

# Tree Species Classification Based on Detailed Shape Evaluation of Bark and Leaf Using Deep Learning

Tomohiro Mizoguchi

Department of Informatics and Data Science, Faculty of Engineering, Sanyo-Onoda City University, 756-0884, Yamaguchi, Japan.  
mizoguchi@rs.socu.ac.jp

**KEYWORDS:** Tree species, Bark, Leaf, Deep Learning, Majority voting.

## ABSTRACT:

In Japan, many urban park trees are becoming large and aged, increasing the risk of structural failures caused by extreme weather events and biological deterioration. Effective management therefore requires reliable risk assessment, for which accurate tree species identification is essential. However, species identification still depends heavily on visual assessment by skilled professionals, creating challenges in efficiency and objectivity. This issue is particularly significant for broad-leaved trees because of their high species diversity and morphological variability. Labor shortages have also increased the demand for automated classification techniques. This study proposes a tree species classification method for broad-leaved trees using ground-level images captured with an RGB camera and deep learning. The method extracts small local patches containing species-specific visual features, such as leaf shape and bark texture, commonly used by professional arborists. These local features are independently evaluated using deep learning models, enabling effective use of fine-scale visual characteristics. To improve robustness against outdoor imaging variations, including illumination changes, shadows, and moss attachment, multiple patches are classified independently and integrated through majority voting. Experiments were conducted on seven tree species commonly found in Japanese urban parks: cherry, ginkgo, zelkova, konara oak, sawtooth oak, plane tree, and flowering dogwood. Results showed a maximum classification accuracy of approximately 95% under real-world conditions, demonstrating the effectiveness of the proposed method for practical urban tree management.

## 1. INTRODUCTION

### 1.1 Background

In urban parks in Japan, many trees planted during the period of rapid economic growth have now become large and aged, increasing the risk of accidents such as tree falls and branch failures due to extreme weather and biological deterioration. To prevent such accidents and enhance the safety of park users, accurate assessment of tree condition and appropriate management are essential.

In recent years, handheld mobile laser scanners (MLS) based on simultaneous localization and mapping (SLAM) technology have emerged. With such devices, operators can acquire dense point cloud data of the surrounding environment simply by walking with the scanner in hand. Typical MLS systems can scan objects located tens of meters away with an accuracy of approximately 1–2 cm. As illustrated in Figure 1, the acquired point cloud enables visual understanding of both the overall park landscape and the condition of individual trees on a computer. Compared with tripod-mounted terrestrial laser scanners (TLS), MLS offers superior mobility and can be practically applied to urban parks spanning several hectares (Balenovic, 2020; Cabo, 2018; Perez-Martin, 2021; Zhou, 2019).

Accordingly, recent studies have focused on utilizing MLS-derived point clouds to extract individual trees and to accurately estimate key management indicators such as tree height and diameter at breast height (DBH). Furthermore, by linking various attributes to three-dimensional models, long-term monitoring and information management of individual trees are expected to become standard practice.

Tree species information is another critical attribute for effective tree management. Tree species strongly influence growth

characteristics, wood properties, and decay processes, and therefore provide the basis for evaluating failure risk and establishing appropriate maintenance strategies. Accurate species identification is thus an essential prerequisite for reliable urban tree management. However, tree species identification still relies heavily on visual assessment by skilled experts. Such manual approaches are inefficient, and Japan faces a shortage of professionals capable of accurate species identification. In particular, broad-leaved trees, which are widely distributed in urban parks, comprise more than one hundred species in Japan and exhibit high morphological variability. Therefore, the development of techniques that enable accurate and automatic classification of broad-leaved tree species is strongly needed.

### 1.2 Related Works

Species classification of trees based on deep learning has recently attracted considerable attention and has been actively studied. A wide variety of data sources have been employed for this purpose, including satellite remote sensing data, aerial imagery, UAV (unmanned aerial vehicle)-acquired images and point clouds, as well as other sensing modalities (Zhong, 2024). In this paper, we restrict the scope to data acquired on the ground and provide an overview of related studies in this domain.

