Klasifikasi Kayu Japanese Fagaceae menggunakan Area Sampel Parsial dan Convolutional Neural Networks

Tugas Akhir
diajukan untuk memenuhi salah satu syarat
memperoleh gelar sarjana
dari Program Studi Informatika
Fakultas Informatika
Universitas Telkom

1301160790
Taufik Fathurahman

UNIVERSITAS
Telkom

Program Studi Sarjana Informatika
Fakultas Informatika
Universitas Telkom
Bandung

2020

LEMBAR PENGESAHAN

Klasifikasi Kayu Japanese Fagaceae menggunakan Area Sampel Parsial dan Convolutional Neural Networks

Wood Classification of Japanese Fagaceae using Partial Sample Area and Convolutional Neural Networks

NIM: 1301160790

Taufik Fathurahman

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh gelar pada Program Studi Sarjana Informatika
Fakultas Informatika
Universitas Telkom

Bandung, 20 Juni 2020
Menyetujui

Pembimbing I	Pembimbing II
Dr. Putu Harry Gunawan, S.Si., M.Si., M.Sc.	Dr. Esa Prakasa, MT.
NIP: 16860043	NIP: 19760322 200502 1003

Ketua Program Studi
Sarjana Informatika,

Niken Dwi Wahyu Cahyani, S.T., M.Kom.
NIP: 00750052

LEMBAR PERNYATAAN

Dengan ini saya, Taufik Fathurahman, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul " **Klasifikasi Kayu Japanese Fagaceae menggunakan Area Sampel Parsial dan Convolutional Neural Networks** " beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang belaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika dikemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya.

Bandung, 20 Juni 2020

Yang Menyatakan,

Taufik Fathurahman

Wood Classification of Japanese Fagaceae using Partial Sample Area and Convolutional Neural Networks

Taufik Fathurahman[1,3], P.H.Gunawan[2], Esa Prakasa[3], and Junji Sugiyama[4,5]
[1,2]School of Computing, Telkom University, Bandung, Indonesia
[3]Computer Vision Research Group, Research Center for Informatics,
Indonesian Institute of Sciences, Bandung, Indonesia
[4]Division of Forestry and Biomaterials Science Faculty / Graduate School of Agriculture,
Kyoto University Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan
[5]College of Materials Science and Engineering, Nanjing Forestry University, Nanjing 210037, Jiangsu, China
Email: [1]taufikfathurahman@student.telkomuniversity.ac.id, [2]phgunawan@telkomuniversity.ac.id,
[3]esa.prakasa@lipi.go.id, [4]sugiyama.junji.6m@kyoto-u.ac.jp

***Abstract*—Wood identification is generally performed manually by observing the anatomy of wood such as colour, texture, fibre direction, and other characteristics. The manual process could take quite a long time, especially when identification work is required at high quantity. By considering this condition, the convolutional neural networks (CNN) based program is applied to improve the image classification results. The research focuses on the algorithm accuracy and efficiency , and the dataset limitations handling. Therefore, it is proposed to do the sample selection process or only take a small portion of the existing image. Still, it can be expected to represent the overall picture to maintain and improve the generalization capabilities of the CNN method in the classification stages. The experiments yielded incredible accuracy of up to 100% for large sample area sizes (300 x 300%) on each CNN architecture (VGG16, ResNet50, MobileNet, DenseNet, and Xception based). Whereas Xception-based architecture becomes an architecture that can maintain its model generalization for each sample area size (100, 200, and 300 pixels). The experimental results show that the proposed algorithm can be an accurate and reliable solution.**

***Index Terms*—Wood, microscopic image, sample selection, classification, convolutional neural network**

I. Introduction

Wood is the most dominant forest product for commercial use, used in various industries as raw material and supporting the daily life of humans, such as building materials, furniture, craft arts, and many others. Indonesia has about 4000 species of trees that could potentially be used as building timber [1]. Wood species are very diverse, but still have unique characteristics that can be distinguished from one another. Therefore, to realize this research, an accurate and effective method of identifying wood will be proposed.

