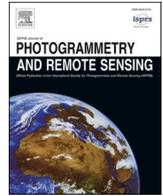




Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Characterizing and classifying urban tree species using bi-monthly terrestrial hyperspectral images in Hong Kong

Sawaid Abbas^{a,1}, Qian Peng^{a,1}, Man Sing Wong^{a,b,*}, Zhilin Li^{a,c}, Jicheng Wang^d, Kathy Tze Kwun Ng^e, Coco Yin Tung Kwok^a, Karena Ka Wai Hui^a

^a Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region

^b Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region

^c Faculty of Geosciences and Environmental Engineering & State-Province Joint Engineering Laboratory in Spatial Information Technology for High-Speed Railway Safety, Southwest Jiaotong University, Chengdu, China

^d Key Laboratory of Ministry of Education on Land Resources Evaluation and Monitoring in Southwest China, Sichuan Normal University, Chengdu, China

^e Landscape Division, The Highways Department, HKSAR Government, China

ARTICLE INFO

Keywords:

Urban tree
Hyperspectral library
Tree species
Seasonality
Deep learning
SPECIM-IQ

ABSTRACT

Urban trees exhibit a wide range of ecosystem services that have long been unveiled and increasingly reported. The ability to map tree species and analyze tree health conditions would become vividly essential. Remote sensing techniques, especially hyperspectral imaging, are being evolved for species identification and vegetation monitoring from spectral response patterns. In this study, a hyperspectral library for urban tree species in Hong Kong was established comprising 75 urban trees belonging to 19 species. 450 bi-monthly images were acquired by a terrestrial hyperspectral camera (SPECIM-IQ) from November 2018 to October 2019. A Deep Neural Network classification model was developed to identify tree species from the hyperspectral imagery with an overall accuracy ranging from 85% to 96% among different seasons. Representative spectral reflectance curves of healthy and unhealthy conditions for each species were extracted and analyzed. The hyperspectral phenology models were developed to achieve high accuracy and optimization of data acquisition. The bi-monthly canopy signatures and vegetation indices revealed different seasonality patterns of evergreen and deciduous species in Hong Kong. We explored the utility of terrestrial hyperspectral remote sensing and Deep Neural Network for urban tree species identification and characterizing. This provides a unique baseline to understand hyperspectral characteristics and seasonality of urban tree species in Hong Kong that can also contribute to hyperspectral imaging and database development elsewhere in the world.

1. Introduction

Urban trees benefit our environment and human inhabitants in multi-facet dimensions, such as reducing the urban heat island effect, enhancing biodiverse habitat, increasing the aesthetic value of the street view and relieving mental distress. It has long been unveiled and increasingly reported by numerous researchers (Alonzo et al., 2014; Bolund and Hunhammar, 1999; Escobedo et al., 2011; Escobedo and Nowak, 2009; Gómez-Baggethun et al., 2013; Liu et al., 2017; Lyttimäki et al., 2008; Nowak et al., 2008; Tratalos et al., 2007). Yet, the benefit of services may vary according to tree species, structures, and locational

contexts, which represent essential criteria to measure the service quality of urban trees, thereby maintaining the ecosystem sustainably at a vibrant level (Escobedo et al., 2011). As such, the ability to map tree species and analyze tree health conditions would become vividly essential (Degerickx et al., 2018; Leckie et al., 2005). However, tree information collection and periodic updates against the biophysical conditions of tree data is essential to keep track of the tree health status which in turn, poses challenges to the current tree health monitoring and species identification using conventional visual tree assessment, alongside the rise of many associated costs, including time-consuming, labour-intensive, costly fieldworks, and sometimes, unable to provide

* Corresponding author at: Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region.
E-mail addresses: sabbas@polyu.edu.hk (S. Abbas), qian.peng@connect.polyu.hk (Q. Peng), ls.charles@polyu.edu.hk (M.S. Wong), dean.ge@swjtu.edu.cn (Z. Li), wangjicheng123@sicnu.edu.cn (J. Wang), cla.lsc@hyd.gov.hk (K.T.K. Ng), yt-coco.kwok@connect.polyu.hk (C.Y.T. Kwok), karena.kw.hui@polyu.edu.hk (K.K.W. Hui).

¹ Joint first authorship.

<https://doi.org/10.1016/j.isprsjprs.2021.05.003>

Received 31 January 2021; Received in revised form 29 April 2021; Accepted 4 May 2021

Available online 25 May 2021

0924-2716/© 2021 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

entire coverage within the heterogeneous urban environment (Alonzo et al., 2014).

Remote sensing techniques can map urban trees to monitor complex urban environment changes effectively. A dozen of research showed that the hyperspectral data obtained from space (Clark, 2020; Ferreira et al., 2019; Ghosh et al., 2012; Liu et al., 2017; Zhang et al., 2013), airborne (Ben-Dor et al., 2002; Dadon et al., 2019; Lin et al., 2018; Osco et al., 2020; Schiefer et al., 2020; Wu et al., 2021; Yan et al., 2020; Zarco-Tejada et al., 2000; Zarco-Tejada et al., 2001) and ground-based platforms (Cho and Skidmore, 2006; Cochrane, 2000; Delegido et al., 2014; Gong et al., 1997; Jensen et al., 2012). Kothari et al. (2018) have been widely used for species identification and vegetation monitoring. Furthermore, detecting changes in spectral response patterns can help to identify variability in plant health indicators (Zhang et al., 2020), such as leaf area index (Halme et al., 2019), the amount of live and senesced biomass (Laurin et al., 2014), moisture content (Cotrozzi et al., 2017; Shen et al., 2021), plant pigment contents (e.g. chlorophylls and anthocyanins) (Cho and Skidmore, 2006; Delegido, 2014; Thenkabail et al., 2018), canopy nitrogen concentration (Townsend et al., 2003), non-structural carbohydrates (Asner and Martin, 2015); tree species diversity and distribution mapping (Paz-Kagan et al., 2017), functional and biological diversity (Asner et al., 2014). Nevertheless, retrieval accuracies from airborne and spaceborne imagery are greatly influenced by the complex atmospheric conditions as well as highly mixed urban landscapes (Yan et al., 2020). Although various non-imaging and hyperspectral imaging sensors are available, the measuring process and the handling of most of these sensors are rather complex (Behmann et al., 2018). Airborne acquisition methods rely on whisk or push broom systems. The spatial referencing is assisted with an inertial measurement unit and global navigation satellite systems. And light-weight platforms, like UAVs, have the advantage of lower flight altitudes but do not provide such high-quality correction signals (Aasen et al., 2015; Rossini et al., 2015). Therefore, researchers adopted in-situ or ground-based spectroscopy measurements to achieve high spatial resolution plus the advantage to minimize the complex atmospheric correction for the images (Katkovsky et al., 2018).

Several spectral libraries consisted of vegetation information have been developed which are available online, for instance, the Global Hyperspectral Imaging Spectral-library of Agricultural (GHISA) provides a hyperspectral library of the five major crops (e.g., winter wheat, rice, corn, soybeans, and cotton) for the Conterminous United States (Thenkabail and Aneece, 2019), the ASTER Spectral Library established by the National Aeronautics and Space Administration (NASA)'s Jet Propulsion Laboratory (JPL) (Baldrige et al., 2009), the SPECCHIO Spectral Library maintained by the Remote Sensing Laboratories in the Department of Geography at the University of Zurich (Bojinski et al., 2003); the Vegetation Spectral Library developed by the Systems Ecology Laboratory at the University of Texas (Goswami, 2011; Goswami and Matharasi, 2015); the Spectranomics (Asner and Martin, 2016); and the Ecological Spectral Information System (EcoSIS, 2014; Kothari et al., 2018). Although the current libraries have wide coverage of spectral information, the spectral signatures of different tree types are of a limited amount; whilst many native tree species in the subtropical and tropical zones, like *Celtis senensis* and *Macaranga tanarius* var. *tomentosa*, are rather unlikely to be found in any of these libraries. Located in the subtropical zone with a high diversity of tree species, Hong Kong is well-known for its unique landscape features where most of the trees are grown in an intense urban environment along with hilly topography, compact skyscraper buildings and a high population density. Therefore, this study with a focus on the abundant tree species in Hong Kong significantly contributes the hyperspectral imaging and database development in the world to bridge the gap of studying plant species using hyperspectral data.

This study built a hyperspectral library for urban tree species in Hong Kong, comprising of 19 species, six families, 75 urban trees, and 450 images, acquired by a terrestrial hyperspectral camera from November

2018 to October 2019 in different seasons. A Deep Neural Network is developed to identify tree species and the hyperspectral phenology models are developed to achieve high accuracy and optimization of data acquisition. The organization of this paper is as follows: in Section 2, the study area, datasets and the adapted methodology are presented. Section 3 presents the results and discussions. Section 4 presents the conclusions and possible future works.

2. Data and methods

2.1. Data

2.1.1. Target tree selection

The selection criteria of targeted trees include but not limited to the tree species, tree health condition, types of defects, tree height, size of a tree canopy, proximity to the road, traffic flow at the nearby highway, and spatial distribution of the samples. Highways trees are nurtured to stabilize slopes, establish a green backdrop to the highway system, and enhance the local ecological habitats in the urban area selected for this study. The selected trees (Table 1) were located in 12 different sites in Hong Kong's urban landscape (Figure S1). Among all of these selected trees, 55 trees were healthy, and 20 trees were identified as unhealthy in the database provided by the Hong Kong Government.

2.1.2. Image acquisition plan

In-situ measurements were adopted to retrieve high spatial resolution and to minimize the atmospheric disturbance. A novel hyperspectral camera, named SPECIM IQ (Model No.: 0604675, Oulu, Finland), was procured to acquire in-field hyperspectral images. The SPECIM IQ can capture a full hyperspectral image without external movement and the slight temporal delay of capturing different parts (spatial or spectral) of the image. The camera acquires hyperspectral imagery in 204 narrow bands with a spectral resolution of 7 nm in Visible and Near Infra-Red wavelengths (400 - 1000 nm) of the electromagnetic spectrum approximately. The sampling interval is around 2.94 nm, and the central wavelength of each band is given in supplementary material Table 5S. The bandwidth can be binned by 2x (102 bands) or 3x (68 bands). The camera provides a spatial sampling of 512 by 512 pixels covered by a field of view of 31° by 31°. An image acquired by the camera at a 1m distance from the target captures an area of 0.55 m by 0.55 m and divides the target into the ground sampling unit of 1.07 mm. It can capture a target from a minimum distance of 150 mm to an infinite distance. The camera weight is 1300 g with a portable size of 207 mm (length) by 91 mm (width) by 125.5 mm (depth - without lens is 74 mm).

