CLASSIFICATION OF TREES IN HYPERSPECTRAL CANOPY DATA USING MACHINE LEARNING: COMPARATIVE ANALYSIS OF FOREST STRUCTURE COMPLEXITY

F. Galdames[1], P. González[1], F. Magni-Pérez[1], S. M. Funk[1], F. Lepín[4], R. Saavedra[2,3], H. J. Hernández[4]*

[1] Lemu Earth SpA, Avenida Manquehue Sur N°520, of. 205,
Las Condes, Santiago - (francisco, pablo, franco, stephan)@le.mu
[2] Arauco S.A, Concepción, Chile - rodrigo.saavedra.p@arauco.com
[3] E.T.S.I. Montes, Forestal y del Medio Natural, Universidad Politécnica de Madrid, Madrid, España.
[4] Geomatics and Landscape Ecology Lab, Forestry and Nature Conservation Faculty,
Universidad de Chile - felipe.lepin@ug.uchile.cl, jhernand@uchile.cl

KEY WORDS: Hyperspectral, Species Classification, Forest Structure, Machine Learning

ABSTRACT:

The classification of tree species by remote sensing is an important task with a broad range of applications, including forest management, environmental monitoring, and climate change studies. Hyperspectral imaging has proven to be a valuable tool for this classification. Additionally, deep learning techniques have obtained outstanding results in hyperspectral classification. In this study, we apply a neural network to the classification of aerial hyperspectral images of trees. The study was conducted at a research station in southern Chile with 32 tree species. Our database has 3080 tree canopies that have been manually segmented and classified. The goal of the work was to study the correlation between forest structure complexity and classification performance across three different forest conditions: native forest, plantation of native species, and plantation of exotic species. The results show that classification performance is higher when forest structure and composition are simpler. We used a ResNet neural network as classifier and compared its performance with support vector machine and random forest. The best performance was obtained using ResNet in exotic plantations, the forest condition with the simplest structure, achieving an F1-score of 85.23%.

1. INTRODUCTION

Remote sensing is a powerful tool for classifying individual trees in forests. However, the effectiveness of different methods can vary depending on the type of data available, the imagery's resolution, and the forest environment's complexity. Methods fall into one of the following groups according to the data type they use: (i) spectral data only (Onishi et al., 2022), (ii) structural data only, i.e. point clouds obtained via photogrammetry or LiDAR scanning and the metrics derived from them (Tinkham and Swayze, 2021), and (iii) combining both types of data (Zhong et al., 2022). Nevertheless, gathering structural data from LiDAR scanners remains highly expensive for local applications. Therefore, we decided to use only spectral data in this study.

Hyperspectral images have been widely used in various fields, including remote sensing, for object classification and analysis (Datta et al., 2022, Osco et al., 2021). On the other hand, deep learning methods have shown very good results in hyperspectral image classification (Paoletti et al., 2019). Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have proven to be effective in mapping forests and trees using hyperspectral images (Zhong et al., 2022). By harnessing the capabilities of artificial intelligence (AI) and deep learning, researchers have developed robust models for accurate vegetation mapping. These models leverage the rich spectral information captured by hyperspectral sensors and are able to extract discriminative features from the data. Several studies have demonstrated that hyperspectral data can be used to identify tree species (Fricker et al., 2019, Modzelewska et al., 2020, Zhao et al., 2021). However, one limitation of these methods is the requirement for a large number of labeled samples during the training stage (Jia et al., 2021, Wambugu et al., 2021). In order to overcome this limitation, we adopted a simple approach based on the classification of spectra using ResNet, and compared these results with two traditional methods, Support Vector Machine and Random Forest. We hypothesise that the classification performance will be higher the simpler the forest structure and the simpler its composition (i.e. fewer species per unit area). We assesssed the classification performance on three different forest conditions present at the Frutillar Research Station in southern Chile (Figure 1). First, the Winter Rainfall-Valdivian native forest (7 species): A complex forest of mixed tolerant species and multilayer canopy structure with a high horizontal species variation. Second, plantation of native species (12 species): Small stands of one dominant native species (i.e. low horizontal variation) with some woody understory vegetation. Third, exotic species plantations (13 species): Small stands of one exotic dominant species with sparse understory vegetation composed of herbs and ferns (i.e. low horizontal and vertical variation).

