

Fine-Grained Wood Species Identification Using Convolutional Neural Networks

Shustrov Dmitrii, Eerola Tuomas, Lensu Lasse, Kälviäinen Heikki, Haario Heikki

This is a Post-print version of a publication
published by Springer, Cham

in Proceedings of the 21st Scandinavian Conference on Image Analysis (SCIA 2019). Lecture Notes in Computer Science.

DOI: 10.1007/978-3-030-20205-7_6

Copyright of the original publication: © Springer Nature Switzerland AG 2019

Please cite the publication as follows:

Shustrov, D., Eerola, T., Lensu, L., Kälviäinen, H., Haario, H., Fine-Grained Wood Species Identification Using Convolutional Neural Networks, Image Analysis, Springer Lecture Notes in Computer Science, LNCS Vol. 11482, pp. 67–77, 2019, Proceedings of the 21st Scandinavian Conference on Image Analysis (SCIA 2019), Norrköping, Sweden, 2019. DOI: https://doi.org/10.1007/978-3-030-20205-7_6

**This is a parallel published version of an original publication.
This version can differ from the original published article.**

Fine-Grained Wood Species Identification Using Convolutional Neural Networks

Dmitrii Shustrov¹, Tuomas Eerola¹[0000-0003-1352-0999], Lasse Lensu¹[0000-0002-7691-121X], Heikki Kälviäinen¹[0000-0002-0790-6847], and Heikki Haario²[0000-0002-0538-5697]

¹ Lappeenranta-Lahti University of Technology LUT, School of Engineering Science, Department of Computational and Process Engineering, Machine Vision and Pattern Recognition Laboratory, P.O.Box 20, FI-53851 Lappeenranta, Finland

`firstname.lastname@lut.fi`

² Lappeenranta-Lahti University of Technology LUT, School of Engineering Science, Department of Computational and Process Engineering, Inverse Problems Research Group, P.O.Box 20, FI-53851 Lappeenranta, Finland

`firstname.lastname@lut.fi`

Abstract. This paper considers the wood species identification from images of boards. The identification using only visual features of the surface is a challenging task even for an expert. The task becomes especially difficult when the wood species are from the same family. We propose a CNN based framework for the fine-grained classification of wood species. The framework includes a patch extraction procedure where board images are divided into image patches. Each patch is separately classified using the CNN resulting in multiple classification results per board. Finally, the patch classification results for a single board are combined. We evaluate various CNN architectures using the challenging data, consisting of species from the *Pinaceae* family. In addition, we propose three alternative decision rules for combining the patch classification results. By selecting a suitable amount of image patches, the proposed framework was able to achieve over 99% identification accuracy and real-time performance.

Keywords: wood species identification · convolutional neural networks · fine-grained classification · visual inspection · machine vision application.

1 Introduction

In a sawmill, it is important to be able to control that tree species do not get mixed in the sawing and quality grading process. This is due to the fact that proper and accurate wood processing, for example, drying or storage, often directly depends on the wood species. More importantly, supplying boards of incorrect wood species to customers results in reclamations and financial loss for the sawmill. In the modern highly automated sawmills with large production capacities, manual verification of every board is impossible. This calls for automated solutions for the tree species identification of wooden boards.

A typical sawmill specializes in a certain type of timber, for example, conifers. Because of this, the sawmill typically processes only species from the same wood family with very similar appearance and the identification method should be able to perform the fine-grained classification between them. Fig. 1 shows example board images of three species from the *Pinaceae* family: Pine (*Pinus*), Spruce (*Picea*), and Fir (*Abies*). It can be seen that the intra-species variation, for example, in color is larger than between the species, and designing image features that distinguish the species from each other is very challenging. Convolutional neural networks (CNNs), however, are able to learn the features from data making them an attractive approach for the task.