For maintaining the consistency of data acquisition under varying weather conditions and seasonal changes, the simultaneous method of White Reference is preferred for this study. The most critical parameter in field reflectance measurements is the irradiance changes due to atmospheric effects and variation in solar illumination. Since most reflectance measurements continue to be single Field of View (FOV) measurements, the time between the reference and the target measurement should be minimized to eliminate the possible atmospheric-induced changes. This optimization is also required because the illumination characteristics must be the same for the reference and target measurements. Any change in position of observer or illuminating geometry between a reference and a target scan introduces a bidirectional reflectance distribution function (BRDF) relating to the change of the signal. Besides, when using any field reflectance standard, the correction for the standard's non-Lambertian behaviour must be taken into account.

By incorporating any field standard ρ_{is} , the reference measurement can be modified based on Eq. (1) as follows:

$$\rho(\lambda)_{corr} = \Phi_{\lambda r, t_1} \frac{\rho_{is}}{\Phi_{\lambda i, t_0}} \quad (1)$$

Table 1
Description of selected trees.

No.	Species	Family	No. of Healthy	No. of Unhealthy	Abbreviation
1	<i>Acacia auriculiformis</i>	Fabaceae	3	Inaccessible	a.aur
2	<i>Acacia confusa</i>	Fabaceae	5	1	a.con
3	<i>Albizia lebbek</i>	Fabaceae	3	Inaccessible	a.leb
4	<i>Aleurites moluccana</i>	Euphorbiaceae	7	3	a.mol
5	<i>Bauhinia variegata</i>	Fabaceae	1	1	b.var
6	<i>Broussonetia papyrifera</i>	Moraceae	2	Inaccessible	b.pap
7	<i>Casuarina equisetifolia</i>	Casuarinaceae	4	1	c.equ
8	<i>Celtis sinensis</i>	Cannabaceae	4	2	c.sin
9	<i>Cinnamomum camphora</i>	Lauraceae	3	1	c.cam
10	<i>Cinnamomum parthenoxylon</i>	Lauraceae	1	1	c.par
11	<i>Delonix regia</i>	Fabaceae	3	Inaccessible	d.reg
12	<i>Ficus hispida</i>	Moraceae	4	Inaccessible	f.his
13	<i>Ficus microcarpa</i>	Moraceae	2	1	f.mic
14	<i>Ficus variegata (var. chlorocarpa)</i>	Moraceae	3	Inaccessible	f.var
15	<i>Ficus virens (var. sub lanceolata)</i>	Moraceae	2	Inaccessible	f.vir
16	<i>Litsea monopetala</i>	Lauraceae	1	Inaccessible	l.mon
17	<i>Macaranga tanarius var. tomentosa</i>	Euphorbiaceae	2	6	m.tan
18	<i>Machilus chekiangensis</i>	Lauraceae	3	Inaccessible	m.che
19	<i>Mallotus paniculatus</i>	Euphorbiaceae	2	3	m.pan
Total			55	20	

where $\rho(\lambda)_{corr}$ is the spectral reflectance corrected for non-Lambertian field reference, $\rho_{\lambda s}$ is the spectral reflectance factor of the field standard. Since in most cases, $t_0 - t_1 \neq 0$ with $\Delta t < 15\text{min}$, $\rho_{\lambda s}$ must be applied to $\Phi_{\lambda i}$ before multiplying it with the reflected radiance. Quantification of atmospheric stability in selected wavelengths can be given using sun-photometer measurements. The measurement time affects the irradiance since it is associated with varying optical transmittance. The Labsphere Spectralon panel is used as a reference in the field measurement for its near-perfect reflectance over 250–2500 nm region and thermal stability. Stratified random sampling is performed in the field measurements. When the observation height of the target is confined to two meters and FOV is about 2°, the representative sample size is considered to be 30–40 (Schaepman, 1998).

The hyperspectral images of the 75 selected trees comprising 19

species were acquired in six rounds, each round every two months, from November/December 2018 to October 2019. The Round 1 (R1) was carried in November-December 2018, followed by Round 2 (R2) in February 2019, Round 3 (R3) in April 2019, Round 4 (R4) in June 2019, Round 5 (R5) in August 2019 and Round 6 (R6) in October 2019.

A detailed flow of step by step procedures to acquire and analyze hyperspectral images from the terrestrial hyperspectral camera is illustrated in Fig. 1.

2.2. Methods

2.2.1. Masking canopies and homogenous regions

Most of the captured images included background areas and canopies of adjacent trees that were mandatory to be excluded before automatic computing of mean canopy signatures and the images'

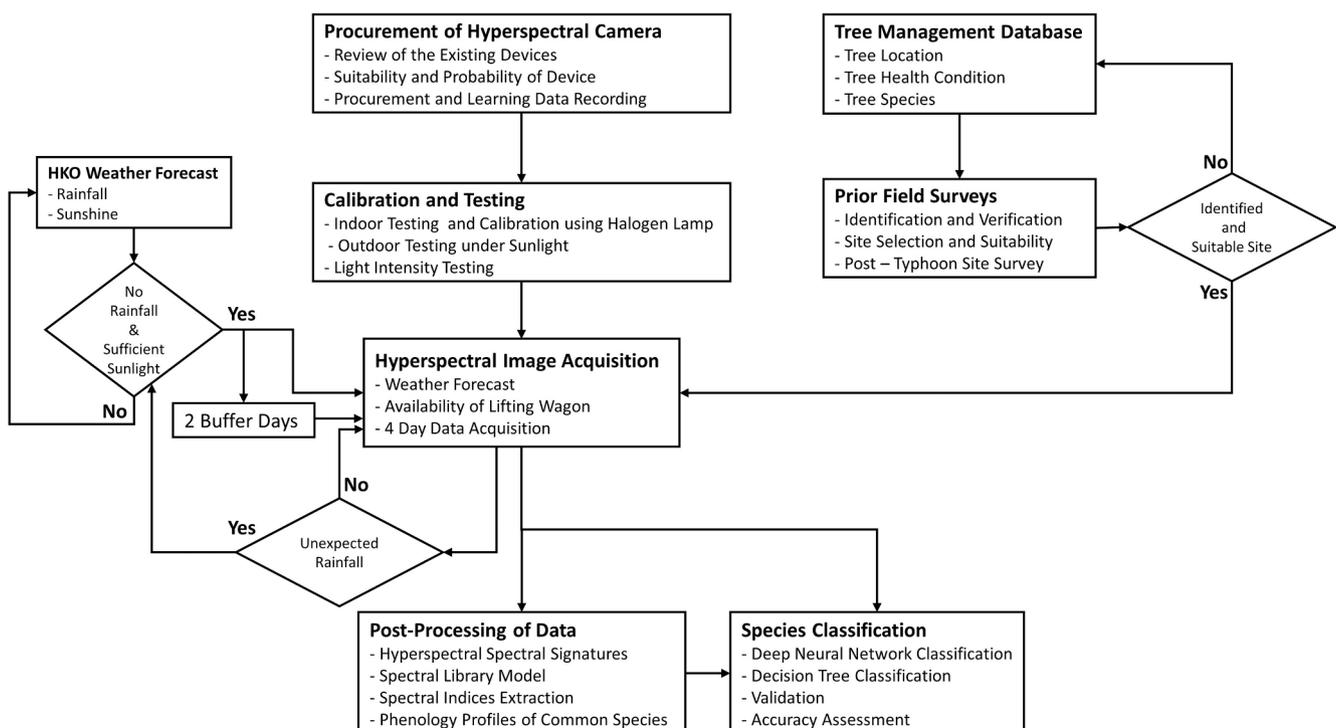


Fig. 1. The flow diagram of the study. HKO refers to Hong Kong Observatory.

classification. Therefore, a masking procedure was applied to every image to mask out background areas, including roads, footpaths, buildings, and adjacent tree canopies, and retain homogenous canopy regions of the images acquired during different rounds of data collection. A few typified examples of the masking procedure are shown in Fig. 2.

2.2.2. Unsupervised clustering of images

A tree canopy is comprised of varying arrays of leaves – healthy, unhealthy, sunlit, shadowed, new, young, and mature leaves. Also, tree branches, trunk, and canopy shadow contribute to the accumulative reflectance. Therefore, prior to the classification of the hyperspectral images, an unsupervised k-means clustering algorithm was applied to all the images to group an image's pixels into 6 clusters, indicating leaf conditions, tree trunk and branches, as well as to exclude canopy background and saturated pixels. As hyperspectral remote sensing images often appear as “synonyms spectrum” and “foreign body with the spectrum”, the same class may have different spectral features (Fig. 3).

These clusters represent naturally similar regions by minimizing within-cluster variability (sum of the squared distance between each pixel) and maximizing cluster homogeneity. The clustering process starts with the analyst's input for a number of required clusters, ($n = 6$), with other parameters to be located in the dataset. A set of 'n' clusters are randomly created in the multidimensional feature space (with 204 bands), and each pixel is allocated to the nearest class with a minimum distance between the cluster centroid. Later, these clusters are repetitively re-clustered and updated based on the inclusion of new pixel until maximum spectral separability is obtained (and no new cluster formation is generated) based on the closest distance-to-centre decision rule (Im and Jensen, 2005). The unsupervised clusters require to be assigned an appropriate class, such as healthy, unhealthy, healthy shadowed, healthy sunlit, trunk or branches, background and saturated. A mean canopy signature was computed by excluding all the clusters or pixels labelled as non-vegetated, background or adjacent vegetation, and saturated. It was calculated for all the 75 trees in each round and a unified signature plot is created to understand changes in spectral response during a different part of the year. The combined plot of mean canopy signatures of each tree, acquired during different seasons (December, February, April, June, August, and October) indicates unique variations in spectral characteristics (see supplementary

material section S7).

2.2.3. Vegetation indices

Remote sensing of vegetation monitoring generally relies on vegetation indices (VIs) which compare the reflectance of vegetation in multiple spectral regions (Nijland et al., 2014). The VIs derived from RS provides efficient and straightforward means for qualitative and quantitative mapping of vegetation cover, health, leaf chlorophyll content, leaf area, canopy cover, and structure (Xue and Su, 2017). The Normalized Difference Vegetation Index (NDVI) is the most widely applicable VI to monitor vegetation's greenness or health condition. It is derived from converse spectral response patterns of vegetation in the NIR and Red portion (Equation (2)) of the solar electromagnetic spectrum (Maselli, 2004; Tucker, 1979).

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2)$$

The NDVI shows a high correlation with vegetation health, green biomass, and vegetation productivity (Pettorelli et al., 2005). Since the mesophyll layer's collapse occurs earlier than a decline in chlorophyll and before any visible changes, the NDVI is an early indicator of plant stress. NDVI values range from -1 to $+1$, with values above 0 generally representing vegetation (Tucker and Sellers, 1986).