* Corresponding author

2. MATERIALS AND METHODS

2.1 Image acquisition

Hyperspectral images obtained at the Frutillar research station were acquired with NEO hyperspectral cameras using HySpex VNIR-1800 and SWIR-384 (HySpex, Norsk Elektro Optikk AS, Oslo, Norway) attached to a Cessna 162 aircraft operated by Chile's Heligraphics. The data were acquired in the summer of 2022 and pre-processed using the PARGE and ATCOR-4 programs for orthorectification and atmospheric correction of the scenes, respectively. The Minimum Noise Fraction (MNF) algorithm was applied to suppress redundant noise in the hyperspectral cube. Both cameras' images were combined to obtain an image with 466 channels ranging from 406 nm to 2515 nm. Since the bordering channels contain noise, only 455 channels from 466 nm to 2455.7 nm were used. The spatial resolution of the final image is 0.5 m.

2.2 Database

A ground truth (GT) database was built by manually segmenting and labeling the crowns of individual trees, or areas of vegetation with the same specie, on a screen. Experts identified the tree species using the images, and in field visits where GPS equipment was used to identify the corresponding trees. The resulting database contains 3080 polygons. The species were divided into 3 different forest conditions:

(1) Native forest (NF): *Caldcluvia paniculata* (CAPA), *Chusquea quila* (CHqu), *Eucryphia cordifolia* (EUCO), *Laureliopsis philippiana* (LAPH), *Luma apiculata* (LUAP), *Nothofagus dombeyi* (NODO), *Podocarpus nubigenus* (PONU).

(2) Plantation of native species (NP): *Amomyrtus luma* (AMLU), *Amomyrtus meli* (AMME), *Drimys winteri* (DRWI), *Fitzroya cupressoides* (FICU), *Gevuina avellana* (GEAV), *Laurelia sempervirens* (LASE), *Lomatia hirsuta* (LOHI), *Nothofagus alpina* (NOAL), *Nothofagus alessandri* (NOAS), *Nothofagus obliqua* (NOOB), *Persea lingue* (PELI), *Pilgerodendron uviferum* (PIUV).

(3) Plantation of exotic species (EP): *Acacia dealbata* (ACDE), *Acacia melanoxylon* (ACME), *Alnus glutinosa* (ALGL), *Calocedrus decurrens* (CADE), *Castanea sativa* (CASA), *Chamaecyparis lawsoniana* (CHLA), *Eucalyptus camaldulensis* (EUCA), *Eucalyptus globulus* (EUGB), *Pinus banksiana* (PIBA), *Pinus ponderosa* (PIPO), *Quercus robur* (QURO), *Sequoiadendron giganteum* (SEGI), *Sequoia sempervirens* (SESE).

Figure 1 shows the geographic location of the Frutillar research station, and the manually delineated polygons of the vegetation. To perform the experiments the database was divided into train, validation, and test sets, as explained in section "3. Experiments".