Several previous studies related to wood identification either in macroscopic and microscopic levels are presented in the papers of [2], [3], [4], and [5]. The four papers above identify wood using different methods with their advantages and disadvantages. In the paper [2], identify wood by applying existing methods within the scope of computer vision, namely the Daubechies Wavelet (DW) and Local Binary Pattern (LBP) method to extract wood patterns which are then classified using the Support Vector Machine (SVM) method. In paper [3], the wood identification process still applies the same classification method as previously, namely SVM method, but uses the Histogram of Oriented Gradient (HOG) as a feature extractor. Furthermore, in the paper [4] using the LBP and Hough Transform methods to improve the wood identification system. Lastly, in the paper [5] Identify wood with a method that is more or less the same as the method used in the paper [2], that is DW and LBP methods to extract wood patterns which are then classified using SVM method.

The classification has become one of the main topics that has attached much attention in recent years. Many methods have been developed for classification purposes. Now days, CNN emerged as a powerful visual model with outstanding performance in various visual recognition and classification problems. For example, as presented in the papers of [6], and [7]. Shiqi Yu et al. [6] has developed an efficient CNN architecture to boost hyperspectral image classification discriminative capability, with an outstanding result in HSI classification on three popular datasets. Whereas Gil Levi and Tal Hassner [7] has proposed a simple convolutional net architecture that can be used even when the amount of learning data is limited. Their method was evaluated with the Adience benchmark for age and gender estimation and successfully outperformed the other current methods. The CNN method can provide excellent classification results, but has a challenge in the number of data sets needed to make the classifier well trained [6], [7], [8], [9].

In this research, we attempt to identify the microscopic wood image. By following the successful example in the paper discussed earlier, the CNN method is one of the potential approaches to solve the microscopic wood image classification. To overcome the limitations of the dataset, and the efficiency, a sample selection process will be proposed before the wood microscopic image dataset enters the classification stage using the CNN method. Inside the sample selection process, the image of wood will be cropped into several sections with a certain size. We assume that even though only used a certain segment of the image, it has characteristics that can more or less represent the overall picture, such as vessel size, the density of vessels, color, transverse wood fiber, and other characteristics.

The remaining part of the paper is organized as follows. The data sets used in the research, the sample selection process, the proposed CNN architectures, and the proposed algorithm are presented in Section II. Experimental results for sample selection, training, and testing process are presented in Section III. The last section, Section IV, conclusion, and future work.

II. Research Methodology

A. Wood Dataset

The wood species is usually recognized by examining a small piece of wood as a sample. There are two main characteristics of wood that can be used to recognize wood species, namely physical (macroscopic) and structural (microscopic) properties [10]. The physical properties (macroscopic) are properties that can be identified directly without using tools. Whereas the structural properties (microscopic) are properties that can only be identified by using tools, such as a loupe [10]. The structural (microscopic) commonly observed are vessels, parenchyma, rays, and others [11] [12].

This research uses a wood dataset obtained from the Research Institute of Sustainable Humanosphere (RISH), Kyoto University, Japan. The dataset consists of microscopic images of nine species of Fagaceae woods. The list of wood are displayed in TABLE. I.

The dataset will consist of nine species of wood. Each image in the dataset is stored as a TIFF file with 4140x3096 pixels dimension. The existing dataset will be divided into three groups, namely data train, data validation, and test data. Data test will consist of 27 images (3 images from each species), data validation will consist of 18 images (2 images from each species), and data train will consist of 120 images (is the rest of images).

B. Sample Selection

The proposed method uses a small segment of the image as an input. This procedure is defined by considering the size of the original image and limited dataset availability. The image has a dimension of 4140x3096 pixels, with a size of around 38 MB. Implementation of an image processing algorithm on an image with a large size could cost the computer resources and computational time. Moreover, it is known from the previous section that the number of datasets is very limited, this can result in overfitting and poor generalization capabilities of the CNN model.

The process of taking certain segments of the input image will consist of 3 categories based on the sample area size. The first category is small, that will crop the image with a size of 100x100 pixels and produce 1230 sample area for each image.