To minimize the changes associated with the diurnal angle of the sun, the applicability of the Photochemical Reflectance Index (PRI) (Gamon et al., 1992) was also explored to study the phenological changes of the 19 species. The PRI is a narrow-band reflectance-based photosynthetic index derived by the normalized difference of reflectance around the wavelength regions of 531 and 570 nm (Equation (2)) (Gamon et al., 1992). It is an effective indicator of photosynthetic radiation use efficiency, CO_2 uptake and nutrient deficiency in plants (Gamon et al., 1997).

$$PRI = \frac{R531 - R570}{R531 + R570} \quad (3)$$

2.2.4. Deep Neural Network classification for species identification

A Deep Neural Network framework was developed to train and validate the species-level classification using hyperspectral signatures. Neural networks are a set of algorithms, modelled loosely after the human brain, are designed to recognize patterns. With the rapid

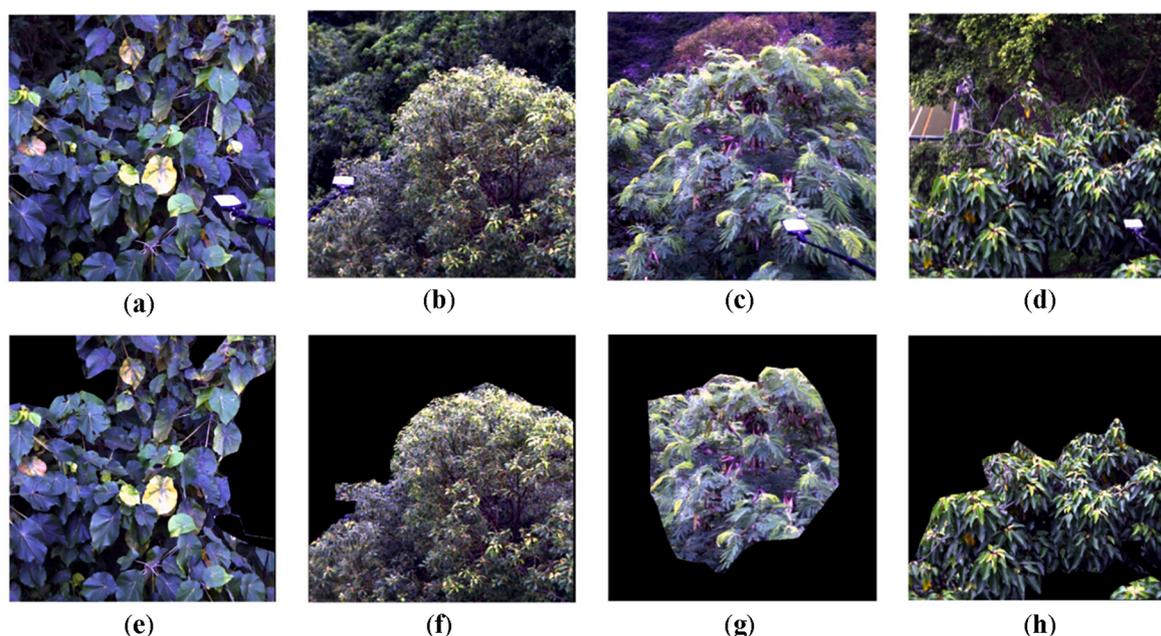


Fig. 2. Typified examples of masking canopies and homogenous regions: (a-d) are original images; (e-h) are corresponding masked images.

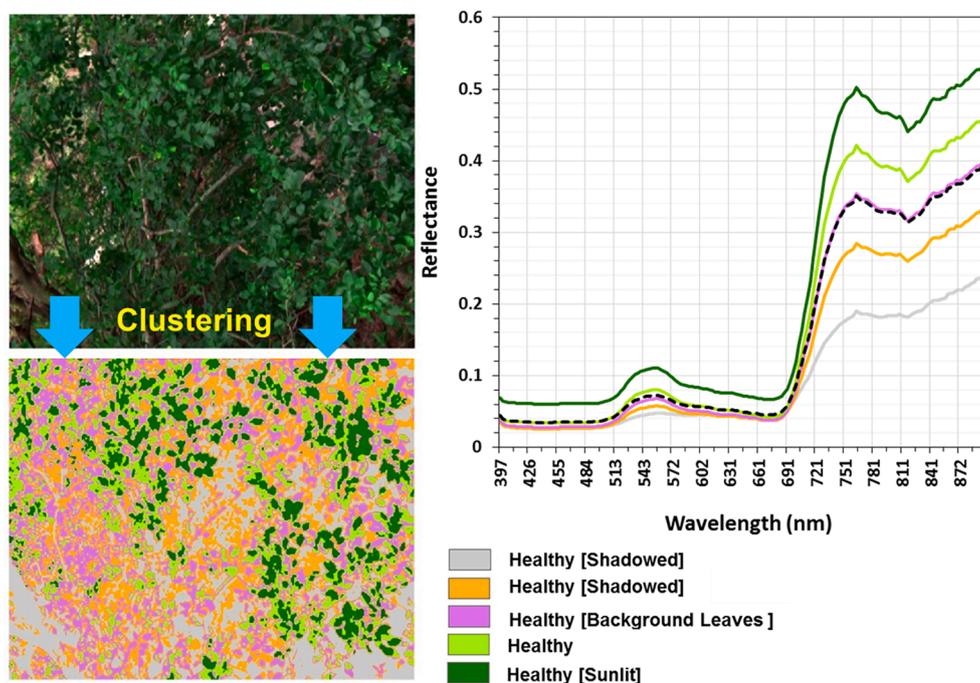


Fig. 3. An example of image clustering and corresponding spectral signatures of classes. The shadow class in grey represents canopy shadow and/or branches, and the shadow class in orange indicates shaded leaves. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

development of computer hardware and increased data, neural networks have been shown as one of the most effective tools for data analysis. A Deep Neural Network was developed and applied to identify the 19 species from hyperspectral images of the 75 trees in the urban landscape of Hong Kong. The structure of the network is shown in Fig. 4. The network inputs are the spectral values (spectral curve or bands), and outputs are the species' probabilities. It contains one input layer (512 neurons), one output layer (19 neurons), and three hidden layers (1024, 1024, 512 neurons separately). The SoftMax function was applied at the

output layer to generate the probabilities.

For the hyperspectral image (spectral cube), each pixel has a spectral signature that contains 204 bands. The size of each image is 512×512 (262, 144) pixels. It was not reasonable to use the whole image to train the neural networks. Some parts of the images do not belong to the target tree canopies, e.g., background objects and adjacent canopies of neighbouring trees. Therefore, all the images are first masked and clustered to exclude irrelevant parts from the images. To train and develop the Deep Neural Networks, a sample of $\sim 50,000$ pixels for each

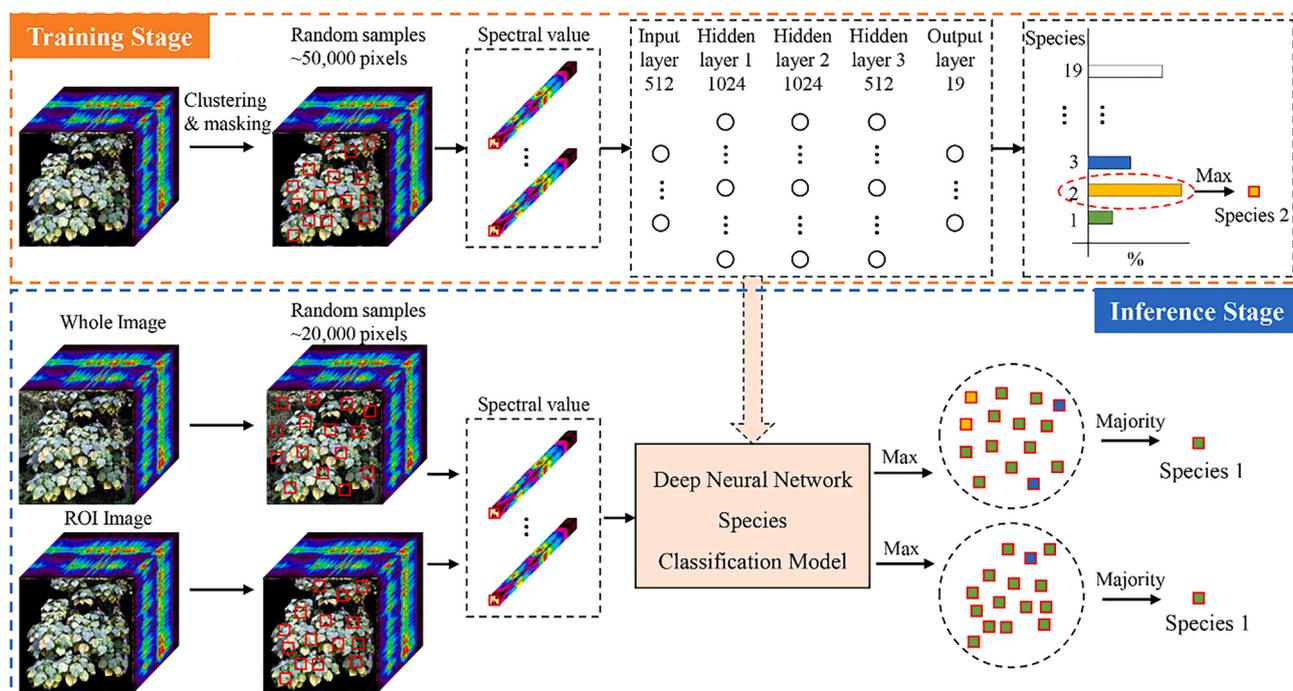


Fig. 4. The overall workflow of species classification framework using the Deep Neural Network modelling.

image were randomly extracted. The model was comprised of two stages, i.e., the Training Stage and the Inference Stage. In the training stage, 3/4th of the sample data from each image was undertaken to construct the training dataset. Each sample of the training dataset contains an input vector and an output vector. The input vector was the spectral curve (the original spectral signatures of the pixels) and the output vector was the 'one-hot coding' vector of species. The remaining 1/4th of the sample data was applied for the individual test and validate the training models of the Deep Neural Networks. In the inference stage, another set of samples, comprising ~20, 000 pixels, were selected from each image. The trained model was operated over these new samples. The model produces a probability of each sample to every species, for example, a sample pixel may have an 80% probability of matching with species A, 10% probability of matching with species B and 5% probability of reaching with species C. Thus, each sampled-pixel was labelled with the species name of the species with the highest probability. In this way, each sampled pixel from an image was classified as a related species name, where the image was classified into a corresponding species based on the class of the majority of the 20, 000 sampled pixels. These samples were then separately taken from "whole image" and from "masked or ROI image", and a similar prediction and classification procedure was also applied to the masked or ROI images to avoid the influence of irrelevant scene elements (background object, saturated pixels and pixels from adjacent tree canopies). In addition to training and validating six models for each round (R1, R2, R3, R4, R5, and R6), another set of four models were trained and tested by combining the datasets from different rounds, for instance, model R13 represents a model trained and tested for a combined dataset of R1 and R3. Similarly, the models R24, R35, R1234 were developed by combining the data acquired in corresponding rounds. The rounds were combined by skipping a consecutive round to avoid model overfitting due to the least phenological difference (Fig. 7) as most of the tree species are evergreen.

