,k}}{m}$; i is the number of spectral samples; $k = 1, 2, 3, \dots, m$, where k is the number of wavelength points; and $m = 1, 2, 3, \dots, n$, where n is the sample size.

Exploratory Data Analysis

Hyperspectral images contain an abundance of high-dimensional data, so it is very difficult to conduct qualitative analysis of spectral data directly, and it is difficult to mine the fine features in spectral data. PCA is one of the most commonly used unsupervised chemometric tools to explore hidden information in large amounts of data (Zhang et al. 2013). It enables an overview of complex multivariate data and has been widely adopted to process hyperspectral imaging data (Bro and Smilde 2014). PCA converts a group of possibly correlated variables into a group of linearly uncorrelated variables through orthogonal transformation, which are called principal components (PCs). After PCA, spectral data usually generate several PCs to reveal the internal structure of multiple variables, so as to retain and extract the original spectral curve features as much as possible. In PCA, for m hyperspectral weed samples $\{x^1, x^2, \dots, x^m\}$, each sample i had n -dimensional characteristics $X^i = (x_1^i, x_2^i, \dots, x_n^i)$. The covariance matrix corresponding to each dimension feature $X_j = (x_j^1, x_j^2, \dots, x_j^m)$, ($j = 1, 2, \dots, m$) has m eigenvalues λ_j and eigenvector u_j , and $\{(\lambda_j^1, u_j^1), (\lambda_j^2, u_j^2), \dots, (\lambda_j^k, u_j^k)\}$ can be obtained by selecting the first k largest eigenvalues. For each dimension feature, the new variable of m samples $\{x_j^1, x_j^2, \dots, x_j^m\}$ after projection is $\{y_j^1, y_j^2, \dots, y_j^k\}$. The formula for calculating the new variable is as follows:

$$\begin{pmatrix} y_j^1 \\ y_j^2 \\ \vdots \\ y_j^k \end{pmatrix} = \begin{pmatrix} (x_j^1, x_j^2, \dots, x_j^m) \bullet u_j^1 \\ (x_j^1, x_j^2, \dots, x_j^m) \bullet u_j^2 \\ \vdots \\ (x_j^1, x_j^2, \dots, x_j^m) \bullet u_j^k \end{pmatrix} \quad [2]$$

where $\{x_1^i, x_2^i, \dots, x_n^i\}$, ($i = 1, 2, \dots, k$) constitutes a PC, and the wavelength of maximum u_{max}^i and minimum u_{min}^i ($i = 1, 2, \dots, k$) represents the characteristic wavelength. In this study, $k = 3$, that is, three principal components (PC1, PC2, and PC3) are extracted.

KPCA

In this study, KPCA was used to visualize weed spectral data. KPCA is similar in principle to PCA, but it aims to capture higher-order statistics and deal with complex nonlinear features that are widespread in hyperspectral images. It is a method of processing nonlinear data using kernel mapping; the original data are mapped to a high-dimensional space using a kernel function, and then the corresponding linear operation is carried out in the space. Here, the kernel is $k(x_i, x_j)$:

$$k(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad [3]$$

where x_i, x_j ($i, j = 1, 2, 3, \dots, m$) represents the original data set; $\phi(x_i), \phi(x_j)$ ($i, j = 1, 2, 3, \dots, m$) represents the high-dimensional data set after the mapping function $\phi(x)$; and m represents the total amount of data.

After the original data are mapped to the high-dimensional space, a dimensionality reduction operation is carried out on the original data according to the eigenvalue λ and eigenvector v of the covariance matrix M of the high-dimensional data set to determine the PC. The formula is as follows:

$$Mv = \lambda v, \text{ subject to } \|v\|^2 = 1 \quad [4]$$

where $M = \begin{pmatrix} C(x_1, x_1) & C(x_1, x_2) & \dots & C(x_1, x_m) \\ C(x_2, x_1) & C(x_2, x_2) & \dots & C(x_2, x_m) \\ \vdots & \vdots & \ddots & \vdots \\ C(x_m, x_1) & C(x_m, x_2) & \dots & C(x_m, x_m) \end{pmatrix}$,

and $C(x_i, x_j)$, ($i, j = 1, 2, 3, \dots, m$) represents the covariance of $\phi(x_i)$ and $\phi(x_j)$.

Because the classes in hyperspectral image data are usually very close to gaussian distribution (Huber-Lerner et al. 2016), this study adopts a gaussian kernel function for KPCA, and the formula is as follows:

$$k(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad [5]$$

where σ represents the standard deviation of x_i and x_j .

Vegetation Index

The vegetation index is an important index of vegetation growth. Over the past 40 years, many spectral vegetation indices have been developed, such as the simple vegetation index, differential environmental vegetation index, normalized differential vegetation index (NDVI), greenness vegetation index, and soil-adjusted vegetation index (SAVI) (Giovos et al. 2021). An ideal vegetation index should contain information that maximizes the specific physical characteristics of the plants (Ji et al. 2014). NDVI is the most widely used vegetation index for estimating the physical and growth status of vegetation (Abbas et al. 2021). It adopts the mid- and near-infrared bands of the hyperspectral spectrum. The formula is as follows:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad [6]$$

where ρ_{NIR} represents the spectral value of the near-infrared band and ρ_{RED} stands for the infrared spectrum value.

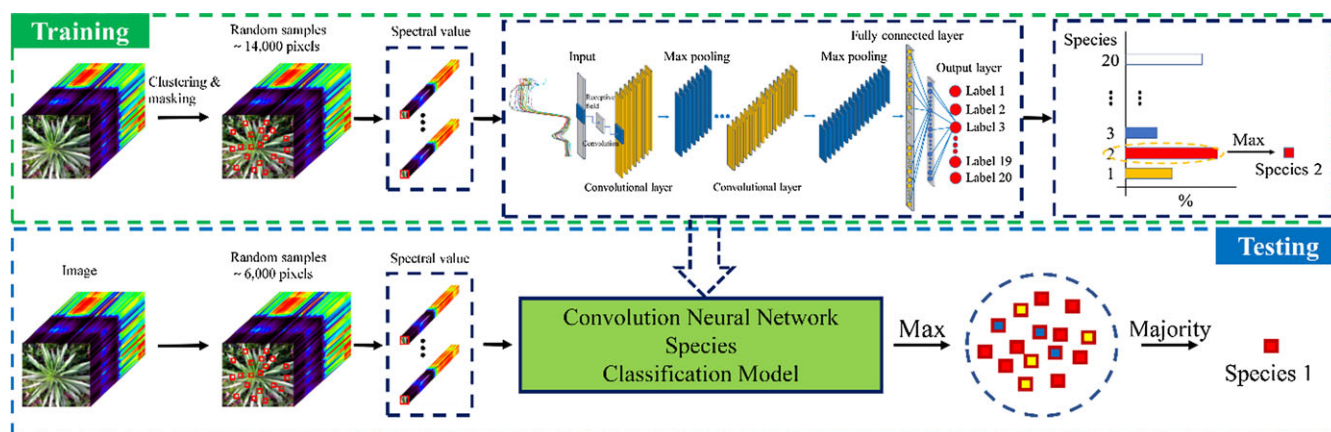


Figure 3. Flowchart of the species classification framework modeled using a convolutional neural network.

The NDVI is highly correlated with vegetation productivity, plant cover, and amount of green vegetation (Pettoirelli et al. 2005), and it has a certain relationship with changes in vegetation quality (Hamel et al. 2009). In addition, the NDVI is an indicator of vegetation vigor. Its values range from -1 to $+1$, while general green vegetation values range from 0.2 to 0.8 .

To minimize variations in camera angle and solar illumination (Mottus and Rautiainen 2013), we also explored the suitability of using the photochemical vegetation index (PRI) to study the differences and associations between 20 weed species. The PRI is based on spectral radiance derived from the normalization of vegetation reflectance near wavelengths of 531 and 570 nm (Equation 3) and is closely related to the photosynthetic intensity of the canopy under nitrogen stress and the lutein cycle (Gamon et al. 1992).

$$\text{PRI} = \frac{\rho_{531} - \rho_{570}}{\rho_{531} + \rho_{570}} \quad [7]$$

where ρ_{531} and ρ_{570} represent the spectral values at 531 and 570 nm, respectively.

CNN Classification for Weed Species Recognition

A CNN framework was established to train and validate hyperspectral weed data sets for weed classification. Neural networks are a group of mathematical algorithm models that roughly mimic the human brain. They are designed to process complex data and are being used for an increasing number of functions, with classification being the most important. A shallow CNN has especially good generalization ability and image edge feature recognition ability, and because of its simple structure, it does not require a great deal of computing power (Yang et al. 2020). In this study, a convolution neural network was established to identify hyperspectral images of 40 weed species from 23 families in northeast China. The network structure is shown in Figure 3. The input is the spectral value (spectral curve or spectral band), and the output is the species probability. The CNN consists of an input layer, an output layer, two convolution layers, two pooling layers, and a fully connected layer. The *SoftMax* function is applied to the output layer to produce probabilities.

