

Short Communication: Geometric morphometric analysis of leaf venation in four *Shorea* species for identification using Digital Image Processing

ISHAK ARIAWAN^{1,♥}, YENI HERDIYENI^{1,♥}, ISKANDAR ZULKARNAEN SIREGAR^{2,♥♥♥}

¹Department of Computer Science, Faculty of Mathematics and Natural Sciences, Institut Pertanian Bogor. Jl. Meranti, Kampus IPB Darmaga, Bogor 16680, West Java, Indonesia. Tel./fax.: +62-251-8625481, ♥email: ishak_ariawan@apps.ipb.ac.id, ♥♥yeniherdiyeni@gmail.com

²Department of Silviculture, Faculty of Forestry, Institut Pertanian Bogor. Jl. Ulin Lingkar Akademik, Kampus IPB Darmaga, Bogor 16680, West Java, Indonesia. Tel./fax.: +62-251-8621677, ♥♥♥email: siregar@apps.ipb.ac.id

Manuscript received: 30 September 2019. Revision accepted: 26 June 2020.

Abstract. Ariawan I, Herdiyeni Y, Siregar IZ. 2020. Short Communication: Geometric morphometric analysis of leaves venation in *Shorea* spp. for identification using Digital Image Processing. *Biodiversitas* 21: 3303-3309. *Shorea* is one of the genera of the *Dipterocarpaceae* family which consists of more than 190 species. Massive exploitation of forests has threatened the sustainability of *Shorea* in nature. A total of 156 species has been listed on the IUCN (International Union for Conservation of Nature) red list. From the 156 species, 59.6% are in the critically endangered category, so urgent conservation is needed. However, during collection of *Shorea* at the seedling phase for conservation purposes, it is often difficult to distinguish among them that can cause errors in their collection process. To avoid these errors, identification needs to be done, usually based on plant leaf and flower morphology. Leaves are easier because they have the main features that distinguish each plant species, one of which is the venation structure. Geometric morphometric techniques are a modern approach recognized as useful for the identification of species in many plants. Geometric morphometrics analyzes the position of the venation point using coordinate geometry values. This research was aimed to extract venation features of *Shorea* leaves using a geometric morphometric approach. The extraction process result in some features, such as straightness, different angle, length ratio, scale projection, and secondary nerves. On extracted features, an analysis was then performed to find out the best features in classifying species of *Shorea* spp. The results of this study indicated that the geometric morphometric approach could extract the value of the features of straightness, different angle, length ratio, scale projection, and secondary nerves. The secondary nerve feature is the best feature because it can distinguish between four commonly planted species of *Shorea* spp. (*S. acuminata*, *S. leprosula*, *S. ovalis*, and *S. selanica*). By using the support vector machine classification technique to identify species of *Shorea* spp., the classification results obtained an average accuracy of 84.46%.

Keywords: Feature extraction, geometric morphometrics, identification, leaf venation, *Shorea*

INTRODUCTION

Shorea is one of the genera of the *Dipterocarpaceae* family which consists of 194 species that grow in the tropics and spread from Borneo, Thailand, Malaysia, to Sumatra (Newman et al. 1999; Cao et al. 2006). Massive forest exploitation has threatened the preservation of *Shorea* in nature (Eni et al. 2018). According to the IUCN (International Union for Conservation of Nature) of 194 *Shorea* species 156 species have been listed on the IUCN red list where 59.6% are in the critically endangered and 25% in the endangered category.

Ex-situ conservation is one of the effective strategies to save plants from extinction (Dodo and Wawangningrum 2018). The type of planting stock that allows it to be carried or moved is uprooted wildlings (Bastoni 2005). However, during the collection process, the seedling phase of *Shorea* species are often difficult to be distinguished and identified from each other so that errors in the collection might result in unexpected *Shorea* species. To avoid this error, careful identification needs to be carried out. The identification process can be done using plant morphology (Ambarwari et al. 2018). Plant morphology is the study of

the physical form and body structure of plants (Raven et al. 2005). One part of the plant that is easily found and has different characteristics between one plant and another and is often used to identify plant species is leaves (Le et al. 2014). Leaves have the main features that distinguish each type of plant, including features of shape, venation structure, color, and leaf surface condition (Rahmadani and Herdiyeni 2010). Among the features, leaf venation has a unique diversity that can describe plant characteristics in more detail, although some plant species do not exhibit clear patterns of venation (Wahyumianto et al. 2011). Besides, leaf venation also has a consistent pattern that genetically determined (Roth-Nebelsick et al. 2001).