2.2.5. Optimal bands for species differentiation

Visual interpretation of the hyperspectral signature may be helpful, but it could suffer from analyst bias. Therefore, a recursive partitioning (rpart) of the multivariate regression tree approach is applied to build a classification tree for the species according to the hyperspectral characteristics of the species (De'Ath, 2002). The general structure of the algorithm is constructed based on a two-stage procedure and the resulting models are represented as binary trees (see [supplementary material](#) section S6).

3. Results and discussions

3.1. Spectral reflectance curves of healthy and unhealthy

After collecting field data, two groups of categories were identified, i.e., healthy and unhealthy for each species. Representative spectral reflectance curves of each healthy and unhealthy species were extracted and plotted for comparison (Fig. 5).

The graph associated with the spectral profiles of *Acacia confusa* is shown in Fig. 5, including a sample of healthy leaves and a sample of extremely stressed yellow leaves. The yellow leaves show a higher reflectance in the red and green regions of the electromagnetic spectrum, indicating a decline of the photosynthetic process such that the plant is less able to absorb the red light. On the other hand, many unhealthy leaves reflect a lower spectral intensity of the NIR region, speculating a collapse of the mesophyll structure within the leaf itself.

Similarly, spectral profiles of *Delonix regia* and *Macaranga tanarius* showed varying spectral responses in the leaves of the healthy and unhealthy canopy. These health-related changes in spectral profiles in visible regions could be visually determined as a decline of photosynthetic activity, but as if the change of internal leaf structure caused by abiotic stress, such as water-deficit stress, could be detected and analyzed by the increasing/decreasing reflectance in NIR portion.

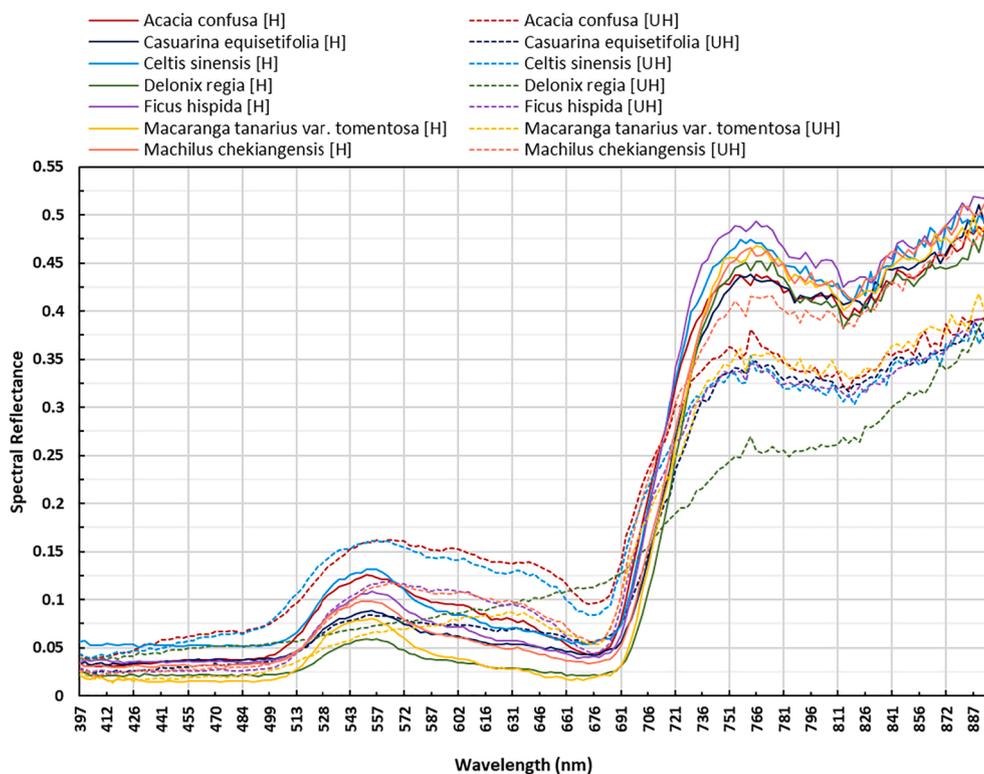


Fig. 5. All spectral profiles of healthy and unhealthy trees among different species. 50, 000 pixels were randomly selected from the hyperspectral photograph (after masking tree canopy or removing background) to calculate the mean canopy signature for each sample tree. Please refer to [Table 1](#) for the number of tree samples for the corresponding species.

Furthermore, a decline in the photosynthetic process could also be indicated by reducing fluorescence energy around 760 nm wavelength. When absorbed by chlorophyll, solar energy is used for carbon fixation, heat dissipation, followed by the release of the emittance source at a longer wavelength in the form of chlorophyll fluorescence (Krause and Weis, 1991). Photosynthesis is the process through which plants make food to supply nutrients for their growth. This process of photosynthesis involves three steps; firstly, the plant converts chemicals into carbohydrates; secondly, heat is released as a by-product; and the third process is related to chlorophyll fluorescence in which plants release low energy photons during photosynthesis, which can give us information about the health of the plants. Spectral profiles of *Acacia confusa*, *Delonix regia* and *Macaranga tanarius*, a significant dip around 760 nm wavelength in plants indicated lesser fluorescence released because of the declining photosynthetic process. On the other hand, healthy plants showed a sharp rise, and a very subtle variation, in the fluorescence region. This drastic change in reflectance at ~760 nm can be clearly observed in the signature profile of healthy *Acacia confusa*. However, the fluorescence level in healthy plants does change along with diurnal and seasonal variations and varies across the species. It is interesting to note that the dotted green line (an unhealthy sample of the species *Delonix regia*) showed the lowest reflectance in the NIR region but with high fluorescence in the 760 band which might be related to the oxygen absorption band. It is difficult to explain the association between high fluorescence and lower NIR reflectance which might be an unusual pattern. It is hard to say if this is associated with the oxygen absorption band. During field visits, it has been observed that the spectral profiles of unhealthy canopies or deciduous tree canopies at the senescing stage are often influenced by a strong reflectance from under canopy vegetation. It required

more research and experiments to fully understand and explain such unusual patterns.

Another set of examples of healthy and unhealthy spectral profiles of *Casuarina equisetifolia*, *Celtis sinensis*, *Ficus hispida*, and *Machilus chekiangensis* is provided. Higher reflectance in the visible wavelength region means that the leaves of unhealthy plants are not performing photosynthesis actively. In contrast, the second clue was also indicated by a dip in the fluorescence region (~760 nm) (an observable sharp dip in the spectral profile of *Celtis sinensis*) (Migliavacca et al., 2017). However, the leaf structure was still intact, which indicates a possibility of significant moisture deficiency or photosynthetic decline. Another feasible explanation could be the plants' seasoning cycle as shown in the phenology plots of each species using numerous vegetation indices (Fig. 7).

Among all of these species, the healthy *Acacia confusa* has the lowest reflection in the NIR region, while the unhealthy *Delonix regia* has the higher reflectance. In the visible portion of the electromagnetic spectrum, healthy *Macaranga tanarius* is found to have the minimum reflectance and the unhealthy *Acacia confusa* shows the highest reflectance an equal amount of green and red reflectance, which is an indicative sign of severely stressed yellow leaves.

3.2. Seasonality analysis

3.2.1. Bi-monthly mean canopy signatures

All the images were split into six clusters and an appropriate label was assigned to each cluster. Later, a mean canopy signature was computed by excluding all the clusters or pixels labelled as non-vegetated, background or adjacent vegetation, and saturated. It was

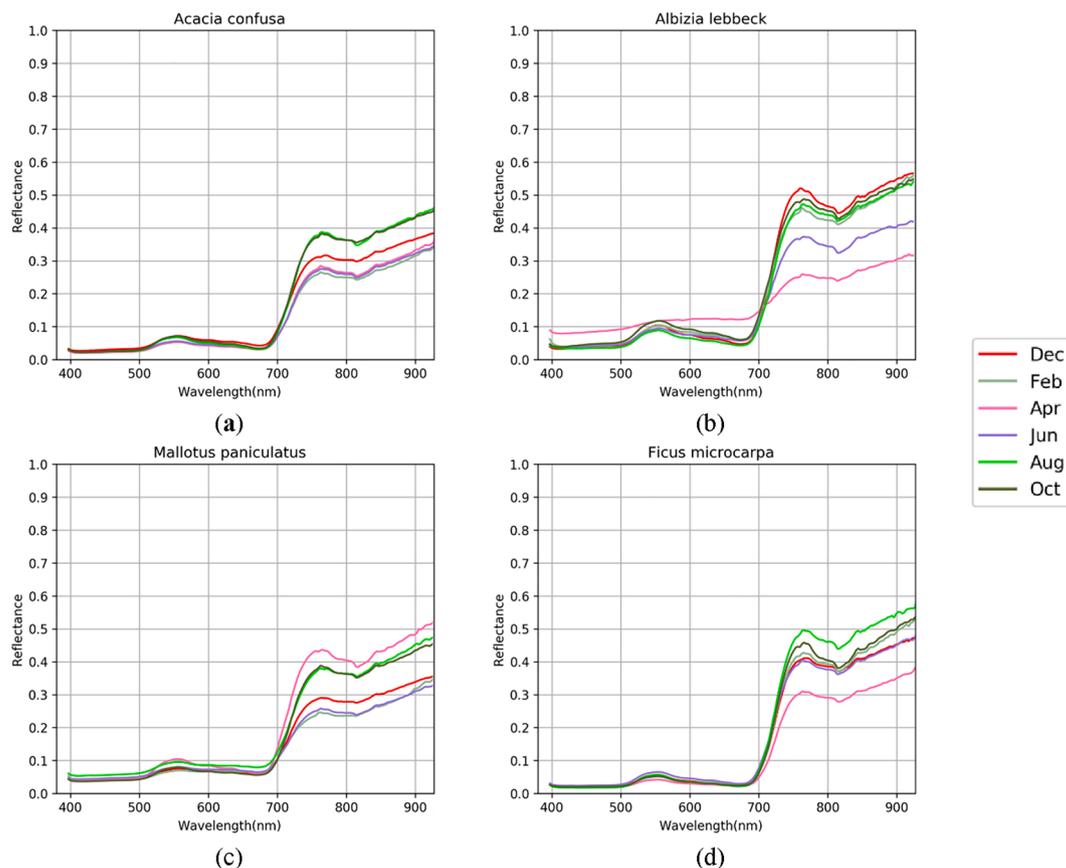


Fig. 6. Mean canopy spectral signature of different species in each round of in-situ data acquisition, the species are: (a) *Acacia confusa* (N = 6); (b) *Albizia lebeck* (N = 3); (c) *Mallotus paniculatus*; (N = 5); (d) *Ficus macrocarpa* (N = 3). N indicates the number of tree samples for the corresponding species; 50, 000 pixels were randomly selected from the hyperspectral photograph (after masking tree canopy or removing background) to calculate the mean canopy signature for each sample tree.

calculated for all the 75 trees in each round and unified signature profiles were created to understand changes in spectral response. The combined plot of mean canopy signatures of each tree, acquired during different seasons (December, February, April, June, August and October) indicated unique variations in spectral characteristics (Fig. 6 and supplementary material section S7). Fig. 6 represents the spectral signature profiles of representative trees of the four species – (a) *Acacia confusa*, (b) *Albizia lebbek*, (c) *Mallotus paniculatus* and (d) *Ficus hispida*. It is interesting to note that the *Acacia confusa* and *Ficus hispida* showed a gradual increase in reflectance, especially in the NIR region of the spectrum, from winter to summer though both species possess different spectral signatures but show a relatively similar pattern in seasonality. This can be attributed to the nature of both of these species that are evergreen with no significant leaf-off season (Fig. 6 a & d). On the other hand, *Albizia lebbek* and *Mallotus paniculatus* showed a decline in photosynthetic activity due to the leaf-off period during February and April, respectively. It was also observed during the data acquisition that the *Albizia lebbek* started to shed leaves in February and all the leaves were fell-off by April (Fig. 6 b). Nonetheless, this species starts to regrow leaves with the start of the summer season, and as a result, reflectance started to increase in the images acquired in June and August with the presence of younger and newer leaves (Fig. 7). *Mallotus paniculatus* (unhealthy tree) started leaves-shedding earlier in February. The leaves' regrowth starts in April and higher reflectance is observed in June due to full canopy coverage (Fig. 6 c).