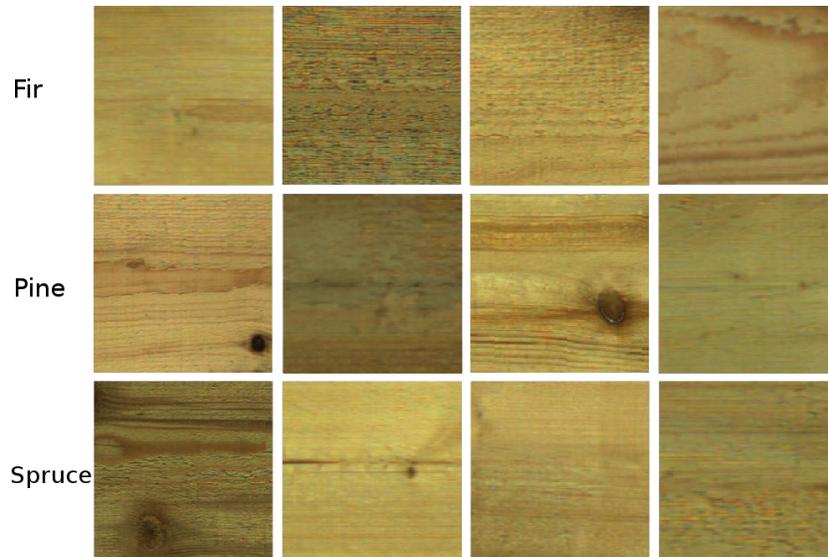


Fig. 1. Example images of the board surface images from the species considered in this study. Each row contains four image patches from one species.

In this paper, a framework for fine-grained classification between wood species is proposed. The framework starts with a patch extraction step where a high resolution image of a wooden board is divided into multiple image patches. Each image patch is fed into CNN which classifies the patch resulting in multiple classification results for each board. Finally, the classification results are combined using a selected decision rule to obtain the final identification result for the whole board. Extensive experiments are performed to select the suitable CNN architecture and the decision rule for the board identification.

2 Related work

With the development of visual inspection systems in the sawmill industry, automated systems have been actively used to identify species and defects in wooden material. For example, Tou et al. [12, 10] proposed a recognition and classification method for microscopic images. Gray-level co-occurrence matrix (GLCM) based features were used to train a multi-layer perceptron (MLP). As a result, the obtained accuracy for the classification of five wood species ranged from 60 to 72%. In [11], multiple features extraction methods were compared for the task and the best results were obtained using Gabor filter-based covariance matrix features. For the classification, k-NN was used. Khalid et al. [5] presented a classification system capable for recognizing 20 different species of wood. The system uses images of wood samples that were prepared by boiling and cutting them into thin sections using a microtome. For the feature extraction, GLCM was used and the classification was performed using MLP. The obtained recognition accuracy was 95%. Hafemann et al. [2] utilized a shallow CNN to solve wood species identification task. Using the database from [7], the recognition accuracy was 95% with a dataset consisting of macroscopic images of 41 wood species. Slightly higher accuracy was obtained for microscopic images.

The existing methods contain various drawbacks. The methods proposed in [12, 10, 11] do not provide accuracy high enough for the industrial environment. Also, the approach is based on microscopic images, so the image acquisition step may need sample preparation and calibration actions. The main drawback of the approach in [5] is the complexity and high cost of preparing the wood samples which makes this work inapplicable for practical applications. The method proposed in [2] obtained rather high accuracy on both macroscopic and microscopic images of the wood. However, it should be noted that the databases used for the experiment contained species that are clearly visually different from each other and the variance inside individual species was low.

3 Proposed method

The scheme of the proposed method is illustrated in Fig. 2. The images are acquired from the both faces of the board. To cover the whole board with reasonable accuracy, multiple slightly overlapping images are captured. Multiple image patches are extracted from each image. A single image patch represents a square image extracted from the board excluding the background. The subsequent identification of individual image patches is done using a CNN-based classifier. Finally, a decision rule is used to combine the image patch classification results into the final identification of the wood species of the whole board.

3.1 Image patches extraction

To extract image patches the board should be separated from the background. To do this, the image is first binarized using thresholding to separate bright

