

Detection of Mechanical Damages in Sawn Timber Using Convolutional Neural Networks

Rudakov Nikolay, Eerola Tuomas, Lensu Lasse, Kälviäinen Heikki, Haario Heikki

This is a Post-print version of a publication
published by Springer, Cham
in German Conference on Pattern Recognition. Lecture Notes in Computer Science.

DOI: 10.1007/978-3-030-12939-2_9

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

Please cite the publication as follows:

Rudakov N., Eerola T., Lensu L., Kälviäinen H., Haario H. (2019) Detection of Mechanical Damages in Sawn Timber Using Convolutional Neural Networks. In: Brox T., Bruhn A., Fritz M. (eds) Pattern Recognition. GCPR 2018. Lecture Notes in Computer Science, vol 11269. Springer, Cham

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

Detection of Mechanical Damages in Sawn Timber Using Convolutional Neural Networks

Nikolay Rudakov^{1,3}, 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 University of Technology, School of Engineering Science, Machine Vision and Pattern Recognition Laboratory, P.O.Box 20, FI-53851 Lappeenranta, Finland

`firstname.lastname@lut.fi`

² Lappeenranta University of Technology, School of Engineering Science, Inverse Problems Research Group, P.O.Box 20, FI-53851 Lappeenranta, Finland

`firstname.lastname@lut.fi`

³ FinScan Oy, Lukupurontie 2, FI-02200 Espoo, Finland

`firstname.lastname@finscan.fi`

Abstract. The quality control of timber products is vital for the sawmill industry pursuing more efficient production processes. This paper considers the automatic detection of mechanical damages in wooden board surfaces occurred during the sawing process. Due to the high variation in the appearance of the mechanical damages and the presence of several other surface defects on the boards, the detection task is challenging. In this paper, an efficient convolutional neural network based framework that can be trained with a limited amount of annotated training data is proposed. The framework includes a patch extraction step to produce multiple training samples from each damaged region in the board images, followed by the patch classification and damage localization steps. In the experiments, multiple network architectures were compared: the VGG-16 architecture achieved the best results with over 92% patch classification accuracy and it enabled accurate localization of the mechanical damages.

1 Introduction

Automated quality control plays an important role in the sawmill process where end products, i.e., boards and planks need to be efficiently sorted into different grades. Higher board grades can be sold at a higher price, resulting in more profit for the sawmill. At the same time, selling low quality boards as high grade products causes reclamations from the customers. Therefore, it is vital to be able to define the grade as accurately as possible. For automating the quality control, computer vision techniques provide an attractive tool for grading of the boards and planks.

Timber can be mechanically damaged in the sawing process. The most common cause for the mechanical damages is that the feed rollers excessively press

the board with spikes while moving it through the sawing machine [11]. Such defects appear as series of marks or traces along the board. However, several other mechanical damages may occur during the process, resulting in large variations in the appearance of the damages. Examples of mechanical damages on sawn timber are shown in Fig. 1. It should be emphasized that the defect type affects the grading of the boards and the end use of the product. Therefore, it is essential to distinguish mechanical damages from the other defects.

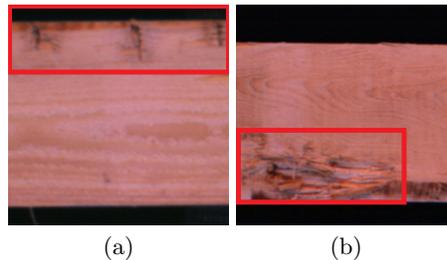


Fig. 1. Examples of the mechanical damages: (a) Periodic mechanical damage caused by a feed roller; (b) Another type of non-periodic mechanical damage.

Convolutional neural networks (CNNs) are the leading technique for image classification and object detection [17]. Being a universal function approximator, CNN can be potentially trained to extract almost any high-level features from the input image including features of mechanical damages of sawn timber. However, applying an existing state-of-the-art CNN-based general object detection method such as YOLO [14] or Faster R-CNN [16] for the detection of defects, especially mechanical damages, is challenging due to the following characteristics of the data: 1) the board images have a very high width to height ratio, 2) annotated data are difficult to obtain resulting in a small amount of training data, and 3) defect (bounding box) sizes and aspect ratios vary considerably.