2.3 Method

We classified the pixels of the image according to the tree species. Then, we studied the classification performance depending on the forest conditions. To perform the classification, three different methods were tested: Random Forest (RF), Support Vector Machine (SVM) (Géron, 2022), and a one-dimensional ResNet convolutional neural network (He et al., 2016). The input to the classifiers was the spectrum in each pixel. In order to remove the brightness variations, the spectra were normalized by dividing the spectrum in each pixel by its mean. The hyperparameters of the SVM and RM were adjusted by using a halving grid search with a 5-fold cross-validation in the training and validation sets. The following parameters were considered in the adjustment of the hyperparameters of the RF: (i) balanced weight for the classes on training; (ii) 100, 500 or 1000 estimators; (iii) Gini impurity, or Shannon information gain for the split criterion; (iv) consider all the features (image bands), or a reduction to 50, 100, 150, or 200 features by PCA. The following parameters were considered in the adjustment of the SVM: (i) balanced weight for the classes in training; (ii) $C = \{2^i \mid i \in \{-5, -3, -1, 0, 1, 3, 5, 7, 9, 10, 13, 15\}\}$; (iii) lineal, polynomial, sigmoid, or radial basis function kernel; (iv) $\gamma = \{1/number\_of\_features \cdot variance\_of\_the\_features, 1/number\_of\_features\}$; (v) one-vs-one or one-vs-the-rest classification strategies; (vi) dimensionality reductions by PCA to 50, 100, or 150 features; (vii) standard normalization or a min-max scaler to scale the features. The parameters of the ResNet were chosen by assessing its performance using the training and validation sets. The ResNet was built with three residual blocks, each one with 3 convolutional layers followed by batch normalization layers before applying the activation function (ReLu). The number of filters used in the three residual blocks was 64, 128, and 128, respectively. In each residual block, the size of the filter kernels in the three convolutional layers were 24, 15, and 9, respectively. The output of the last residual block was connected with 3 fully connected layers of sizes 64, 32, and the number of classes.

3. EXPERIMENTS

The polygons of the species in the database were divided into 3 groups, according to the different forest conditions (NF, NP, EP). A 5-fold cross-validation scheme was used to evaluate the classification methods (Russell and Norvig, 2009) in the different forest conditions. Accordingly, five different sets for training, validation, and testing were built with the polygons of each forest condition group. The partitions were mostly performed randomly, with the only constraint that no polygon could appear in more than one test set. The following ratio was used for the sets: 80% for training, 10% for validation, and 10% for testing. Finally, all the spectra within each polygon were assigned to their respective training, validation, or test set, and used to evaluate the classification methods. Figure 2 shows the mean spectra of the different species.

All classification methods were trained in the train sets of each forest condition. Then, the methods were tested in the corresponding test sets. The precision, recall, and F1-score were used to measure the performance of the methods:

$$presicion = \frac{TP}{TP + FP}, \tag{1}$$

$$recall = \frac{TP}{TP + FN}, \tag{2}$$

$$F1 = 2 \cdot \frac{presicion \cdot recall}{presicion + recall}, \tag{3}$$

where TP, FP, FN are the true positives, false positives, and false negatives, respectively. The performance metrics were computed for every specie within each set. Subsequently, the average of each performance metric across the species was computed in the cross-validation sets. Finally, the mean and the standard deviation across the cross-validation sets were obtained for each forest condition.

To explore the importance of each band in the classification, an RF classifier was trained without a reduction of the number of features (image bands). The same alternatives considered in the previous experiments were used during the hyperparameter tuning process, except for using PCA for dimensionality reduction. This training was performed on all the cross-validation training sets of each forest condition. The importance of the bands