Each hyperspectral image had a pixel size of 512×512 (262,144), and each pixel had a spectral reflectance feature of 204 bands. It was unreasonable to train the CNN using all the pixels of

an entire image, as an image contains not only the target weed samples with different growth states but also the background regions and other weed spectral images. Therefore, before the ENVI software was used to extract the spectral curve from hyperspectral images, regions of interest (ROIs) (see Supplementary Figure S1) of target weeds were selected (Tang et al. 2017). An ROI mainly includes weed leaf regions to extract representative target weed spectral curves while excluding the influence of non-weed sample areas on spectral curves. To use the CNN to identify and classify weeds, 20 ROIs were selected for each image, each containing about 1,000 pixels. The neural network consists of two parts: training and testing. In training, a training data set was built with 7/10 sample data from all images. The training data set's samples each had an input vector and an output vector. The input vector was the spectral curve, and the output vector was the "one-hot coding" vector of the species. First of all, the features of the spectral curve were extracted through the convolution layer. To eliminate the adverse impacts of bad data on the whole sample, batch normalization was adopted. A pooling layer was adopted to reduce redundant data and retain key features. Finally, the full-connection layer and *SoftMax* function were used to classify and output the features. In testing, the remaining 3/10 weed sample data points were used for separate verification, and the trained model of the CNN was verified. The CNN generated the probability that a weed sample corresponded to a particular weed species. For example, after neural network train and test, weed pixel sample A had a 70% probability of being weed species A, a 20% probability of being weed sample B, and a 5% probability of being weed sample C. So, each pixel was labeled with the name of the species with the highest probability. This allowed each sampled pixel in an image to be assigned to a certain species, with the image as a whole being assigned to a particular species based on the majority of sampled pixels.

Precision Evaluation

In this study, six performance evaluation indexes commonly used in deep learning were used to verify the performance of the model to judge the classification performance of various weeds, including: train accuracy (TA), test accuracy (TEA), average accuracy (AA), kappa coefficient (Kappa), producer's accuracy (PA), and user's accuracy (UA). TA and TEA represent the ratio between the correctly classified samples and the total number of samples in the train set and the test set, respectively. AA represents the average of

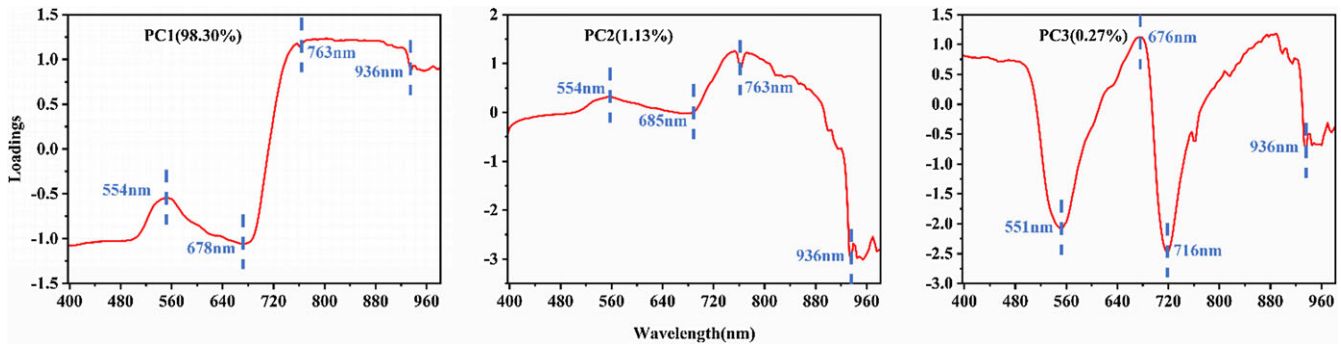


Figure 4. The results of the principal component analysis in the wavelength range of 397–981 nm.

the ratio of the number of samples correctly predicted for all weed species in the test set to the actual number of samples. Kappa is often used to measure the degree of match between the predicted sample and the actual sample (Zahisham et al. 2023). PA represents the probability that a class is correctly identified, while UA represents the probability that the classifier correctly classifies the samples belonging to a particular category (Cao et al. 2019). The formulas are as follows:

$$TA = \frac{\sum_{i=1}^n x_{ii}}{\sum_{i=1}^n \sum_{j=1}^n x_{ij}} \quad [8]$$

$$TEA = \frac{\sum_{i=1}^n y_{ii}}{\sum_{i=1}^n \sum_{j=1}^n y_{ij}} \quad [9]$$

$$AA = \frac{1}{n} \sum_{i=1}^n \frac{y_{ii}}{\sum_{j=1}^n y_{ij}} \quad [10]$$

$$Kappa = \frac{\sum_{i=1}^n \sum_{j=1}^n y_{ij} \times \sum_{i=1}^n y_{ii} - \sum_{i=1}^n (\sum_{i=1}^n y_{ij} \times \sum_{j=1}^n y_{ij})}{(\sum_{i=1}^n \sum_{j=1}^n y_{ij})^2 - \sum_{i=1}^n (\sum_{i=1}^n y_{ij} \times \sum_{j=1}^n y_{ij})} \quad [11]$$

$$PA_i = \frac{y_{ii}}{\sum_{j=1}^n y_{ij}} \quad [12]$$

$$UA_j = \frac{y_{jj}}{\sum_{i=1}^n y_{ij}} \quad [13]$$

where n represents the total number of categories; $i = 1, 2, 3, \dots, n$, represents the number of categories of real samples; $j = 1, 2, 3, \dots, n$, represents the number of categories of prediction samples; x_{ij} represents the number of samples in the train set where the real class is class i , but the prediction class is class j ; x_{ii} represents the number of samples in the train set that are both true and predicted to be of class i ; y_{ij} represents the number of samples of the prediction set where real class is class i , but the

prediction class is class j . y_{ii} represents the number of samples in the prediction set that are both true and predicted to be of class i . PA_i represents PA of class i weeds, UA_j represents the UA of class j weeds.

Results and Discussion

Analysis of the Spectral Reflectance Curves of Different Weeds

After weed sample data were collected in the field, the species of weeds and the families to which they belonged were determined. ENVI software was used to extract and plot spectral reflectance curves of the weeds, and PCA was performed in the wavelength range of 397 to 981 nm.

The loading plot in Figure 4 shows the spectral characteristics of various weeds. The first three PCs accounted for 99.7% of the total sample variance, indicating that they represent most of the information on weed types. According to the component score coefficient matrix in Supplementary Table S1, among the three PCs, PC1 was closely related to four weed species, namely, Indian strawberry [*Duchesnea indica* (Andrews) Teschem.], stony stonecrop [*Hylotelephium spectabile* (Boreau) H. Ohba], *P. supina*, and Chinese violet (*Viola philippica* (Cav.). PC2 was mainly affected by horseweed [*Conyza canadensis* (L.) Cronquist], *H. spectabile*, woodland sage [*Salvia x sylvestris* L. (pro sp.) [*nemorosa* x *pratensis*]; syn.: *Salvia nemorosa* L.], and white clover (*Trifolium repens* L.), while PC3 mainly reflected the spectral characteristics and related characteristic peaks of huoxiang [*Agastache rugosa* (Fisch. & C.A. Mey.) Kuntze], *H. spectabile*, bai hua ma lin (*Iris lactea* Pall.), and garden sorrel (*Rumex acetosa* L.). The peaks and troughs in the figure provide the main characteristic wavelengths of the weed spectrum. The loading diagram for PC1 shows that the loading gradually increases from blue-violet light to infrared light and peaks at 936 nm, indicating that weeds mainly absorb blue-violet light and absorb little of longer wavelengths, especially in the near-infrared region. The PC1, PC2, and PC3 loading diagrams show obvious peaks and troughs at 554 nm, 678 nm, 763 nm, and 936 nm, as well as in nearby wavelength regions, and they show different forms in different loading diagrams. In particular, the PC3 loading pattern is more prominent, which shows that the characteristics of *A. rugosa*, *H. spectabile*, *I. lactea*, and *R. acetosa* are more special than other weeds at 554 nm, 678 nm, 763 nm, and 936 nm.