In this paper, the above problems are tackled by proposing an efficient CNN-based method for detecting mechanical damages from board images by utilizing a one-dimensional sliding window approach. The proposed method is based on the fact that the damage localization is relevant only along the longitudinal direction of the board since the defected parts of the board are sawn off. This makes it possible to implement the sliding window based detection method with only one scale and with a relatively small amount of image patches per image to be classified with CNN. Since most of the individual mechanical damages are large, each of them is represented by numerous image patches resulting in a large amount of positive examples for the training. We further utilize data augmentation and transfer learning to make it possible to train CNNs with a relatively small amount of data. Finally, we evaluate multiple CNN architectures for the classification task, and demonstrate the efficiency of the approach.

2 Related Work

Numerous computer vision based methods for surface inspection of sawn timber including defect detection can be found in the literature. However, most of the existing studies consider the detection of natural timber defects such as different kinds of knots, bark pockets, wane, fungus, worm holes, cracks, and resin leaving out mechanical damages [6].

The existing approaches vary between the different defect types, but three common stages can be highlighted [6]: 1) The defects are localized, 2) features are extracted from each defect, and 3) feature-based classification of the defects is performed. Several reviews on the existing solutions for the timber surface inspection exist [6, 23, 15]. They categorize the approaches based on the methods used for defect localization, feature extraction, and classification, and compare the method performances.

The feature localization stage can be implemented with the combination of various image filtering, segmentation, thresholding, connected components labeling, region merging, and many other image processing techniques. Most of the previous studies have used texture features for the timber defect classification [23]. The most common feature extraction methods are local binary patterns (LBP) [12], gray-level co-occurrence matrix (GLCM) [5], scale-invariant feature transform (SIFT) [10], speeded up robust features (SURF) [1] and Tamura texture features [22].

With the increase of computing performance, deep learning approaches have become more popular in image processing [21]. The most recent articles considering defect detection rely on CNNs. In [15], a generic deep-learning-based approach for automated surface inspection was introduced. The method utilizes a pretrained Decaf CNN [3] to generate a heatmap of defects on the surface. The heatmap is further binarized and segmented using the graph-based Felzenszwalbs segmentation method [4]. In [2], a method for crack detection on concrete surface using a single CNN was introduced. The image is processed using a sliding window, and the corresponding part of the image is fed to CNN for each position of the window.

3 Proposed Method

The scheme of the proposed method is illustrated in Fig. 2. The method starts with the patch extraction procedure where the board image is divided into overlapping image patches. The patches are classified using CNN according to the defect type located within the patch. Finally, the mechanical damages are localized based on the coordinates of the patches.

The sawn timber boards are long and narrow, and that is why their surface images have the high width-to-height ratio. At the same time, the existing trained CNN architectures require input images to be typically scaled to a fixed size with the aspect ratio of 1. Since the mechanical damages caused by feed rollers consist of quite narrow notches scaling a board image to the required

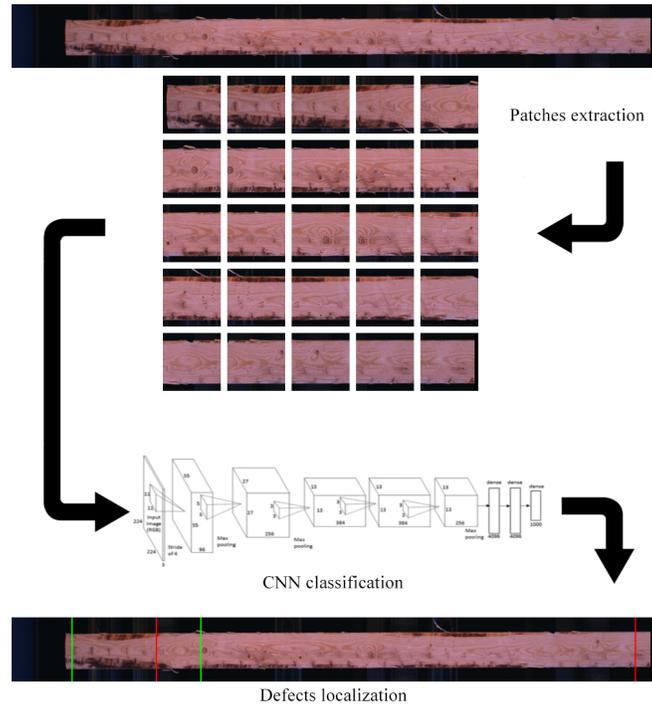


Fig. 2. The scheme of the proposed method for mechanical damage detection from sawn timber images: the board image is split into overlapping patches, the patches are fed to CNN, and then the classified patches and their original locations determine the beginning (green line) and the end (red line) positions of mechanical damages in the longitudinal direction.

dimensions (approximately 8 times in the horizontal dimension) leads to losing the significant information about defect appearance and causes difficulties in the learning process as a consequence.

