

Machine learning technique for close-range tree species classification

Abstract:

Tree species detection and identification are essential to recognize and predict natural disasters and protect tree species. Tree species detection was a difficult task in the past because of the time and human resources for the measurements. However, technological improvements contributed to different strategies and techniques for species identification and detection, like Terrestrial Laser Scanners (TLS), Photogrammetry devices, Lidars, and so on. Compared to the previous ages, the efficiency and accuracy are increased. However, utilizing some advanced techniques still need to pay higher amounts. Photogrammetry devices and techniques are user-friendly and cost-effective for species detection, but they will produce a large amount of data. To deal with a large amount of data, we need to utilize advanced computational techniques like Artificial Intelligence (AI) and Machine learning (ML).

Moreover, tree species classification is a challenging task, especially considering the classification based on two-dimensional digital images. We believe bark images have more consistency for the classification because of the usual consistency of the tree trunk pattern over different seasons.

In our study, we utilized data from a Sony (ILCE-7M2) camera with lens: E PZ 16-50 mm to analyse four species of trees in Slovakia. The Slovak dataset consisted of 1369 cropped images (cropping the entire bark) and 527 precisely cropped images (small part of the bark) of European beech (*Fagus sylvatica* L.), Sessile Oak (*Quercus petraea* (Matt.) Liebl.), Norway Spruce (*Picea abies* (L.) H. Karst.), and European Silver Fir (*Abies alba* Mill.). However, the ordinary photographs without cropping from the Slovak dataset were not used in the research due to additional markings and numbers on the barks.

Initially, Gray-Level Co-Occurrence Matrix was used as a feature extractor, and the features are used with different machine learning algorithms, including Support Vector Machine (SVM). However, it failed to provide better accuracy on SVM. Therefore we used Convolutional Neural Networks (CNN) for the classification task. The research followed by pre-processing the dataset and applying feature extraction. Then applying traditional machine learning algorithms and neural networks. The final results show that CNN achieved higher accuracy of 94% on Slovak extract cropped dataset and 81% with Random forest algorithm. As a future work, we used different feature selection methods to boost SVM accuracy. When we successfully identify the best approach to classify the tree species, we will work on the implementation to the workflows of terrestrial laser scanning, mobile laser scanning and terrestrial photogrammetry.

1. Introduction

The accurate identification of tree species within diverse natural forests plays a pivotal role in comprehending forest functionality and forecasting their reactions to both natural phenomena and human-induced impacts. The continuous progression of technology, exemplified by Terrestrial Laser Scanners (TLS), Photogrammetry devices, and LIDARs, has vastly broadened the spectrum of options available for species identification and detection (Liang, X. et al., 2018). While these methods bring forth enhanced efficiency and precision compared to earlier approaches, they also generate substantial volumes of data that necessitate automated or semi-automated processing capabilities. Recent advancements in cutting-edge computational techniques, notably Artificial Intelligence (AI) and Machine Learning (ML), exhibit significant potential in addressing this challenge.

AI algorithms have demonstrated their effectiveness in deciphering data collected through laser scanners and photogrammetry equipment. Nonetheless, current technology limitations restrict estimations to a forest stand or plot scale, falling short of the required resolution and computational capacity for evaluating individual trees. The classification of tree species presents a particularly formidable challenge, with two predominant analytical approaches in use. AI techniques also find application in forest management, encompassing tasks like the scrutiny and surveillance of forested regions, the detection of forest fires, and the identification of wildlife habitats (space-of-innovation, n.d.). To expedite and enhance monitoring efforts, cameras and Unmanned Aerial Vehicles (UAVs), such as drones, are deployed. Additionally, even compact computing devices now possess the capability to carry out fundamental image processing tasks, streamlining the complexity of these operations and reducing the dependence on human resources.

Open-access databases for tree species identification are currently scarce, despite their pivotal role in research endeavors. To illustrate, the AFF dataset stands as a private compilation comprising 1082 images spanning 11 tree species (Stefan, F. et al., 2010). In contrast, the publicly available Trunk12 dataset encompasses only 360 images representing 12 distinct species. Nevertheless, the sample collection process for the Trunk12 dataset exhibited flaws, rendering it less suitable for rigorous research purposes. Yet another dataset, Bark101 (Boudra, S. et al., 2022), derived from PlantCLEF 2017, presents its own set of challenges, primarily stemming from an uneven distribution of images across different classes. This dearth of accessible datasets constrains the breadth of research and contributions within this field.