Plants exhibit diverse physical characteristics that can be utilized for species classification, such as bark, leaves, flowers, and fruits. Among these, bark is the most representative feature. Because bark is less affected by seasonal variation and is relatively easy to access, numerous studies have focused on bark-based classification (Kim, 2022, Wu, 2021, Carpentier, 2018, Misra, 2020). For example, (Carpentier, 2018) proposed a method that emphasizes the detailed morphology of bark. In their approach, tree trunks were imaged from multiple angles at distances of approximately 20–60 cm, and multiple small patches were extracted from each image. Each patch was classified using ResNet, and



(a) Image in the park



(b) Colored point cloud

Figure 1. Point cloud captured by MLS in park

the final species decision was made by majority voting. This method was applied to 23 tree species and achieved classification accuracies of 93.88% with single patches and 97.81% with majority voting across multiple patches. On the other hand, (Misra, 2020) proposed a method in which a single bark image was re-scaled, multiple patches were extracted, and each patch was classified using a CNN, with the results integrated by voting. Their study targeted 101 tree species and reported a maximum classification accuracy of 57.22% when using VGG16. These results suggest that bark is a useful feature for tree species classification; however, because morphological differences in bark among species are often subtle, it remains difficult to achieve high accuracy in classifying the more than 100 broad-leaved species found in Japan based solely on bark images.

Leaves are another effective cue for species identification. Although they undergo seasonal changes, leaves are relatively accessible, easy to capture images, and exhibit clear morphological differences among species, making them a frequent focus of research (Sun, 2017, Okuda, 2022). For instance, (Sun, 2017) conducted species classification of 100 plant species using leaf images captured under natural conditions with mobile devices, employing a modified convolutional neural network (CNN) based on ResNet. Their method achieved a maximum classification accuracy of 91.78%. In another study, (Okuda, 2022) proposed a deep learning-based classification method using RGB images of leaves. In addition to whole-leaf images, they created leaf-shape images capturing only the outline and leaf-vein images extracted from 10 local patches to represent venation patterns. These three types of images were evaluated integratively. The classification accuracy using whole-leaf images alone was 87.91%, but by incorporating leaf-shape and leaf-vein images, the final accuracy improved to 92.07%. Nevertheless, when considering practical applications in outdoor environments such as parks, where leaves are imaged while moving, obtaining high-resolution and clear leaf-vein images remains technically and practically challenging.

In addition, (Hui, 2023) proposed a method that uses point cloud obtained from laser scanning of trees to classify species based on the overall shape of the tree, including the treetop. However, when applied to real park environments, this method is limited by the scanning range and frequently suffers from occlusions caused by foreground trees, leaving issues of data completeness unresolved. As a result, its applicability to general outdoor environments remains constrained.

Considering these circumstances, integrative approaches that evaluate multiple indicators have recently attracted attention as effective methods. In the aforementioned study by (Okuda, 2022), three different perspectives, whole-leaf, leaf-shape, and leaf-vein images, were classified individually using CNNs, and the results were comprehensively evaluated. This integrative approach achieved higher classification accuracy compared to methods relying on a single indicator. Similarly, (He, 2016) proposed an approach that combines three physical characteristics, leaves, flowers, and fruits, while (Bertrand, 2018) introduced a method integrating bark and leaves. In both cases, improvements in classification accuracy were confirmed compared with single-indicator methods, demonstrating the effectiveness of multimodal inputs. However, these studies did not employ deep learning models such as CNNs. It is therefore expected that combining multiple indicators with high-performance algorithms like CNNs could lead to further improvements in classification performance.

### 1.3 Characteristics of Species from a Botanical Perspective

From a botanical perspective, it is known that tree species can be classified to a certain extent based solely on the morphological characteristics of bark (Hayashi, 2006). The primary features used for such classification include bark texture, specifically the presence or absence of fissures and striations, exfoliation, color tone, and the presence and shape of lenticels. On the other hand, it has also been reported that detailed analysis of the morphological characteristics of a single leaf enables the classification of many tree species (Amakawa and Osada, 1988). In leaf-based classification, in addition to leaf shape, several attributes are evaluated, including the presence or absence of climbing habit, whether the leaf is simple or compound, entire or lobed, phyllotaxis (alternate or opposite), and leaf margin characteristics (entire or serrated). Among these morphological features, those that can be relatively easily obtained from images captured with a general-purpose camera in outdoor environments are leaf shape and bark texture.