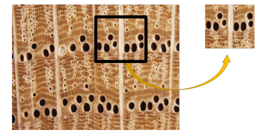

Fig. 1: The illustration of the sample selection process on image.

The second is medium category, that will crop images with a size of 200x200 pixels and produce 300 sample area for each image. And the last is a large category, that will crop images with a size of 300x300 pixels and produce 130 sample areas for each image.

The features provided by the selected area sample as in Fig 1 are more specific and detailed. The CNN model that is trained on specific and detailed features, resulting in a CNN model that has excellent generational capabilities even though the dataset provided is limited.

C. Implementation of Convolutional Neural Networks

CNN is one of the deep learning models that is widely used for image and visual analysis [13]. CNN is the standard neural network architecture used for prediction when the input observations are images, which is the case in a wide range of neural network applications [14]. In principle, CNN mimics the visual cortex, this is based on studies conducted like Neocognitrons in 1980 [15], which gradually evolved into what we now call convolutional neural networks [16]. CNN is very similar to an ordinary neural networks. They still consist of neurons with weights that can be learned from data. Each neuron receives several inputs and performs a dot product. In general, there are three main layers in CNN, the convolution layer, the pooling layer, and the fully connected layer [17].

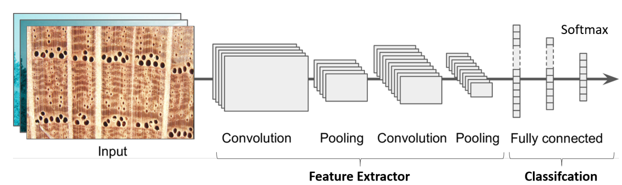

Fig. 2: Convolutional Neural Networks (CNN) architecture.

Fig. 2 shows the typical CNN architecture piles up several convolutional layers (followed by activators), then pooling layer, then several convolutional layers (followed by activators), then other pooling layers, and so on until the final layer produces predictions (for example, the softmax layer that outputs estimated class probabilities) [16].

In practice nowadays, very few people train entire convolutional networks from scratch. This is due to several reasons such as requiring a lot of data to train CNN, lots of parameters (weights) must be trained and avoid overfitting. The right solution to overcome this problem is to do transfer learning. Five CNN architectures has been well design by applying transfer learning to fit in the sample selected data set.

TABLE I: Nine species of Japanese Fagaceae

Japanese Common Name	Academic Name	Number of Individual Wood	Number of Image	Data and Kyoto University ID (KYOw number)				
Abemaki	Quercus variabilis	7	13	1620	6576	8332	8333	10328
				14129	17782			
Akagashi	Quercus accuta	10	19	342	1615	2867	2957	4920
				8312	9277	13837	14477	14679
Arakashi	Quercus glauca	10	15	5536	5664	9280	12846	12847
				12942	13743	15619	17789	18774
Buna	Fagus crenata	10	12	458	972	1116	1294	8305
				458	972	1116	1294	8305
Ichiigashi	Quercus gilva	10	26	2958	4973	5535	5663	8317
				9279	11593	13839	14829	18722
Inubuna	Fagus japonica	9	12	354	1613	5368	8306	8308
				13836	13955	17510	18594	
Kunugi	Quercus actissima	10	30	60	1120	1617	5540	5668
				8314	8315	12092	15353	17763
Kuri	Castanea crenata	10	20	10246	10293	10294	10997	10997
				11596	12944	13755	13760	13830
Mizunara	Quecus crispula	10	18	62	411	462	2963	8203
				8320	10297	11421	13841	