To resolve this problem, the patch extraction method is used. The general idea is to segment the board from the image and to split the segmented region into a series of small overlapping patches with the required size. The proposed patch extraction technique turns the problem of the mechanical damage detection into a patch classification task.

3.1 Patch extraction

Before the image patch extraction, the board needs to be detected from the image. Typical board images captured in the sawmill process consist of a bright board with a dark background. For the board segmentation, the image is con-

verted from RGB to grayscale, and the Otsu's thresholding method [13] is applied. After this, the largest connected component is searched.

The coordinates of the bounding box for the board are computed using the coordinates of the top-, bottom-, left- and right-most pixels of the connected component. The board in the image could be located imprecisely in the horizontal direction and its orientation could be different from the one of the image. The non-ideal orientation and possible presence of slivers makes the direct use of connected component extreme points unreliable for accurate extraction of the patches. Nevertheless, the left and right sides of the bounding box can be used as reference points.

To locate the board more accurately, the following steps are performed. First, the board centroid is calculated as the mean value of the connected component pixels coordinates. Then, the angle to the horizontal direction θ and the minor axis length of the ellipse with the same normalized central moment as the connected component are computed. To find the ellipse, the origin of the pixels' coordinates is shifted to the position of the centroid. If the height of the bounding box is smaller than the minor ellipse axis multiplied by $\cos \theta$ then the bounding box is assumed to be correct and the upper board edge is approximated using the coordinates of one of the top bounding box corners as a reference. The lower board edge is a parallel line that goes under the upper edge at the approximated board height distance. If the height of the bounding box is larger than the minor ellipse axis multiplied by $\cos \theta$ then the reference is the central line passing through the centroid with the angle to the horizontal direction θ . The upper and lower edges are at the distance of half board height from the central line.

The patch positions are computed using the top and bottom board edges, and the left and right bounding box sides. Moreover, the top and bottom parts of all the patches include the background margins with the size of 10% of the board height. This is done because the mechanical damages are often located close to or at the edges of the board. Also, the background margin of 10 pixels is added to the left and right ends of the board. The patches have a square shape and adjacent patches have 50% overlap with each other. Finally, the patches are scaled to the input size of the CNN architecture. An example of localized patches is shown in Figure 3.



Fig. 3. Extraction the overlapping image patches from the segmented board image. The rightmost patch is aligned with the end of the board and has, therefore, a larger overlap with the neighboring patch.

3.2 Patch classification and damage localization

After the patches have been extracted they are fed to the CNN network in which binary classification between the defected and non-defected patches is made. During the extraction step, the patch coordinates of the entire image are preserved. The patch is considered to contain the mechanical damage if the damage covers more than the half of the patch in the longitudinal direction of the board. Also, the patch overlap is half of their width. To deduce the localization, the left border of the defect is considered to be at $x+0.25w$ of the first patch in a sequence of defective patches while the right border is considered to be at $x+0.75w$ of the last patch in a sequence of defective patches where x is the longitudinal coordinate of the top left corner of the patch and w is the width of the patch. The damage localization is visualized in Figure 4.

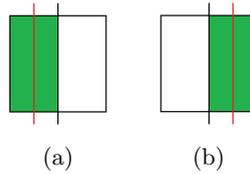


Fig. 4. Damage localization: (a) The first defective patch in a sequence; (b) The last defective patch in a sequence.

4 Experiments

4.1 Data

The data consists of images of 127 sawn timber boards (see Fig. 5). Six overlapping images (three from the top and three from the bottom sides) were taken from every board to cover the whole length of the board. The images were manually annotated for the following defect types: periodic mechanical damages caused by feed rollers and non-periodic mechanical damages. The defect examples are shown in Fig. 1.

