ScienceDirect

Ecological Informatics
Volume 61, March 2021, 101181

# Improving the classification of invasive plant species by using continuous wavelet analysis and feature reduction techniques

Abdulla A. Omeer [a] [b], Ratnadeep R. Deshmukh [a]

Show more

Share Cite

https://doi.org/10.1016/j.ecoinf.2020.101181
Get rights and content

## Abstract

The impacts of invasive plant species on the environment and economy make effectively detecting and mapping them crucial. Using leaf spectral reflectance and the advantages of continuous wavelet analysis (CWA), we aimed to utilize the CWA and features reduction techniques (principal component analysis (PCA), regularized random forest (RRF), and guided regularized random forest (GRRF)) and two famous classifiers (random forest (RF) and support vector machine (SVM)) to discriminate between five invasive plant species. The sample used in the study consisted of 562 leaves representing five species (*Senna uniflora*, *Hyptis suaveolens*, *Parthenium hysterophorus*, *Prosopis juliflora*, and *Xanthium strumarium*), which were collected from two sites. Both spectra (smoothed and original) were analyzed using CWA with different scales. 120 models of feature reduction methods (PCA, RRF, and GRRF) were established, combined with two classifiers (RF and SVM) and then compared. 90% of the smoothed CWA models (54 models) showed improvements in the overall accuracy values [1.18%, 19.38%] as compared to the smoothed spectra models alone. 94% of the non-smoothed CWA models (54 models) showed improvements in the overall accuracy values [0.18%, 19.38%] as compared to the non-smoothed spectra models alone. The highest overall accuracy was achieved at 98.87% with a model of CWA at scale 16 by using the GRRF and SVM; whereas, the models of smoothed and non-smoothed spectra without CWA had overall accuracies of 90.22% and 89.87%, respectively.

Moreover, the models of CWA coupled with GRRF or RRF had better performance rates than the models of CWA with PCA. We concluded that the classification accuracy is improved when CWA with appropriate scales are used, and the feature selection process with the GRRF or RRF methods is also recommended for improving the classification performance.

## Introduction

Invasive plant species pose a threat to the ecosystem because they reduce biological diversity (Binggeli, 1996) and wildlife habitat and replace native species with new, non-native species (Kohli et al., 2004). The economic and environmental impact of such species has been reported by several researchers (Henderson et al., 2006; Mack et al., 2000). Nevertheless, some invasive plant species have beneficial uses; they serve as medicinal plants (Kothapallia et al., 2017; Talakal et al., 1995) and as a source of fuelwood energy (Oduor and Githiomi, 2013). However, the harmful effects on natural resources and the environment have prompted governments across the world to strive to control the spread of invasive plant species through various chemical, mechanical, and biological means (Simberloff, 2002). About 25% percent of Indian flora comprises of invasive plants that have spread across the Indian subcontinent (Singh, 2005). The five plants that have been chosen in this study are the most eminent, destructive, and problematic invasive plant species in India and globally (Kumari and Prasad, 2018; Sharma et al., 2017; Timsina et al., 2011; Wade et al., 2017).

Due to the high cost of detection using field survey or aerial photograph interpretation methods (Zhang et al., 2006), obtaining the location of invasive plants in the early stage when their spatial propagation is small reduces the cost and time of effective eradication (Rejmánek and Pitcairn, 2002). Remote sensing has been used to detect and map invasive plants with rapid, large-scale, and cost-effective monitoring methods (Lawrence et al., 2006). Multispectral satellites such as IKONOS (Gil et al., 2013), QuickBird (Laba et al., 2008), WorldView-2 (Lantz and Wang, 2013), and MODIS (Landmann et al., 2020) have been used to map invasive plants on regional scales, as these devices have high spatial and spectral resolutions.

