Three-step Approach for Localization, Instance Segmentation and Multi-facet Classification of Individual Logs in Wooden Piles

Christoph Praschl[1] [a] and Gerald Adam Zwettler[1,2] [b]

[1] *Research Group Advanced Information Systems and Technology (AIST), University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, Austria*

[2] *Department of Software Engineering, School of Informatics, Communications and Media, University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, Austria*

Keywords: Instance Localization, Segmentation, Classification, Wooden Piles, Logs, Cross Faces, Deep Learning, Neural Networks.

Abstract: The inspection of products and the assessment of quality is connected with high costs and time effort in many industrial domains. This also applies to the forestry industry. Utilizing state-of-the-art deep learning models allows the analysis automation of wooden piles in a vision-based manner. In this work a three-step approach is presented for the localization, segmentation and multi-facet classification of individual logs based on a client/server architecture allowing to determine the quality, volume and like this the value of a wooden pile based on a smartphone application. Using multiple YOLOv4 and U-NET models leads to a client-side log localization accuracy of 82.9% with low storage requirements of 23 MB and a server-side log detection accuracy of 94.1%, together with a log type classification accuracy of 95% and 96% according to the quality assessment of spruce logs. In addition, the trained segmentation model reaches an accuracy of 89%.

1 INTRODUCTION

With the recent improvements in computer vision due to availability of good deep learning paradigms, machine learning frameworks and improved GPU hardware, the automated vision-based measurement becomes feasible in many industrial areas such as the printing industry (Villalba-Diez et al., 2019), the food industry (Kakani et al., 2020), the construction industry (Xu et al., 2020) or the steel production industry (Chen et al., 2020) among many others. This development is also one of the fundamentals in autonomous driving (Janai et al., 2020). However, in forestry industry, for the domain of wood pile trading, key aspects for the price such as cross-sections of the logs, quality and type of the wood are still mostly manually assessed. As this is a very time-consuming process, digitization is a key factor in cost reduction. With a smartphone app, the log front faces can be automatically detected and segmented, besides precisely assessing wood type and quality in an objective and reproducible way, resulting in the possibility for determining a first price estimation of the felled trees.

[a] https://orcid.org/0000-0002-9711-4818
[b] https://orcid.org/0000-0002-4966-6853

1.1 Research Questions and Approach

In this research paper we therefore introduce a client/server based localization and instance segmentation, as well as classification approach allowing live application in the woods with subsequent server-side refinement of the results. Besides it is analyzed, if splitting into localization and segmentation allows to achieve a higher level of accuracy overall.

2 RELATED WORK

Multiple publications are available in the context of the individual, proposed pipeline steps for the analyzation of wooden piles using localization, segmentation or classification approaches. Regarding the localization of individual logs in wooden piles Gutzeit and Voskamp (Gutzeit and Voskamp, 2012) as well as Auersperg-Castell (Auersperg-Castell, 2018) are able to identify instances using Haar Cascade classifications. Multiple authors (Chiryshev and Atamanova, 2016; Knyaz et al., 2004; Kruglov, 2016) use geometric form based approaches by detecting circles or ellipses to detect logs in wooden piles. In con-

Praschl, C. and Zwettler, G.
Three-step Approach for Localization, Instance Segmentation and Multi-facet Classification of Individual Logs in Wooden Piles.
DOI: 10.5220/0010892100003122
In *Proceedings of the 11th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2022)*, pages 683-688
ISBN: 978-989-758-549-4; ISSN: 2184-4313

trast to the previous mentioned publications, Herbon et al. (Herbon et al., 2014) combine local binary patterns, together with histograms of oriented gradients (HOG) for detecting individual logs in a wooden pile. Based on the localized logs, the authors use the watershed algorithm based on a Gaussian mixture model to segment the logs. Chiryshev and Atamanova (Chiryshev and Atamanova, 2016) continue the idea of using HOG features and combine this approach with the Random Forest learning method.

In the context of segmenting an individual log, classic approaches such as the watershed algorithm as proposed by Herbon et al. (Herbon et al., 2014) or similarity based methods as described by Schraml and Uhl (Schraml and Uhl, 2014) are used. In contrast to that, Decelle and Jalilian (Decelle and Jalilian, 2020) juxtapose multiple neural network architectures to separate the image foreground (log) from the background based on images of wooden cross-sections. Like in the present work, the authors use among others an U-NET architecture for the separation of the logs' cut face from the remaining background with a comparable accuracy of 92% based on a non-public image dataset of 2381 wooden logs.

According to the last remaining step of the proposed pipeline, in the context of the classification of the wood type and quality, recent approaches have been published. Kryl et al. (Kryl et al., 2020) compare different methodologies in this area using classic computer vision methods such as gray level covariance matrices or k-nearest neighbor, but also more advanced methodologies using deep learning. This review, shows that previous publications differ from the present approach by carrying out the classification using images of the tree's bark (Carpentier et al., 2018; Fiel and Sablatnig, 2010), using cut out wooden boards instead of the raw cross-sections (Shustrov, 2018), based on macroscopic images of the wood (Khalid et al., 2008; Tou et al., 2009; Tou et al., 2007; Paula Filho et al., 2014; Paula Filho et al., 2010; Gunawan et al., 2018; Seng and Guniawan, 2018; Yusof et al., 2013; Yadav et al., 2013; Nasirzadeh et al., 2010; Urbonas et al., 2019) or utilizing infrared (Cao et al., 2017) and even x-ray images (Mu et al., 2008) instead of RGB images of the log's cross section.

To the best of the authors' knowledge, none of the related works have proposed such a client/server based architecture using deep learning approaches in the context of individual log cross-sections of wooden piles for localizing, segmenting and tree type as well as quality classification and for this reason differ from the present approach.

