LignoIndo: image database of Indonesian commercial timber

R Damayanti¹, E Prakasa², Krisdianto¹, L M Dewi¹, R Wardoyo², B Sugiar², H F Pardede², Y Riyanto², V F Astutiputri¹, G R Panjaitan³, M L Hadiwidjaja³, Y H Maulana³ and I N Mutaqin⁴

¹Forest Product Research and Development Center, Research, Development and Innovation Agency, Ministry of Environment and Forestry, Jl. Gunung Batu No. 5 Bogor 16610 Indonesia
²Research Center for Informatics, Indonesian Institute of Sciences (LIPI), Jl. Cisitu No. 21/154D Bandung, Indonesia
³STT TELKOM, Bandung, Indonesia
⁴Universitas Islam Nusantara, Bandung, Indonesia

E-mail: ratih_turmuzi@yahoo.com

Abstract. Macroscopic image database of Indonesian commercial timber “LignoIndo” has been developed to support the development of a computer vision-based wood identification system. A total of 809 wood species were taken their macroscopic images by using digital loupe at two magnifications: 50 times and 200 times. All of these timber species are Indonesian commercial timber listed in LIU (License Information Unit) in November 2016. Cross section images of 186 wood species listed in the Ministry of Forestry Degree No 163/2003 were also collected by using mobile phone camera that was attached with 60x magnify lens, as the basis data to the mobile application of automatic wood identification tool (AIKO). The number of wood species in this database will continue to grow as the number of commercial wood species increases. Besides, the number of replicates of specimens for each species of wood can also continue to be added to accommodate wood variations in the same species, the same species from different trees, and different geographic locations. This paper presents data collection methods and summaries of wood data that have been collected including trade name, botanical name, family, durability classes, strength classes, commercial class classification, conservation status, and recommended utilization.

1. Introduction
Development of database is required for various uses and purposes. A database is any collection of related data. It is a persistent, logically coherent collection of inherently meaningful data, which is relevant to some aspects in the real world. In the database management system, there is a collection of programs that enable users to create and maintain the database. The database management systems should provide functions in addition to simple file management such as allowing concurrency, controlling the data security, maintaining data integrity, providing back up and recovery, controlling redundancy, allowing data independency, providing non-procedural query language, and performing automatic query optimization [1].

Various database has been created to supply data and information in the specific field such as data set of images for acute lymphoblastic leukemia database, which provides images for leukemia interpretation diagnose [2]. The examples of datasets that have been used for the development of science
and commercial purposes are UCID: Uncompressed Color Image Dataset, FERET: Face Recognition Technology, ImageNet: A Large-Scale Hierarchical Image Database, and PKLot - A Robust Database for Parking Lot Classification. The dataset can be shared in a limited way by registering online. One of the examples is https://www.kaggle.com/c/diabetic-retinopathy-detection.

In the health area, a dataset that has been applied is ALL-IDB: The acute lymphoblastic leukemia image database [3]. The dataset that is maintained and continually developed to support image processing and analysis, as well as machine vision in various fields, is the USC-SIPI image database [4]. This database contains various colored data and monochrome images to the texture of objects such as bark, water, wood fiber shapes, trees, and tank cars. The users have to follow a series of procedures and payments to use the image.

One of the datasets that are important to be developed in wood identification is the dataset of both macroscopic and microscopic image of wood species. Wood identification is a process of determining wood species based on the anatomical properties. Identifying wood species is important not only to determine the utilization of wood for industrial purposes but also to support bioforensic activities in handling legal cases where wood is used as evidence.

At present, the identification of wood can be undertaken macroscopically and microscopically using thinned sections of wood in the Laboratory of Lignocellulose Anatomy, Forest Products Research and Development Center (P3HH), as the only laboratory that has been accredited by the National Accreditation Committee (KAN), with reference to Xylarium Bogoriense 1915 authentic wood collections. The authentic wood collections of Xylarium Bogoriense 1915 have been collected from most regions in Indonesia since 1914. Nowadays, the Xylarium Bogoriense 1915 wood collection has become the first largest Xylarium in the world with a collection of more than 190,000 specimens consisting of 110 families, 785 genera, and 3,667 authentic wood species.