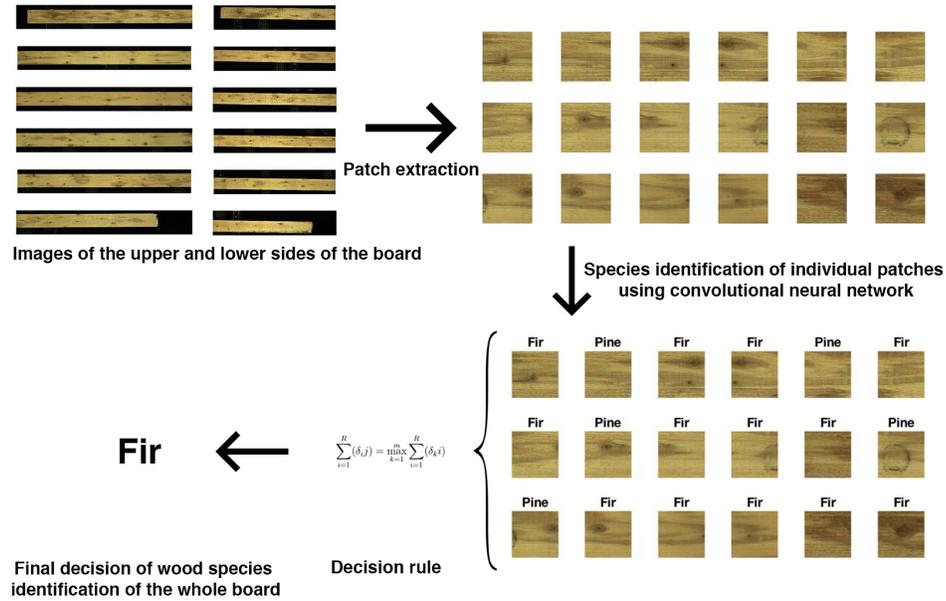


Fig. 2. Proposed method for wood species identification.

board from the dark background. The largest connected component is assumed to correspond to the board. The rotation of the board is computed and compensated by rotating the image, and a bounding box is drawn around the connected component. Finally, overlapping image patches are extracted from the bounding box.

3.2 Decision rule

The image patches are classified using CNN producing multiple identification results for each board. To be more specific, each patch classifier produces a class probability to each class that need to be combined to obtain the final species identification. For this purpose we utilize decisions rules. The following decision rules are considered: 1) Majority vote rule, 2) Mean probability rule, and 3) Max rule.

Let p_{ij} be the class probability for the wood species i obtained from the j th batch from a board Z . Then *Majority vote rule* is defined as follow. Assign Z to class label (species) \hat{C} as

$$\hat{C} = \arg \max_{i=1}^K \sum_{j=1}^N c_{ij}, \quad (1)$$

where

$$c_{ij} = \begin{cases} 1 & p_{ij} = \max p_{ij} \\ 0 & \text{otherwise} \end{cases}, \quad (2)$$

K is the number of species, and N is the number of image patches. In practice, each patch has one vote and the final classification is made based on the number of votes for each species.

Mean probability rule is defined as follows. Assign Z to class label \hat{C} as

$$\hat{C} = \arg \max_{i=1}^K \frac{1}{N} \sum_{j=1}^N p_{ij}. \quad (3)$$

Finally, *Max rule* is defined as follows. Assign Z to class label \hat{C} as

$$\hat{C} = \arg \max_{i=1}^K (\max_{j=1}^N p_{ij}). \quad (4)$$

4 Experiments and results

4.1 Data

The data consist of board images obtained during the sawmill process. The data were collected from several sawmills using the existing camera-based board sorting systems. Imaging systems including illumination varied between sawmills resulting in a challenging and realistic dataset. Each board was imaged with multiple cameras from both faces to capture the full surface of the board. Depending on the length of the board, the data contains 3 to 5 images with resolution varying from 2480×280 to 2480×480 per one face of a board. Fig. 3 shows an example image of a board.

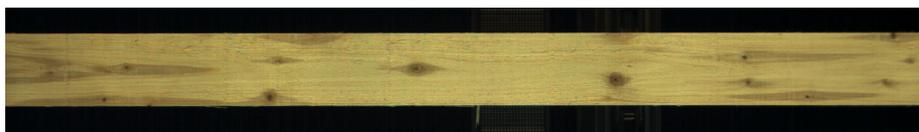


Fig. 3. Example of a board image.

The image patches were extracted from each image. Fig. 4 shows examples of extracted image patches. It should be noted that due to, for example, defects in boards and varying types of the background, some patches were not extracted accurately (see Fig. 5). These patches were included into the dataset to evaluate the robustness of the proposed method.

Table 1 shows the numbers of boards and extracted image patches for each wood species. The total number of all extracted image patches from all boards



Fig. 4. Example of high quality image patches.



Fig. 5. Example of inaccurately extracted image patches.

was 255724 where 242938 patches were extracted accurately. On the average, approximately 250 image patches were extracted from each board. The boards were randomly splitted into the training, validation, and testing sets, i.e., all patches from one board were allocated to one of the three sets. Table 2 shows the number of boards in each sets.