3.2.2. Vegetation indices (VIs) analysis of seasonality

High spectral resolution (higher number of narrow bands) images acquired by the SPECIM IQ provides sufficient spectral information, particularly in visible and NIR spectral wavelength region. These narrow bands can effectively be used to identify species with varying growing and health conditions in the urban landscape of Hong Kong. Therefore, in addition to 204 spectral bands of the hyperspectral images, 60 VIs were computed to enhance the spectral feature space by combining information of different bands, which were obtained from the online index database (<https://www.indexdatabase.de/>) and the wavelengths were substituted by corresponding bands of the hyperspectral images (please

see Table S1 for the list of VIs).

Image acquisition was followed by the development of phenology of the species using conventional multispectral broadband vegetation indices (e.g., Simple Ratio, Normalized Difference Vegetation Index, and Enhanced Vegetation Index) and advanced hyperspectral narrow-band indices (e.g., Hyperspectral Vegetation Index, Greenness Index, Normalized Difference Infrared Index, Red-Edge Vegetation Stress Index and Photochemical Reflectance Index). A representative pattern of vegetation changes along the year using NDVI and PRI is shown in Fig. 7 and S4. For example, the first six asterisks showed the trend of NDVI for *Acacia auriculiformis* all year round. The NDVI of *Acacia auriculiformis* keeps increasing and reaches a peak in June. But for *Acacia confusa*, the NDVI keeps increasing and a slight drop exists around June for its flowering season. The yellow flowers had a significant influence on the index. For *Albizia lebbek*, there is a significant drop in April and the trends before and after April are quite different due to its flowering season is around May, and its fruiting season is July to December. *Casuarina equisetifolia* is an evergreen tree, so the trend of NDVI keeps increasing after winter. The flowering season of *Machilus chekiangensis* is observed in December and in the last third of July, its fruits turn from green to black, so there was a drop in August. *Mallotus paniculatus* is an evergreen tree and its lower surface of leaves is white or rusty white.

In peak growing season, almost all species showed high reflectance at the upper canopy and under canopy levels during the summer period, which made it difficult to differentiate among different species. On the other hand, during the start (February) and end (October) of the growing season, upper canopies of trees vary due to different phenological cycles, which can distinguish among different species. Also, during these periods, under canopy vegetation (grasses or shrubs) were relatively less green or dry. Thus, the spectral response of trees exclusively comprised of upper canopies and showed variation among different species.

3.3. Deep Neural Network classification for species identification

Six models were trained and tested separately for each round of the datasets (Fig. 4). Then, four additional models were trained and tested

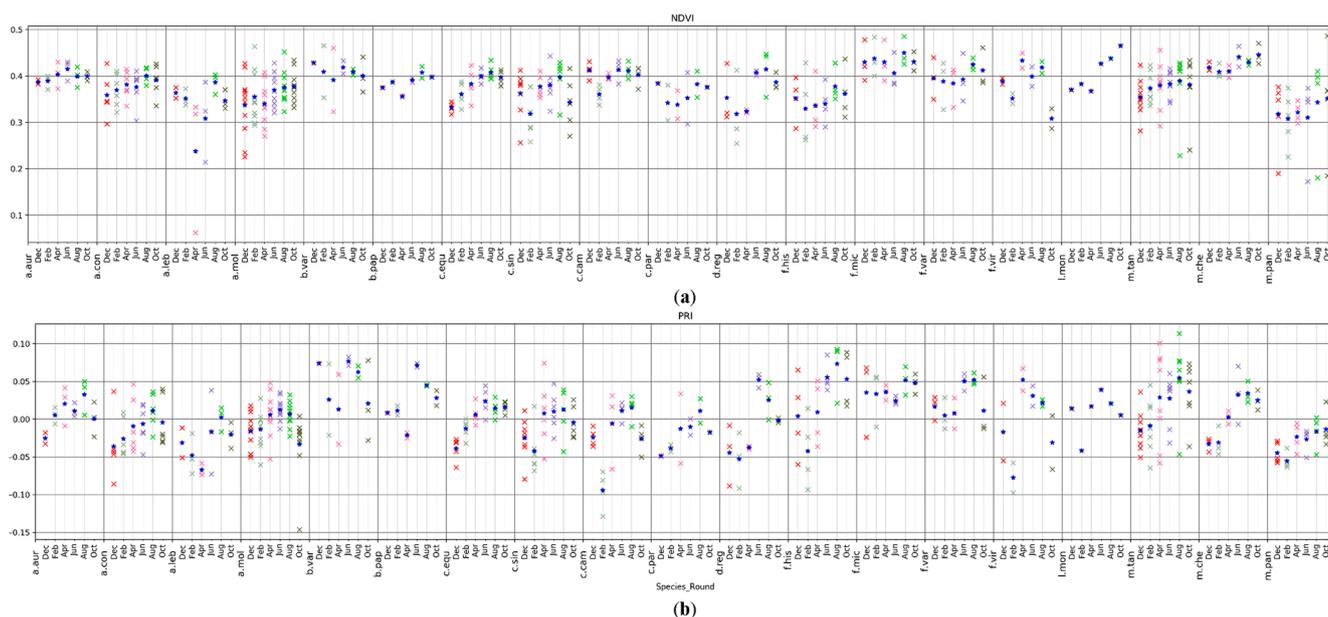


Fig. 7. Phenology patterns (changes in the vegetation greenness or productivity) of all the species using the Normalized Difference Vegetation Index (NDVI) (a) and Photochemical Reflectance Index (PRI) (b). For each species, an average value of an index (shown with blue asterisk *) is calculated from all the samples of a species (shown with different colours for each round), major x-axis label is showing the species name (please refer to Table 1 for full names of the species) while the minor x-axis label is representing each round (Dec (R1), Feb (R2), Apr (R3), Jun (R4), Aug (R5), and Oct (R6)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

by combining the datasets from different rounds. Besides, models trained on each round were cross-validated with the data from all other rounds. For example, the model developed for R1, was also validated with data collected during R2, R3, R4, R5 and R6. Similarly, the combined models (R13, R24, R35, and R1234) were also validated. The models' training accuracies (T), validation accuracies (V), overall accuracies, producer's accuracies (omission errors) and user's accuracies (commission errors) of classification inferences using the full image (All) and/or masked image (ROI), are presented in Table 2 and Table 3.

All the models showed very high accuracy for individual model testing, validation and inferences. The individual models' testing accuracy ranged from 0.856 to 0.946 and the validation accuracy ranged from 0.878 to 0.956, with the highest accuracy achieved for the model R6 (October). The accuracy of species classification from the inferences ranged from 0.853 to 0.932, for samples taken from the whole images, and ranged from 0.88 to 0.96, for samples taken from the masked images. Specifically, the maximum accuracy is achieved for R6 (October) for both full and masked images while the lowest accuracy is obtained for R3 (April) and R4 (June). The testing and validation accuracies were the highest for R5 (0.946 and 0.956) and R1 showed the lowest accuracies (0.856 and 0.878).

For the combined models, testing accuracy ranged from 0.891 to 0.923, and validation accuracy is ranged from 0.914 to 0.937. The highest testing and validation accuracy was achieved for R35 (combined datasets of Round 3 - April and Round 5 - August), in 0.923 and 0.937, respectively. It is interesting to note that the inferential classification results were higher for the rounds integrated with the training model. For example, for the combined model of R13, the Round 1 and Round 3 testing data showed the highest accuracy. These results were also indicative of seasonality in the datasets as the seasonality pattern was selected by the training models, which was depicted as a higher accuracy of the results obtained from the corresponding models that had been combined for the training. The producer's accuracies (omission errors) and user's accuracies (commission errors) for tree species identification using the Deep Neural Network classification along with detailed confusion matrices are given in Table 3 and supplementary material section S3. Most of the species show high (100%) producer's and user's accuracies. Nevertheless, poor accuracies for some of the species is indicative of sample size (Table 1) for species identification. This implies that future studies could consider increasing the sample size either by adding more samples for each species or by grouping species by their spectral families if it is improbable to increase the sample size. Another alternative could be benefitting from high spatial resolution hyperspectral imagery acquired from UAVs. It is important to note that the extraction of 50,000 pixels from each of the 75 samples might lead to autocorrelation problem. However, the Deep Neural Network

classification adopted in this study can self-learn to distinguish between the samples for classification; and the higher accuracies can be achieved by increasing samples size and amount of input data. Nevertheless, special attention should be paid if traditional classification approaches are used for species identification.