The limitation of existing science, technology, and human resource capacity has caused experienced and trained researchers or officers can only undertake the process of wood identification. The experience and ability of personnel greatly influence the result and duration of wood identification. Misidentification of wood can cause financial losses for country or company so that there is no tolerance to make a mistake in identifying wood [5]. The possibility of error identification can indeed be reduced if the availability of expert is fulfilled. However, providing experts in wood identification with recognized capabilities requires a very long base of knowledge and training. A method that can be considered to facilitate the identification of wood species in the field and industry is to develop a system of wood identification using computer vision. The experience and knowledge of wood identification experts are used as reference material in the development of a system so that this system is expected to have similar ability to the experts in identifying wood species.

The development of this dataset has an important role in the development of a mobile application for automatic wood identification and cloud-based database which can be widely utilized. The application stored in the cloud can be utilized for the training process of wood identification system by using the database as media storage for wood images. A recent study on the classification of Indonesian commercial wood species shows that from about 4,000 of total wood species in Indonesia [6], approximately 1,044 wood species have been traded. About 809 wood species have already been identified for their properties so that they can be classified based on their quality [7].

This paper presents the method of data collection and summary of supporting data such as trade name, botanical name, strength class, durability class, commercial classification, and the utilization of wood species for development an Automatic Wood Identification System based on computer vision and as the basis of Indonesian commercial timber classification. In the future, the dataset can integrate wood species data from many regions in Indonesia which can be used as a guideline for collecting data and mapping the distribution of wood species in Indonesia.

2. Materials and methods

2.1. Materials

809 wood collections of Xylarium Bogoriense 1915 of Forest Products Research and Development Center (FPRDC), Research, Development and Innovation Agency, Ministry of Environment and
Forestry Republic of Indonesia were used in this research. Examples of wood specimen are presented in figure 1. All of these timber species are Indonesian commercial timber listed in the LIU (License Information Unit) in November 2016.

Figure 1. Prepared wood samples.

2.2. Methods
The period of data collection was divided into two batches depending on the image capturing device. Firstly, Dynolite® and laptop were used for capturing the macroscopic wood surface (figure 2). Digital zoom magnifications were 50 times and 200 times. The LignoIndo database obtained by using Dynolite was applied for the classification of Indonesian commercial timber and developing WoodId, an automatic wood identification system. The second batch was collecting wood cross section images by using a mobile phone camera. To obtain more details structure of wood, we attached 60x magnification lens. Some lens were tried to find the best pictures.

Figure 2. Images capturing process by using a portable digital microscope.

2.2.1. Sample preparation. The woods were prepared in accordance with Indonesian National Standard SNI 8491-2018 [8] as follows:

1. The wood observation fields were specified, in which the sections were used to observe the wood anatomy. Wood sections are cross section, radial and tangential.
2. Cross section was used for macroscopic photo taking in this study. Cross section is the surface perpendicular to the axis of the stem of the tree, which is characterized by the presence of pores like and parallel lines. This section is the “face” of wood.
3. The observation field was made by flat cutting off the wood using a cutter or other sharp tools. The cutting process can be done several times until the desired field is obtained. The cutting should be done with one movement.

The macroscopic anatomy structure of wood cross section was taken using digital loupe in focus of 50 times and 200 times. In the application, the magnification of 50 times is used to obtain a wide field of observation for large timber such as logs or sawn timbers, while the magnification of 200 times is
used if the available sample for identification is too small (less ideal) such as plywood or other products which are not possible to be widely sliced.

2.2.2. Coding in database. The naming for each picture in the Lignoindo database was:

\[ V_{1000} \text{ _Litsea robusta }_1 \text{ _050} \]

Remarks:
V: Commercial timber classification classes according to Djarwanto et al. (2017) namely I: Commercial decorative Class I; II: Commercial decorative Class II; III: Commercial Class I; IV: Commercial decorative Class II; and V: Commercial decorative Class III)
1000: Wood collection code (ID number) in Xylarium Bogoriense 1915; the use of this unique code is intended if there are additional data of the same species but come from different geographical locations.
\text{Litsea robusta}: Botanical name
1: First repetition (There were three repetitions: 1, 2 and 3 which represents three different capturing photo locations in cross section of every sample with a magnification of 50 times; and three repetitions: 4, 5 and 6 which represents three different capturing photo locations with a magnification of 200 times)
050: Code that shows magnification of digital loupe (50 times); and ‘200’ is used for the magnification of 200 times.

Pictures equipped with unique coding were saved in a folder named commercial classification classes and wood collection number, like the example above, \( V_{1000} \).