Table 1. Number of the boards and extracted image patches for each wood species.

| Wood species | Number of boards | Number of image patches |
|--------------|------------------|-------------------------|
| Fir | 455 | 104353 |
| Pine | 256 | 58713 |
| Spruce | 404 | 92656 |
| Total | 1115 | 255724 |

Table 2. Number of boards for the training, validation, and test sets.

| Wood species | Training set | Validation set | Test set |
|--------------|--------------|----------------|----------|
| Fir | 318 | 65 | 72 |
| Pine | 178 | 35 | 43 |
| Spruce | 282 | 59 | 62 |
| Total | 778 | 159 | 177 |

4.2 Experimental arrangements

The following CNN architectures were selected for the experiments: AlexNet [6], VGG-16 [8], GoogLeNet [9], and ResNet-50 [3]. Transfer learning with ImageNet pretrained models were used for the following convolutional neural networks:

VGG-16, GoogLeNet, and ResNet-50. AlexNet was trained from the scratch. The dropout rate was 0.5 for the CNN regularization to avoid the overfitting.

All experiments were performed using the Caffe deep learning framework [4] which utilizes the CUDA 9.0 deep learning library with CUDNN 7.1 [1]. All experiments were performed on the computer with the following computational resources: two NVIDIA GeForce GTX TITAN Black GPU, Intel Xeon CPU E5-2680, and 128 gygabytes of random access memory.

4.3 Results

The following performance measures were used: the classification accuracy (ACC) defined as the percentage of the correctly classified samples, and the F-measure defined as

$$F = \frac{2PPV \cdot TPR}{PPV + TPR} \quad (5)$$

where PPV is the positive predictive value and TPR is the true positive rate.

Table 3 shows the comparison between the CNN architectures for the task of the species identification from individual patches. Fig. 6 shows the corresponding confusion matrices. As it can be seen, the identification performance varies between the architectures. For VGG-16 there is a lot of misclassifications between pine and spruce when the true label is pine. Also fir and spruce are mixed when the true label is fir. Within spruce the wood species is predicted more accurately. ResNet-50 shows the opposite. The misclassifications occurred mainly in the spruce class confusing spruce with pine. AlexNet and GoogleNet show more accurate predictions without strong confusion in any combination of the classes. The shallower architectures appear to outperform the very deep ones. One reason for this could be the fact that high-level knowledge is not needed for the species identification and the method does not have to be invariant, for example, to view-point. Therefore, the task favors the shallower networks with less down-sampling operations that do not destroy the gradient information essential for the identification.

Table 3. Performance comparison of the different architectures in case of the wood species identification based on single patches.

| Architecture | ACC | F-measure |
|--------------|-------|-----------|
| AlexNet | 0.915 | 0.915 |
| VGG-16 | 0.687 | 0.681 |
| GoogLeNet | 0.947 | 0.940 |
| ResNet-50 | 0.715 | 0.685 |

Table 4 compares the performance of the wood species identification for the whole boards using the selected decision rules. Fig. 7 represents the confusion