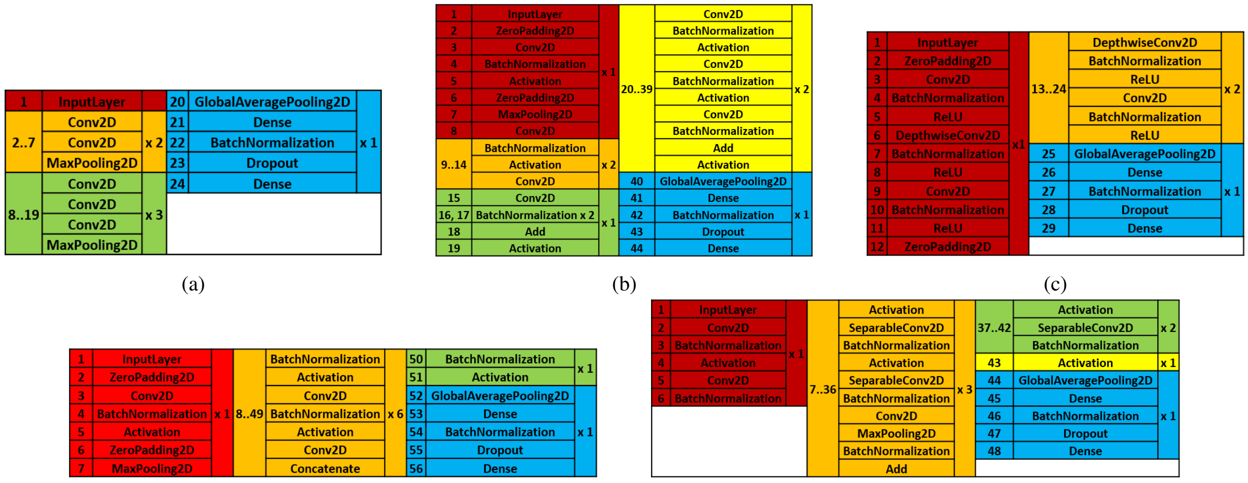

Fig. 3: Five architectures of CNN, VGG16 (a), ResNet50 (b), MobileNet (c), DenseNet121 (d), and Xception (e) based architecture. For each architecture, the first column is the layer number, the second column is the layer type, the third column is the number of blocks, respectively. Color differences indicate differences in architectural blocks, blue blocks indicate fully connected.

Fig. 3 illustrates architecture of each model. Each base architecture is cut to a certain network depth, such as VGG16 [18] at the 24th layer, ResNet50 [19] at the 39th layer, MobileNet [20] at the 23rd layer, DenseNet121 [21] at the 51st layer, and Xception [22] at the 42nd layer. Furthermore, a new fully-connected layer will be used to fit the nine species/class of wood. The use of shallow network depth aims to prevent the network from losing important features of the wood. As it is known that the deeper the network the features produced will be more detailed, this can affect the sample selected image whose features are sufficiently detailed to be lost.

D. Proposed Algorithm

To solve the problems in this research, the solution will be divided into the training process (Fig. 8), and the testing process (Fig. 5).

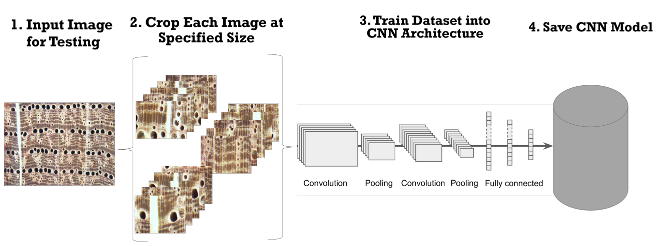

Fig. 4: The training process in getting the CNN model

The training process begins with a sample selection process as described in subsection B. Furthermore, the result of the sample selection process will be trained with the prepared architecture as in subsection C. The training process will produce 15 CNN models, which will then be tested in the testing process.

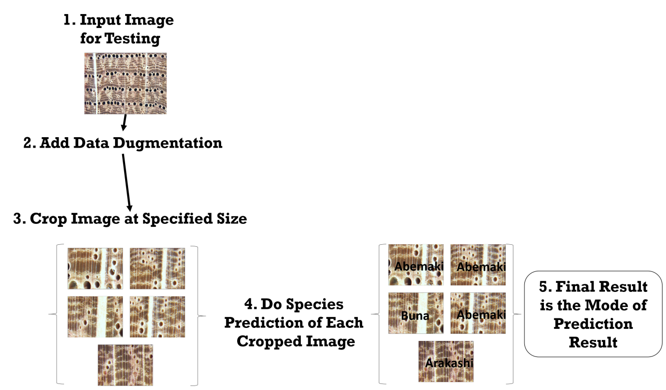

Fig. 5: The testing process in getting the classification result.