Geometric morphometric techniques are a modern approach recognized as useful for the identification of species in many plants (Viscosi and Cardini 2011). Geometric morphometrics analyze the position of the venation point using coordinate geometry values. This method has been successfully used to explain passion fruit leaves (Plotze and Bruno 2009), venation type of medicinal plants (Salima et al. 2015; Ambarwari et al. 2018), leaf variations in *Passiflora* (Chitwood and Otoni 2017), and olive tree varieties (Bourgeon et al. 2017). This method has

also been successfully applied in the identification of among and different types of cereal grains (Bonhomme et al. 2017), to distinguish dental morphology among domestic equids (Chuang and Bonhomme 2018), identification the types of flies in the cocoon phase (Chaiphongpachara and Tubsamut 2019), and fish diversity in the North Sea (Caillon et al. 2018). Despite available manuals for identification of dipterocarps, deeper information on leaf venation of four selected *Shorea* species is still required to support morphology identification. Therefore, this study was carried out with aims to extract venation features of *Shorea* leaves using a morphometric geometric approach. The feature extraction process results in some features, such as straightness, different angles, length ratio, scale projection, and secondary nerves. On extracted features, an analysis is then performed to find out the best features in classifying *Shorea* spp species.

MATERIALS AND METHODS

Plant materials

Shorea is one of Indonesia's primary wood producing forest trees and is an important commodity. As a family of *Dipterocarpaceae*, *Shorea* dominates lowland rainforest in western Indonesia and is the most important genus that is most exploited in the Asian wet forest region. The average size of growth of each *Shorea* takes more than 30 years to obtain a diameter of 30 cm (Newman et al. 1999). The *Shorea* species used in this study were *S. acuminata*, *S.*

leprosula, *S. ovalis*, and *S. selanica* with leaf characteristics as described by Newman et al (1996).

Shorea acuminata

Leaves elliptical, total of secondary nerves 8 - 13 pairs (Newman et al. 1996), Primary nerves length 9-11.5cm, length of secondary nerves 1.5-2.5cm, distance of secondary nerves 2.5-3 cm, upper angle 61-65°, middle angle 64-69°, and bottom angle 66-68°.

Shorea leprosula

Leaves oblong or oblong-lanceolate, total of secondary nerves 10-15 pairs (Newman et al. 1996), length of primary nerves 14.5-17 cm, length of secondary nerves 2.5-3.5 cm, distance of secondary nerves 3.5-5cm, upper angle 42-50°, middle angle 55-61°, and bottom angle 60-65°.

Shorea ovalis

Leaves oblong, elliptical or obovate, total of secondary nerves 17 - 24 pairs (Newman et al. 1996), length of primary nerves 14-18.5 cm, length of secondary nerves 2.5-3cm, distance of secondary nerves 3.5-4.5cm, upper angle 37-39°, middle angle 55-60°, lower angle 55-60°.

Shorea selanica

Leaves oblong or ovate, total of secondary nerves 19-23 pairs (Newman et al. 1999), length of primary nerves 13-17cm, length of secondary nerves 2.5-3.5cm, distance of secondary nerves 3-4cm, upper angle 58-61°, middle angle 59-65°, lower angle 61-65°.