Hyperspectral imagery is more effective than multispectral imagery because it has a larger number of contiguous spectral wavelengths that are related to the spectral feature's characteristics. Thus, it is found more efficient for the discrimination of plant species (Govender et al., 2008). Hyperspectral data that is derived from satellites, for instance, Hyperion (Tsai et al., 2007) and PROBA-1 (Lawrence et al., 2006), airborne devices such as Hymap (Andrew and Ustin, 2008; Hestir et al., 2008), CASI, AVIRIS (Asner et al., 2008), and AISA (Narumalani et al., 2009), and field spectroradiometers (Tesfamichael et al., 2018) have been used for the discrimination of invasive plants in the forest, wildlife habitat, and wetlands. Recent advancements in unmanned aerial vehicles (UAVs) have also been used and have achieved potential success (Alvarez-Taboada et al., 2017; Hill et al., 2017; Wu et al., 2019).

A major challenge posed by hyperspectral data is its high dimensionality (Hughes, 1968), which reduces the performance of the supervised learning process. This is because the number of spectral features is larger than the sample size. Hence, dimensionality reduction is a necessary preprocessing step to transform the large dimension space into a smaller feature space (Melgani and Bruzzone, 2004). Dimensionality reduction, through feature extraction such as principal component analysis (PCA) and feature selection methods such as regularized random forest (RRF) and guided regularized random forest (GRRF), is used to select sensitive spectral features related to the information of plant species (Adam et al., 2017; Deng and Runger, 2012; Uddin et al., 2020).

Continuous wavelet analysis (CWA) can enhance the spectral details of hyperspectral data as well as reduce the noise (Guzmán et al., 2018). Therefore, CWA of remote sensing data has been utilized in many study cases, such as for the retrieval of plant traits (Cheng et al., 2014) and discrimination of mangrove species (Xu et al., 2019). Further, a few studies have used CWA for the enhancement of the classification accuracy of trees in tropical dry forests (Harrison et al., 2018) and herbaceous wetlands (Gross and Heumann, 2014). Nevertheless, the merits of CWA of hyperspectral data have been rarely explored for the classification of invasive plants.

This study aimed to explore the advantages of employing CWA to improve the classification performance of invasive plant species through hyperspectral reflectance at the leaf level. Moreover, the study sought to investigate the feature selection methods (RRF and GRRF) in relation with hyperspectral data. The results obtained at the leaf level may provide the basis for further studies using hyperspectral images captured by satellites and UAVs.

---

## Section snippets

### Materials and methods

Previous studies used preprocessing methods to improve the performance of the classification model since the raw spectral reflectance of plants is characterized by the similarities between the plant species (Prabhakar and Geetha, 2016). In addition to that, the dimensionality reduction process is necessary to remove redundant features and improve performance. In terms of the selection of features, two feature reduction methods were used in this study, namely, RRF and GRRF. These methods have…

### Analyzing the mean spectral reflectance and CWA power spectra of invasive plant leaves

Considering *Su* (*Senna uniflora*) as an example, the other species can be found in supplementary figures (Figs. S-1, S-2), the mean reflectance of spectra and CWA power of nine scales have been illustrated in Fig. 6.

The mean reflectance spectra of smoothed reflectance had 8 peaks and 7 troughs (Fig. 6 a-1) and those of non-smoothed reflectance had 9 peaks and 8 troughs (Fig. 6 b-1). The number of peaks and troughs for the mean wavelet power spectra of smoothed and non-smoothed reflectance was the …

## Continuous wavelet analysis for the classification of invasive plants

Continuous wavelet analysis plays a more important role in the efficiency of the classification process of the invasive plant species than the smoothed or non-smoothed spectra. Previous studies revealed the effect of the pre-processing methods such as the first derivative method and CWA on the prediction of the leaf trait (Blackburn and Ferwerda, 2008), classification of species, and identification of the spectral features on the comparison of spectral libraries (Harrison et al., 2018).

In this…

## Conclusion

This study investigated the potential of CWA coupled with feature reduction techniques in the classification of invasive plant species. Considering the results, the following conclusions can be articulated:

- No matter the feature reduction technique, CWA at suitable scales is a promising tool to improve the accuracy of the classification of invasive plant species.…
- GRRF and RRF are recommended to reduce the dimensionality of the hyperspectral data that leads to high classification accuracy.…

…

## Declaration of Competing Interest

None.…

## Acknowledgment

The authors would like to thank DST-FIST for the support provided under the sanction no SR/FST/ETI-340/2013 to the Department of Computer Science and Information Technology, DR.BAMU. Further, the authors would like to thank the anonymous reviewers for their valuable comments that contributed toward the improvement of this work.…