The objective of this study is to investigate the capabilities of both conventional machine learning algorithms and advanced deep learning algorithms, including Convolutional Neural Networks and Multi-layer Perceptrons, in the context of tree species classification. We seek to assess their applicability to small datasets and their feasibility for real-time implementation.

2. Data Acquisition

In the course of this research, we generated two distinct datasets consisting of images of four tree species. These images were captured using Sony (ILCE-7M2) camera. Dataset encompasses a total of 1369 images, which were divided into two categories: cropped and exact cropped. The tree species featured in this dataset include European beech (*Fagus sylvatica* L.), sessile oak (*Quercus petraea* (Matt.) Liebl.), Norway spruce (*Picea abies* (L.) H. Karst.), and European silver fir (*Abies alba* Mill.). It's worth noting that the original images could not be used due to the presence of additional markings and numbering on the tree stems. As a result, Dataset comprises these two segmented categories, as demonstrated in Figure 1.



Figure 1. Example of an image in Dataset (normal image, cropped, and exact cropped)

For the machine learning phase, we partitioned each dataset into training and testing sets. Approximately 75 to 80% of the images for each species were allocated for training purposes, while the remaining images were reserved for testing and validation.

2.1. Methodology

In the field of forestry, machine learning and deep learning algorithms are extensively employed, often relying on pre-defined architectures and scenarios borrowed from various domains. Our research, however, is centered on the quest to identify the optimal scenario and architecture specifically tailored for tree species classification based on bark images. We have harnessed both traditional machine learning methods, including K-nearest neighbor, Random Forest, Decision tree, and Support vector machine, as well as neural network approaches, encompassing Multi-layer perceptron and Convolutional Neural Networks, to undertake this investigation.

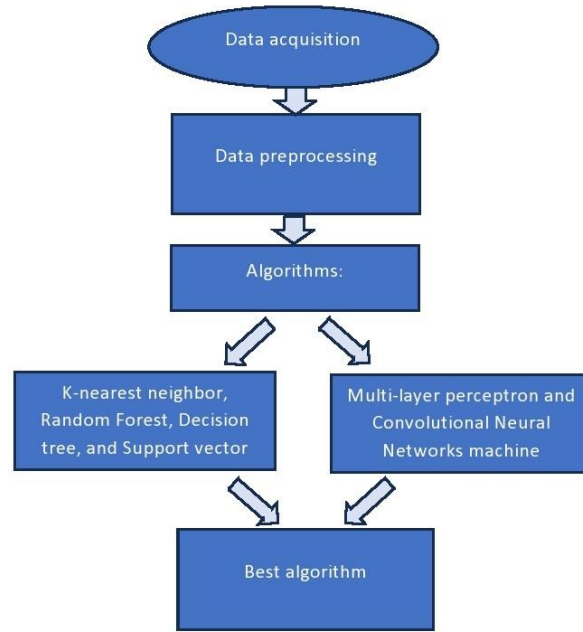


Figure 2. Proposed architecture

The figure displayed above illustrates the proposed architecture of this study. It entails the initial preprocessing of data, followed by the application of various machine learning algorithms and neural networks. Ultimately, our aim is to discern the most effective algorithm for the intended classification task.

2.1.1. Pre-processing: Segmentation

To begin, we initiated a meticulous examination of the datasets to eliminate images featuring numerical elements and other extraneous objects situated on the tree bark. We refrained from considering automated segmentation algorithms for this study due to their inherent difficulties in accurately identifying the bark's boundaries and accommodating peculiar markings and anomalies during the segmentation process. Additionally, the manner in which the images were acquired played a pivotal role; images captured at a considerable distance from the tree stem or with a high aperture value exceeding 7-8 (typically seen with beginner cameras) posed significant challenges for automated segmentation, as exemplified in Figure 3.



Figure 3. Tree barks with occlusion.

Consequently, we embarked on a manual segmentation process, utilizing the open-source software Krita. During this process, we cropped the bark images based on the presence of extraneous elements and diligently removed these elements from the image. Conversely, any images where extraneous elements covered more than 60% of the bark area were excluded from our sample.