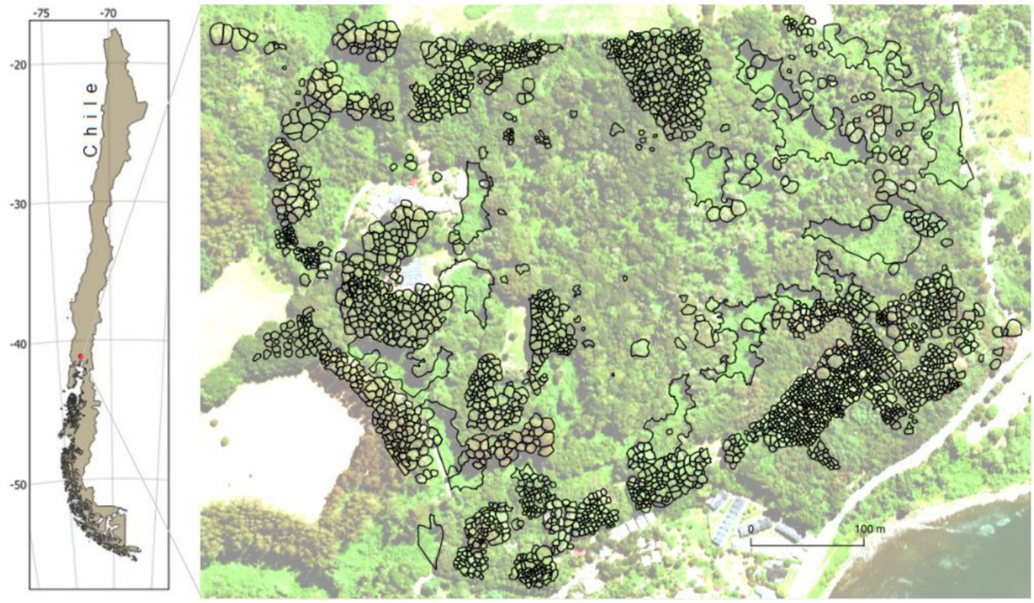

Figure 1. Study area (Frutillar Research Station). The image shows the manual segmentations of tree canopies or areas of vegetation with the same species

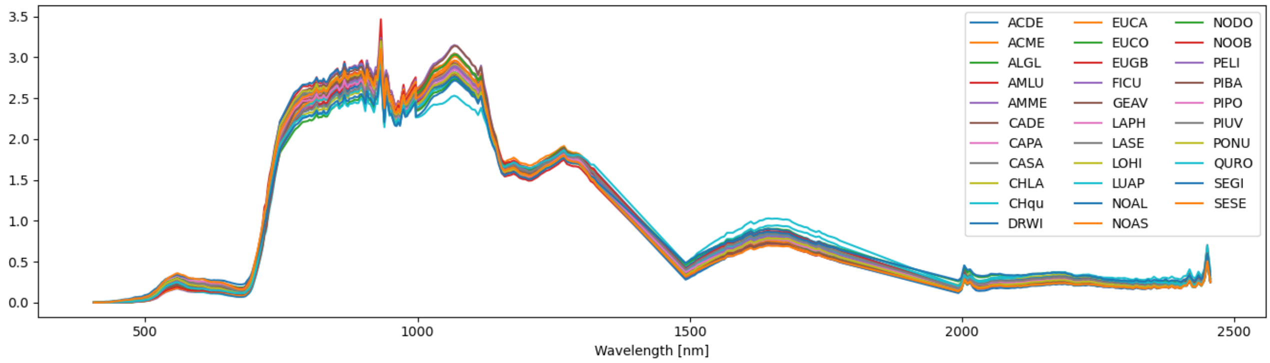

Figure 2. Spectra of the different species in the database.

was computed using the mean decrease in impurity (Breiman, 1984). Finally, the mean importance of each band and its standard deviation across the cross-validation sets were computed for each forest condition.

4. RESULTS AND DISCUSSION

Table 1 shows the results obtained in each forest condition using different classification methods. The results were computed for the test and training sets. The method that achieved the highest performance on the test set across all forest conditions was ResNet. The best F1-score in test (85.23%) was achieved using ResNet on EP. The precision of RF was higher than those of ResNet and SVM in NF and NP, nevertheless its recall was very low. The method with the second-best performance was SVM, reaching a F1-score of 82.35% on EP.

When comparing the performance metrics obtained in each forest conditions, the performance in exotic plantation is higher than that of native plantation, while the performance of native plantation surpasses that achieved in native forest. A possible explanation for this disparity lies in the complexity of the forest stand. Exotic plantations, composed of broadleaf and coniferous species, involve carefully planned and managed cultivation techniques, resulting in a more uniform forest structure. Native plantations, also tend to maintain a level of organization and management resulting in better performance of the classification methods used. Contrarily, native forests possess inherent characteristics of natural and diverse ecosystems resulting in a complex intermingling of various species within their structure. This intricate composition presents a significant challenge when attempting to assess them using either conventional or machine learning techniques accurately.