Recommended articles

## References (89)

E. Adam *et al.*
Spectral discrimination of papyrus vegetation (*Cyperus papyrus* L.) in swamp wetlands using field spectrometry
ISPRS J. Photogramm. Remote Sens. (2009)

A.M. Ali *et al.*
Retrieval of forest leaf functional traits from HySpex imagery using radiative transfer models and continuous wavelet analysis
ISPRS J. Photogramm. Remote Sens. (2016)

M.E. Andrew *et al.*
The role of environmental context in mapping invasive plants with hyperspectral image data
Remote Sens. Environ. (2008)

G.P. Asner *et al.*
Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR
Remote Sens. Environ. (2008)

G.A. Blackburn *et al.*
Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis
Remote Sens. Environ. (2008)

T. Cheng *et al.*
Deriving leaf mass per area (LMA) from foliar reflectance across a variety of plant species using continuous wavelet analysis
ISPRS J. Photogramm. Remote Sens. (2014)

H. Deng *et al.*
Gene selection with guided regularized random forest
Pattern Recogn. (2013)

D. Harrison *et al.*
Classification of tree species based on longwave hyperspectral data from leaves, a case study for a tropical dry forest
Int. J. Appl. Earth Obs. Geoinf. (2018)

E.L. Hestir *et al.*

Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem

Remote Sens. Environ. (2008)

C.H. Koger *et al.*

Wavelet analysis of hyperspectral reflectance data for detecting pitted morningglory (*Ipomoea lacunosa*) in soybean (*Glycine max*)

Remote Sens. Environ. (2003)

View more references

---

## Cited by (15)

On the importance of integrating convolution features for Indian medicinal plant species classification using hierarchical machine learning approach

2024, Ecological Informatics

Show abstract

BivalveNet: A hybrid deep neural network for common cockle (Cerastoderma edule) geographical traceability based on shell image analysis

2023, Ecological Informatics

Show abstract

Identification of suitable location to cultivate grape based on disease infestation using multi-criteria decision-making (MCDM) and remote sensing

2023, Ecological Informatics

Show abstract

Species classification from hyperspectral leaf information using machine learning approaches

2023, Ecological Informatics

Show abstract

A stochastic conceptual-data-driven approach for improved hydrological simulations

2022, Environmental Modelling and Software

*Citation Excerpt :*

...As IVS is an important feature of RF, this study utilizes Guided Regularized Random Forests (GRRF) since it has been shown to select input variables that result in improved prediction models when compared to those selected by RF (Deng and Runger, 2013). Interestingly, GRRF has yet to be used in the hydrology or water resources domains, although it has been used with success for different applications including: genetic population assignment (Sylvester et al., 2018), classification of invasive plant species (Omeer and Deshmukh, 2021), and remote sensing (Izquierdo-Verdiguier and Zurita-Milla, 2020). The GRRF input variable selection process selects a subset of relevant and non-redundant inputs by assigning penalties through regularization to information retrieved from previous nodes in the decision tree....

Show abstract

## Identifying optimal waveband positions for discriminating Parthenium hysterophorus using hyperspectral data

2021, Ecological Informatics

*Citation Excerpt :*

...This variation in spectral depth and waveband positions could be attributed to the differences in biophysical (i.e., leaf area, leaf structure and orientation, leaf surface characteristics, amount of biomass etc.), biochemical (i.e., pigments, nitrogen, cellulose, fibers, potassium, magnesium etc.) and structural properties (i.e., planophile, erecctrophile, cuticle layers and internal layer arrangements etc.) of different species (Ullah et al., 2012a; Ullah et al., 2012b). The variations in spectral reflectance (shape, size and wavelength position of the absorption feature) is thus vital and provide the foundation for discerning different plant species using reflectance spectra (Omeer and Deshmukh, 2021; Ullah et al., 2012b). The integration of Genetic Algorithm and Spectral Angle Mapper (GA-SAM) discerned Parthenium Hysterophorus and four co-occurring species with high classification accuracy (see Tables 2 & 3)....

Show abstract

View all citing articles on Scopus

View full text