As shown in Figure 2, this study targets seven species commonly planted in Japanese parks: *Prunus × yedoensis* (cherry), *Ginkgo biloba* (ginkgo), *Zelkova serrata* (zelkova), *Quercus serrata* (konara oak), *Quercus acutissima* (sawtooth oak), *Platanus × acerifolia* (plane tree), and *Cornus florida* (flowering dogwood). These species can be grouped into six categories (B1–B6) based on bark texture, and similarly into six categories (L1–L6) based on leaf shape, as illustrated in Figure 3. In terms of leaf morphology, cherry and zelkova are classified into the same group L4. Regarding bark texture, ginkgo, konara oak, and sawtooth oak are classified into group B4, while zelkova and plane tree are classified into group B6. Among the seven target species, none belong to the same group when both leaf shape and bark texture are considered simultaneously. This suggests that combining features from different perspectives, leaf shape and bark texture, enables more accurate classification of tree species.

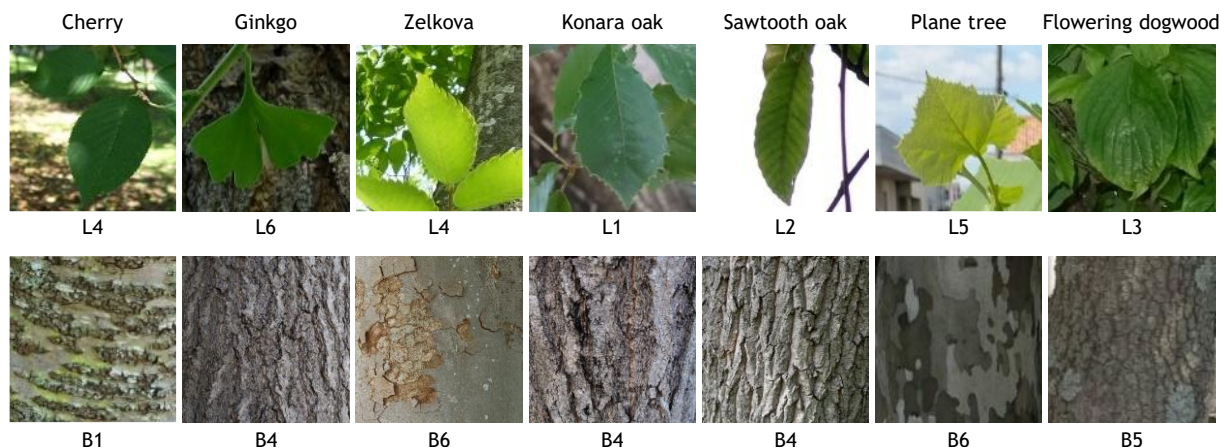


Figure 2. Indicators for Classification Based on Morphological Features

#### 1.4 Research Purpose

In this study, we develop an accurate species classification method for broad-leaved trees by utilizing sets of images of individual trees captured from multiple directions with a standard RGB camera, in combination with deep learning and a voting scheme. The originality of this study lies in the explicit integration of expert-inspired local morphological feature evaluation and a voting-based deep learning framework using both bark and leaf images. Furthermore, through a series of experiments, we not only verify the effectiveness of the proposed method but also discuss the classification results from a botanical perspective, thereby providing a multifaceted evaluation of its validity and potential general applicability. The proposed method is characterized by the following features:

##### 1. Evaluation of local morphological features

In expert-based visual classification, local morphological features that are prominent for each species, such as leaf shape and bark texture, are critical decision criteria. Focusing on this aspect, our method extracts small patches containing bark or leaf regions from RGB images and evaluates these local features with high accuracy using a deep learning model that excels at texture analysis.