The patch extraction algorithm was applied to the labeled data. As a result, 10808 image patches were extracted from the board images. All patches overlapping with a bounding box of a mechanical damage at least 50% of its width were labeled as containing a defect. Most of the patches represent defectless parts of the board, and therefore, there is a large class imbalance in the dataset. For example, in the case of periodic mechanical damages, there are 8683 clear patches and only 2125 defective. The number of defective patches was increased by generating augmented image patches with horizontal and vertical mirroring and 180° rotation. As a result, the number of defective patches was 8500. The

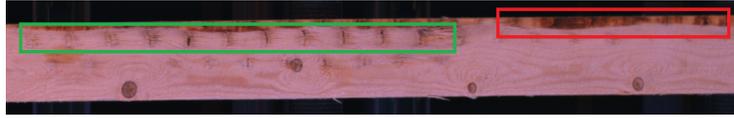


Fig. 5. An example of feed roller traces as mechanical damage (green bounding box) and wane not considered as mechanical damage (red bounding box).

patches were board-wise randomly split into the training (70%) and test sets (30%). This means that all the patches from one board were used either in the training phase or in the test phase. The augmented image patches were used only in the training phase. The vast majority of mechanical damages belong to the group of periodic mechanical damages. Therefore, for the further experiments, two datasets were constructed. In the first dataset, all mechanical damages were combined into one defect class and, in the second dataset, non-periodic mechanical damages were eliminated causing the defect class to contain only periodic mechanical damages. Table 1 shows the numbers of images and patches in the datasets.

Table 1. The number of images in the datasets.

Defects	Training					Test				
	Boards	Images	Patches			Boards	Images	Patches		
			Clear	Defect	Total			Clear	Defect	Total
Combined	89	597	5490	7840	13330	38	229	2667	691	3358
Periodic	89	534	6045	6164	12209	38	228	2638	584	3222

4.2 Evaluation criteria

The accuracy (ACC) of image patch classification was measured as the percentage of correctly classified patches over all test images. Additionally, the damage localization accuracy was measured using the Jaccard metric. The Jaccard metric is typically defined as follows:

$$S_{\text{Jaccard}} = \frac{|A_P \cap A_{GT}|}{|A_P \cup A_{GT}|} \quad (1)$$

where A_P is the area of the predicted object bounding box and A_{GT} is the area of the ground truth bounding box. In our experiments, comparing detections and ground truth this way is problematic since a portion of the test images does not contain any mechanical damages causing the intersection $|A_P \cap A_{GT}|$ to be zero even if the method works correctly, i.e. it does not detect any false positives. This further results in the Jaccard metric values to be unreliable. To

avoid this, the Jaccard metric was used to measure the accuracy of detecting the non-damaged parts of the boards, i.e., how well the method predicts the regions that do not contain mechanical damages. This means that A_{GT} in (1) defines the ground truth for non-damaged areas and A_P defines the areas where no mechanical damages were detected. It should be noted that the dataset does not contain any image that is fully covered with mechanical damages, and therefore, the intersection is zero only if the method fails, i.e. does not detect any non-damaged regions. An example of the intersection and the union of the ground truth, and the predicted mechanically damaged one-dimensional regions of the board is shown in Figure 6. The average Jaccard metric over all the test images was used as the performance measure.

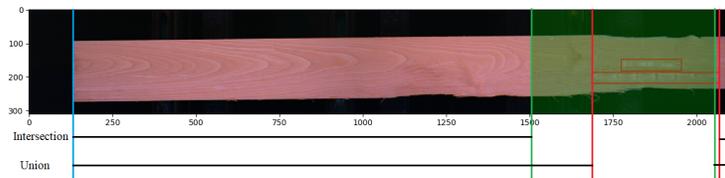


Fig. 6. Jaccard metric in case of one-dimensional mechanical damage localization from a board image. The ground truth defect borders are red and the predicted defect borders are green.

4.3 Results

Two experiments were carried out. In the first experiment, the method was trained to detect any mechanical damages, i.e., the both periodic and non-periodic mechanical damages were considered to belong to the same class (combined). In the second experiment, the method was trained to detect the periodic mechanical damages only. The purpose of the first experiment was to evaluate the ability of the selected CNN architecture to distinguish the mechanically damaged parts of the board regardless of the damage type. The non-periodic mechanical damages may significantly affect the accuracy because they are more variable in their appearance and less common in the existing dataset. That is why the second experiment was carried out to evaluate the performance of the CNN to classify periodic mechanical damages. Since the non-periodic mechanical damages are considerably less common, it was not possible to train a CNN to detect them alone.