Figure 5 shows the mean spectra of various weeds after QUAC and the spectral curves after various pretreatment methods, including the true characteristics of weed samples after removing

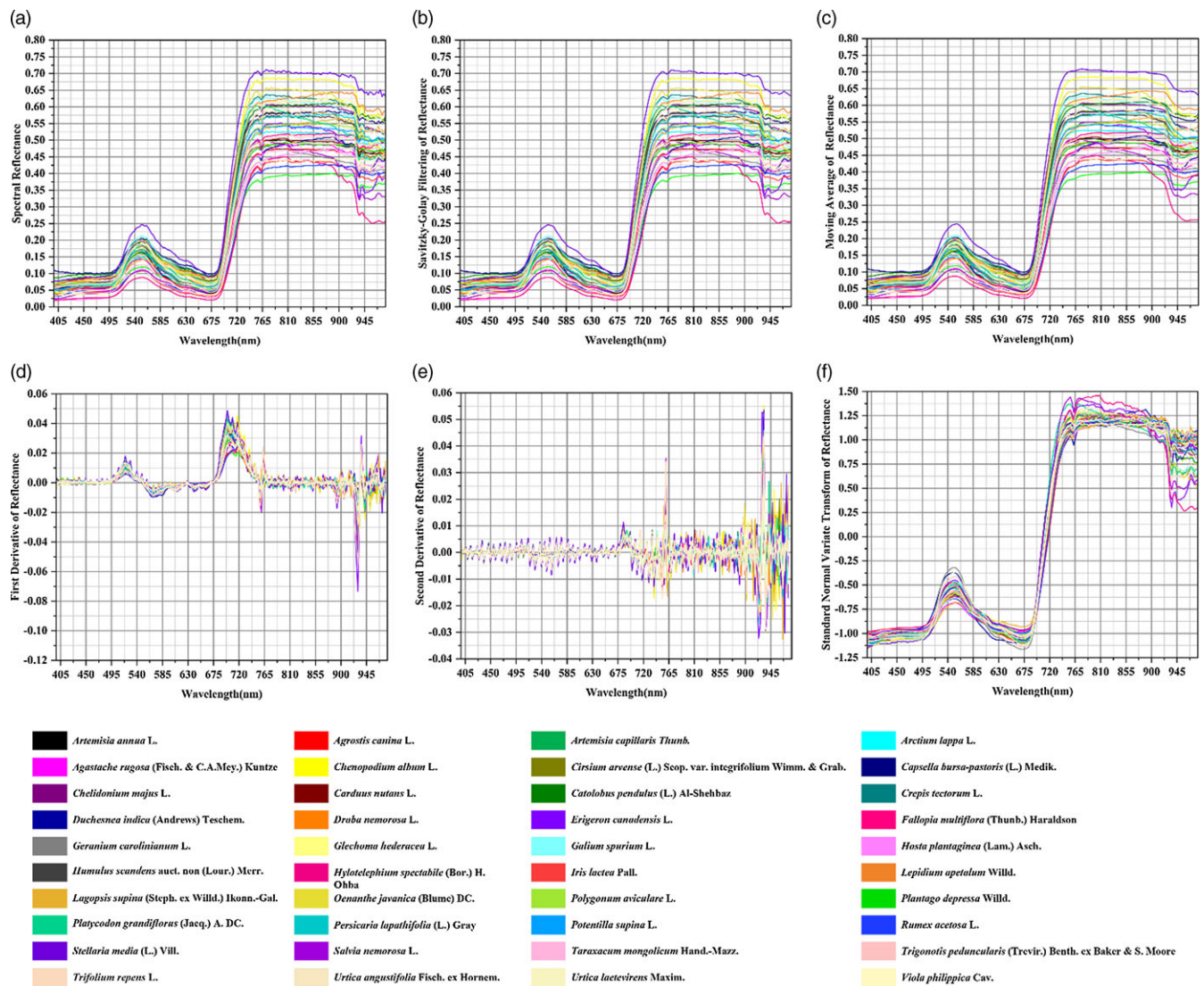


Figure 5. Spectral curves for all weed species collected in Heilongjiang Province, northeast China: (A) original and with pretreatment, (B) Savitzky-Golay smoothing, (C) moving averages, (D) first derivative spectrum, (E) second derivative spectrum, and (F) standard normal variate.

radiation errors caused by atmospheric effects such as atmospheric absorption and scattering. It can be seen that the spectral shapes of the different weed samples are almost the same, indicating that different weed species absorb similar light in different bands. This may be due to the green color of weed leaves. As shown in Figure 5, the characteristic wavelengths of various weeds appear at 551 nm, 678 nm, 760 nm, 935 nm, and nearby wavelength regions, which is consistent with the PCA results and is mainly related to the structure of the weeds' leaf tissues (Oerke et al. 2016). In most normal green plants, the pigments in the leaves are mainly chlorophyll and carotenoid, with few to no other flavonoid pigments. Chlorophyll mainly absorbs blue-violet and red light and partially absorbs green light, so the overall color of the leaves appears green. In addition, in the original spectral curve, there are two absorption valleys in the bands of 397 to 504 nm and 678 nm, while there is a small reflection peak near 551 nm. The 678-nm band and its vicinity is one of the most commonly used bands for discriminating crop classes and has proven to be important in the study of crops (Cho et al. 2010; Eddy et al. 2014; Fassnacht et al. 2014; Mariotto et al. 2014). The FDS of the 721-nm band and

nearby areas is particularly significant, as it is located in the red margin area of 680 to 780 nm and is closely related to chlorophyll content and contains rich physiological information (such as water content) (Farquhar et al. 1980). In addition, in the reflectance curve, the infrared light band has high reflectance, which is related to the fact that infrared light does not readily produce photosynthesis. In the near-infrared region, the spectral reflectance of most weed species shows a trend of gradual increase and dynamic stability. The gradual declines in Carolina geranium (*Geranium carolinianum* L.), *H. spectabile*, balloon-flower [*Platycodon grandiflorus* (Jacq.) A. DC.], and other species are speculated to be due to changes in leaf structure or the collapse of mesophyll structure caused by abiotic stress (such as water-deficit stress) (Li et al. 2022; Lou et al. 2022). This leads to increased infrared light absorption by plants. Furthermore, reduced fluorescence energy near the 760-nm wavelength may also be a sign of a decline in the photosynthetic process. When solar energy is absorbed by chlorophyll, it is used to fix carbon and dissipate heat before emitting at a longer wavelength as chlorophyll fluorescence (Krause and Weis 1991). On the other hand, there

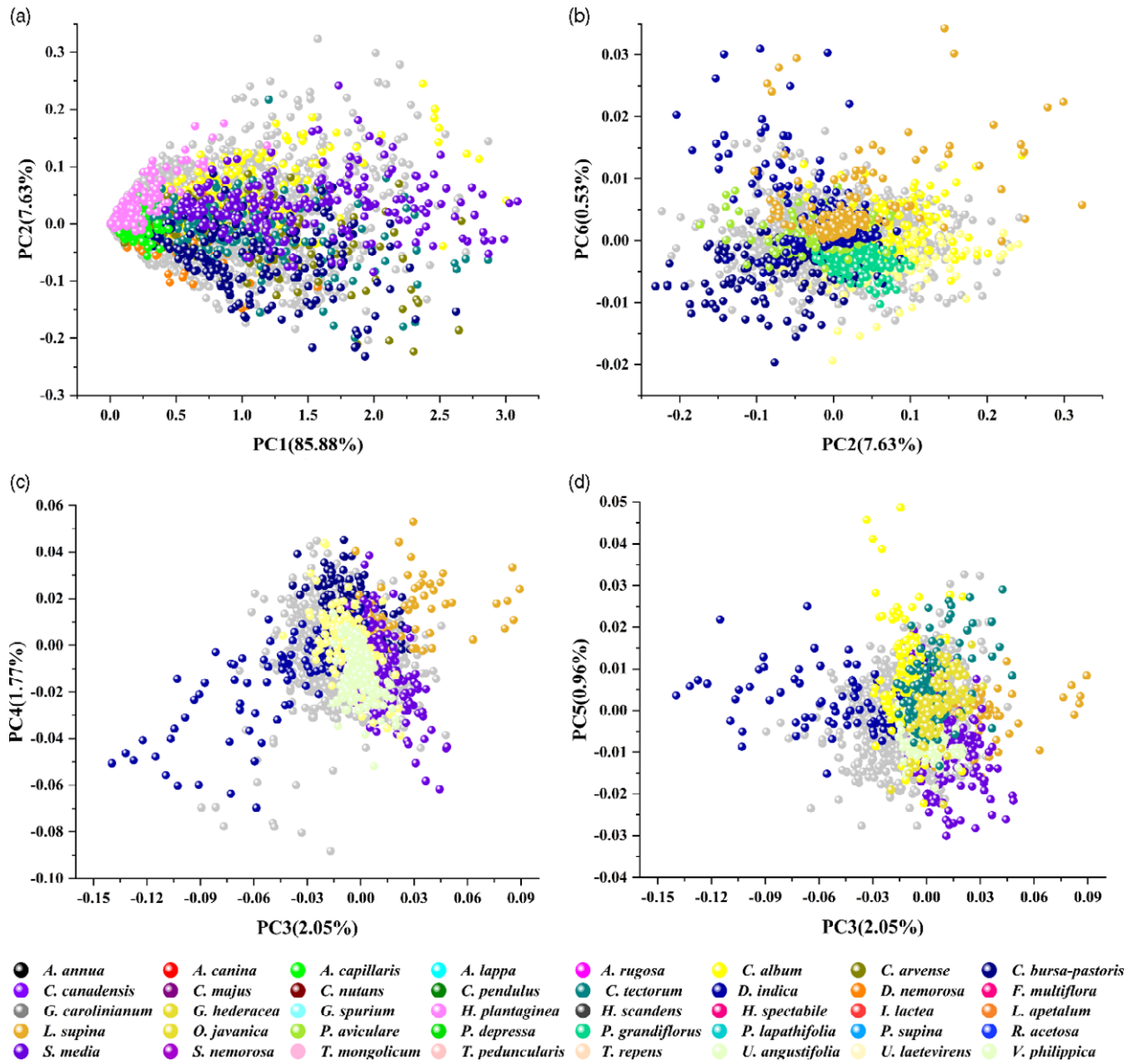


Figure 6. Hyperspectral images highlighting the classes clustered in the corresponding kernel principal component analysis (KPCA) score plots: (A) PC1–PC2, (B) PC2–PC6, (C) PC3–PC4, and (D) PC3–PC5. See Figure 1 for full species names.