##### 2. Classification via voting

In outdoor tree imaging, variations in illumination, shadows cast by branches and leaves, and moss growth often degrade image quality and negatively impact classification accuracy. To address this issue, the proposed method leverages the capability of mobile image acquisition, capturing multiple images of a tree from various positions and angles. Multiple local patches are then extracted from each image and individually classified using a deep learning model. The patch-level predictions are subsequently integrated using a majority voting scheme to derive the final tree-level classification. Compared with conventional approaches that rely on single images, this strategy effectively reduces the influence of imaging conditions (e.g., camera position and lighting) as well as inter-individual variations and local morphological differences in bark and leaves, thereby enabling more accurate species identification.

##### 3. Integrated evaluation of bark and leaves

In many previous studies, tree species classification has primarily been conducted based on a single physical feature, such as either bark or leaves. However, when different species share similar leaf shapes or bark patterns, relying on a single indicator has been shown to be insufficient for achieving accurate classification

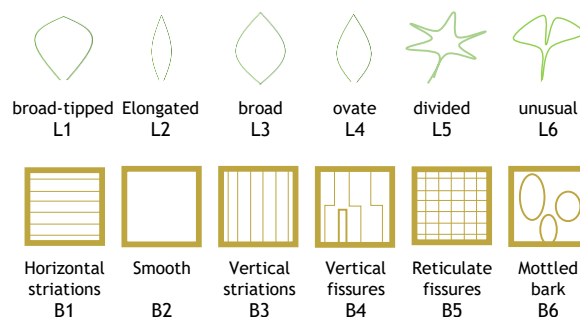


Figure 3. Leaf and bark classification of the target tree species

across multiple species. Moreover, in real-world environments, it is often difficult to obtain suitable images: leaves may be located high in the canopy and thus hard to capture images, or bark surfaces may be obscured by mosses and lichens, reducing their visibility. To address these limitations, the present study adopts an integrative approach that evaluates both bark and leaves as classification criteria, rather than depending on either feature alone. This strategy enables complementary information to be utilized, making it possible to classify species that are difficult to distinguish using only one type of feature, and thereby improving the overall classification accuracy.

## 2. OUR PROPOSED METHOD

### 2.1 Dataset

For image acquisition, we used SONY digital cameras RX100II and RX100VII. The shooting distance was set to within approximately 3 m in order to clearly capture the detailed morphology of leaves and bark. For each tree, multiple images were taken from different positions and orientations to obtain information from diverse viewpoints. From the captured images, small regions (patches) considered effective for classification were manually selected, and approximately 3-9 patches were extracted from each image. For leaves, square patches were extracted to include at least one complete leaf.

For bark, discussions with expert arborists confirmed that a wider area is more appropriate for capturing textural features such as fissures and exfoliation, as illustrated in Figure 3. Therefore, some extracted patches were as large as  $3,000 \times 3,000$  pixels, which is considerably larger than the  $224 \times 224$  pixel input size used in the CNN model by (Carpentier, 2018). To ensure sufficient numbers of patch images when the available images were

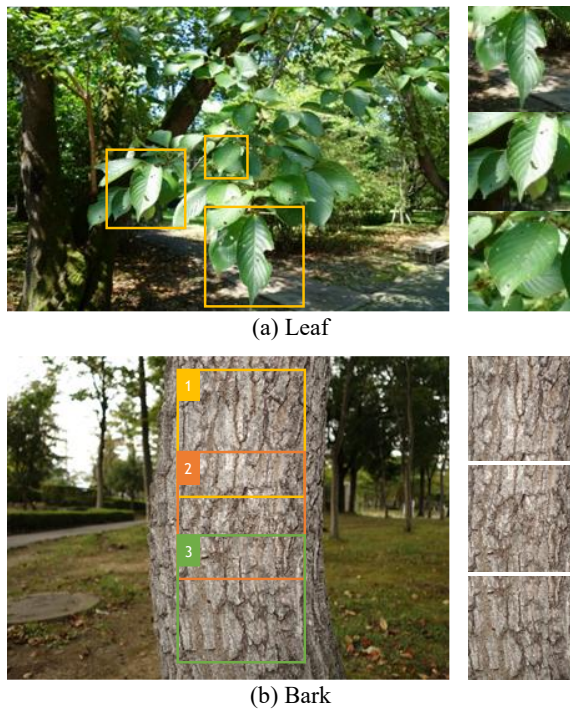


Figure 4. Illustration of patch image generation

Table 1. Number of training and test images

Species	training	validation
Cherry blossom	450	225
Ginkgo	300	165
Zelkova	360	195
Konara oak	330	180
Sawtooth oak	330	165
Plane tree	360	195
Flowering dogwood	390	195

limited, up to approximately 50% overlap between patches was permitted. An example of the patch extraction procedure is shown in Figure 4. Image acquisition was conducted between June and September, under predominantly clear weather conditions. For subsequent training with the CNN model (VGG16), all patch images were resized to  $224 \times 224$  pixels.