3.4. Separability of the species from the hyperspectral space and influence of seasonality

Results from the Deep Neural Network classification for species identification indicated that the data acquired in R6 (i.e., during October, autumn season) provided maximum prediction accuracy for species separability. It is difficult to visually differentiate spectral characteristics of species to find optimal wavelengths (or bands) for species identification. Therefore, a decision tree classification was performed for each round of data (supplementary material section S6). Like the Deep Neural Network classification (section 3.5, Table 2; and supplementary material section S2), the Decision Tree classification (supplementary material section S6) showed that the data acquired during the autumn season (R6 – October) can provide the maximum accuracy (overall accuracy = 0.80, kappa = 0.784) for species differentiation. Generally, the species can be categorized into six spectral groups distinguished by the bands B1 (400 nm), B7 (405 nm), B87 (649 nm), B91 (661 nm), B111 (721 nm), B174 (911 nm), and B204 (1000 nm) which can be further segregated to individual species level by incorporating other relevant bands (Figure S9, supplementary material section S6.3). *Aleurites moluccana*, *Casuarina equisetifolia*, *Acacia confuse*, *Bauhinia variegata*, and *Mallotus paniculatus* can be distinguished and isolated by the bands B17 (405 nm), B174 (911 nm), B111 (721 nm), B181 (933 nm), B40 (510 nm) and B199 (988 nm) while another group of species (*Cinnamomum camphora*, *Aleurites moluccana* and *Acacia confuse*) can be detected by including bands B87 (649 nm), B108 (712 nm), and B83 (637 nm). The third group of species (*Broussonetia papyrifera*, *Ficus variegata* (var. *chlorocarpa*), *Ficus macrocarpa*, *Acacia auriculiformis*, and *Macaranga tanarius* var. *tomentosa*) can be recognized by including B204 (1000 nm), B113 (727 nm), B22 (458 nm), B1 (400 nm), and B5 (409 nm). It is important to note that the bands in the blue wavelength region played an important role in differentiating among the species. The information in the blue bands is often neglected in vegetation monitoring from satellite remote sensing due to a higher atmospheric perturbation, however, the wavelength region might play a pivotal in the identification of tree species from remote sensing datasets. The right branch of the decision tree characterized the other three groups of species. B7 (405 nm), B204 (1000 nm) and B1 (400 nm) can help to label *Ficus virens* (var. *sublanceolata*) and *Celtis sinensis*. *Mallotus paniculatus* and *Albizia lebeck* were distinguished with additional bands – B42 (516 nm), B9 (420 nm)

Table 2
Overall accuracies of all the models for species identification.

Model	R1	R2	R3	R4	R5	R6	R13	R24	R35	R1234	
T ¹	0.856	0.892	0.919	0.910	0.946	0.923	0.905	0.891	0.923	0.897	
V ²	0.878	0.908	0.934	0.923	0.956	0.936	0.927	0.914	0.937	0.923	
R1	All	0.918	0.178	0.123	0.082	0.147	0.082	0.932	0.205	0.137	0.932
	ROI	0.932	0.164	0.123	0.164	0.164	0.068	0.945	0.192	0.164	0.945
R2	All	0.147	0.920	0.133	0.133	0.137	0.120	0.933	0.187	0.933	
	ROI	0.147	0.960	0.160	0.173	0.164	0.133	0.227	0.947	0.227	0.947
R3	All	0.067	0.187	0.853	0.160	0.147	0.080	0.867	0.267	0.827	0.893
	ROI	0.067	0.200	0.880	0.160	0.133	0.093	0.867	0.253	0.853	0.893
R4	All	0.147	0.200	0.080	0.853	0.187	0.093	0.160	0.853	0.173	0.867
	ROI	0.133	0.200	0.080	0.880	0.200	0.107	0.133	0.920	0.200	0.893
R5	All	0.147	0.107	0.107	0.147	0.907	0.187	0.253	0.200	0.907	0.227
	ROI	0.133	0.067	0.120	0.160	0.933	0.173	0.187	0.173	0.920	0.200
R6	All	0.081	0.095	0.149	0.068	0.216	0.932	0.108	0.149	0.203	–
	ROI	0.081	0.135	0.176	0.095	0.284	0.959	0.135	0.162	0.230	–

¹ T is Training accuracy and

² V is Validation accuracy; The training accuracy is the correct rate of training dataset while validation accuracy is the correct rate of the testing dataset. All stands for the testing points sampled from the full image, while ROI represents the testing points sampled from the masked images.

Table 3

Producer's accuracies (omission errors) and user's accuracies (commission errors) for tree species identification using the Deep Neural Network classification.

Species	Producer's Accuracy						User's Accuracy					
	R1	R2	R3	R4	R5	R6	R1	R2	R3	R4	R5	R6
a.aur	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.75	1.00	1.00
a.con	1.00	1.00	1.00	0.83	1.00	1.00	0.63	0.86	0.86	0.71	0.75	1.00
a.leb	1.00	1.00	0.00	0.67	0.67	1.00	1.00	1.00	0.00	1.00	1.00	1.00
a.mol	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.91	0.91
b.var	0.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00
b.pap	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
c.equ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.83	0.83	1.00	1.00	1.00
c.sin	0.88	1.00	1.00	0.83	1.00	1.00	1.00	1.00	0.75	1.00	0.86	0.86
c.cam	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.80	0.80	0.80
c.par	1.00	0.50	1.00	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d.reg	1.00	1.00	1.00	1.00	1.00	1.00	0.75	0.75	0.60	1.00	1.00	1.00
f.his	0.75	1.00	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
f.mic	1.00	1.00	1.00	1.00	1.00	1.00	0.75	1.00	1.00	1.00	1.00	0.75
f.var	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.75	1.00	1.00
f.vir	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.50
l.mon	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
m.tan	1.00	1.00	1.00	0.88	0.88	0.71	1.00	1.00	0.73	0.70	1.00	1.00
m.che	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.75	1.00	0.75
m.pan	0.60	0.60	0.20	0.20	0.40	0.60	1.00	1.00	1.00	1.00	1.00	1.00

and B1 (400 nm). Both the species changes leaves during February and April, therefore, were easy to identify but at the same time absence of leaves can give false reflection from under canopy vegetation which can lead to confusion among other species and reduce prediction accuracy. Bands B115 (733 nm), B94 (670 nm), and B185 (945 nm) helped to further segregate another group of species including *Delonix regia*, *Casuarina equisetifolia*, *Celtis sinensis* and *Cinnamomum parthenoxylon*. These optimal spectral wavelengths or bands may vary in different seasons due to the inherent seasonality of the species and the emergence of under-canopy vegetation. Nevertheless, an iterative method of ensemble machine learning approaches can be adapted to identify a spectrally unique group of species to enhance prediction accuracy with limited spectral data or in the regions with a higher number of spectrally similar species. A multivariate regression tree (MRT) can be developed to group species with similar spectral behaviour and to determine indicator species for each spectral group. The MRT assembles similar species in association with environmental or spectral constraints (Abbas et al., 2021). Grouping species in their biological families may not provide optimal results as the biologically similar species might not be similar in leaves and/or spectral traits. However, the grouping may change based on the available spectral space.

Overall classification accuracy of each round, confusion matrices, temporal spectral profiles and vegetation indices indicated a strong influence of phenology or seasonality on the identification of species from the hyperspectral images. Models developed on multitemporal hyperspectral and multispectral data have proven their ability to enhance the accuracy of classifying forest types and plantation tree species composition in a tropical landscape (Fagan et al., 2015). Seasonality can influence negatively or positively in the recognition and prediction capability of classification. In this study, the bi-monthly data was explored to determine which season is the best to separate most of the species. On the other hand, data acquired during the start (Round 3 - April) or peak (Round 4, 5 - June, August) of the growing season showed the least separability among the species. This is probably since most of the species in Hong Kong's urban landscape are evergreen, although there are spectral differences among the species, especially when using the full spectral space, it is difficult to distinguish at the peak of the growing seasons. During the leave off-season (in April) of the *Albizia lebbek*, it is difficult to distinguish it from other species including *Celtis sinensis* and *Delonix regia*. Although during this time of the season these species are mixed due to similar canopy traits (absence or reduction in the proportion of green leaves), the similar spectral and seasonal property can help to distinguish these species among the rest of the species. However, due mixing of these three species overall accuracy (user's and

producer's accuracy) drops significantly during round three (R3 - April). Therefore, for the highest accuracy of these species' identification, the autumn season is the best, during which the maximum accuracy can be achieved around 100%. These three species can be distinguished by the bands B7 (405 nm), B204 (1000 nm), B91 (661 nm), B1 (400 nm), B115 (733 nm), B42 (516 nm), B185 (945 nm) from the hyperspectral data acquired in October (R6). *Aleurites moluccana* was the most distinguishable species during all the seasons, probably due to its leaf structure, big and evergreen leaves. However, it is important to note the number of sampling of this species was also the highest, which indicates that the user's and producer's accuracies of other species might be increased by increasing the number of samples. The detection accuracy of *Macaranga tanarius* var. *tomentosa* varied from 55% to 100% in different seasons, seasonal analysis of the detectability of the species indicated that it can be easily distinguished with hyperspectral data acquired in the winter season (R1 - December) and it can be categorized by using the bands B41 (513 nm), B107 (709 nm), B189 (957 nm), B26 (470 nm), B60 (569 nm), B7 (405 nm), B174 (911 nm), and B87 (649 nm). *Macaranga tanarius* var. *tomentosa* is a common species in the urban landscape of Hong Kong, it is an evergreen species but its leaves turn yellow during the leaf changing season where new leaves sprout and older leaves shed down.

Apart from the seasonality and phenological traits of a plant, the spectral characteristics of plants can vary along different growth stages and environmental conditions. In evergreen ecosystems, changes in vegetation phenology and productivity are driven by the emergence of new leaves and the shedding of older leaves, whereas seasonality in photosynthesis is explained by leaf development and demography (Wu et al., 2016). Hyperspectral analysis of the 20 tropical species experimental plots showed that a significant correlation between growth rates and spectral properties (Caughlin et al., 2016). Fluorescence and NDVI both are strongly associated with surface soil moisture which varies with the vegetation class (grass, shrub and forest) and phenological stages (Shen et al., 2021). Buddenbaum et al. (2005) achieved 75% accuracy to classify coniferous tree species and age classes from hyperspectral remote sensing in a forest of western Germany. Another study in Hong Kong's secondary forest documented a significant difference in fluorescence emissions (which is a proxy variable for plant biomass or carbon flux measurement) among different age groups of Hong Kong's recovering tropical forests (Irteza et al., 2021). They analyzed hyperspectral data obtained from the Hyperion satellite and the ground-based hyperspectral radiometer. They observed higher productivity up to the age of 61 years old and then a decline in productivity in the old-growth forests. This accords with fairly recent pronouncements by scientists that

older forests are not as efficient as younger rapidly growing forests as carbon sinks. Integration of imaging spectroscopy and light detection and ranging (LiDAR) have enormous potential of combined assessment of tree foliage traits along with structural attributes (Asner et al., 2015; Dalponte et al., 2012; Shi et al., 2021; Trier et al., 2018). The current study did not incorporate the growth stages of the 75 trees in the urban landscape of Hong Kong due to limited accessibility. This research gap can be fulfilled in future by targeting a few most common species but with varying degrees of the growth stage.