Figure 1. Variation in morphology of leaf shape of *S. acuminata*



Figure 2. Variation in morphology of leaf shape of *S. leprosula*

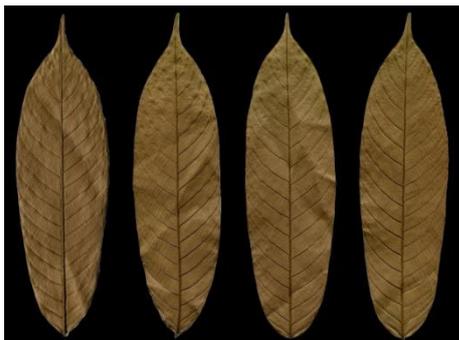


Figure 3. Variation in morphology of leaf shape of *S. ovalis*



Figure 4. Variation in morphology of leaf shape of *S. selanica*



Figure 5. Leaf picking technique

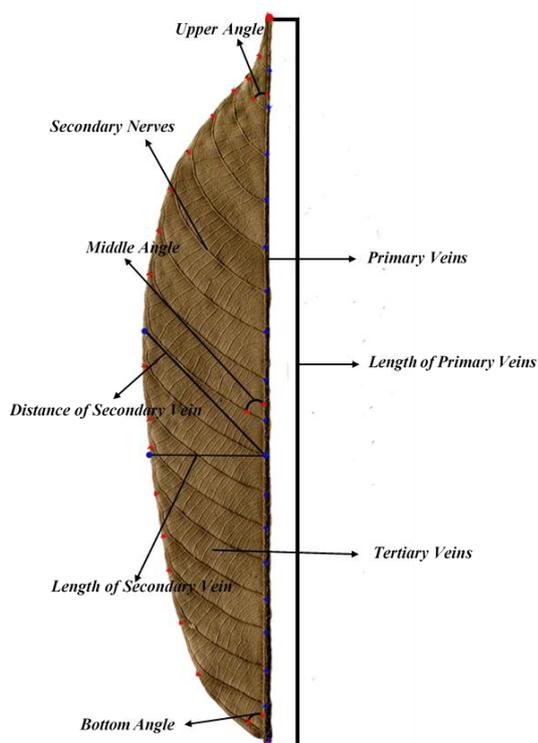


Figure 6. Morphometric characteristics of venation of leaves of *Shorea* spp.

Leaf collection and preparation

Collection of *S. acuminata*, *S. leprosula*, *S. ovalis*, and *S. selanica* leaves was carried out in 3 nursery locations in Bogor, West Java, Indonesia. The three nurseries were the Faculty of Forestry-IPB University, Forest Research and Development Center - Bogor, and Dramaga Research Forest near CIFOR Headquarter - Bogor. The collected leaves were from four *Shorea* species which were still in the seedling phase. The leaves were collected by taking 5-7 outer leaves (Figure 5). Total leaves were 257 where each species consisted of 61 samples of *S. acuminata*, 68 samples of *S. leprosula*, 66 samples of *S. ovalis*, and 62 samples of *S. selanica*, respectively. The leaves were then brought to the IPB Faculty of Forestry for a herbarium process. The herbarium process was carried out to obtain the clear leaf venation.

Segmentation and venation detection

At this stage segmentation was done on the original image of the leaves using the Sobel operation technique (Al-amri et al. 2010), the results obtained then followed thresholding with a range of values from 1 to 255. The results of thresholding were processed to obtain binary images using the Fast algorithm Parallel Thinning (Zhang and Suen 1984). After that, vein detection was applied where binary image data that has been collected from the segmentation process is determined to coordinate bifurcation, intermediary, and extreme point. Bifurcation is the beginning of venation branching. Bifurcation is a connection between primary, secondary, and tertiary venations. The last point of a branch is called the extreme point. The connection between the extreme point and bifurcation is named intermediary (Figure 7).

Feature analysis

The leaf venation feature extraction is obtained from the calculation of straightness, different angle, length ratio and scale projection. To calculate the value of the feature, it takes the position of the corresponding point coordinates (one segment) of the segment extraction process as shown in Figure 8. While the total of secondary nerves was calculated from total of segments produced during the venation detection process of leaves.

Straightness

Straightness is the measurement of the straightness value of a segment. From Figure 8, the straightness value can be calculated using equation 1 (Ambarwari et al. 2018).

$$Straightness = \frac{l_j}{d_j} \tag{1}$$

Where; l_j is the length of the segment j represented by the number of pixels connected by the points j and point l , d_j is the distance between the pixel (x_s^j, y_s^j) coordinates with the pixel (x_e^j, y_e^j) coordinates.

Different angle

The different angle is the measurement of the angular difference between segments that coincide. From Figure 8, different angle (δ) values can be calculated using equation 2 (Ambarwari et al. 2018).

$$\delta_1 = |\alpha_i - \alpha_j| \tag{2}$$

Length ratio

The length ratio value is measured by comparing the length of each segment to the maximum length of the segment in a leaf venation image. From Figure 8, the value of the length ratio can be calculated using equation 3 (Ambarwari et al. 2018).

$$R_i = \frac{l_i}{\max(\vec{l})} \tag{3}$$

Where; R_i is the length ratio of the segment i , l_i is the length of the segment i and \vec{l} is the segment vector in the leaf image.