## 2.2 Deep Learning

Species classification was performed on the extracted patch images using a CNN. The model employed was VGG16 (Simonyan, 2015), which is widely used in deep learning research. Although VGG16 is considered a classical CNN architecture, it remains effective for evaluating detailed texture information as required in this study, and stable classification performance can be expected. In particular, its simple and deep hierarchical structure is well suited for capturing fine-grained texture patterns such as bark surfaces. For training, transfer learning was applied using weights pretrained on ImageNet. This design enables efficient and highly accurate classification even with a limited number of training samples.

## 2.3 Classification based on a Single Indicator

In this study, we adopted a majority voting approach with the aim of improving classification accuracy. This method involves inputting multiple patch images into a single classifier and

determining the class of the entire target image by majority voting based on the individual inference results. In general, the simple majority rule is employed, in which the class that obtains the largest number of votes (predictions) among all patches is assigned as the final label. However, ties can occur when multiple classes receive the same maximum number of votes. To address this tie-breaking issue in majority voting, various methods have been proposed. Among them, this study employed an approach based on softmax accumulation (Kokkinos, 2014). The detailed procedure adopted in this study is described below.

From multiple images of a tree,  $N$  patch images  $\{r_i\}$  are generated. For each patch image  $r_i$ , the CNN output is passed through a softmax function to obtain the probability value  $p_{i,j}$  for each class  $j$ . Using these outputs from multiple images, the average voting score for class  $j$  is calculated as shown in Equation (1).

$$P_j = \frac{1}{N} \sum_{i=1}^N p_{i,j} \quad (1)$$

Let the average voting score for class  $j$  obtained from leaf images and bark images be denoted as  $P_j^{leaf}$  and  $P_j^{bark}$  respectively. The final classification is determined based on these values, assigning the labels to the class for which either  $P_j^{leaf}$  or  $P_j^{bark}$  attains the maximum value.

## 2.4 Classification based on Integrated Evaluation of Bark and Leaves

Next, we describe the method for integrating the classification results obtained from leaves and bark to derive the overall classification outcome. Using the voting scores for each class obtained in the previous section, a weighted sum of the outputs from leaves and bark is calculated according to Equation (2).

$$P_j^{overall} = \alpha P_j^{leaf} + (1 - \alpha) P_j^{bark} \quad (2)$$

Here, the weighting coefficient  $\alpha$  is defined within the range of 0 to 1, and this value controls the influence of the classification results from leaves and bark on the final decision. When both are evaluated integratively, the label is assigned to the class whose score, computed by Equation (2), attains the maximum value.

## 3. EXPERIMENTAL RESULTS AND DISCUSSION

### 3.1 Training and Test Data

In this study, the training dataset for deep learning consisted of 15 cherry trees, 10 ginkgo trees, 12 zelkova trees, 11 konara oaks, 11 sawtooth oaks, 12 plane trees, and 13 flowering dogwoods. From each tree, 15 image patches were extracted from different angles. Data augmentation was further performed by horizontal flipping, resulting in a total of 2,520 training images.

For the test dataset, 15 cherry trees, 11 ginkgo trees, 13 zelkova trees, 12 konara oaks, 11 sawtooth oaks, 13 plane trees, and 13 flowering dogwoods were used, from which 15 image patches per tree were extracted and employed for evaluating classification accuracy. The detailed numbers of training and test images for each species are summarized in Table 1.