In the testing process, augmentation data will be added to the test sets. The addition of data augmentation aims to add data variation, and make testing process can describe a more general situation. Furthermore, test sets will enter the sample selection stage as in the training process. However, the difference is that not all sample areas for each image will be used. The optimal number of sample areas will be sought, which can produce high accuracy while still considering computational costs. For example, for one image a sample area of one size (100, 200, or 300) will be taken as many as 5 images. Furthermore, all sample area images will be classified by each CNN model. From all prediction results, the mode value will be searched to be the final prediction result for each individual image (the real image). An illustration of this process can be seen in Fig. 5.

III. Result and Discussion

A. Sample Selection Process Analysis

The first step is to process all wood data sets that will be used into the sample selection process. Three categories of sample areas (small, medium, and large) were created, with the sample results in Fig. 6.

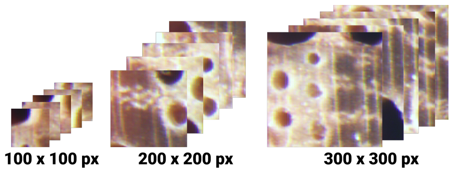

Fig. 6: Sample selection results for Abemaki species (small, medium, and large sample area).

The training process will use all sample areas from each original image (1230 sample areas for small size, 300 sample areas for medium size, and 130 sample areas for large size). Meanwhile, for the testing process, we will look for the most optimal number of sample areas with accuracy and computational cost. Ten trials using different numbers of data samples (1, 3, 5, 7, 9, 11, 13, 15, 21, 25) are presented in Fig. 7.

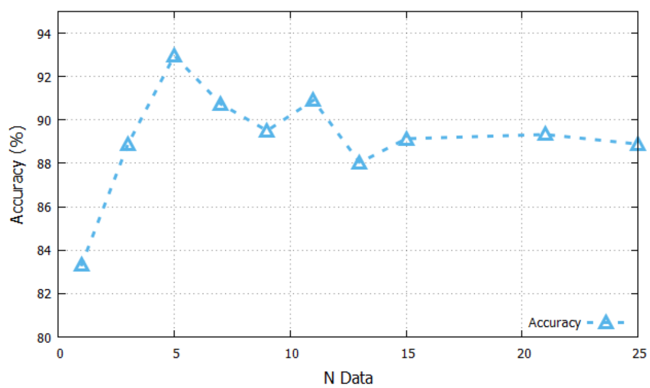

Fig. 7: The correlation between the number of test sets and the resulting accuracy

Large sample area sizes and VGG16 architecture are used in this process. The results of this process analysis will be used in the next process. With the most optimal result are achieved when the number of samples in the test set is 5. The next process is to evaluate 15 CNN models using the number of sample areas defined in this section.

B. Training Process Analysis

The training process for all CNN models uses the same parameters to produce valid benchmarks. Starting from the number of epoch (30 epoch), learning rate (0.0001), and the optimizer (Adam optimizer). If the training process has not increased in 10 epochs, the training process will be terminated without waiting for the maximum number of epochs. Furthermore, if in 5 epochs training process does not increase, the learning rate will be reduced by multiplying it by a factor value of 0.6. By lowering the learning rate, the process of updating the weight becomes smoother to increase accuracy. As a sample of the total 15 training processes, a history of training from the VGG16 architecture with a sample size area of 200 was chosen with the results shown in Fig. 8