For batch 2, the images which meet the criteria set were stored according to the commercial classification classes in a folder of wood ID number, for examples \( Carallia borneensis \ 10107_\text{Indah 1} \). \( C. \text{ borneensis} \) is species name, 10107 is wood ID number, and Indah 1 means that the timber belongs to Commercial Decorative Class I classification. A hundred images from many locations, and some positions of the lens were taken for each wood specimen, totally 80,900 were obtained.

2.2.3. Data analysis. Supporting data of each species covering trade name, botanical name, family, durability classes, strength classes, commercial class classification, conservation status, and recommended utilizations were added to complete the ‘LignoIndo’ database. Publications (books and journals) were used as sources of the data [9–17]. Furthermore, the data then were stored in ‘cloud’ so it can be accessed by the public in limited use. Registration is required for each data usage. This data became the basis of the development of automatic wood identification system by using computer vision, WoodId and AIKO [18, 19]. The prototype of WoodId and AIKO is presented in figure 3 and figure 4.

![Figure 3. Prototype of WoodId.](image_url)
3. Results and discussion

The activity of taking image data for the development of an automatic wood identification system was initially carried out on six species of wood namely durian burung (*Durio carinatus* Masters), eboni beragis (*Diospyros celebica* Bakh), jati (*Tectona grandis* Linn. f), patin (*Musaendopsis beccariana* Baill.), pelawan (*Tristania ferruginea* (C.T.White) P.G.W.,) and ulin (*Eusideroxylon zwageri* Teijsm. & Binnend.). These species were taken randomly from Commercial Decorative Class I. This activity was carried out at the Lignocellulose Anatomy Laboratory of FPRDC. From this activity, 48 images of the wood cross section have been successfully collected. The data consist of six species of wood in which each species is available from seven to nine images. At this stage, coding of images and number of replications of each species of wood has not been established. Data were obtained by using Dynolite digital loupe. Collected image data later became the beginning of a dataset that is indispensable in the development of wood identification images systems [18]. Examples of macroscopic images of the six wood species are presented in figure 5. Image size is 1280 x 1204 pixels. The arrangement of wood image dataset can be seen in figure 6.

![Figure 4. Prototype of AIKO, mobile application of wood identification.](image-url)

![Figure 5. Top row: Durian burung, Eboni beragis, Jati; Bottom row: Patin, Pelawan, Ulin.](image-url)
Further data collection activities were then carried out according to the procedures described above. This procedure can then be applied in the addition of the dataset. Macroscopic photo data collection of wood cross sections of 809 species with 4,854 images was carried out for 71 days from 7 July 2017 to 29 September 2017 and continued from 1-27 November 2017. Examples of storage in folders for each wood species are presented in figure 7. Examples of wood species that have been supplemented with supporting information are presented in figure 8. Image data have been uploaded and stored to GRID LIPI (high-performance computing facility provided by LIPI).
Figure 8. Example of selected commercial timber completed with supporting information for the development of WoodId and AIKO in ‘LignoIndo’ database.

Data collection by using a mobile phone camera that was attached with a 60x magnification lens (the second batch) was conducted from October 2018 to April 2019. Experiments were conducted several times to obtain the best device in capturing wood anatomical images (figure 9 – figure 11). The images will not be accepted if the source light is not at maximum intensity for data collection; there is a significant blur area; the resulted image does not cover wood section; and the cutting surface is not flat. From the result, it was then decided to use Samsung S9 for further data collection.
**Figure 9.** Image capturing by using Samsung Note.

**Figure 10.** Image capturing by using Samsung S8+.

**Figure 11.** Image capturing by using Samsung S9 without additional magnification.
4. Conclusion

A total of 4,854 (using Dynolite) and 80,900 cross-section images (using a mobile phone and additional lens) have been collected from 809 species of commercial timber for the development of Indonesian commercial timber image datasets ‘LignoIndo’. The image data were also completed with supporting information, namely trade name, botanical name, family name, durability classes, strength classes, commercial class classification, conservation status, and recommended utilization. This data can continue to be added for another wood species and from the same species of different geographical location according to the development of Indonesian timber trade by following the established macroscopic photo data retrieval procedures.

Acknowledgment

This research is a collaborative work between Forest Products Research and Development Center (FORDA) and Research Center for Informatics, Indonesian Institute of Sciences (LIPI). The work was funded by the Ministry of Research, Technology and Higher Education under Research Grant INSINAS 2018. The collected wood images were stored and analyzed by using high-performance computing infrastructure, namely Grid LIPI. All authors are equally contributed to paper writing and data collection.