is an absorption valley in the original spectral curve at 935 nm. This sharp decline in spectral reflectance indicates that the reflectance of the band with wavelengths >935 nm is not affected by the structure of the leaf itself. The higher the proportion of water in plants, the lower the spectral reflectance in the near-infrared ray (NIR) band (780 to 1,300 nm) (Zhang et al. 2022). It is inferred that the decrease in spectral reflectance at 935 nm may be caused by cell fluid, the cell membrane, absorbed water, and carbon dioxide emissions from the leaves (Qu et al. 2018).

Within a certain range, the higher the chlorophyll content, the greater the efficiency of light energy conversion. At the same time, chlorophyll mainly absorbs blue-violet and red light and absorbs almost no green light. The green region reflectance, which indicates the chlorophyll content, is the basis for inferring the strength of photosynthesis. Of all the weed species, common chickweed [*Stellaria media* (L.) Vill.] has the highest reflectivity in the green area, with stickywilly (*Galium spurium* L.) second and *H. spectabile* having the least. However, this does not directly reflect the strength of photosynthesis, which is not only affected by

chlorophyll but also by other factors such as enzyme activity, water, and leaf structure (Zhang et al. 2015; Zhang et al. 2019a). The spectral curves of the weeds show that almost all the green and infrared areas show high reflectance, while the blue-violet and red-orange areas show low reflectance, which makes it difficult to distinguish between species. Hence, further clustering and visualization of the different weed species using PCA are needed.

KPCA

In this study, KPCA was used to extract six PCs from the high-dimensional hyperspectral data to evaluate the relationship between the original hyperspectral data samples. The results show that the six PCs accounted for 98.83% of the total sample variance, indicating that they represent most of the characteristic information of the weed samples. Figure 6 shows the KPCA score plots of the hyperspectral images of various weeds clustered under different PC combinations. The PC1–PC2 score plot (Figure 6A) shows that various types of weeds clustered in a certain range, but all weed

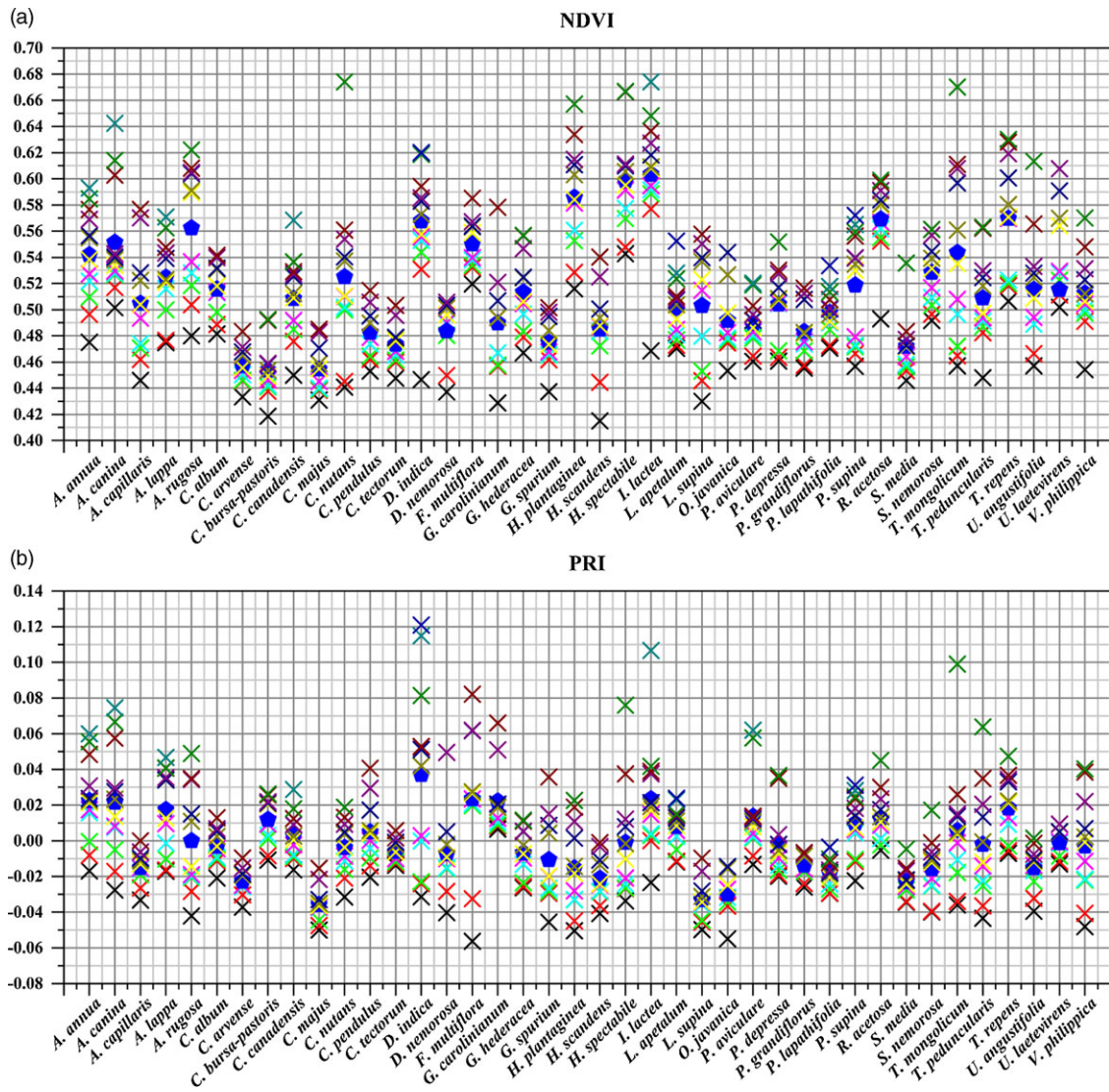


Figure 7. Vegetation indices for all weeds collected in in Heilongjiang Province, northeast China: (A) normalized difference vegetation index (NDVI) and (B) photochemical reaction index (PRI). For each species, the mean vegetation index (represented by blue pentagons) is derived from all samples of a species (each image is represented by a different color). The x-axis label shows the species names (see Figure 1 for the full names), and the y-axes represent NDVI or PRI.

clusters show differences. Meanwhile, PC2 and PC3 are linearly correlated for most weeds. In the PC2–PC6 score plot (Figure 6B), the clusters of lambsquarters (*Chenopodium album* L.), shepherd’s purse [*Capsella bursa-pastoris* (L.) Medik.], *D. indica*, ground ivy (*Glechoma hederacea* L.), lagopsis supina [*Lagopsis supina* (Steph.) Ikon. -Gal.], prostrate knotweed (*Polygonum aviculare* L.), and *P. grandiflorus* are not well separated. The clusters of *D. indica* and *L. supina* are mainly concentrated in the first and second quadrants, while those of *C. bursa-pastoris*, *G. hederacea*, and *P. grandiflorus* are clustered in the negative area of PC6. The PC3–PC4 score plot (Figure 6C) shows the clustering separation among the categories of *C. bursa-pastoris*, *D. indica*, *G. hederacea*, *L. supina*, *S. media*, and *V. philippica*. The clusters of *C. bursa-pastoris* and *L. supina* are mainly clustered in the first and second quadrants, which are effectively distinguished from those of *D. indica*. In the PC3–PC5 score plot (Figure 6D), clusters are identified for the categories of *C. album*, narrowleaf hawksbeard (*Crepis tectorum* L.), *D. indica*, *L. supina*, Java waterdropwort [*Oenanthe javanica* (Blume) DC.], *S. media*, and *V. philippica*. In this case, their clusters are very close and are not clearly separated. Based on the this information and

KPCA score maps, we can determine that all species of weeds show differences. But the clusters of various weeds are concentrated in the same areas, so it is not possible to distinguish all categories clearly. It is possible to reduce the dimensions and visualize the existing data. However, it is not possible to mine further information from the KPCA about the growth status and health status of the weeds, so it is necessary to analyze the existing spectral characteristics from many aspects, such as the vegetation index.