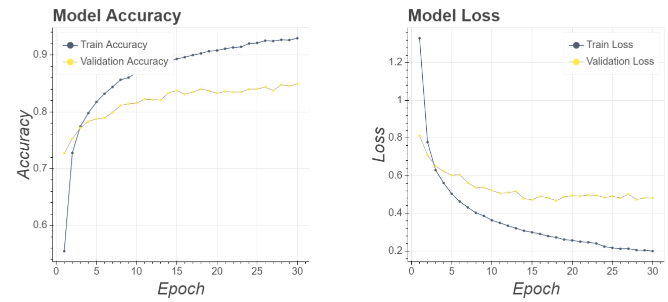

Fig. 8: Accuracy and Loss Curves During CNN Training

Fig. 8 shows the architecture that has been designed to be able to accept wooden datasets properly. From each epoch, score accuracy continues to increase, and score loss continues to decrease. Scores between the validation and the training are not far apart showing that the model does not overfit. Furthermore, to test the performance and generalization of the models, a testing process will be carried out on the wood testing dataset that has been prepared.

C. Testing Process Analysis

The testing process was done by evaluating each CNN models with the prepared testing data set. In detail, the results of the testing process are presented in Table. II.

TABLE II: CNN models performance evaluation

		VGG-16	**Res-Net50**	**Mobile-Net**	**Dense-Net121**	**Xcep-tion**
Small	*Precision*	0.92	0.93	0.87	0.95	0.94
	Recall	0.93	0.94	0.87	0.94	0.93
	F1	0.92	0.95	0.86	0.94	0.92
Medium	*Precision*	0.94	0.96	0.89	0.98	0.97
	Recall	0.93	0.94	0.87	0.98	0.96
	F1	0.92	0.94	0.87	0.98	0.96
Large	*Precision*	0.95	0.91	0.87	0.98	0.98
	Recall	0.94	0.89	0.85	0.98	0.98
	F1	0.94	0.89	0.84	0.98	0.98
Weight Size		60 Mb	4.5 Mb	1.3 Mb	5.9 Mb	30 Mb

Based on test results, in general, the accuracy score is increases following the size of the input image. With a sample area size = 100, the average F1 scores of VGG16, ResNet50, MobileNet, DenseNet121, and Xception are (92, 95, 86, 95, and 92%) respectively. With a sample area size = 200, the average F1 scores increase to (92, 94, 87, 98, and 96%) respectively. Finally for sampe area size = 300, the average F1 scores is increase for VGG16, DenseNet121, and Xception to (94, 98, and 98%) respectively. For the results obtained by ResNet and MobileNet are decreased to (89, and 84%).

Test result data in the Table. II provides a correlation between the size of the sample area and the capabilities of the resulting CNN model. The size of the sample area is related to the number of features contained in the image. If we look back at Fig. 6, the smaller the sample area, the surface features of the wood disappear due to the cutting process. Medium category area samples appear as the most optimal size in maintaining features contained in wood images. Medium area sample areas can create samples that contain the most detailed and specific wood features compared to 2 other categories. Proved by the average F1 score of the medium size for each CNN architecture is achieve 93.4%. Although the large category sample area is capable of more features, it is not more specific than the medium category sample area. Thus, the results from large area samples do not produce a more general CNN model compared to medium area samples.

Other information obtained from the Table. II is the DenseNet121-based model being the most general CNN architecture for all sample area sizes and considered to be the best CNN model for Japanese Fagaceae wood identification. DenseNet121-based produces F1 scores above 94% for every size of the sample area (100, 200, and 300 pixels). Meanwhile, MobileNet-based is the CNN architecture that provides the lowest accuracy. But it should be taken into account to use, with the given size of the resulting weight to be the lightest. Although accuracy is not as high as other architectures, it can be the most portable architecture option in use.

IV. Conclusion and Future Work

CNN has become a popular deep learning model that is widely used for image and visual analysis. In the applications, CNN requires a large number of datasets to learn a powerful classifier. In this research, the microscopic images taken from the laboratory are in a limited number and have a large size of around 38 Mb and a dimension of 4140x3096 pixels. Therefore, a well-designed algorithm has been proposed, which can handle dataset limitations and identify wood species accurately, and efficiently. The proposed algorithm principle is to implement the sample selection process before the dataset enters the classification stage using VGG16, ResNet50, MobileNet, DenseNet121, and Xception based architectures.