5. References

[1] R. J. Robbins, Database fundamentals John Hopkins University, Baltimore, Maryland, p. 31, 1995
[2] R. D. Labati, V. Piuri, and F. Scotti, ALL-IDB: The Acute Lymphoblastic Leukimia Database for Image Processing, The 18th IEEE International Conference on Image Processing, 2011, pp. 2045–2048
[3] R. D. Labati, V. Piuri, and F. Scotti, ALL-IDB: The acute lymphoblastic leukemia image database for image processing, Thirteenth IEEE International Conference on Image Processing, 2011
[4] G. A. Weber, USC-SIPI Report #244 The USC-SIPI Images Database: Version 4, Los Angeles, 1993
[5] R. Damayanti and S. Rullyati, The first in the world: Automatic Wood Detector as a Solution for Preventing Illegal Logging (in Indonesian), War. Has. Hutan, 7(1), 13–14, 2012
[6] I. Kartasujana and A. Martawijaya, Indonesia's Commercial Wood: Characteristics and Its Usage Announcement No. 3 Year 1973 and No. 56 Year 1975 (in Indonesian) Bogor: Lembaga Penelitian Hasil Hutan. Badan Penelitian dan Pengembangan Departemen Pertanian Republik Indonesia, 1979
[7] D. Djarwanto, R Damayanti, J Balfas, E Basri, Jasni, I M Sulastiningsih, Andianto, D Martono, G Pari, A Sopandi, Mardiansyah and Krisdianto *Grouping of Commercial Wood Indonesia* (in Indonesian) Bogor: Forda Press, 2017

[8] R. Damayanti, Mu’min, Andianto, and N. Herlina, *Macroscopic wood identification* (in Indonesian) Jakarta, Indonesia: Badan Standardisasi Nasional, 2018, p. 31

[9] Y. I. Mandang and I. K. N. Pandit, *Guidelines for Wood Type Identification in Field* (in Indonesian) Bogor: Yayasan Prosea, Bogor and Pusat Diklat Pegawai SDM Kehutanan, 2002

[10] D. S. Oey, *Specific gravity of Indonesian woods and its significance for practical use* (in Indonesian). Bogor: Pusat Penelitian dan Pengembangan Hasil Hutan, 1964

[11] A. Martawijaya, I. Kartasujana, K. Kadir, and S. A. Prawira, *Indonesian Wood Atlas* (in Indonesian), vol. I Bogor: Badan Penelitian dan Pengembangan Kehutanan, 2005

[12] A. Martawijaya, I. Kartasujana, Y. I. Mandang, S. A. Prawira, and K. Kadir, *Wood Indonesian Atlas Vol. II* (in Indonesian) Bogor: Badan Penelitian dan Pengembangan Kehutanan, 2005

[13] S. Abdurrohim, Y. Mandang, and U. Sutisna, *Indonesian Wood Atlas Vol III* (in Indonesian). Bogor: Pusat Penelitian dan Pengembangan Hasil Hutan, 2006

[14] M. Muslich, M. Wardani, T. Kalima, S. Rulliaty, R. Damayanti, N. Hadjib, G. Pari, S. Suprapti, M.I. Iskandar, Abdurachman, E. Basri, I. Heriansyah and H. L. Tata *Indonesian Wood Atlas Vol IV* (in Indonesian) Bogor: Pusat Penelitian dan Pengembangan Keteknikan Kehutanan dan Pengolahan Hasil Hutan, 2013

[15] I. Soerianegara and R. Lemmens, *Plant Resources of South-East Asia 5 (1). Timber trees: Major commercial species*. Leiden: Backhuys Publisher, 1994

[16] R. Lemmens, I. Soerianegara, and W. Wong, *Plant Resources of South-East Asia 5 (2). Timber trees: Minor commercial species*. Leiden: Backhuys Publisher, 1995

[17] M. Sosef, L. Hong, and S. Prawirohatmodjo, *Plant Resources of South-East Asia 5 (3). Timber trees: Lesser-Known species*. Leiden: Backhuys Publisher, 2008

[18] B. Sugianto, E. Prakasa, R. Wardoyo, R. K. Damayanti, H. F. Dewi, Listya Mustika Pardede, and Y. Rianto, Wood identification based on histogram of oriented gradient (HOG) feature and support vector machine (SVM) classifier, in 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2018, pp. 336–340

[19] M. Hadiwidjaja et al., Developing wood identification system by local binary pattern and hough transform method, *2nd International Conference on Data and Information Science (ICoDIS) Proceeding*