2.1.2. Grey Level Co-occurrence Matrix (GLCM)

The Grey-Level Co-occurrence Matrix (GLCM) is a well-established technique for extracting features from images, originally devised by Haralick, Shanmugam, and Dinstein in 1973. Its purpose is to capture spatial information from grayscale images by studying the connections between the brightness levels of the central pixel and its neighboring surroundings. Through second-order texture computations, GLCM analyzes the correlation and distinction between pairs of pixels within the input image. This methodology considers both the spatial layout and the textural attributes of these pixels. In the context of this specific investigation, we examined the following five textural features:

- Contrast: the difference between neighbouring pixel values (its immediate neighbor in a specific direction) where weights increase exponentially, considering the equation below;

$$Contrast = \sum_{i,j=0}^{N-1} P_{i,j}(i - j)^2 \quad (1)$$

if both i and j are equal, there is no contrast between these pixels. Example $i-j=0$, pixels are similar.

- Dissimilarity: the linear difference between pixel values, where the dissimilarity weights increase linearly;

$$Dissimilarity = \sum_{i,j=0}^{N-1} P_{i,j} |i - j| \quad (2)$$

- Homogeneity: contrast and homogeneity values are inversely proportional to each other an effect known as Inverse Difference Moment. The homogeneity measurement of images assumes larger values for more minor gray tone differences in the pair and reaches the maximum when all elements in the image are the same;

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (3)$$

- Correlation: the linear relationship between neighboring pixel values;

$$Correlation = \sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (4)$$

- Energy: the textural uniformity of pixel pairs, also known as Uniformity or Angular second moment;

$$Energy = \sum_{i,j=0}^{N-1} P(i,j)^2 \quad (5)$$

2.1.3. Machine learning and Deep learning Algorithms

In this study, we have chosen a comprehensive set of well-established traditional machine learning algorithms. Additionally, for the deep learning component, we have opted for Convolutional

Neural Networks (CNNs), owing to their recognized robustness and widespread adoption in classification tasks.

The considered algorithms are listed below,

- **Support Vector Machine (SVM):** SVMs stand out as a widely utilized algorithm for both classification and regression tasks. This method identifies a hyperplane within an N-dimensional space to categorize data points. The hyperplane acts as a boundary between different classes, and its dimensionality is contingent upon the number of features within the N-dimensional space.
- **K-nearest Neighbors (KNN):** This algorithm seeks to identify similar entities in close proximity using various distance metrics. Initially, KNN sets a value for 'k,' representing the number of neighbors, and then computes the distances between data points, grouping them based on similarity values. KNN is distinctive for not utilizing training datasets in the classification or regression processes, earning it the moniker of a "lazy learner."
- **Decision Tree Classifier:** Decision trees exhibit a tree-like structure, making decisions at each node in the tree based on features within the training dataset.
- **Random Forest Classifier:** A random forest consists of a collection of randomly selected decision trees. It begins by assembling a random set of decision trees from the training data, and then it aggregates rankings or votes from these diverse decision trees to arrive at the final prediction. Random forests are esteemed for their popularity and robustness in classification tasks.
- **Multi-layer Perceptron (MLP):** MLPs are feedforward neural networks featuring multiple layers of perceptrons.
- **Convolutional Neural Networks (CNN):** CNNs represent deep learning architectures with more than one hidden layer. They excel at automatically learning from data, resulting in more robust feature extraction compared to other algorithms. CNNs are particularly well-suited for image classification and recognition challenges.

3. Results

The exact-cropped dataset comprises solely the central segments of each tree bark. Dataset included a total of 527 precise-cropped images representing four distinct tree species: European

beech (155), European silver fir (105), Norway spruce (119), and Sessile oak (148). Within this dataset, 75% of the data was allocated for the training phase, while the remaining portion was reserved for testing the dataset, 75% of the data was used for the training process, with the remained used for testing.

Algorithm	Training accuracy	Testing accuracy	Scaling	
			Training	Testing
K-nearest Neighbor	0.82	0.65	0.83	0.70
Decision Tree	0.81	0.71	0.83	0.76
Random Forest	1.00	0.83	0.99	0.77
Support Vector Machine	0.65	0.61	0.65	0.64

Table 1. The performance of traditional algorithms trained on Dataset of exact-cropped images.

We applied a standard scaler for feature standardization, involving the subtraction of the mean and division of all values by the standard deviation. In the initial scenario, let's take the K-nearest neighbor algorithm as an example. Without data scaling, we obtained a training accuracy of 82% (as shown in Table 1). However, the testing accuracy stood at 65%. Following data scaling, there was an enhancement observed, with the algorithm achieving a 1% increase in training accuracy and a 5% increase in testing accuracy. Similar improvements were also noticeable in the results when applying the Decision Tree algorithm after the scaling process.