This study examined the utility of hyperspectral imagery for the precise and accurate classification of urban tree species. Results obtained from the hyperspectral mapping provided higher accuracy (96%) when compared with a similar study using multispectral aerial photos in the urban landscape of Hong Kong (unpublished). Hyperspectral imagery provided better accuracy for tree species identification when compared with results obtained from multispectral imagery as numerous studies has documented an increase in mapping accuracies of tree species with a finer spectral resolution (Arasumani et al., 2021; Laurin et al., 2016; Lee et al., 2016; Miyoshi et al., 2020; Modzelewska et al., 2021; Wan et al., 2021). Ferreira et al. (2016) noted a 14–17% increase in the accuracies when SWIR (shortwave infrared) bands were combined VNIR (visible/near-infrared) bands to discriminate and map tree species in tropical seasonal semi-deciduous Brazilian Atlantic Forest Biome. Overall, the accuracy achieved with hyperspectral data, including narrowband indices, was 15 % higher than the maximum accuracy obtained with multispectral VNIR and SWIR imagery (Ferreira et al., 2016). Cho and Lee (2014) also observed an additive advantage of SWIR for tree species classification. A study in Finnish boreal forest revealed the utility of hyperspectral imagery to retrieve species-specific canopy traits. Likewise, a comparative assessment of hyperspectral (Hyperion and CHRIS-proba) and multispectral (ALI and Landsat 8) satellite images revealed 40–50% better results for classifying stone pine forest in Beirut (Awad, 2018). Future studies of urban tree species identification can be benefited from the flexibility by adding hyperspectral SWIR imagery.

4. Conclusion

In this study, the hyperspectral images of 75 trees belonging to 19 species and six families were acquired using terrestrial hyperspectral remote sensing. The images were acquired every two months from November 2018 to October 2019, with 450 images captured throughout the six rounds of image acquisition. The hyperspectral phenology patterns of the species were developed using the 60 conventional multispectral broad-band, hyperspectral narrow-band, and advanced hyperspectral narrow-band indices, to indicate the phenological trait of particular species to characterize phenology and tree health conditions. A Deep Neural Network framework was developed to estimate and classify species from the hyperspectral images in addition to plants' phenological traits. The results showed that the Deep Neural Network approach achieved high accuracy, ranging from 85% to 96%. The spectral reflectance curves of representative tree species in healthy and unhealthy conditions were extracted and analyzed including *Acacia confusa*, *Delonix regia*, *Macaranga tanarius*, *Casuarina equisetifolia*, *Celtis sinensis*, *Ficus hispida*, and *Machilus chekiangensis*. The species identification results are indicative of different seasonality characteristics for evergreen (e.g., *Acacia auriculiformis*, *Acacia confusa*, *Casuarina equisetifolia*, *Mallotus paniculatus*) and deciduous (e.g., *Albizia lebeck*) species in Hong Kong. The results of this study can not only be recommended to develop a spectral library for territory-wide mapping and monitoring for tree species and health conditions in Hong Kong, and also be a representative case study for a subtropical zone with a high diversity of tree species in an intense urban environment or along with hilly topography, which contribute the hyperspectral imaging and database development in the world.

In current image acquisition procedures, the experiments reveal that

integration time and light intensity are the key parameters for high quality data acquisition. For the future work, it is important that the sample size should be increased by 10–100 folds in order to enhance the producer's and user's accuracies. This could also be achieved by an integrated measurement of UAV-based hyperspectral and LiDAR data for identifying tree species and their structural characteristics (Alonzo et al., 2014; Ballanti et al., 2016; Degerickx et al., 2018; Hartling et al., 2019; Liu et al., 2017). Furthermore, random extraction of a number of pixels (50, 000 in this case) from each of the 75 samples might lead to autocorrelation issue, while applying conventional classification algorithms for species identification. A well-designed stratified sampling method should then be further developed. In addition, by increasing the number of samples with more field spectra data could help on improving the performance of the Deep Neural Network classification model.

Declaration of Competing Interest

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

Acknowledgement

The Highways Department supported this project under the project "Feasibility Study on Setting Up Spectral Library for Common Tree Species in HK". M.S. Wong thanks the support from the Research Institute for Sustainable Urban Development, the Hong Kong Polytechnic University with a project id: 1-BBWD. S. Abbas would also like to thank the support from PolyU (UGC) funding grant (1-ZVUU).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.isprsjprs.2021.05.003>.

References

- Aasen, H., Burkart, A., Bolten, A., Bareth, G., 2015. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: from camera calibration to quality assurance. *ISPRS J. Photogramm. Remote Sens.* 108, 245–259.
- Abbas, S., Nichol, J.E., Zhang, J., Fischer, G.A., Wong, M.S., Irteza, S.M., 2021. Spatial and environmental constraints on natural forest regeneration in the degraded landscape of Hong Kong. *SciEn* 752, 141760.
- Alonzo, M., Bookhagen, B., Roberts, D.A., 2014. Urban tree species mapping using hyperspectral and lidar data fusion. *RSEnv* 148, 70–83.
- Arasumani, M., Singh, A., Bunyan, M., Robin, V.V., 2021. Testing the efficacy of hyperspectral (AVIRIS-NG), multispectral (Sentinel-2) and radar (Sentinel-1) remote sensing images to detect native and invasive nonnative trees. *bioRxiv*.
- Asner, G.P., Martin, R.E., 2015. Spectroscopic remote sensing of non-structural carbohydrates in forest canopies. *Remote Sensing* 7, 3526–3547.
- Asner, G.P., Martin, R.E., 2016. Spectranomics: emerging science and conservation opportunities at the interface of biodiversity and remote sensing. *Global Ecol. Conserv.* 8, 212–219.
- Asner, G.P., Martin, R.E., Anderson, C.B., Knapp, D.E., 2015. Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *RSEnv* 158, 15–27.
- Asner, G.P., Martin, R.E., Carranza-Jimenez, L., Sinca, F., Tupayachi, R., Anderson, C.B., Martinez, P., 2014. Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. *New Phytol.* 204, 127–139.
- Awad, M.M., 2018. Forest mapping: a comparison between hyperspectral and multispectral images and technologies. *J. For. Res.* 29, 1395–1405.
- Baldrige, A.M., Hook, S., Grove, C., Rivera, G., 2009. The ASTER spectral library version 2.0. *RSEnv* 113, 711–715.
- Ballanti, L., Blesius, L., Hines, E., Kruse, B., 2016. Tree species classification using hyperspectral imagery: a comparison of two classifiers. *Remote Sensing* 8, 445.
- Behmann, J., Acebron, K., Emin, D., Bennertz, S., Matsubara, S., Thomas, S., Bohnenkamp, D., Kuska, M.T., Jussila, J., Salo, H., 2018. Specim IQ: evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection. *Sensors* 18, 441.
- Ben-Dor, E., Patkin, K., Banin, A., Karnieli, A., 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data—a case study over clayey soils in Israel. *IJRS* 23, 1043–1062.
- Bojinski, S., Schaepman, M., Schläpfer, D., Itten, K., 2003. SPECCHIO: a spectrum database for remote sensing applications. *Comput Geosci* 29, 27–38.