An analysis of variance (ANOVA) and a multiple comparison test were conducted to verify the statistical significance ($p < 0.05$) of the differences between the classification performance of the different methods applied to the three forest conditions. The Tukey's HSD significant difference method was used in the multiple comparison test. The F1-score was used to compute the statistics because it is an overall performance metric that takes into account both precision and recall. The ANOVA analysis showed that the difference in performance obtained in each forest condition and by using the different methods is statistically significant ($F = 52.86$, $p = 1.3e - 17$). The Tukey's HSD analysis showed that the performance difference between NP and EP obtained with the three methods is not statistically significant. Nevertheless, the differences are very close to being statistically significant with p-values of 0.99, 0.12, and 0.08 for

Method	Native forest [%]		Native plantation [%]		Exotic plantation [%]	
Presicion						
	train	test	train	test	train	test
RF	100.00 ± 0.00	65.91 ± 13.32	100.00 ± 0.00	82.48 ± 3.13	100.00 ± 0.00	84.94 ± 3.60
SVM	95.70 ± 7.74	54.38 ± 9.77	99.06 ± 1.87	74.87 ± 2.09	99.39 ± 1.22	84.97 ± 4.47
ResNet	99.99 ± 0.01	61.78 ± 6.70	99.87 ± 0.13	75.30 ± 2.97	99.99 ± 0.02	85.54 ± 3.92
Recall						
	train	test	train	test	train	test
RF	100.00 ± 0.00	33.11 ± 5.91	100.00 ± 0.00	59.69 ± 2.97	100.00 ± 0.00	61.93 ± 2.95
SVM	99.21 ± 1.49	50.21 ± 5.76	99.59 ± 0.81	74.12 ± 3.63	99.88 ± 0.24	82.08 ± 3.53
ResNet	99.98 ± 0.02	56.04 ± 2.13	99.86 ± 0.15	78.53 ± 4.43	99.98 ± 0.03	86.61 ± 2.49
F1-score						
	train	test	train	test	train	test
RF	100.00 ± 0.00	37.77 ± 6.72	100.00 ± 0.00	64.86 ± 2.54	100.00 ± 0.00	66.43 ± 2.66
SVM	97.17 ± 5.19	49.45 ± 6.95	99.31 ± 1.37	73.31 ± 2.98	99.62 ± 0.76	82.35 ± 4.01
ResNet	99.99 ± 0.01	56.70 ± 3.02	99.86 ± 0.14	75.89 ± 3.64	99.99 ± 0.02	85.23 ± 3.37

Table 1. Results of the classification methods in the three different forest conditions. The 5-fold cross-validation results were computed in the train and test sets of each forest condition. The results are presented in the format [mean ± std].

RF, SVM, and ResNet, respectively. Figure 3 shows a boxplot of the F1-score achieved in each forest condition by using the different methods.

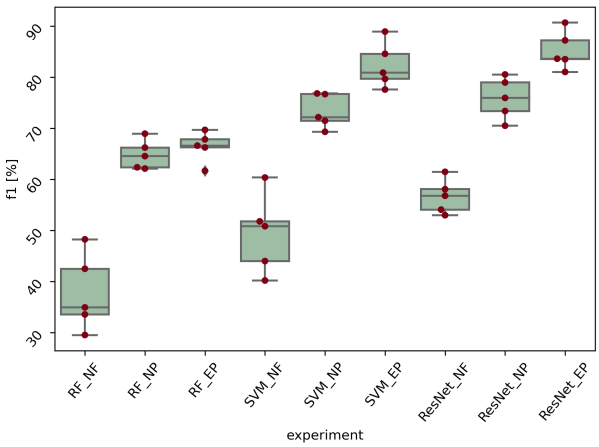

Figure 3. Boxplot of the F1-score obtained in the test sets for each forest condition, utilizing different methods.