Vegetation Index Analysis

Vegetation indices were used to study the differences in the growth status and physiological activity of various weeds. Figure 7 shows the changes in NDVI and PRI among different weed species. In Figure 7A, the NDVI values reflect plant growth and physiological conditions (Bai et al. 2019), indicating that the various weeds have different growth levels. The NDVI of *I. lactea* is the highest, followed by those of *H. spectabile* and fragrant plantain lily [*Hosta plantaginea* (Lam.) Asch.], while those of *C. bursa-pastoris* and celandine (*Chelidonium majus* L.) are the lowest, indicating that

Table 1. Overall accuracy of all species identification models.^a

Model	AA	TA	TEA	Kappa
Original	97.58%	100.00%	97.45%	0.9738
FDS	98.05%	100.00%	98.05%	0.9800
SDS	97.90%	100.00%	97.88%	0.9782
SNV	98.16%	100.00%	98.15%	0.9810
MA	95.62%	100.00%	95.63%	0.9551
SG	95.40%	100.00%	95.32%	0.9519

^aAbbreviations: AA, average accuracy; FDS, first derivative spectrum; Kappa, kappa coefficient; MA, moving averages; SDS, second derivative spectrum; SG, Savitzky-Golay smoothing; SNV, standard normal variate; TEA, test accuracy; TA, train accuracy.

the physiological status of *I. lactea* was the highest. *Capsella bursa-pastoris* and *C. majus* were the weakest, which may be related to the appearance and growth environment of these weeds. The NDVIs of *C. album*, *G. hederacea*, *Urtica angustifolia* (*Urtica angustifolia* Fisch. ex Hornem.), *Urtica laetevirens* (*Urtica laetevirens* Maxim.), and *V. philippica* are similar, indicating that the growth of these weeds was similar. In Figure 7B, the PRI indicates the lutein content (Frechette et al. 2016). Compared with the NDVI in Figure 7A, the PRI values are more stable, indicating that the lutein content in the various weeds was extremely low. This may be because the chlorophyll content of weeds increases sharply with vigorous growth, while the rate of light utilization increases gradually, thus inhibiting the secretion of lutein. It is worth noting that in the PRI diagram of all weed species, the lutein of *D. indica* is the highest, which is speculated to be related to the regulation of pigment regeneration and the size of the circulating pigment pool (Xiao et al. 2018). However, due to the influences of environmental and other factors in the process of data collection, enzyme activities cannot be completely distinguished according to the vegetation index, so it was necessary to use the CNN to classify the different kinds of weeds.

CNN Classification of Weed Species

Tables 1 and 2 show the model TA, TEA, AA, Kappa, PA, and UA of classification inference under the different pretreatment methods. All models show very high accuracy in train and test. The individual model has a TA of 100% and the TEA values range from 95.32% to 98.15%. The TEA values for all pretreatment methods and original data were >95%, among which the TEA values for SNV were the highest. Specifically, among all the pretreatment methods, the accuracy using the FDS, SDS, and SNV methods was higher than accuracy achieved using the original data, while the accuracy using the two smoothing algorithms (MA and SG) was lower, and the AA and kappa coefficients showed similar trends. It should be noted that the overall classification accuracy generated by SG was the lowest among the five pretreatment methods, being 2.13% lower than the original data, while the TEA of MA was 1.82% lower than the original data. The results show that the smoothing algorithms can easily cause distortion in the process of processing spectral data, especially SG. At the same time, it can be seen that the precision of the differential spectrum is higher than that of the original spectral data, indicating that the differential spectrum can extract the features of weed spectral data very well.

Table 2 shows the PA (omission error) and UA (delegation error) for species classification using CNNs. Most species have high PA and UA (100%). However, the lower accuracy of some species is consistent with a smaller sample size used for species identification (Figure 1). This means that larger samples or more samples per species could be considered in future studies. Another

option would be to use high-resolution hyperspectral images obtained from UAVs. It is crucial to note that extracting 20,000 pixels from each sample of 435 images can cause autocorrelation problems. However, the CNN method adopted in this study has the ability to self-learn to discriminate between samples for classification. Meanwhile, increasing the sample size and amount of input data can lead to higher accuracy. However, care should be taken when identifying species using conventional classification methods.

In this study, we combined data on the physiological activity of weeds with hyperspectral characteristics of weeds and obtained spectral data with low noise and effective information using different pretreatment methods. This allowed 40 species of weeds in an urban ecosystem to be distinguished. The results of the CNN for species identification demonstrated that the spectral library based on weed reflectance showed a good ability to identify weed species. The TEA of the original spectral data of weeds was 97.45%, in which SNV was used to obtain the largest TEA and Kappa coefficients of 98.15% and 0.9810, respectively. The classification results highlight the potential use of ground-based hyperspectral camera data for vegetation research, particularly for weed identification and the establishment of specific spectral libraries for various vegetation and crop species. In recent years, researchers have developed advanced techniques to improve the accuracy of weed identification. Farooq et al. (2019) proposed a feature extraction method, FCNN-SPLBP, using a multilayer fused convolution neural network (FCNN) and a superpixel-based local binary pattern (SPLBP). FCNN was used to extract textural features from superpixels as input of a support vector machine for weed classification. The recognition accuracy was 89.7%, and the performance was better than that of CNN, LBP, FCNN, and SPLBP. Tang et al. (2017) constructed a weed recognition model based on K-means feature learning combined with CNNs. The accuracy is 92.89%, 1.82% higher than that of the randomly initialized CNN and 6.01% higher than that of the two-layer neural network without fine tuning. Wei et al. (2014) used canonical discriminant analysis and a partial least squares-discriminant analysis (PLS-DA) model to identify broadleaf grass species with an accuracy of 90.91%. It should be noted that broadleaf species are easier to identify than narrow-leaved ones, because of more uniform spectral data collection (Li et al. 2021). This study not only distinguishes broadleaf grass species, but also identifies narrow-leaved ones. The accuracy of weed identification using hyperspectral imaging technology and a CNN was 98.15%, which is 8.45% higher than identification using FCNN-SPLBP, and 5.26% higher than identification using a weed recognition model based on K-means feature learning combined with a CNN. It is 7.24% higher than identification using canonical discriminant analysis and a PLS-DA model. The results show that the convolutional network classification model established in this study is superior to other techniques for vegetation identification. It is worth noting that shallow CNN has better classification performance than deep neural networks, and it is easier to apply in practical situations due to its simple structure and fast running speed (Han et al. 2020).

The PA, UA, and overall accuracy of each classification model shows that the preprocessing method has a great influence on species identification based on hyperspectral images. In this paper, the three preprocessing methods based on the original data—FDS, SDS, and SNV—obtained the highest predictive accuracy, while the two smoothing algorithms were less accurate. The results show that, compared with FDS, SDS, and SNV, the smoothing algorithms were prone to distortion and not suitable for

Table 2. Producer's accuracy (omission error) and user's accuracy (delegation error) using convolutional neural network classification to identify species collected in Heilongjiang Province, northeast China.^a