### 3.2 Evaluation of Classification Performance by Voting

The effect of voting on improving classification performance was evaluated. Figure 5 shows the changes in the F-score as the number of voting images varied. For leaf images, the average F-score was 0.746 when only a single image was used, but it improved as the number of images increased, reaching a maximum of 0.892

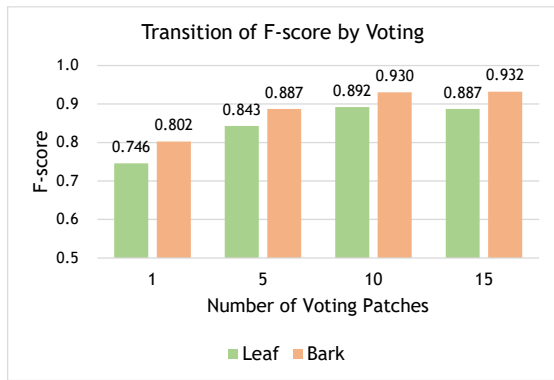


Figure 5. Transition of F-score with Voting

with 10 images. A similar trend was observed for bark images: the F-score was 0.802 with a single image, but increased to 0.932 when 15 images were used, confirming that voting enhanced classification performance.

Table 2(a) presents the confusion matrix obtained from 15-image voting using leaf images, while Table 2(b) shows the corresponding confusion matrix for bark images. In the classification results based on leaf images, a total of nine misclassifications were observed. Among them, two zelkova trees (s3) were misclassified as cherry (s1). Since both species are categorized as ovate (L4) in the morphological classification table shown in Figure 3, their similarity in shape is considered a likely cause of this error. In addition, one zelkova (s3) was misclassified as konara oak (s4), which belongs to the broad-tipped group (L1); this also reflects morphological similarity between the two groups. On the other hand, six sawtooth oaks (s5) were misclassified as konara oak (s4). Although sawtooth oak is classified as elongated (L2) and konara oak as broad-tipped (L1), which are morphologically distinct, this misclassification is presumed to be attributable to non-shape factors, such as similarities in color or texture. In contrast, in the classification results based on bark images, a total of six misclassifications were observed. Among these, three cases involved ginkgo (s2) and sawtooth oak (s5) being misclassified as konara oak (s4). Since all three species are categorized as vertical fissures (B4), their similar bark textures are considered the main cause of these errors.

From these results, it was confirmed that most of the misclassifications exhibited trends consistent with botanical morphological classifications, indicating that the classification outcomes can be explained from an academic perspective.

### 3.3 Classification Performance Based on Integrated Evaluation of Leaves and Bark

Subsequently, we describe the results of classification using an integrated approach that combines leaf and bark images. Fifteen patches were used, and the integration weight coefficient  $\alpha$  varied from 0 to 1 in increments of 0.1 for classification. Figure 6 shows the transition of the F-score as  $\alpha$  changed. The results indicated that the highest F-score was achieved when  $\alpha = 0.5$ , yielding higher accuracy than evaluations based on either leaves or bark alone. This confirms the effectiveness of integratively utilizing both types of information.

Furthermore, Table 2(c) presents the confusion matrix obtained with  $\alpha = 0.5$ , where the highest F-score was recorded. The results revealed multiple cases in which samples misclassified when using either leaves or bark alone were correctly classified through the integrated evaluation. These findings suggest that

Table 2. Evaluation of Classification Accuracy Using 15 Patches (s1: Cherry, s2: Ginkgo, s3: Zelkova, s4: Konara oak, s5: Sawtooth oak, s6: Plane tree, s7: Flowering dogwood)

(a) Leaf

		Prediction							Recall
		s1	s2	s3	s4	s5	s6	s7	
Ground truth	s1	15	0	0	0	0	0	0	1.00
	s2	0	11	0	0	0	0	0	1.00
	s3	2	0	10	1	0	0	0	0.77
	s4	0	0	0	12	0	0	0	1.00
	s5	0	0	0	6	5	0	0	0.45
	s6	0	0	0	0	0	13	0	1.00
	s7	0	0	0	0	0	0	13	1.00
	Precision	0.88	1.00	1.00	0.63	1.00	1.00	1.00	avg F-score
F-score	0.94	1.00	0.87	0.77	0.63	1.00	1.00	0.89	