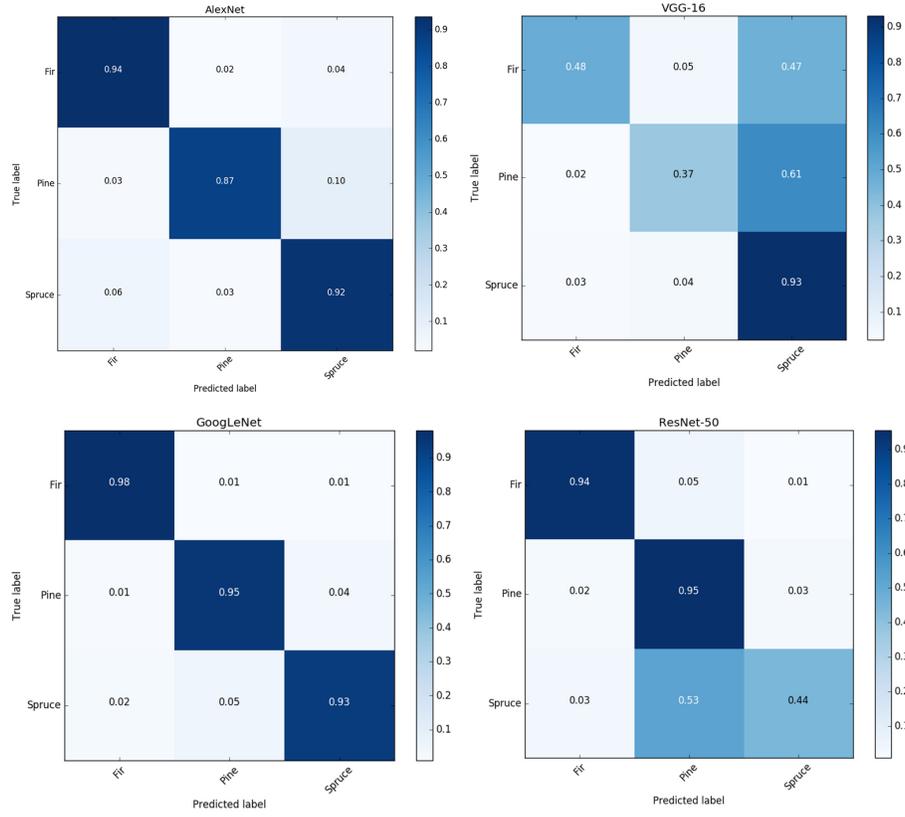


Fig. 6. Confusion matrices of the classification models in case of the wood species identification based on single patches.

matrices for the board identification when the best decision rule is used for each CNN architecture.

When all the patches (on the average 250) extracted from the board are used for the identification the computation time is too long for the industrial environment even with architectures with smaller number of parameters. To address this issue, an additional experiment was carried out to determine how the number of patches affects the accuracy of the identification. Fig. 8 shows the mean accuracy with respect to the number of image patches. For each number of patches the experiment was repeated 100 times and the means of the obtained accuracies were computed. As it can be seen, the accuracy close to the maximum can be obtained with only 25 image patches per board. Table 5 represents the accuracies and the inference times τ_B of the board identification when 25 image patches per board are used.

Table 4. Identification accuracies for the different architectures and decision rules in case of board species identification. The rows in bold are the architectures with the best performance.

| Architecture | Majority vote | Mean probability | Max rule |
|------------------|---------------|------------------|--------------|
| AlexNet | 0.989 | 0.989 | 0.989 |
| VGG-16 | 0.721 | 0.721 | 0.721 |
| GoogLeNet | 0.994 | 0.994 | 0.994 |
| ResNet-50 | 0.783 | 0.764 | 0.764 |

Table 5. The accuracy and the inference time when only 25 image patches are used per board.

| Architecture | ACC | τ_B , seconds |
|--------------|-------|--------------------|
| AlexNet | 0.986 | 0.26 |
| VGG-16 | 0.721 | 0.72 |
| GoogLeNet | 0.994 | 0.47 |
| ResNet-50 | 0.780 | 0.95 |

5 Conclusion

In this paper a novel CNN-based framework to identify wood species from board images was proposed. The four CNN architectures, namely, AlexNet, VGG-16, GoogLeNet, and ResNet were compared on the task of classification of board image patches based on the wood species. The highest accuracy was obtained with the GoogLeNet architecture that correctly classified 94.7% of the patches. Furthermore, three decision rules were proposed to be used to make the final species identification for the board based on multiple patch classification results. No significant difference on performance were found between the proposed decision rules. Using any of the rules and GoogLeNet, close to perfect accuracy of 99.4% was obtained. Finally, the effect of the number of image patches used for the identification was studied and it was shown that over 99% identification accuracy can be achieved by using only 25 patches per board, providing inference time of 0.47 seconds. These results indicate that it seems to be possible to sort the boards in real-time in the sawmill environment.

Acknowledgements

The research was carried out in the DigiSaw project (No. 2894/31/2017) funded by Business Finland. The authors would like to thank FinScan Oy for providing the data for the experiments.

References

1. Chetlur, S., Woolley, C., Vandermersch, P., Cohen, J., Tran, J., Catanzaro, B., Shelhamer, E.: cudnn: Efficient primitives for deep learning. arXiv preprint arXiv:1410.0759 (2014)