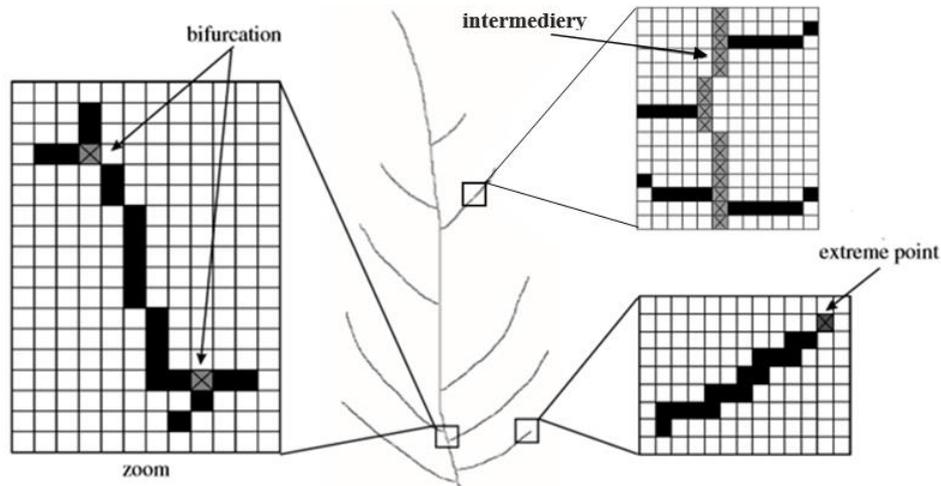


Figure 7. Bifurcation, extreme point and intermediary in leaf venation (Plotze and Bruno 2009)

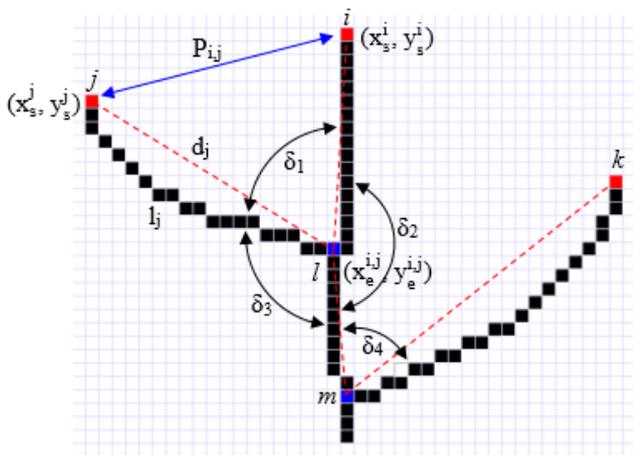


Figure 8. Illustration of *Shorea* leaf venation extraction (Ambarwari et al. 2018)

Scale projection

Scale projection is the measurement of projection length between segments that coincide. In Figure 8, to determine the length of the projection between segments i and j, can use equation 4 (Ambarwari et al. 2018).

$$P_{i,j} = \frac{(x_s^i - x_s^j)(x_s^j - x_s^j) + (y_s^i - y_s^j)(y_s^j - y_s^j)}{(\max(d_i, d_j))^2} \tag{4}$$

After getting the value of these features, boxplot analysis and classification were carried out. Boxplot analysis is conducted to find out the distribution of data from each feature in order to figure out the most important feature (Thirumalai et al. 2017), while the classification is applied to determine the accuracy in identifying the species

of *Shorea* spp. (*S. acuminata*, *S. leprosula*, *S. ovalis*, and *S. selanica*). The classification technique that was used is the support vector machine technique (Parikh and Shah 2016).

RESULTS AND DISCUSSION

Segmentation

The results of segmentation in the form of binary images could extract primary venation and secondary venation explicitly as shown in Figure 9.

From the results of the 257 leaf venation segmentation, the best and most suitable segmentation results were selected for the next process. The feasibility of the leaf image segmentation results could be seen from the resulting venation form.

Venation detection

The leaf image data from venation segmentation showed that out of 251 leaf images data that were successfully processed until vein detection, there were 6 leaf images data classified as not successful in vein detection. Leaf images data that did not succeed in vein detection were for 3 leaves of *S. acuminata*, 1 leaf of *S. selanica* and 2 leaves of *S. leprosula*. So, from 251 leaf images data that were successfully processed at this stage, 245 of them had good vein detection results. The results of venation detection were binary images that have bifurcation and extreme point as shown in Figure 10.