3.1. Neural Networks trained on Dataset

In this research, we evaluated the performance of two algorithms: Multi-layer Perceptron (MLP) and Convolutional Neural Networks (CNN). These two algorithms are particularly well-suited for image classification tasks, with CNNs excelling in capturing spatial relationships.

The parameters for the MLP were predefined and detailed in the Appendix. When initially applied to the data without scaling, the MLP yielded less favorable results, with only a 67% training accuracy and 59% testing accuracy. However, after the data was scaled, there was a significant

improvement in both training and testing accuracy, achieving remarkable values of 93% and 92%, respectively.

For the CNN, we fine-tuned the parameter values based on insights gained from our previous experiments. The specific parameter values and corresponding accuracy scores are provided below,

CNN Parameter Values and Results												
Validation ratio	Batch size	Epochs	No. Conv. layers	Filters	Kernel	Pooling (max. pool)	Activation	Dense layer	Drop out	Final activation function	Train accuracy	Test accuracy
0.2	12	40	5	32,64,64,128,512	3*3	5 (2*2)	Relu	64	0.5	Softmax	0.97	0.94
0.2	12	40	5	32,64,64	3*3	3 (2*2)	Relu	64	0.5	Softmax	1.00	0.94
0.2	40	65	3	32,64,128	3*3	3 (2*2)	Relu	64	0.5	Softmax	0.91	0.88

4.Conclusion

This study aimed to delve deeper into the effectiveness of traditional machine learning and deep learning algorithms for the classification of tree species. It sought to explore how these algorithms could be optimally harnessed for this purpose. Specifically, the research tackled questions regarding the potential advantages of using region-specific datasets, such as the exact-cropped dataset, and whether combining various dataset types could yield improved outcomes. The investigation encompassed a diverse array of sub-datasets and their combinations.

In the realm of traditional machine learning, the study initially employed predefined parameters. However, it became evident that these parameter values were often insufficient for many

algorithms. The study also delved into the significance of data scaling when working with these datasets and algorithms, finding that scaling played a crucial role in enhancing the performance of traditional machine learning algorithms.

The research extended to neural networks, including Convolutional Neural Networks (CNNs). Remarkably, in the case of neural networks, the structure of the dataset did not pose a significant constraint, and, with the inclusion of dropout, achieved an average accuracy of approximately 90% across all datasets.

References

- A Community for Naturalists · iNaturalist. (n.d.). Retrieved August 09, 2022, from <https://www.inaturalist.org/>
- Boudra, S., Yahiaoui, I., & Behloul, A. (2022). Tree trunk texture classification using multi-scale statistical macro binary patterns and CNN. *Applied Soft Computing*, 118. <https://doi.org/10.1016/j.asoc.2022.108473>
- Detecting forest threats with Artificial Intelligence – AZO – Space of Innovation. (n.d.). Retrieved August 09, 2022, from <https://space-of-innovation.com/detecting-forest-threats-with-ai/PlantNet>, <https://identify.plantnet.org/>
- Fiel, Stefan & Sablatnig, Robert. (2010). Automated identification of tree species from images of the bark, leaves or needles, *Proceedings of the 16th Computer Vision Winter Workshop*, pp. 67–74, 2011
- Haralick, R.M., Shanmugam, K. and Dinstein, I. (1973) Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3, 610-621
- Liang, X., Hyypä, J., Kaartinen, H., Lehtomäki, M., Pyörälä, J., Pfeifer, N., Holopainen, M., Brolly, G., Francesco, P., Hackenberg, J., Huang, H., Jo, H. W., Katoh, M., Liu, L., Mokroš, M., Morel, J., Olofsson, K., Poveda-Lopez, J., Trochta, J., ... Wang, Y. (2018). International benchmarking of terrestrial laser scanning approaches for forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144, 137–179. <https://doi.org/10.1016/j.isprsjprs.2018.06.021>
- M. Carpentier, P. Giguère and J. Gaudreault (2018). Tree Species Identification from Bark Images Using Convolutional Neural Networks. 2018 IEEE/RSJ International Conference on Intelligent

Robots and Systems (IROS), Madrid, Spain, 2018, pp. 1075-1081,
<https://doi.org/10.1109/IROS.2018.8593514>

Zhou, H., Yan, C., & Huang, H. (2016). Tree species identification based on convolutional neural networks. Proceedings - 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2016, 2, 103–106. <https://doi.org/10.1109/IHMSC.2016.144>