Figure 4 shows the importance of each band of the hyperspectral image for classifying the different forest conditions. Certain wavelengths exhibited greater importance for classifying specific forest conditions. For example, the wavelengths near 1500 nm were more important in the classification of NP than in the classification of the NF and EP. The structural and physiological characteristics, including pigment absorption, play a significant role in differentiating between tree species based on their reflectance properties (Jones and Vaughan, 2010). Our results confirm that visible and near-infrared spectral bands are particularly important in this regard. These spectral bands contain the most prominent features that contribute to the discrimination of tree species, especially between native and exotic ones. In the visible range, various pigments, such as chlorophylls and carotenoids, absorb specific wavelengths of light, resulting in distinct colorations of leaves. These differences in pigment absorption contribute significantly to the visual discrimination between tree species. In the wavelength range of 455 to 475 nm, chlorophyll b exhibits relatively low absorbance, resulting in the reflection of light within that range. This phenomenon contributes to the greenish coloration of leaves by allowing some blue light to be reflected rather than absorbed. The importance of bands between 540 and 570 nm is related to changes in the efficiency of the photosynthetic light reaction caused by xanthophyll pigments of the leaves. The spectral bands between 700-730 nm reflex the influence of chlorophyll a, where it exhibits a low absorbance, causing the light in that range to be reflected rather than absorbed. This characteristic results in the low reflectance of light in the near-infrared region by tree leaves, making them appear dark or almost black in that specific range. Within the Infrared spectral range, there are also some important bands contributing to the discrimination among species. In particular, we can distinguish two specific ranges, from 950 to 990 nm and between 1060 and 1100 nm. The reflectance in these bands is primarily influenced by the internal structure of the leaf known as the leaf mesophyll. These spaces exist within the leaf mesophyll and are composed of air-filled cavities or gaps between the cells. The presence of these air spaces affects the reflectance of light in the near-infrared region.

5. CONCLUSIONS

Tree classification by means of hyperspectral images is a valuable tool that can be utilized for environmental studies and forest management, among others. In this paper, we have shown that the performance of this classification is affected by the complexity of the forest structure. The simpler the forest structure and its composition, the better the classification will perform. The study was conducted at a research station in southern Chile, where we classified 32 species belonging to the following forest conditions: native forest, plantation of native species, and plantation of exotic species. The hyperspectral image used has 455 channels from 466 nm to 2455.7 nm. We built a database for experimentation by manually segmenting and labeling 3080 crowns of individual trees or vegetation areas with the same specie. Three classification methods were assessed: two traditional methods, SVM and RF; and a ResNet neural network. The best classification performances were achieved in the simplest forest condition, plantation of exotic species. The best result was obtained by using ResNet in exotic plantations, reaching an F1-score of 85.23%. The ResNet classifier also achieved the best results on the other two forest conditions. RF was the worst-performing classifier in all the forest conditions.

We also studied the importance of each band in the classification. The importance of the bands was estimated by computing the mean decrease in impurity on the RF classifier. The results indicate that visible and near-infrared spectral bands are particularly important in all the forest conditions. Additionally, two