The experimental results show that the sample selection process can produce CNN models that have good generalization capabilities. The experimental results prove that by making sample areas more specific and detailed, it can produce a powerful CNN model. The medium area sample category (200 x 200 pixels) appears as the most optimal sample area. While the CNN DenseNet121 architecture is the most optimal architecture when used with medium sample areas.

To improve results to be more reliable in future use, more microscopic images will be needed. The results of this research are expected to be widely used. With a small size of the CNN model, its use will be even more flexible, as is mobile devices.

Acknowledgment

The author would like to thank the Indonesian Institute of Sciences (Lembaga Ilmu Pengetahuan Indonesia) and the Research Institute of Sustainable Humanosphere (RISH), Kyoto University, Japan for the assistance and provision of datasets that support this research.

References

[1] M. PLAN, "Sorting kayu menggunakan plc cj1m fuzzy logic design and implementation on wood sorting mini plant using plc cj1m."

[2] S. Salma, P. Gunawan, E. Prakasa, B. Sugiarto, R. Wardoyo, Y. Rianto, R. Damayanti, L. M. Dewi *et al.*, "Wood identification on microscopic image with daubechies wavelet method and local binary pattern," in *2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*. IEEE, 2018, pp. 23–27.

[3] M. L. Hadiwidjaja, P. Gunawan, E. Prakasa, Y. Rianto, B. Sugiarto, R. Wardoyo, R. Damayanti, K. Sugiyanto, L. M. Dewi, and V. F. Astutiputri, "Developing wood identification system by local binary pattern and hough transform method," in *Journal of Physics: Conference Series*, vol. 1192, no. 1. IOP Publishing, 2019, p. 012053.

[4] B. Sugiarto, E. Prakasa, R. Wardoyo, R. Damayanti, L. M. Dewi, H. F. Pardede, Y. Rianto *et al.*, "Wood identification based on histogram of oriented gradient (hog) feature and support vector machine (svm) classifier," in *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*. IEEE, 2017, pp. 337–341.

[5] S. Salma, P. Gunawan, E. Prakasa, R. Damayanti, J. Sugiyama *et al.*, "Classification of japanese fagaceae wood based on microscopic image analysis," in *2019 7th International Conference on Information and Communication Technology (ICoICT)*. IEEE, 2019, pp. 1–6.

[6] S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification," *Neurocomputing*, vol. 219, pp. 88–98, 2017.

[7] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 34–42.

[8] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 645–657, 2016.

[9] D. Marmanis, M. Datcu, T. Esch, and U. Stilla, "Deep learning earth observation classification using imagenet pretrained networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 105–109, 2015.

[10] B. Wijanarko, "Mengenal karakteristik kayu," *https://www.vedcmalang.com/pppptkboemlg/index.php/baru/42-bangunan/597-mengenal-karakteristik-kayu*, 2013.

[11] W. Schoch, I. Heller, F. H. Schweingruber, and F. Kienast, *Wood anatomy of central European Species*. Swiss Federal Institute for Forest, 2004.

[12] P. Prislan, J. Gričar, and K. Čufar, *Wood sample preparation for microscopic analysis*. University of Ljubljana, Department of Wood Science and Technology, 2014.

[13] R. Primartha, *Belajar Machine Learning Teori dan Praktik*. Informatika Bandung, 2018.

[14] S. Weidman, *Deep Learning from Scratch*. O'Reilly Media, Inc, 2019.

[15] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.

[16] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition*. O'Reilly Media, Inc, 2019.

[17] M. Sewak, M. R. Karim, and P. Pujari, *Practical Convolutional Neural Networks: Implement advanced deep learning models using Python*. Packt Publishing Ltd, 2018.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[20] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[21] G. Huang, Z. Liu, K. Weinberger, and L. van der Maaten, "Densely connected convolutional networks. arxiv 2017," *arXiv preprint arXiv:1608.06993*.

[22] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.