Species	PA						UA					
	Original	FDS	SDS	SNV	MA	SG	Original	FDS	SDS	SNV	MA	SG
<i>Artemisia annua</i> L.	0.96	0.96	1.00	1.00	0.96	1.00	0.95	1.00	0.98	0.98	0.98	0.98
<i>Agrostis canina</i> L.	1.00	0.98	0.98	0.96	0.96	0.96	1.00	0.96	0.96	0.98	0.94	0.96
<i>Artemisia capillaris</i> Thunb.	1.00	0.97	0.97	1.00	0.95	0.92	0.91	0.92	0.94	0.97	0.95	0.94
<i>Arctium lappa</i> L.	0.90	0.96	1.00	1.00	1.00	0.98	1.00	1.00	0.96	1.00	0.90	1.00
<i>Agastache rugosa</i> (Fisch. & C.A. Mey.) Kuntze	0.98	0.96	1.00	0.96	0.88	0.95	0.91	0.96	1.00	0.98	0.98	0.95
<i>Chenopodium album</i> L.	1.00	0.97	1.00	0.94	0.95	0.93	1.00	1.00	1.00	1.00	0.97	1.00
<i>Cirsium arvense</i> (L.) Scop. var. <i>integrifolium</i> Wimm. & Grab.	1.00	0.94	0.93	0.96	0.82	0.76	1.00	0.94	0.93	0.98	0.86	0.84
<i>Capsella bursa-pastoris</i> (L.) Medik.	1.00	0.94	0.89	0.98	0.89	0.85	1.00	0.94	0.89	0.95	0.81	0.76
<i>Coryza canadensis</i> (L.) Cronquist	0.97	0.92	0.93	0.96	0.96	0.89	1.00	0.92	0.96	1.00	0.93	0.85
<i>Chelidonium majus</i> L.	1.00	1.00	0.98	0.98	0.92	0.98	1.00	0.97	0.96	1.00	0.97	0.96
<i>Carduus nutans</i> L.	0.86	1.00	1.00	1.00	0.95	0.95	0.95	0.98	0.93	0.98	0.87	0.95
<i>Catolobus pendulus</i> (L.) Al-Shehbaz	1.00	0.98	1.00	1.00	1.00	0.93	1.00	1.00	1.00	0.97	1.00	1.00
<i>Crepis tectorum</i> L.	1.00	1.00	0.99	1.00	1.00	0.99	1.00	1.00	1.00	0.99	1.00	0.97
<i>Duchesnea indica</i> (Andrews) Teschem.	1.00	0.98	0.95	0.95	0.88	0.84	1.00	0.98	0.95	0.95	0.93	0.95
<i>Draba nemorosa</i> L.	1.00	0.95	0.87	0.95	0.93	0.94	0.86	0.95	0.93	0.95	0.90	0.85
<i>Fallopia multiflora</i> (Thunb.) Haraldson	0.98	0.93	0.97	0.97	0.92	0.88	1.00	0.95	0.97	0.94	0.95	0.91
<i>Geranium carolinianum</i> L.	1.00	0.98	1.00	1.00	0.97	0.97	0.89	0.97	0.97	0.94	0.94	0.91
<i>Glechoma hederacea</i> L.	1.00	0.98	0.98	0.98	0.95	0.95	1.00	0.98	1.00	1.00	0.98	0.93
<i>Galium spurium</i> L.	1.00	1.00	0.98	0.98	0.95	0.97	1.00	0.98	1.00	1.00	0.98	0.97
<i>Hosta plantaginea</i> (Lam.) Asch.	0.92	1.00	1.00	0.98	1.00	0.98	0.95	0.98	0.98	0.97	0.92	0.95
<i>Humulus scandens</i> auct. non (Lour.) Merr.	1.00	0.99	1.00	0.97	0.96	0.96	0.97	0.99	1.00	1.00	0.99	0.99
<i>Hylotelephium spectabile</i> (Boreau) H. Ohba	1.00	0.98	1.00	1.00	0.94	0.98	1.00	0.98	1.00	0.98	1.00	0.96
<i>Iris lactea</i> Pall.	0.96	1.00	1.00	1.00	0.96	0.98	0.98	1.00	1.00	1.00	0.92	0.98
<i>Lepidium apetalum</i> Willd.	1.00	1.00	1.00	0.98	1.00	0.97	1.00	0.98	1.00	0.98	0.97	1.00
<i>Lagopsis supina</i> (Steph. ex Willd.) Ikonn.-Gal.	1.00	0.98	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
<i>Oenanthe javanica</i> (Blume) DC.	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00
<i>Polygonum aviculare</i> L.	1.00	1.00	1.00	0.98	1.00	1.00	0.91	0.98	0.95	0.98	1.00	0.93
<i>Plantago depressa</i> Willd.	0.93	0.95	0.93	0.93	0.90	0.98	0.95	0.95	1.00	0.95	0.95	0.98
<i>Platycodon grandiflorus</i> (Jacq.) A. DC.	1.00	1.00	1.00	1.00	1.00	0.96	1.00	1.00	1.00	1.00	1.00	0.96
<i>Persicaria lapathifolia</i> (L.) Gray	1.00	1.00	1.00	0.96	0.96	0.98	1.00	1.00	0.96	1.00	1.00	1.00
<i>Potentilla supina</i> L.	0.91	1.00	0.99	0.99	0.97	0.99	0.97	0.97	0.97	0.97	0.99	1.00
<i>Rumex acetosa</i> L.	1.00	1.00	0.98	0.96	0.96	0.96	0.88	1.00	1.00	1.00	0.98	0.94
<i>Stellaria media</i> (L.) Vill.	0.97	1.00	1.00	1.00	0.98	0.98	1.00	1.00	1.00	1.00	1.00	1.00
<i>Salvia xylvestris</i> L. (pro sp.) [<i>nemorosa</i> × <i>pratensis</i>]	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00
<i>Taraxacum mongolicum</i> Hand.-Mazz.	0.93	0.98	0.93	1.00	0.91	0.93	0.95	1.00	0.93	0.93	0.85	0.98
<i>Trigonotis peduncularis</i> (Trevir.) Benth. ex Baker & S. Moore	0.97	0.97	0.95	0.97	0.98	1.00	0.95	0.98	0.98	0.98	0.92	0.97
<i>Trifolium repens</i> L.	0.92	0.97	0.98	1.00	0.97	0.94	1.00	1.00	1.00	1.00	1.00	0.98
<i>Urtica angustifolia</i> Fisch. ex Hornem.	1.00	1.00	1.00	1.00	0.98	0.97	1.00	1.00	1.00	1.00	0.98	1.00
<i>Urtica laetevirens</i> Maxim.	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
<i>Viola philippica</i> Cav.	0.96	0.99	0.97	0.99	0.96	1.00	0.99	0.99	1.00	0.97	0.96	0.96

^aAbbreviations: FDS, first derivative spectrum; MA, moving averages; PA, producer's accuracy; SDS, second derivative spectrum; SG, Savitzky-Golay smoothing; SNV, standard normal variate; UA, user's accuracy.

distinguishing weed species or challenging classification scenarios, and the difference in spectral characteristics was not significant. Other pretreatment methods can be used to improve the predictive accuracy. On the other hand, the PA and UA were low for many weeds, especially *D. nemorosa*, probably because it was the least-sampled of all weeds. This indicates that increasing the number of samples may improve the PA and UA for other species. Of all species, *C. canadensis* was the easiest to identify, probably because of its leaf structure (Li et al. 2021). It is important to note that this species had the most samples collected overall. According to the data obtained using the original spectrum, PCA, and the differential spectrum, 554 nm, 678 nm, 763 nm and 936 nm and their adjacent spectral bands are considered to be the most prominent bands reflecting the characteristics of weeds. They can indirectly reflect the physiological activity of weeds and could play a key role in weed recognition.

In addition to the weeds' species and physiological activity, their spectral characteristics may also vary according to environmental conditions. In urban ecosystems, different soil and climate

conditions can lead to different growth conditions and the presence of different weed species (Carlesi et al. 2013; Steponavičienė et al. 2021). For example, hyperspectral analysis of 48 experimental plots of temperate species showed a significant relationship between growth characteristics and spectral characteristics at different soil nitrogen concentrations, with broadleaf grasses being particularly responsive to nitrogen (Jabran and Doğan 2020; Waheed et al. 2006). Another study showed that there were significant differences in the canopy reflectance characteristics of weeds in environments with different temperatures. Higher temperatures increased the amplitude and variability of leaf reflectance in the 480- to 670-nm region, while the opposite effect occurred in the 720- to 810-nm region (Zhang and Slaughter 2011). As reported in a recent study, higher temperatures cause plants to distribute more biomass to their stems, causing their leaves to expand and promoting light capture and assimilation (Zhang et al. 2014). In addition, the vegetation index may change with the change of weed growth stage. A vegetation index shows statistical differences in different weed phenological stages and

provides a valuable reference for the differentiation of weeds in different growth periods (Pena-Barragan et al. 2006). Due to the limitations of the conditions, the current study did not consider the physiological activity of various weeds in different growing environments and growth periods. Future studies can close this gap by targeting different growth stages of a few more common weed species that grow in different environments.

This study investigated the application of hyperspectral images to accurately classify urban weeds. The model generated can be altered and updated to effectively change species composition, community structure, and functional characteristics of weeds through continuous detection and control to adapt to the process of urbanization. It can be applied to identify relationships between urbanization and ecological/environmental impacts and can be used to support intensive urbanization while protecting the ecological environment. It also can be used to identify invasive plants and prevent them from adversely affecting native flora and fauna, public health, and ecosystem services. In addition, the results acquired from hyperspectral imagery give superior accuracy (97.29%) when compared with results from studies that used multispectral data in urban contexts (Hahn et al. 2021). Many studies have shown that hyperspectroscopy improves the accuracy of weed mapping by providing finer spectral resolution (Che'Ya et al. 2021; Lauwers et al. 2020). In addition, Ferreira et al. (2016) showed that when short-wave infrared (SWIR) bands are combined with visible/near-infrared (VNIR) bands to map vegetation, the accuracy is improved by 14% to 17%, while the accuracy of hyperspectral data is 15% higher than that possible with multispectral VNIR and SWIR images. Cho and Lee (2014) also noted that SWIR offers an additional benefit for classifying vegetation. Therefore, through the addition of hyperspectral SWIR imaging, future investigations of urban weed species identification could profit from this versatility.