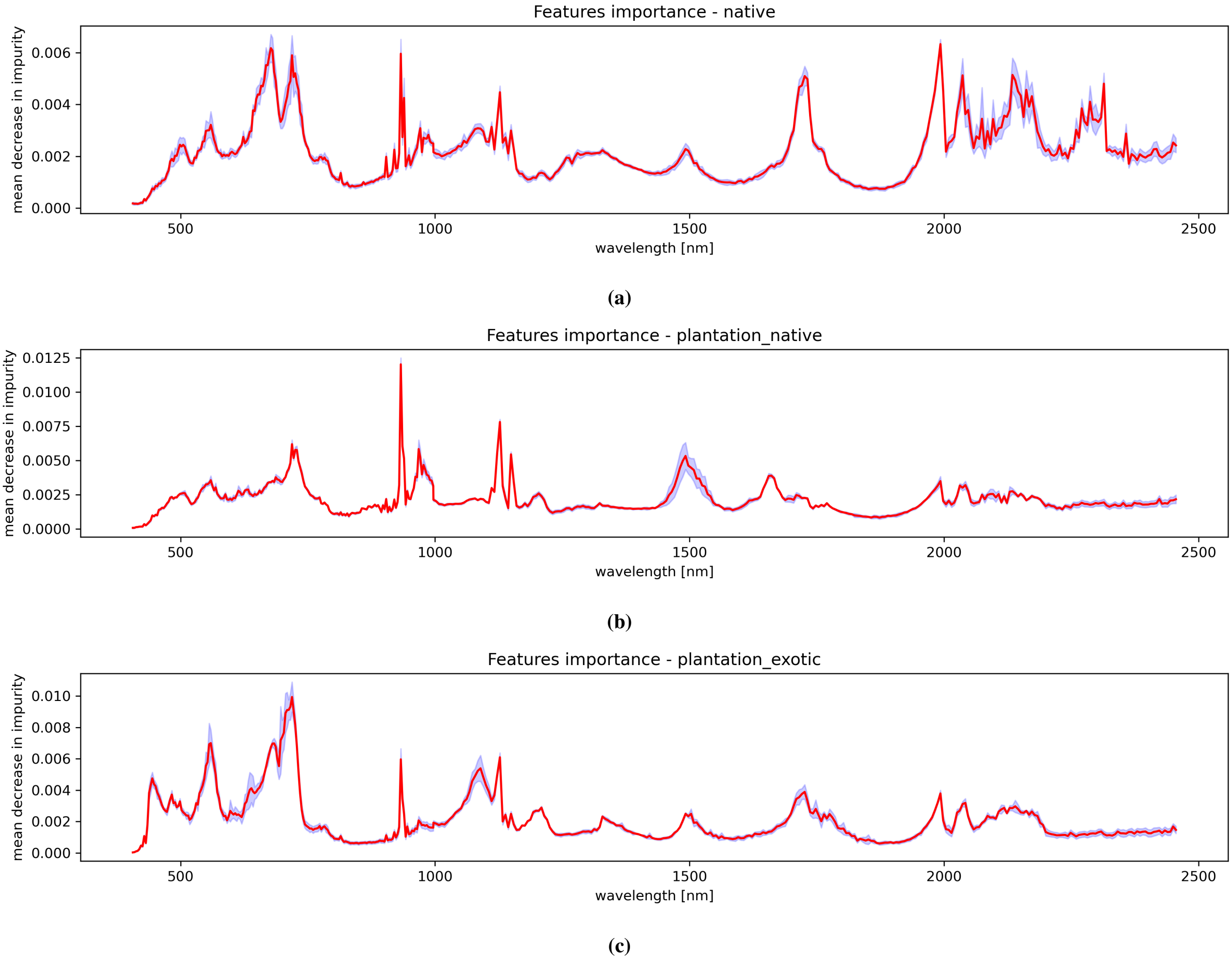

Figure 4. Importance of the different wavelengths in the classification of each forest condition. The red lines represent the mean across the cross-validation sets for each forest condition, while the blue strips indicate the standard deviations. (a) Native forest. (b) Native plantations. (c) Exotic plantations.

ranges related to the internal structure of the leaf (from 950 to 990 nm, and from 1060 to 1100 nm) were also found to be important. The wavelengths near 1500 nm were particularly important for classifying native plantations.

We are enhancing our database by adding more species and segmented tree canopies to improve the accuracy of the data. In addition, to improve the results, we are exploring more sophisticated deep learning techniques or hybrid models that could effectively handle the increased complexity and variability of forest conditions. One promising approach is to use instance classification methods based on neural networks. These methods have the capability to incorporate both spectral information and spatial texture from the image, enabling them to consider the structural characteristics of tree canopies. The use of point clouds can also further enhance the accuracy and effectiveness of our data analysis.