- Bolund, P., Hunhammar, S., 1999. Ecosystem services in urban areas. *Ecol. Econ.* 29, 293–301.
- Buddenbaum, H., Schlerf, M., Hill, J., 2005. Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. *IJRS* 26, 5453–5465.
- Caughlin, T.T., Graves, S.J., Asner, G.P., Van Breugel, M., Hall, J.S., Martin, R.E., Ashton, M.S., Bohlman, S.A., 2016. A hyperspectral image can predict tropical tree growth rates in single-species stands. *Ecol. Appl.* 26, 2369–2375.
- Cho, H., Lee, K.-S., 2014. Comparison between hyperspectral and multispectral images for the classification of coniferous species. *Korean J. Remote Sensing* 30, 25–36.
- Cho, M.A., Skidmore, A.K., 2006. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. *RSEnv* 101, 181–193.
- Clark, M.L., 2020. Comparison of multi-seasonal Landsat 8, Sentinel-2 and hyperspectral images for mapping forest alliances in Northern California. *ISPRS J. Photogramm. Remote Sens.* 159, 26–40.
- Cochrane, M., 2000. Using vegetation reflectance variability for species level classification of hyperspectral data. *IJRS* 21, 2075–2087.
- Cotrozzi, L., Couture, J.J., Cavender-Bares, J., Kingdon, C.C., Fallon, B., Pilz, G., Pellegrini, E., Nali, C., Townsend, P.A., 2017. Using foliar spectral properties to assess the effects of drought on plant water potential. *Tree Physiol.* 37, 1582–1591.
- Dadon, A., Mandelmilch, M., Ben-Dor, E., Sheffer, E., 2019. Sequential PCA-based classification of mediterranean forest plants using airborne hyperspectral remote sensing. *Remote Sensing* 11, 2800.
- Dalponde, M., Bruzzone, L., Gianelle, D., 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *RSEnv* 123, 258–270.
- De' Ath, G., 2002. Multivariate regression trees: a new technique for modeling species–environment relationships. *Ecology* 83, 1105–1117.
- Degerickx, J., Roberts, D.A., McFadden, J.P., Hermy, M., Somers, B., 2018. Urban tree health assessment using airborne hyperspectral and LiDAR imagery. *IJAEO* 73, 26–38.
- Delegido, J., Van Wittenbergh, S., Verrelst, J., Ortiz, V., Veroustraete, F., Valcke, R., Samson, R., Rivera, J.P., Tenjo, C., Moreno, J., 2014. Chlorophyll content mapping of urban vegetation in the city of Valencia based on the hyperspectral NAOAC index. *Ecol. Indicators* 40, 34–42.
- Delegido, J.a., 2014. Chlorophyll content mapping of urban vegetation in the city of Valencia based on the hyperspectral NAOAC index. *Ecol. Indicators* 40, 34–42.
- EcoSIS, N., 2014. Ecological Spectral Information System, 2021.04.15 ed.
- Escobedo, F.J., Kroeger, T., Wagner, J.E., 2011. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environ. Pollut.* 159, 2078–2087.
- Escobedo, F.J., Nowak, D.J., 2009. Spatial heterogeneity and air pollution removal by an urban forest. *Landscape Urban Plann.* 90, 102–110.
- Fagan, M.E., DeFries, R.S., Slesnie, S.E., Arroyo-Mora, J.P., Soto, C., Singh, A., Townsend, P.A., Chazdon, R.L., 2015. Mapping species composition of forests and tree plantations in northeastern Costa Rica with an integration of hyperspectral and multitemporal landsat imagery. *Remote Sensing* 7, 5660–5696.
- Ferreira, M.P., Wagner, F.H., Araújo, L.E.O.C., Shimabukuro, Y.E., de Souza Filho, C.R., 2019. Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. *ISPRS J. Photogramm. Remote Sens.* 149, 119–131.
- Ferreira, M.P., Zortea, M., Zanotta, D.C., Shimabukuro, Y.E.a., 2016. Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *RSEnv* 179, 66–78.
- Gamon, J., Penuelas, J., Field, C., 1992. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *RSEnv* 41, 35–44.
- Gamon, J., Serrano, L., Surfus, J., 1997. The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia* 112, 492–501.
- Ghosh, G., Kumar, S., Saha, S., 2012. Hyperspectral satellite data in mapping salt-affected soils using linear spectral unmixing analysis. *J. Indian Soc. Remote Sens.* 40, 129–136.
- Gómez-Baggethun, E., Gren, Å., Barton, D.N., Langemeyer, J., McPhearson, T., O'farrell, P., Andersson, E., Hamstead, Z., Kremer, P., 2013. Urban ecosystem services, Urbanization, biodiversity and ecosystem services: Challenges and opportunities. Springer, Dordrecht, pp. 175–251.
- Gong, P., Pu, R.L., Yu, B., 1997. Conifer species recognition: an exploratory analysis of in situ hyperspectral data. *RSEnv* 62, 189–200.
- Goswami, S., 2011. Monitoring ecosystem dynamics in an Arctic tundra ecosystem using hyperspectral reflectance and a robotic tram system.
- Goswami, S., Matharasi, K., 2015. Development of a web-based vegetation spectral library (VSL) for remote sensing research and applications. *PeerJ PrePrints*.
- Halme, E., Pellikka, P., Mttus, M., 2019. Utility of hyperspectral compared to multispectral remote sensing data in estimating forest biomass and structure variables in Finnish boreal forest. *IJAEO* 83, 101942.
- Hartling, S., Sagan, V., Sidike, P., Maimaitijiang, M., Carron, J., 2019. Urban tree species classification using a WorldView-2/3 and LiDAR data fusion approach and deep learning. *Sensors* 19, 1284.
- Im, J., Jensen, J.R., 2005. A change detection model based on neighborhood correlation image analysis and decision tree classification. *RSEnv* 99, 326–340.
- Irteza, S.M., Nichol, J.E., Shi, W., Abbas, S., 2021. NDVI and fluorescence indicators of seasonal and structural changes in a tropical forest succession. *Earth Syst. Environ.* 5, 127–133.
- Jensen, R.R., Hardin, P.J., Hardin, A.J., 2012. Classification of urban tree species using hyperspectral imagery. *Geol* 27, 443–458.
- Katkovsky, L.V., Martinov, A.O., Siliuk, V.A., Ivanov, D.A., Kokhanovsky, A.A., 2018. Fast atmospheric correction method for hyperspectral data. *Remote Sensing* 10, 1698.
- Kothari, S., Cavender-Bares, J., Bitan, K., Verhoeven, A., Wang, R., Montgomery, R., Gamon, J., 2018. Community-wide consequences of variation in photoprotective physiology among prairie plants. *Psyn* 56, 455–467.
- Krause, G., Weis, E., 1991. Chlorophyll fluorescence and photosynthesis: the basics. *Annu. Rev. Plant Biol.* 42, 313–349.
- Laurin, G.V., Chen, Q., Lindsell, J.A., Coomes, D.A., Del Frate, F., Guerriero, L., Pirotti, F., Valentini, R., 2014. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* 89, 49–58.
- Laurin, G.V., Puletti, N., Hawthorne, W., Liesenberg, V., Corona, P., Papale, D., Chen, Q., Valentini, R., 2016. Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. *RSEnv* 176, 163–176.
- Leckie, D.G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F.A., Cloney, E., Paradine, D., 2005. Issues in species classification of trees in old growth conifer stands. *CaJRS* 31, 175–190.
- Lee, J., Cai, X., Lellmann, J., Dalponde, M., Malhi, Y., Butt, N., Morecroft, M., Schonlieb, C.B., Coomes, D.A., 2016. Individual Tree Species Classification from Airborne Multisensor Imagery Using Robust PCA. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9, 2554–2567.
- Lin, C., Chen, S.-Y., Chen, C.-C., Tai, C.-H., 2018. Detecting newly grown tree leaves from unmanned-aerial-vehicle images using hyperspectral target detection techniques. *ISPRS J. Photogramm. Remote Sens.* 142, 174–189.
- Liu, X., Coops, N.C., Aven, N.W., Pang, Y., 2017. Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *RSEnv* 200, 170–182.
- Lyytimäki, J., Petersen, L.K., Normander, B., Bezák, P., 2008. Nature as a nuisance? Ecosystem services and disservices to urban lifestyle. *Environ. Sci.* 5, 161–172.
- Maselli, F., 2004. Monitoring forest conditions in a protected Mediterranean coastal area by the analysis of multiyear NDVI data. *RSEnv* 89, 423–433.
- Migliavacca, M., Perez-Priego, O., Rossini, M., El-Madany, T.S., Moreno, G., van der Tol, C., Rascher, U., Berninger, A., Bessenbacher, V., Burkart, A., 2017. Plant functional traits and canopy structure control the relationship between photosynthetic CO₂ uptake and far-red sun-induced fluorescence in a Mediterranean grassland under different nutrient availability. *New Phytol.* 214, 1078–1091.
- Miyoshi, G.T., Imai, N.N., Tommaselli, A.M.G., de Moraes, M.V.A., Honkavaara, E., 2020. Evaluation of hyperspectral multitemporal information to improve tree species identification in the highly diverse atlantic forest. *Remote Sensing* 12, 1–21.
- Modzelewska, A., Kamiska, A., Fassnacht, F.E., Stereczak, K., 2021. Multitemporal hyperspectral tree species classification in the Bia l owie . z a Forest World Heritage site. *Forestry: An Int. J. Forest Res.* 1–13.
- Nijland, W., De Jong, R., De Jong, S.M., Wulder, M.A., Bater, C.W., Coops, N.C., 2014. Monitoring plant condition and phenology using infrared sensitive consumer grade digital cameras. *Agric. For. Meteorol.* 184, 98–106.
- Nowak, D.J., Crane, D.E., Stevens, J.C., Hoehn, R.E., Walton, J.T., Bond, J., 2008. A ground-based method of assessing urban forest structure and ecosystem services. *Aboriculture & Urban Forestry.* 34 (6): 347–358. 34.
- Osco, L.P., Arruda, M.d.S.d., Marcato Junior, J., da Silva, N.B., Ramos, A.P.M., Moryia, É. A.S., Imai, N.N., Pereira, D.R., Creste, J.E., Matsubara, E.T., Li, J., Gonçalves, W.N., 2020. A convolutional neural network approach for counting and geolocating citrus-trees in UAV multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 160, 97–106.
- Paz-Kagan, T., Caras, T., Herrmann, I., Shachak, M., Karnieli, A., 2017. Multiscale mapping of species diversity under changed land use using imaging spectroscopy. *Ecol. Appl.* 27, 1466–1484.
- Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* 20, 503–510.
- Rossini, M., Nedbal, L., Guanter, L., Ac, A., Alonso, L., Burkart, A., Cogliati, S., Colombo, R., Damm, A., Drusch, M., 2015. Red and far red Sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geol* 42, 1632–1639.
- Schaepman, M.E., 1998. Calibration of a field spectroradiometer, Department of Geography, University of Zurich.
- Schiefer, F., Kattenborn, T., Frick, A., Frey, J., Schall, P., Koch, B., Schmidlein, S., 2020. Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks. *ISPRS J. Photogramm. Remote Sens.* 170, 205–215.
- Shen, Q., Liu, L., Zhao, W., Yang, J., Han, X., Tian, F., Wu, J., 2021. Relationship of surface soil moisture with solar-induced chlorophyll fluorescence and normalized difference vegetation index in different phenological stages: a case study of Northeast China. *Environ. Res. Lett.* 16.
- Shi, Y., Wang, T., Skidmore, A.K., Holzwarth, S., Heiden, U., Heurich, M., 2021. Mapping individual silver fir trees using hyperspectral and LiDAR data in a Central European mixed forest. *IJAEO* 98, 102311.
- Thenkabail, P.S., Anece, I., 2019. Global Hyperspectral Imaging Spectral-library of Agricultural Crops (GHISA) released through USGS and NASA's Land Processes Distributed Active Archive Center (LP DAAC).
- Thenkabail, P.S., Lyon, J.G., Huete, A., 2018. Biophysical and biochemical characterization and plant species studies. CRC Press.
- Townsend, P.A., Foster, J.R., Chastain, R.A., Currie, W.S., 2003. Application of imaging spectroscopy to mapping canopy nitrogen in the forest of the central Appalachian mountains using hyperion and AVIRIS. *ITGRS* 41, 1347–1354.

- Tratalos, J., Fuller, R.A., Warren, P.H., Davies, R.G., Gaston, K.J., 2007. Urban form, biodiversity potential and ecosystem services. *Landscape Urban Plann.* 83, 308–317.
- Trier, i.D., Salberg, A.B., Kermit, M., Rudjord, y., Gobakken, T., Nsset, E., Aarsten, D., 2018. Tree species classification in Norway from airborne hyperspectral and airborne laser scanning data. *European Journal of Remote Sensing* 51, 336–351.
- Tucker, C., Sellers, P., 1986. Satellite remote sensing of primary production. *IJRS* 7, 1395–1416.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *RSEnv* 8, 127–150.
- Wan, H., Tang, Y., Jing, L., Li, H., Qiu, F., Wu, W., 2021. Tree species classification of forest stands using multisource remote sensing data. *Remote Sensing* 13, 1–24.
- Wu, J., Albert, L.P., Lopes, A.P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K.T., Guan, K., Stark, S.C., Christoffersen, B., Prohaska, N., 2016. Leaf development and demography explain photosynthetic seasonality in Amazon evergreen forests. *Sci* 351, 972–976.
- Wu, S.B., Wang, J., Yan, Z.B., Song, G.Q., Chen, Y., Ma, Q., Deng, M.F., Wu, Y.T., Zhao, Y. Y., Guo, Z.F., Yuan, Z.Q., Dai, G.H., Xu, X.T., Yang, X., Su, Y.J., Liu, L.L., Wu, J., 2021. Monitoring tree-crown scale autumn leaf phenology in a temperate forest with an integration of PlanetScope and drone remote sensing observations. *ISPRS J. Photogramm. Remote Sens.* 171, 36–48.
- Xue, J.R., Su, B.F., 2017. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors* 2017.
- Yan, W.Y., van Ewijk, K., Treitz, P., Shaker, A., 2020. Effects of radiometric correction on cover type and spatial resolution for modeling plot level forest attributes using multispectral airborne LiDAR data. *ISPRS J. Photogramm. Remote Sens.* 169, 152–165.
- Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H., Noland, T.L., Sampson, P.H., 2000. Chlorophyll fluorescence effects on vegetation apparent reflectance: II. Laboratory and airborne canopy-level measurements with hyperspectral data. *RSEnv* 74, 596–608.
- Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., Sampson, P.H., 2001. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *ITGRS* 39, 1491–1507.
- Zhang, N., Yang, G., Pan, Y., Yang, X., Chen, L., Zhao, C., 2020. A review of advanced technologies and development for hyperspectral-based plant disease detection in the past three decades. *Remote Sensing* 12, 3188.
- Zhang, X.F., Liao, C.H., Li, J.H., Sun, Q., 2013. Fractional vegetation cover estimation in arid and semi-arid environments using HJ-1 satellite hyperspectral data. *IJAEO* 21, 506–512.