(b) Bark

		Prediction							Recall
		s1	s2	s3	s4	s5	s6	s7	
Ground truth	s1	14	0	1	0	0	0	0	0.93
	s2	0	9	0	2	0	0	0	0.82
	s3	1	0	12	0	0	0	0	0.92
	s4	0	0	1	11	0	0	0	0.92
	s5	0	0	0	1	10	0	0	0.91
	s6	0	0	0	0	0	13	0	1.00
	s7	0	0	0	0	0	0	13	1.00
	Precision	0.93	1.00	0.86	0.79	1.00	1.00	1.00	avg F-score
F-score	0.93	0.90	0.89	0.85	0.95	1.00	1.00	0.93	

(c) Leaf and bark with  $\alpha=0.5$

		Prediction							Recall
		s1	s2	s3	s4	s5	s6	s7	
Ground truth	s1	15	0	0	0	0	0	0	1.00
	s2	0	10	0	1	0	0	0	0.91
	s3	2	0	11	0	0	0	0	0.85
	s4	0	0	0	12	0	0	0	1.00
	s5	0	0	0	1	10	0	0	0.91
	s6	0	0	0	0	0	13	0	1.00
	s7	0	0	0	0	0	0	13	1.00
	Precision	0.88	1.00	1.00	0.86	1.00	1.00	1.00	avg F-score
F-score	0.94	0.95	0.92	0.92	0.95	1.00	1.00	0.95	

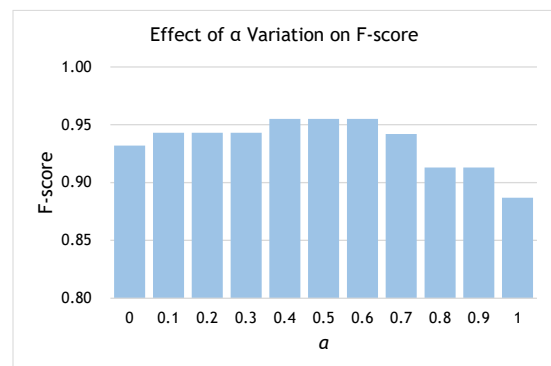


Figure 6. Variation of F-score with Changes in Weight  $\alpha$

combining features from different perspectives contributes to improved classification accuracy.

## 4. CONCLUSION AND FUTURE WORKS

In this work, we proposed an accurate species classification method for broad-leaved trees using images acquired with a standard RGB camera and verified its effectiveness through experiments. The main findings are summarized as follows. First, by incorporating detailed local features of leaves and bark used by expert arborists for species identification into the deep learning evaluation process, accurate classification was achieved. Second, local image patches extracted from images captured at various positions and orientations were individually classified, and the final decision was determined through majority voting; classification accuracy increased as the number of voting images

increased. Furthermore, by using both leaf and bark images as evaluation indicators and performing integrative classification, several species that were misclassified when using either feature alone were correctly identified, resulting in further improvements in overall accuracy.

Although the dataset used in this study was limited in size and collected from a single park environment, the experimental results demonstrated the effectiveness of the proposed framework under real-world conditions. In most cases, the classification outcomes could be reasonably explained from a botanical perspective based on the detailed visual features described in Section 1.3. This suggests that incorporating more taxonomically detailed criteria into the classification framework may further improve accuracy. Moreover, the results indicate that strategies reflecting expert judgment, such as using both leaves and bark as discriminative features or extracting large-area bark patches at reduced resolution, can be effectively integrated into CNN-based models to achieve high classification accuracy. Such empirically grounded decision criteria derived from expert knowledge are also expected to further enhance performance.

Because the proposed framework relies solely on RGB images captured by standard cameras and local patch-based evaluation, it is inherently scalable across diverse tree species and outdoor environments. This scalability enables large-scale vegetation monitoring and species-level mapping, thereby contributing to landscape management and the sustainable planning and maintenance of urban green infrastructure.

Future work includes the development of automatic patch extraction methods, the application of activation maps to visualize and interpret the basis of classification decisions, and the evaluation of end-to-end learning frameworks and their impact on classification accuracy. To further assess the generalization capability of the proposed method, multi-site data collection across different seasons will be conducted. Furthermore, we plan to explore methods for integrating species information with point cloud data acquired by mobile laser scanners, thereby enabling the direct annotation of species information onto 3D models.