Four CNN architectures were selected for the comparison concerning their ability to recognize mechanical damages in sawn timber patches. These architectures were AlexNet [9], GoogLeNet [20], ResNet-50 [7] and VGG-16 [19]. The architectures were trained with the Caffe deep learning framework [8]. To reduce

the training time and to increase the classification accuracy, the transfer learning approach was applied. The selected CNNs were initialized with pretrained models trained for wood species identification using a similar but more extensive dataset [18].

The performance for localization and binary classification of the combined periodic and non-periodic mechanical damages to a single class is provided in Table 2. The performance for localization and binary classification of the periodic mechanical damages is provided in Table 3. The percentages of images that were analyzed correctly, i.e., all the mechanical damages were correctly detected with respect to the Jaccard metric threshold, are shown in Fig. 7. Examples of periodic damage detection results with the GoogLeNet CNN architecture are given in Fig. 8.

Table 2. Performance comparison of different CNN architectures in case of the combined binary classification. Accuracy refers to the classification accuracy for individual patches and $S_{Jaccard}$ measures the defect location accuracy for the full images.

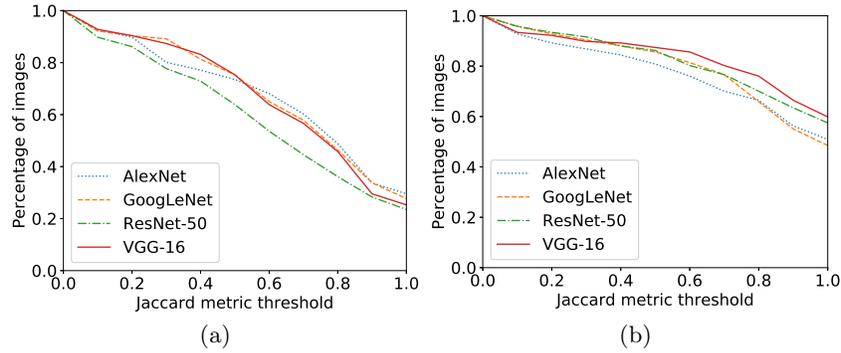
Architecture	Accuracy	$S_{Jaccard}$	Confusion matrix		
				Predicted 0	Predicted 1
AlexNet	0.800	0.692	True 0	0.86	0.14
			True 1	0.42	0.58
GoogLeNet	0.860	0.700	True 0	0.87	0.13
			True 1	0.16	0.84
ResNet-50	0.799	0.624	True 0	0.79	0.21
			True 1	0.17	0.83
VGG-16	0.861	0.696	True 0	0.85	0.15
			True 1	0.11	0.89

The most time consuming stage of the proposed method is the sequential patch classification with CNN. The lower the inference time consumed by a single patch classification is, the more effective is the given CNN architecture. The inference times for all the CNNs were measured on a MSI GE70 laptop with Intel Core i7-4700MQ @ 2.4 GHz processor, NVIDIA GeForce GTX 760M single GPU, and Ubuntu 17.10 operating system. Table 4 contains the average time required to classify one patch of a board image.

As it can be seen, the VGG-16 architecture achieves the best patch classification accuracy of more than 92% for the periodic mechanical damages and 86% for all the mechanical damages. At the same time, the GoogLeNet architecture achieves the lowest false negative and false positive rates and it is four times faster in the single patch processing than the VGG-16. The worst classification and detection accuracy was shown by the AlexNet architecture, and it showed the worst false positive rate of 41%.

Table 3. Performance comparison of different CNN architectures in case of the periodic mechanical damages binary classification.

Architecture	Accuracy	S_{Jaccard}	Confusion matrix		
				Predicted 0	Predicted 1
AlexNet	0.857	0.779	True 0	0.92	0.08
			True 1	0.41	0.59
GoogLeNet	0.907	0.795	True 0	0.91	0.09
			True 1	0.09	0.91
ResNet-50	0.913	0.829	True 0	0.94	0.06
			True 1	0.20	0.80
VGG-16	0.927	0.840	True 0	0.95	0.05
			True 1	0.16	0.84