In this study, hyperspectral images of 40 species of weeds from 23 families were obtained using terrestrial hyperspectral remote sensing technology. A total of 435 hyperspectral images were obtained. Various preprocessing methods (FDS, SDS, SNV, MA, and SG) were used to maximize the retention of spectral characteristics while removing the influence of noise. The spectral profile, PCA, and spectral reflectance curves of all weed species in different bands were analyzed, and the characteristic wavelengths of weeds were obtained, including 554 nm, 678 nm, 763 nm, and 936 nm. These bands reflect the different physiological activities of different weed species, such as *A. rugosa*, *C. nutans*, *T. repens*, and other species. The differences in NDVI and PRI of different species of weeds were analyzed. The results showed that the two vegetation indexes had the same trend, but the wide-band vegetation index was more beneficial for the detection and evaluation of vegetation status than the narrow-band vegetation index. However, using a vegetation index for high-precision vegetation monitoring is a challenging task. Additionally, a framework for a CNN was established to identify species from hyperspectral data. The results show that the CNN method has high accuracy, and the classification results of different pretreatment methods range from 95.32% to 98.15%, in which SNV+CNN achieves the best effect. The results of this study can be the basis for a recommendation to develop a spectral library for monitoring physiological activity and species identification of various weeds in northeast China. It is also a representative study of mesotemperate weed species with high diversity in an intense urban environment with a plains topography.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/wsc.2023.36>

Data availability. Data will be made available on request.

Acknowledgments. This work was supported by China's National Key R & D Plan (2021YFD200060502; 2018YFD0300105; 2016YFD0300909). 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.

References

- Abbas S, Peng Q, Wong MS, Li Z, Wang J, Ng KTK, Kwok CYT, Hui KKW (2021) Characterizing and classifying urban tree species using bi-monthly terrestrial hyperspectral images in Hong Kong. *ISPRS J Photogramm Remote Sens* 177:204–216
- Adegenjo AO, Liu L, Ngadi MO (2020) Non-destructive assessment of chicken egg fertility. *Sensors* 20:5546
- Bai Y, Gao J, Zhang B (2019) Monitoring of crops growth based on NDVI and EVI. *Trans Chin Soc Agric Mach* 50:153–161. Chinese
- Becker BL, Lusch DP, Qi JG (2005) Identifying optimal spectral bands from in situ measurements of great lakes coastal wetlands using second-derivative analysis. *Remote Sens Environ* 97:238–248
- Behmann J, Acebron K, Emin D, Bennertz S, Matsubara S, Thomas S, Bohnenkamp D, Kuska MT, Jussila J, Salo H, Mahlein AK, Rascher U (2018) Specim IQ: evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection. *Sensors* 18:441
- Bretagnolle V, Gaba S (2015) Weeds for bees? A review. *Agron Sustain Dev* 35:891–909
- Bro R, Smilde AK (2014) Principal component analysis. *Anal Methods* 6: 2812–2831
- Brunzel S, Fischer SF, Schneider J, Jetzkowitz J, Brandl R (2009) Neo- and archaeophytes respond more strongly than natives to socio-economic mobility and disturbance patterns along an urban-rural gradient. *J Biogeogr* 36:835–844
- Cao J, Liu K, Liu L, Zhu Y, Li J, He Z (2019) Identifying mangrove species using field close-range snapshot hyperspectral imaging and machine-learning techniques. *Remote Sens* 10:2047
- Carlesi S, Bocci G, Moonen AC, Frumentio P, Barberi P (2013) Urban sprawl and land abandonment affect the functional response traits of maize weed communities in a heterogeneous landscape. *Agric Ecosyst Environ* 166:76–85
- Chen H, Pan T, Chen J, Lu Q (2011) Waveband selection for NIR spectroscopy analysis of soil organic matter based on SG smoothing and MWPLS methods. *Chemometr Intell Lab Syst* 107:139–146
- Chen X, Zhang L, Zhang X, Liu B (2012) Comparison of the sensor dependence of vegetation indices based on hyperion and CHRIS hyperspectral data. *Int J Remote Sens* 34:2200–2215
- Cheoi K, Choi H, Ko J (2020) Empirical remarks on the translational equivariance of convolutional layers. *Appl Sci* 10:3161
- Che'Ya NN, Dunwoody E, Gupta M (2021) Assessment of weed classification using hyperspectral reflectance and optimal multispectral UAV imagery. *Agronomy* 11:1435
- Cho H, Lee KS (2014) Comparison between hyperspectral and multispectral images for the classification of coniferous species. *Korean J Remote Sens* 30:25–36
- Cho MA, Debba P, Mathieu R, Naidoo L, Van Aardt J, Asner GP (2010) Improving discrimination of savanna tree species through a multiple-endmember spectral angle mapper approach: canopy-level analysis. *IEEE Trans Geosci Remote Sens* 48:4133–4142
- Eddy PR, Smith AM, Hill BD, Peddle DR, Coburn CA, Blackshaw RE (2014) Weed and crop discrimination using hyperspectral image data and reduced bandsets. *Can J Remote Sens* 39:481–490.
- Farooq A, Jia X, Hu J, Zhou J (2019) Multi-resolution weed classification via convolutional neural network and superpixel based local binary pattern using remote sensing images. *Remote Sens* 11:1692