Analysis of feature extraction results

After conducting vein detection, feature extraction was then performed by calculating the value of secondary nerves, straightness, different angles, length ratio, and scale projection. The results of the feature extraction of secondary nerves, straightness, different angle, length ratio, and scale projection are shown in Table 1.

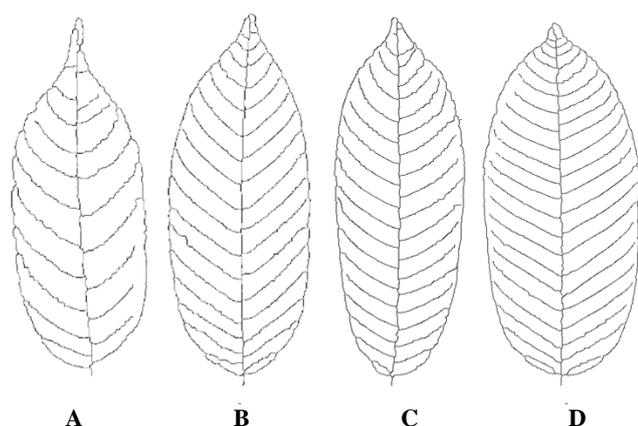


Figure 9. Results of venation segmentation of *Shorea* spp. leaves; A. *S. acuminata*, B. *S. leprosula*, C. *S. ovalis*, and D. *S. selanica*

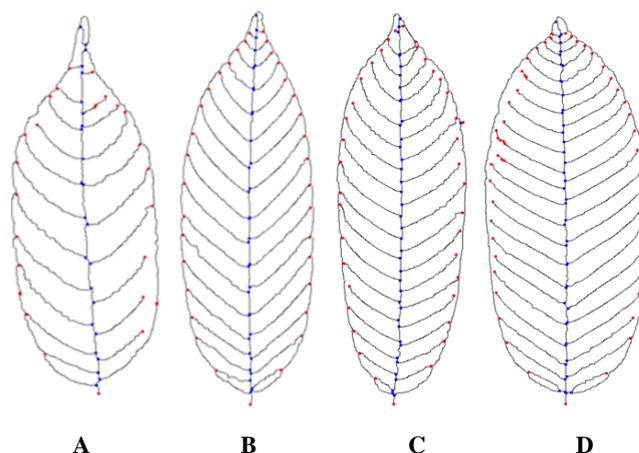


Figure 10. Results of venation detection of *Shorea* spp. leaves; A. *S. acuminata*, B. *S. leprosula*, C. *S. ovalis*, and D. *S. selanica*

Table 1. The values of features for straightness, different angle, length ratio, scale projection and secondary nerves

Venation feature	Value
Mean of Straightness	39.957
Mean of different angle	116.298
Mean of length ratio	0.236
Mean of scale projection	0.004
Total secondary nerves	37

Table 2. The values of features for straightness, different angle, length ratio, scale projection and secondary nerves resulting from Z-Score normalization

Venation feature	Value
Mean of Straightness	-0.3252
Mean of different angle	-0.3137
Mean of length ratio	-0.3253
Mean of scale projection	-0.3254
Total secondary nerves	-0.3216

After obtaining the value of each venation feature of *Shorea* spp. leaves, then normalization was performed using Z-Score normalization. The function of the normalization process is to avoid traits whose values dominate the values of other identifiers (Li and Liu 2011). The results of Z-Score normalization are shown in Table 2.

From Table 2, a boxplot was generated to examine data patterns of each leaf venation feature of the four *Shorea* species. Based on Figure 11, it shows that *S. selanica* has a higher value on the secondary nerves feature but lower in the straightness feature. On the contrary, *S. acuminata* has a lower value on the secondary nerves feature but higher in the straightness feature. Figure 11 also explains that the spread range of the four species of *Shorea* in the secondary nerves feature is better than the straightness feature, in which the secondary nerves feature can distinguish between four species well but the straightness feature can only distinguish *S. ovalis* with *S. selanica* well.

Based on Figure 12, it shows that *Shorea selanica* has a higher value on the scale projection feature but lower in the length ratio feature. While *S. acuminata* has a lower value on the different angle and scale projection feature but higher in the length ratio feature. The length ratio feature shows that there is a higher outlier in *S. ovalis*, this is because the length of the secondary nerves of the leaf is longer than the other leaves, which means that there are some leaves on *S. ovalis* that are wider in size than the

other leaves. Figure 12 also explains that the spread range of the four species of *Shorea* in the length ratio feature is better than the different angle and scale projection feature. Where length ratio feature can distinguish between *Shorea* species well, except *S. ovalis* and *S. selanica*, different angle features can only distinguish *S. acuminata* from *S. selanica* well, and scale projection feature can only distinguish *S. leprosula* from *S. selanica* well.