REFERENCES

Breiman, L., 1984. *Classification and Regression Trees*. Routledge, New York.

Datta, D., Mallick, P. K., Bhoi, A. K., Ijaz, M. F., Shafi, J., Choi, J., 2022. Hyperspectral Image Classification: Potentials, Challenges, and Future Directions. *Computational Intelligence and Neuroscience*, 2022, e3854635. doi.org/10.1155/2022/3854635. Publisher: Hindawi.

Fricker, G. A., Ventura, J. D., Wolf, J. A., North, M. P., Davis, F. W., Franklin, J., 2019. A Convolutional Neural Network Classifier Identifies Tree Species in Mixed-Conifer Forest from Hyperspectral Imagery. *Remote Sensing*, 11(19), 2326. doi.org/10.3390/rs11192326. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.

Géron, A., 2022. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Beijing Boston Farnham Sebastopol Tokyo.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Identity mappings in deep residual networks. *Computer Vision – ECCV 2016*, Amsterdam, The Netherlands, 630–645.

Jia, S., Jiang, S., Lin, Z., Li, N., Xu, M., Yu, S., 2021. A survey: deep learning for hyperspectral image classification with few labeled samples. 448(11), 179–204. doi.org/10.1016/j.neucom.2021.03.035.

Jones, H. G., Vaughan, R. A., 2010. *Remote Sensing of Vegetation: Principles, Techniques, and Applications*. Oxford University Press, Oxford, New York.

Modzelewska, A., Fassnacht, F. E., Stereńczak, K., 2020. Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation*, 84, 101960. doi.org/10.1016/j.jag.2019.101960.

Onishi, M., Watanabe, S., Nakashima, T., Ise, T., 2022. Practicality and Robustness of Tree Species Identification Using UAV RGB Image and Deep Learning in Temperate Forest in Japan. *Remote Sensing*, 14(7), 1710. doi.org/10.3390/rs14071710. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

Osco, L. P., Marcato Junior, J., Marques Ramos, A. P., de Castro Jorge, L. A., Fatholahi, S. N., de Andrade Silva, J., Matsubara, E. T., Pistori, H., Gonçalves, W. N., Li, J., 2021. A review on deep learning in UAV remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 102, 102456. doi.org/10.1016/j.jag.2021.102456.

Paoletti, M. E., Haut, J. M., Plaza, J., Plaza, A., 2019. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158, 279–317. doi.org/10.1016/j.isprsjprs.2019.09.006.

Russell, S. J., Norvig, P., 2009. *Artificial Intelligence: A Modern Approach*. Pearson.

Tinkham, W. T., Swayze, N. C., 2021. Influence of Agisoft Metashape Parameters on UAS Structure from Motion Individual Tree Detection from Canopy Height Models. *Forests*, 12(2), 250. doi.org/10.3390/f12020250. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

Wambugu, N., Chen, Y., Xiao, Z., Tan, K., Wei, M., Liu, X., Li, J., 2021. Hyperspectral image classification on insufficient-sample and feature learning using deep neural networks: A review. *International Journal of Applied Earth Observation and Geoinformation*, 105, 102603. doi.org/10.1016/j.jag.2021.102603.

Zhao, Q., Jiang, P., Wang, X., Zhang, L., Zhang, J., 2021. Classification of protection forest tree species based on UAV hyperspectral data. *Trans. Chin. Soc. Agric. Machinery*, 52(11), 190–199. doi.org/10.6041/j.issn.1000-1298.2021.11.020.

Zhong, H., Lin, W., Liu, H., Ma, N., Liu, K., Cao, R., Wang, T., Ren, Z., 2022. Identification of tree species based on the fusion of UAV hyperspectral image and LiDAR data in a coniferous and broad-leaved mixed forest in Northeast China. *Frontiers in Plant Science*, 13. doi.org/10.3389/fpls.2022.964769.