#### ACKNOWLEDGEMENT

The author gratefully acknowledges Mr. Yasuyuki Sanpei, a certified tree doctor, for his valuable advice and professional guidance throughout this study.

#### REFERENCES

Amakawa, T., Osada, T. (1988). *Trees*, Vol. 1. Hoikusha Publishers Co., Ltd. (in Japanese)

Balenovic, I., Liang, X., Jurjević, L., Hyypä, J., Seletković, A., Kukko, A. Hand-Held Personal Laser Scanning – Current Status and Perspectives for Forest Inventory Application, *Croatian Journal of Forest Engineering*, 42, 11, 2020.

Bertrand, S., Ameur, R. B., Cerutti, G., Coquin, D., Valet, L., Tougne, L., 2018. Bark and Leaf Fusion System to Improve Automatic Tree Species Recognition, *Ecological Informatics*, 46, 57-73.

Cabo, C., Del Pozo, S., Rodríguez-González, P., Ordóñez, C., González-Aguilera, D., 2018. Comparing Terrestrial Laser Scanning (TLS) and Wearable laser Scanning (WLS) for Individual Tree Modeling at Plot Level, *Remote Sensing*, 10, 540.

Carpentier, M., Giguère, P., Gaudreault, J., 2018. Tree Species Identification from Bark Image Using Convolutional Neural Networks, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Hayashi, M., 2006. *The handbook of bark*, Bun-ichi Sogo Shuppan Co., Ltd. (in Japanese)

He, A., Tian, X., 2016. Multi-Organ Plant Identification with Multi-Column Deep Convolutional Neural Networks, *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*.

Hui, Z., Cai, Z., Xu, P., Xia, Y., Cheng, P., 2023. Tree Species Classification using Optimized Features Derived from Light Detection and Ranging Point Clouds Based on Fractal Geometry and Quantitative Structure Model, *Forest*, 14, 1265.

Kim, T. K., Hong, J., Ryu, D., Kim, S., Byeon, S. Y., Huh, W., Kim, K., Baek, G. H., Kim, H. S., 2022. Identifying and Extracting Bark Key Features of 42 Tree Species using Convolutional Neural Networks and Class Activation Mapping, *Scientific Reports*, 12, 4772.

Kokkinos, Y., Margaritis, K.G., 2014: Breaking ties of plurality voting in ensembles of distributed neural network classifiers using soft max accumulations. *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pp. 2028, Springer.

Misra, D., Crispim-Juniro, C., Tougne, L., 2020. Patch-based CNN Evaluation for Bark Classification, *Computer Vision – ECCV 2020 Workshops. ECCV 2020. Lecture Notes in Computer Science*, vol 12540.

Okuda, M., Ohshima, H., 2022. Feature Fusion for Leaf Image Classification, *IEEE International Conference on Big Data and Smart Computing (BigComp)*.

Pérez-Martín, E., López-Cuervo Medina, S., Herrero-Tejedor, T., Pérez-Souza, M.A., Aguirre de Mata, J., Ezquerro-Canalejo, A., 2021. Assessment of Tree Diameter Estimation Methods from Mobile Laser Scanning in a Historic Garden, *Forests*, 12, 1013.

Simonyan, K., Zisserman, A., 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition, *International Conference on Learning Representations*.

Sun, Y., Liu, Y., Wang, G., Zhang, H., 2017. Deep Learning for Plant Identification in Natural Environment, *Computational Intelligence and Neuroscience*.

Wu, F., Gazo, R., Benes, B., Haviarova, E., 2021. Deep BarkID: a Portable Tree Bark Identification System by Knowledge Distillation, *European Journal of Forest Research*, 140, 1391-1399.

Zhong, L., Dai, Z., Fang, P., Cao, Y., Wang, L., 2024: A Review: Tree Species Classification Based on Remote Sensing Data and Classic Deep Learning-Based Methods. *Forest*, 15, 852.

Zhou, S., Kang, F., Li, W., Kan, J., Zheng, Y., He, G., 2019. Extracting Diameter at Breast Height with a Handheld Mobile LiDAR System in an Outdoor Environment, *Sensors*, 19, 3212.