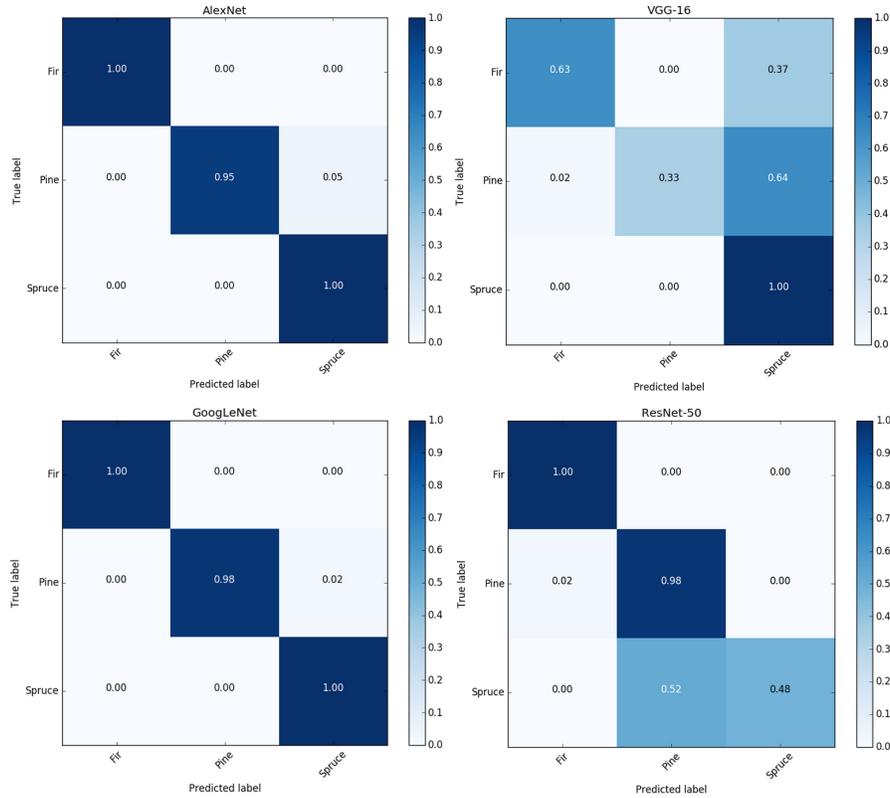


Fig. 7. Confusion matrices of the classification models in case of board identification.

- Hafemann, L.G., Oliveira, L.S., Cavalin, P.: Forest species recognition using deep convolutional neural networks. In: Proceedings of the 22nd International Conference on Pattern Recognition (ICPR). pp. 1103–1107 (2014)
- He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016)
- Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM International Conference on Multimedia. pp. 675–678 (2014)
- Khalid, M., Lee, E.L.Y., Yusof, R., Nadaraj, M.: Design of an intelligent wood species recognition system. *International Journal of Simulation System, Science and Technology* **9**(3), 9–19 (2008)
- Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 1097–1105 (2012)
- Martins, J., Oliveira, L.S., Nisgoski, S., Sabourin, R.: A database for automatic classification of forest species. *Machine Vision and Applications* **24**(3), 567–578

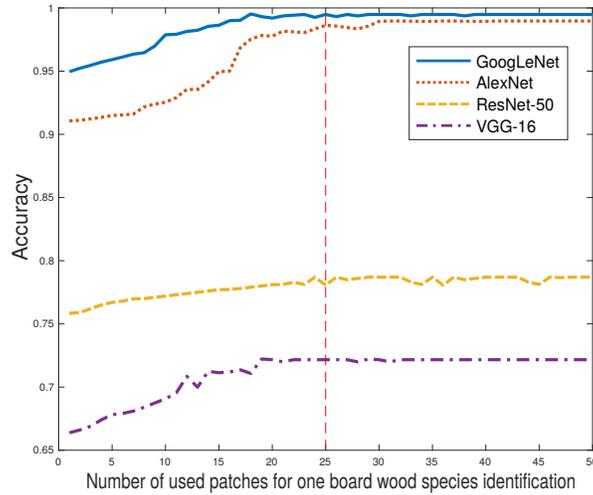


Fig. 8. Influence of the number of image patches on the board identification accuracy.

- (2013)
8. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
 9. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., Rick Chang, J.H., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
 10. Tou, J.Y., Tay, Y.H., Lau, P.Y.: One-dimensional grey-level co-occurrence matrices for texture classification. In: Proceedings of International Symposium on Information Technology (2008)
 11. Tou, J.Y., Tay, Y.H., Lau, P.Y.: A comparative study for texture classification techniques on wood species recognition problem. In: Processings of the Fifth International Conference on Natural Computation (2009)
 12. Tou, J.Y., Lau, P.Y., Tay, Y.H.: Computer vision-based wood recognition system. In: Proceedings of International Workshop on Advanced Image Technology (2007)