- Farquhar GD, Von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* 149:78–90
- Fassnacht FE, Neumann C, Forster M, Buddenbaum H, Ghosh A, Clasen A, Joshi PK, Koch B (2014) Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. *IEEE J Sel Top Appl Earth Obs Remote Sens* 7:2547–2561
- Ferreira GL, Souza MD, de Queiroz GP, Ferreira LR, Ferreira EA, Neto SND, de Souza WM, Silva DV (2016) Physiological characteristics of eucalypts in association with signal grass. *Aust For* 79:203–207
- Ferreira MP, Zortea M, Zanotta DC, Shimabukuro YE, De Souza Filho CR (2016) Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *Remote Sens Environ* 179:66–78
- Fréchette E, Chang CYY, Ensminger I (2016) Photoperiod and temperature constraints on the relationship between the photochemical reflectance index and the light use efficiency of photosynthesis in *Pinus strobus*. *Tree Physiol* 36:311–324
- Gage KL, Schwartz-Lazaro LM (2019) Shifting the paradigm: an ecological systems approach to weed management. *Agriculture* 9:179
- Gamon JA, Peñuelas J, Field CB (1992) A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens Environ* 41:35–44
- Giovos R, Tassopoulos D, Kalivas D, Lougkos N, Priovolou A (2021) Remote sensing vegetation indices in viticulture: a critical review. *Agriculture* 11:457
- Guo J, Xu X, Zhao Q, Liu J (2016) Precise orbit determination for quad-constellation satellites at Wuhan University: strategy, result validation, and comparison. *J Geod* 90:143–159
- Guo L, Qiu J, Li L, Lu B, Olsen K, Fan L (2018) Genomic clues for crop-weed interactions and evolution. *Trends Plant Sci* 23:1102–1115
- Hahn DS, Roosjen P, Morales A, Nijp J, Beck L, Cruz CV, Leinauer B (2021) Detection and quantification of broadleaf weeds in turfgrass using close-range multispectral imagery with pixel- and object-based classification. *Int J Remote Sens* 42:8035–8055
- Hamel S, Garel M, Festa-Bianchet M, Gaillard JM, Cote SD (2009) Spring normalized difference vegetation index (NDVI) predicts annual variation in timing of peak faecal crude protein in mountain ungulates. *J Appl Ecol* 46:582–589
- Han F, Zhu J, Liu B, Zhang B, Xie F (2020) Fish shoals behavior detection based on convolutional neural network and spatiotemporal information. *IEEE Access* 8:126907–126926
- Hu Y, Zhang J, Ma J, An J, Ren G, Li X (2019) Hyperspectral coastal wetland classification based on a multiobject convolutional neural network model and decision fusion. *IEEE Geosci Remote Sens* 16:1110–1114
- Huber-Lerner M, Hadar O, Rotman SR, Huber-Shalem R (2016) Hyperspectral band selection for anomaly detection: the role of data Gaussianity. *IEEE J Sel Top Appl Earth Obs Remote Sens* 9:732–743
- Jabran K, Doğan MN (2020) Elevated CO₂, temperature and nitrogen levels impact growth and development of invasive weeds in the Mediterranean region. *J Sci Food Agric* 100:4893–4900
- Ji L, Zhang L, Rover J, Wylie BK, Chen X (2014) Geostatistical estimation of signal-to-noise ratios for spectral vegetation indices. *ISPRS J Photogramm Remote Sens* 96:20–27
- Kachrimanis K, Braun DE, Griesser UJ (2007) Quantitative analysis of paracetamol polymorphs in powder mixtures by FT-Raman spectroscopy and PLS regression. *J Pharm Biomed Anal* 43:407–412
- Katkovsky LV, Martinov AO, Siliuk VA, Ivanov DA, Kokhanovsky AA (2018) Fast atmospheric correction method for hyperspectral data. *Remote Sens* 10:1698
- Khaliq A, Comba L, Biglia A, Aimonino DR, Chiaberge M, Gay P (2019) Comparison of satellite and UAV-Based multispectral imagery for vineyard variability assessment. *Remote Sens* 11:436
- Khdery G, Yones M (2021) Innovative spectral library for identification common wild plants using hyperspectral technology in Northwestern Coast, Egypt. *Egypt J Remote Sens* 24:485–491
- Korres NE, Norsworthy JK, Tehranchian P, Gitsopoulos TK, Loka DA, Oosterhuis DM, Gealy DR, Moss SR, Burgos NR, Miller MR, Palhano M (2016) Cultivars to face climate change effects on crops and weeds: a review. *Agron Sustain Dev* 36:12
- Krause GH, Weis E (1991) Chlorophyll fluorescence and photosynthesis: the basics. *Annu Rev Plant Physiol Plant Mol Biol* 42:313–349
- Lauwers M, De Cauwer B, Nuytens D, Cool SR, Pieters JG (2020) Hyperspectral classification of cyperus esculentus clones and morphologically similar weeds. *Sensors* 20:2504
- Li Q, Gao M, Li Z (2022) Ground hyper-spectral remote-sensing monitoring of wheat water stress during different growing stages. *Agronomy* 12:2267
- Li Y, Al-Sarayreh M, Irie K, Hackell D, Bourdot G, Reis MM, Ghamkhar K (2021) Identification of weeds based on hyperspectral imaging and machine learning. *Front Plant Sci* 11:611622
- Lou Z, Quan L, Sun D, Li H, Xia F (2022) Hyperspectral remote sensing to assess weed competitiveness in maize farmland ecosystems. *Sci Total Environ* 844:157071
- Maes WH, Steppe K (2019) Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci* 24:152–164
- Manjunath KR, Kumar A, Meenakshi M, Renu R, Uniyal SK, Singh RD, Ahuja PS, Ray SS, Panigrahy S (2014) Developing spectral library of major plant species of western Himalayas using ground observations. *J Indian Soc Remote Sens* 42:201–216
- Mariotto I, Thenkabail PS, Huete A, Slonecker ET, Platonov A (2014) Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HypSPRI mission. *Remote Sens Environ* 139:291–305
- Mottus M, Rautiainen M (2013) Scaling PRI between coniferous canopy structures. *IEEE J Sel Top Appl Earth Obs Remote Sens* 6:708–714
- Oerke EC, Herzog K, Toepfer R (2016) Hyperspectral phenotyping of the reaction of grapevine genotypes to *Plasmopara viticola*. *J Exp Bot* 67:5529–5543
- Pena-Barragan JM, Lopez-Granados F, Jurado-Exposito M, Garcia-Torres L (2006) Spectral discrimination of *Ridolfia segetum* and sunflower as affected by phenological stage. *Weed Res* 46:10–21
- Pettorelli N, Vik JO, Mysterud A, Gaillard JM, Tucker CJ, Stenseth NC (2005) Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol Evol* 20:503–510
- Prati A, Shan C, Wang KIK (2019) Sensors, vision and networks: from video surveillance to activity recognition and health monitoring. *J Ambient Intel Smart Environ* 11:5–22
- Qian Y, Yu J, Jia Z, Yang F, Palidan T (2013) Extraction and analysis of hyperspectral data from typical desert grassland in Xinjiang. *Acta Prataculturae Sin* 22:157–166. Chinese
- Qu J, Sun D, Cheng J, Pu H (2018) Mapping moisture contents in grass carp (*Ctenopharyngodon idella*) slices under different freeze drying periods by Vis-NIR hyperspectral imaging. *LWT Food Sci Technol* 75:529–536
- Rossini M, Nedbal L, Guanter L, Ac A, Alonso L, Burkart A, Cogliati S, Colombo R, Damm A, Drusch M, Hanus J, Janoutova R, Julitta T, Kokkalis P, Moreno J, et al. (2015) Red and far red sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geophys Res Lett* 42:1632–1639
- Rozenberg G, Kent R, Blank L (2021) Consumer-grade UAV utilized for detecting and analyzing late-season weed spatial distribution patterns in commercial onion fields. *Precis Agric* 22:1317–1332
- Segarra J, Buchailot ML, Araus JL, Kefauver SC (2020) Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy* 10:641
- Steponaviciene V, Marcinkeviciene A, Butkeviciene LM, Skinuliene L, Boguzas V (2021) The effect of different soil tillage systems and crop residues on the composition of weed communities. *Agronomy* 11:1276
- Tang J, Wang D, Zhang Z, He L, Xin J, Xu Y (2017) Weed identification based on K-means feature learning combined with convolutional neural network. *Comput Electron Agric* 135:63–70
- Tzamalí E, Akoumianakis G, Argyros A, Stephanedes YJ (2006) Improved design for vision-based incident detection in transportation systems using real-time view transformations. *J Transp Eng A Syst* 132:837–844
- Von der Lippe M, Kowarik I (2007) Do cities export biodiversity? Traffic as dispersal vector across urban-rural gradients. *Divers Distrib* 14:18–25
- Waheed T, Bonnell RB, Prasher SO, Paulet E (2006) Measuring performance in precision agriculture: CART-A decision tree approach. *Agric Water Manage* 84:173–185

- Wang Z, Fu Z, Weng W, Yang D, Wang J (2023) An efficient method for the rapid detection of industrial paraffin contamination levels in rice based on hyperspectral imaging. *LWT Food Sci Technol* 171:114125
- Wei X, Bai J, Zhang J, Wu S (2014) Weed identification from winter rape at seedling stage based on spectral variable selection. *Sensor Lett* 12:530–536
- Xiao Y, He X, Ma Q, Lu Y, Bai F, Bai J, Wu Q (2018) Photosynthetic accumulation of lutein in *Auxenochlorella protothecoides* after heterotrophic growth. *Mar Drugs* 16:283
- Yang W, Liu W, Liu J, Zhang M (2020) Acoustic emission recognition based on a two-streams convolutional neural network. *Comput Mater Contin* 64:515–525
- Yang Z, Tian J, Feng K, Gong X, Liu J (2021) Application of a hyperspectral imaging system to quantify leaf-scale chlorophyll, nitrogen and chlorophyll fluorescence parameters in grapevine. *Plant Physiol Biochem* 166:723–737
- Yeow YL, Leong YK (2005) A general computational method for converting normal spectra into derivative spectra. *Appl Spectrosc* 59:584–592
- Zahisham Z, Lim K, Koo V, Chan Y, Lee C (2023) 2SRS: two-stream residual separable convolution neural network for hyperspectral image classification. *IEEE Geosci Remote Sens* 20:5501505
- Zhang C, Zhou L, Zhao Y, Zhu S, Liu F, He Y (2020) Noise reduction in the spectral domain of hyperspectral images using denoising autoencoder methods. *Chemometr Intell Lab Syst* 203:104063
- Zhang Q, Zhang Y, Peng S, Zobel K (2014) Climate warming may facilitate invasion of the exotic shrub *Lantana camara*. *PLoS ONE* 9:e105500
- Zhang R, Zheng Z, Liu G, Du C, Du C, Lei S, Xu Y, Xu J, Mu M, Bi S, Li J (2021) Simulation and assessment of the capabilities of orbita hyperspectral (OHS) imagery for remotely monitoring chlorophyll-a in eutrophic plateau lakes. *Remote Sens* 13:2821
- Zhang X, Liu F, He Y, Li X (2013) Application of hyperspectral imaging and chemometric calibrations for variety discrimination of maize seeds. *Sensors* 12:17234–17246
- Zhang X, Pu P, Tang Y, Zhang L, Lv J (2019a) C4 photosynthetic enzymes play a key role in wheat spike bracts primary carbon metabolism response under water deficit. *Plant Physiol Biochem* 142:163–172
- Zhang Y, Cao K, Sack L, Li N, Wei X, Goldstein G (2015) Extending the generality of leaf economic design principles in the cycads, an ancient lineage. *New Phytol* 206:817–829
- Zhang Y, Gao J, Cen H, Lu Y, Xu X, He Y, Pieters JG (2019b) Automated spectral feature extraction from hyperspectral images to differentiate weedy rice and barnyard grass from a rice crop. *Comput Electron Agric* 159:42–49
- Zhang Y, Slaughter DC (2011) Hyperspectral species mapping for automatic weed control in tomato under thermal environmental stress. *Comput Electron Agric* 77:95–104
- Zhang Y, Wu J, Wang A (2022) Comparison of various approaches for estimating leaf water content and stomatal conductance in different plant species using hyperspectral data. *Ecol Indic* 142:109278
- Zhong Y, Wang X, Xu Y, Wang S, Jia T, Hu X, Zhao J, Wei L, Zhang L (2019) Mini-UAV-Borne hyperspectral remote sensing: from observation and processing to applications. *IEEE Geosci Remote Sens* 6:46–62
- Zomer RJ, Trabucco A, Ustin SL (2009) Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. *J Environ Manag* 90:2170–2177