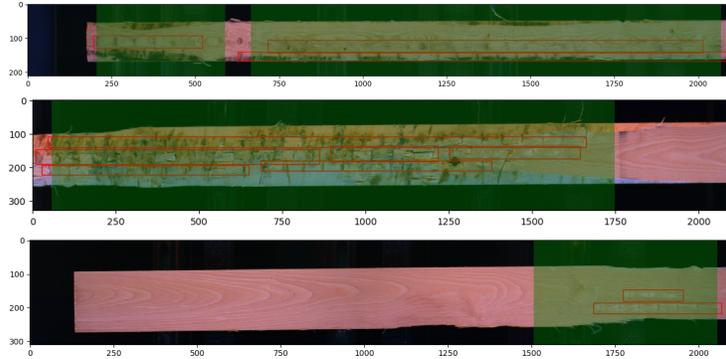
**Fig. 7.** Percentage of images with correctly detected damages with respect to the Jaccard metric threshold: (a) Combined mechanical damages; (b) Periodic mechanical damages only.

5 Conclusion

In this paper, a method for mechanical damage detection on sawn timber images was proposed. The proposed method segments the board on the image, splits the part of the image containing the board into overlapping patches, classifies the patches with the CNN, and, finally, determines the defect location based on classification results and the coordinates of the patches. The experiments compared the performance of the following four CNN architectures: AlexNet, GoogLeNet, VGG-16, and ResNet-50. The VGG-16 architecture produced the best results with a promising classification accuracy of more than 92% for individual patches.

Table 4. Single patch average inference time for each of the trained CNN architectures.

Architecture	Inference time, seconds
AlexNet	0.022
GoogLeNet	0.023
ResNet-50	0.051
VGG-16	0.101


Fig. 8. Examples of the periodic mechanical damage detection with the GoogLeNet CNN architecture: the ground truth damage areas (red bounding boxes) and the predicted defective parts (green overlay).

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

References

1. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features SURF. *Computer Vision and Image Understanding* **110**(3), 346–359 (2008)
2. Cha, Y.J., Choi, W., Büyüköztürk, O.: Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering* **32**(5), 361–378 (2017)
3. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T.: DeCAF: a deep convolutional activation feature for generic visual recognition. In: *Proceedings of the 31st International Conference on Machine Learning (ICML)*. vol. 32, pp. 647–655. PMLR (2014)
4. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. *International Journal of Computer Vision* **59**(2), 167–181 (2004)
5. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3**(6), 610–621 (1973)
6. Hashim, U., Hashim, S., Muda, A.: Automated vision inspection of timber surface defect: a review. *Jurnal Teknologi* **77**(20), 127–135 (2015)

7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778. IEEE (2016)
8. 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 International Conference on Multimedia. pp. 675–678. ACM (2014)
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Proceedings of the 25th Conference on Neural Information Processing Systems (NIPS). pp. 1097–1105 (2012)
10. Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of the 7th International Conference on Computer Vision (ICCV). vol. 2, pp. 1150–1157. IEEE (1999)
11. Nuutinen, Y., Väätäinen, K., Asikainen, A., Prinz, R., Heinonen, J.: Operational efficiency and damage to sawlogs by feed rollers of the harvester head. *Silva Fennica* **44**(1), 121–139 (2010)
12. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition* **29**(1), 51–59 (1996)
13. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* **9**(1), 62–66 (1979)
14. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). pp. 779–788. IEEE (2016)
15. Ren, R., Hung, T., Tan, K.C.: A generic deep-learning-based approach for automated surface inspection. *IEEE Transactions on Cybernetics* **48**(3), 929–940 (2018)
16. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: Proceedings of the 28th Conference on Neural Information Processing Systems (NIPS). pp. 91–99 (2015)
17. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: Imagenet large scale visual recognition challenge. *International Journal of Computer Vision (IJCV)* **115**(3), 211–252 (2015). <https://doi.org/10.1007/s11263-015-0816-y>
18. Shustrov, D.: Species identification of wooden material using convolutional neural networks. Master’s thesis, Lappeenranta University of Technology, Finland (2018)
19. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: Proceedings of the International Conference on Learning Representations (ICLR) (2014)
20. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1–9. IEEE (2015)
21. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2818–2826. IEEE (2016)
22. Tamura, H., Mori, S., Yamawaki, T.: Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics* **8**(6), 460–473 (1978)
23. Tong, H.L., Ng, H., Yap, T.V.T., Ahmad, W.S.H.M.W., Fauzi, M.F.A.: Evaluation of feature extraction and selection techniques for the classification of wood defect images. *Journal of Engineering and Applied Sciences* **12**(3), 602–608 (2017)