Classification

At this stage, the classification was done using the support vector machine classification technique to test the accuracy of each class. The results of the classification are in the form of accuracy in correctly identifying *Shorea* species, which can be seen in Figure 13.

Based on Figure 13, *S. acuminata* is classified as the class with the best level of accuracy, which is 100%, and *S. leprosula* is classified as the class with a less good accuracy level, which is 69.23%. This is due to the presence of *S. leprosula* leaves that were classified into *S. acuminata* and *S. ovalis*, where 1 is *S. acuminata*, and 3 is *S. ovalis*. The cause of *S. leprosula* is more classified in *S. ovalis* due to the number of secondary venations of the two species being the same. The average accuracy obtained from the test results for each *Shorea* species was 84.46%.

Several studies on geometrics morphometrics of leaves are carried out to determine the geographical area of the

plant, including on Oak plant (Viscosi and Cardini 2011). MorphJ method had been used with the leaf-shaped parameter. The usage of MorphJ can ease the botanist to analyse the spread of the Oak plant, since the MorphJ comes with the analysis tool like PCA, ANOVA, Regression and VCA.

The result of the research showed that the leaf-shaped parameters can be used to know the the geographical area from the plant. Therefore, leaf-shaped parameters can be used to distinguish Shorea species. However, the single usage of MorphJ cause the parameters used in this research not useful like the straightness, different angle, and number of secondary nerves. Hence, geometrics morphometrics is still needed to increase the accuracy in Shorea species classification.

Weakness

This study only used four out of 194 Shorea species (*S. acuminata*, *S. leprosula*, *S. ovalis*, and *S. selanica*). Besides, the amount of data from each species is approximately 65 leaves, so it needs more additional leaves

in order to increase the accuracy. Extracting venation is performed only on secondary and primary venation, so that the segmentation images that still have tertiary venation are removed manually using Adobe Photoshop.

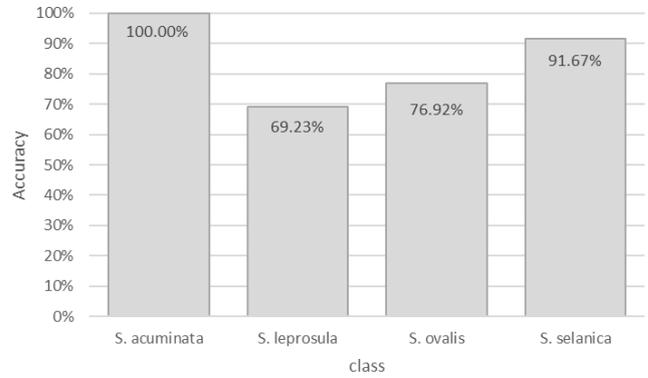


Figure 13. Graph of accuracy comparison of each species

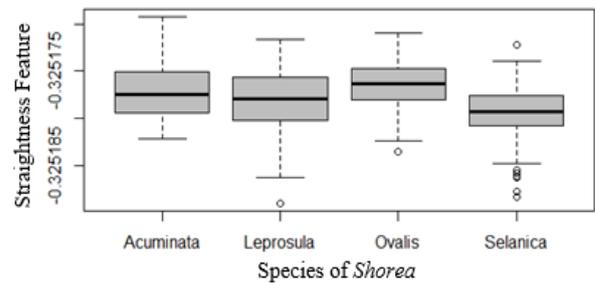
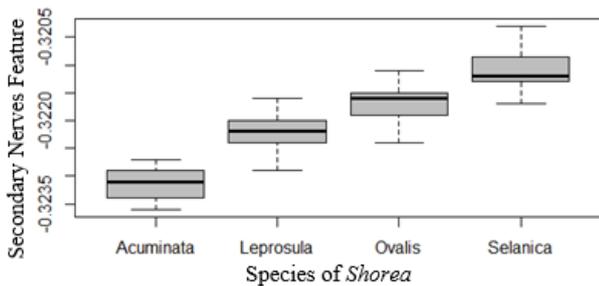


Figure 11. Boxplot secondary nerves and straightness feature

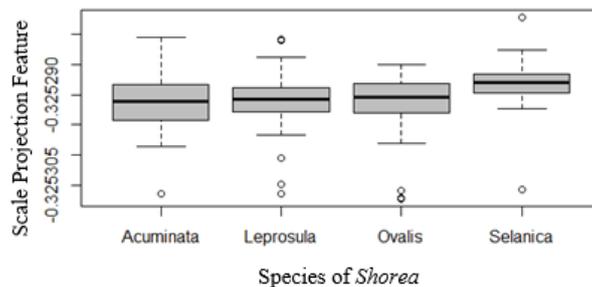
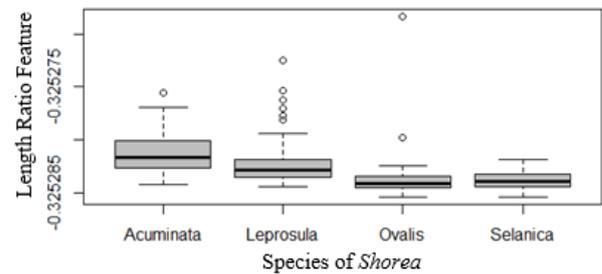
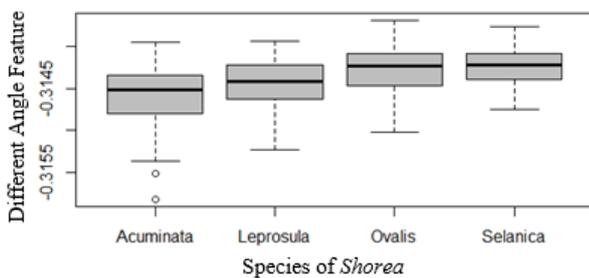


Figure 12. Boxplot different angle, length ratio and scale projection feature

In conclusion, the geometric morphometric approach could successfully extract the values of the features of straightness, different angle, length ratio, scale projection, and secondary nerves. From the five features mentioned, the secondary nerve feature is the best feature because it could clearly distinguish four species. On the other hand, the use of length ratio feature is relatively good compared to the features of different angle and straightness because it can distinguish three species of *Shorea* clearly, while different angle and straightness can only distinguish two *Shorea* species clearly. The classification results using support vector machine (SVM) to identify *Shorea* species based on leaf venation obtained the highest accuracy rate of 84.46%.

ACKNOWLEDGEMENTS

This research is supported by Department of Computer Science, Faculty of Mathematics and Natural Sciences and Department of Silviculture, Faculty of Forestry, IPB University, Bogor, Indonesia.

REFERENCES

- Al-amri SS, Kalyankar NV, Khamitkar SD. 2010. Image segmentation by using edge detection. *Intl J Comput Sci Eng* 2 (3): 804-807.
- Ambarwari A, Herdiyeni Y, Hermadi I. 2018. Biometric analysis of leaf venation density based on digital image. *Telkomnika (Telecommunication Comput Electron Control)* 16 (4): 1735.
- Bastoni. 2005. Kajian ekologi dan silvikultur ramin di Sumatera Selatan dan Jambi. Pusat Penelitian dan Pengembangan Hutan dan Konservasi Alam, ITTO dan PPD, Bogor. [Indonesian]
- Bonhomme V, Forster E, Wallace M, Stillman E, Charles M, Jones G. 2017. Identification of inter- and intra-species variation in cereal grains through geometric morphometric analysis, and its resilience under experimental charring. *J Archaeol Sci* 86: 60-67.
- Bourgeon O, Pagnoux C, Mauné S, Vargas EG, Ivorra S, Bonhomme V, Ater M, Moukhli A, Terral JF. 2017. Olive tree varieties cultivated for the great Baetican oil trade between the 1st and the 4th centuries ad: morphometric analysis of olive stones from Las Delicias (Ecija, Province of Seville, Spain). *Veg Hist Archaeobot* 27 (3): 463-476.
- Caillon F, Bonhomme V, Möllmann C, Frelat R. 2018. A morphometric dive into fish diversity. *Ecosphere* 9 (5): e02220. 10.1002/ecs2.2220.
- Cao CP, Gailing O, Siregar I, Indrioko S, Finkeldey R. 2006. Genetic variation at AFLPs for the Dipterocarpaceae and its relation to molecular phylogenies and taxonomic subdivisions. *J Plant Res* 119 (5): 553-558.
- Chaiphongpachara T, Tubsamut P. 2019. Geometric morphometry of pupae to identify four medically important flies (Order: Diptera) in Thailand. *Biodiversitas* 20 (6): 1504-1509.
- Chitwood DH, Otoni WC. 2017. Morphometric analysis of Passiflora leaves: The relationship between landmarks of the vasculature and elliptical Fourier descriptors of the blade. *Gigascience* 6 (1): 1-13.
- Chuang R, Bonhomme V. 2019. Rethinking the dental morphological differences between domestic equids. *J Archaeol Sci* 101: 140-148.
- Dodo, Wawangningrum H. 2018. Metode penyimpanan cabutan anakan pohon untuk konservasi ex-situ: Beraja (*Shorea guiso* (Blanco) Blume). Caring wildlings methods for ex situ conservation: red balau (*Shorea guiso* (Blanco) Blume). *Pros Sem Nas Masy Biodiv Indon* 4 (2): 139-143. [Indonesian]
- Eni A, Dewantara I, Sisilia L. 2018. Identification of tengkawang (*Shorea* spp.) species as natural days of tenun ikat District Kapuas Hulu Regency West Kalimantan. *J Hutan Lestari* 6 (1): 7-15. [Indonesian]
- Le T-L, Tran D-T, Hoang V-N. 2014. Fully automatic leaf-based plant identification, application for Vietnamese medicinal plant search. *Proceedings of the Fifth Symposium on Information and Communication Technology*. Hanoi, 04-05 December 2014. [Vietnam]
- Li W, Liu Z. 2011. A method of SVM with normalization in intrusion detection. *Procedia Environ Sci* 11 (PART A): 256-262.
- Newman M, Burgess P, Whitmore T. 1996. *Manuals of Dipterocarps for foresters: Sumatra light hardwoods*. Prosea Indonesia, Bogor.
- Newman M, Burgess P, Whitmore T. 1999. *Pedoman identifikasi pohon-pohon Dipterocarpaceae Jawa sampai Nugini*. Prosea Indonesia, Bogor. [Indonesian]
- Parikh KS, Shah TP. 2016. Support vector machine - A large margin classifier to diagnose skin illnesses. *Procedia Technol* 23: 369-375.
- Plotze RDO, Bruno OM. 2009. Automatic leaf structure biometry: computer vision techniques and their applications in plant taxonomy. *Intl J Pattern Recognit Artif Intell* 23 (02): 247-262.
- Rahmadani, Herdiyeni Y. 2010. Shape and vein extraction on plant leaf images using Fourier and B-Spline modeling. *Proceedings of AFITA 2010 International Conference, The Quality Information for Competitive Agricultural Based Production System and Commerce*. Bogor Agricultural University, Bogor, 4-7 October 2010.
- Raven P, Event R, Eichhorn S. 2005. *Biology of Plants*, 7th ed. W.H. Freeman, New York.
- Roth-Nebelsick A, Uhl D, Mosbrugger V, Kerp H. 2001. Evolution and function of leaf venation architecture: A review. *Ann Bot* 87 (5): 553-566.
- Salima A, Herdiyeni Y, Douady S. 2015. Leaf vein segmentation of medicinal plant using Hessian matrix. *Proceeding of international conference on advanced computer science and information systems*. Universitas Indonesia, Depok. 10-11 October 2015. [Indonesian]
- Thirumalai C, Vignesh M, Balaji R. 2017. Data analysis using box and whisker plot for lung cancer. *Proceeding of Innovations in Power and Advanced Computing Technologies*. VIT University, Vellore, India. 21-22 April 2017.
- Viscosi V, Cardini A. 2011. Leaf morphology, taxonomy and geometric morphometrics: A simplified protocol for beginners. *PLoS ONE* 6 (10): e25630. DOI: 10.1371/journal.pone.0025630 10..
- Wahyumianto A, Purnama IKE, Christyowidiasmoro. 2011. Identifikasi Tumbuhan Berdasarkan Minutiae Tulang Daun Menggunakan SOM Kohonen. [Thesis]. Institut Teknologi Sepuluh Nopember, Surabaya. [Indonesian]
- Zhang TY, Suen CY. 1984. A modified fast parallel algorithm for thinning digital patterns. *Commun ACM* 27 (3): 236-239.