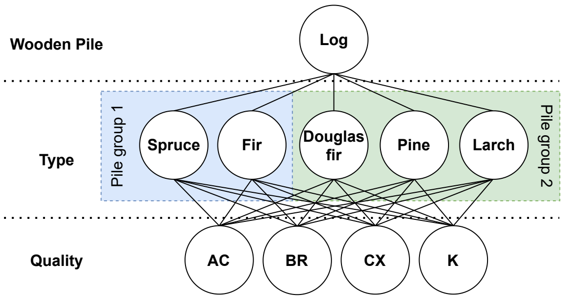

Figure 1: Ontology for the classification of logs of a wooden pile with two levels: (1) the type of tree and (2) the quality. The types of trees are grouped according to their occurrences in the Austrian forestry landscape.

3 ONTOLOGY

In the Austrian wood market, two kinds of wooden piles are typical for construction work or product packaging, such ones with *spruces* and *firs*, and the second group with *douglas firs*, *pines* and *larchs*, as shown in Figure 1. Next to the actual tree type, a log can be classified to its quality based on e.g. cracks or color differences, but always in association with the type of tree. This separate consideration is necessary, because of the situation, that for example a red discoloration of a log may be a bad indicator for most types of trees but is a typical feature of douglas fir logs. For this reason, we have decided to create a classification system based on the wood classes used in Austrian sawmills. We have reduced the original number of classes, which e.g. differs between logs of quality *A* to *C*, to the following four final classes based on the associated price segment:

- **AC:** These are the best qualities *A* through *C* with no or only few flaws and are usually paid the same price, so they are combined into one class.
- **BR:** This quality class contains logs of minor quality with partially rooten spots mostly due to fungus and for this generate lower revenues.
- **CX:** The class of *CX* contains logs of minor quality with cracks or irregular shape that are crooked or knotty. This class can be only partially classified from a photograph of the cutting area, since the crucial features can also occur along the whole remaining log.
- **K**: The last class contains wood that is infested by bark beetle and for this has a blue/gray shade at the rim on the cutting surface. However, the infestation has not progressed that far to impair the structural quality and is for this still sellable.

4 METHODOLOGY

Based on the proposed ontology, a process is suggested as shown in Figure 2, that is designed for foresters and forest owners, who are interested in the value of wooden piles. Due to the situation that a network connection may not be available on the site of a forest, the process is based on a separated architecture using a client, as well as a server application. To do so, two pipelines are considered, with one local pipeline for a first assessment that happens on e.g. a smartphone or tablet, and a complementary pipeline on an external server for a more accurate result. The idea of the local pipeline is to have a very slim and fast method to get a first estimation of the number of logs in a pile, the volume of the logs and a rough estimation of the pile's value. These first results are further refined on the server. Consequently, a two step input process is used. First, the user takes an image of the wooden pile, which is distorted according to the camera model and additional sensor data as tilt and orientation. Second, the user is asked for meta information: (I) the distance to the pile of the image, (II) the forest type (hardwood, softwood) as well as (III) the pile's length, which is defined by the harvester or lumberjack.

4.1 Client Pipeline

In a first step of the local pipeline, a tiny YOLO (Wang et al., 2020) model is used to detect the logs in the image. Based on the resulting number $i \in \mathbb{N}$ of bounding boxes the Grab Cut algorithm is utilized to segment the actual logs from the clipped bounding boxes. The resulting mask images I_i contain 0 for background, and 1 for foreground pixels respectively. The mask images with a width w and a height h are used in a next step to estimate the logs' trunk areas a_i as shown in Equation 1. The areas a_i are in turn

$$a_i = \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} I_i(x,y) \tag{1}$$

used together with the user inputs, the pile's length p and a pixel to mm conversion factor d to approximate the log volume v_i as shown in Equation 2. The conversion factor d can be determined by the size ratio of one sample log i in px o_{px_i} and mm o_{mm_i} as shown in Equation 4. To do so, the log's size in mm can be calculated via the image's metadata and the user input using the distance to the log d_{mm}, the sensor height s_h, the image height h and the focal length fl as shown in Equation 3. Using the determined volume and a forest type specific volume price $price_{ft}$, the pipeline is able to estimate a first rough value v_{pile} for the given pile as shown in Equation 5, where $logs$ is the number of logs in the current pile as detected by the YOLO model.

$$v_i = a_i \cdot p \cdot d \tag{2}$$

$$o_{mm_i} = \frac{d_{mm} * o_{px_i} * s_h}{h * fl} \tag{3}$$

$$d = o_{mm_i} / o_{px_i} \tag{4}$$

$$v_{pile} = \sum_{i=1}^{logs} v_i \cdot price_{ft} \tag{5}$$

4.2 Server Pipeline

The server-side pipeline basically pursues the same goal as the local one, calculating the value of a photographed pile and for this task uses the same input. In difference to the local pipeline, not a tiny YOLO model, but a classic YOLO model is used to detect the logs. The reason for this design decision is that the tiny YOLO model is faster than the classic model and has lower storage requirements, while the classic model leads to a better accuracy. These characteristics make it possible to realize a small and fast local version for a first user preview. In the second step of the server-side pipeline, two tasks are executed in parallel based on the detected log bounding boxes of the classic YOLO. On the one hand an U-Net model is used to segment the individual logs in a more accurate way compared to the Grab Cut method of the client side, and on the other hand a log classification process is executed. Based on the U-Net result, the pile's volume is calculated identically to the local process. The classification is in turn done in two successive steps (I) the tree type classification and (II) the log quality classification. The reason for this separated classification process is attributed to the ontology shown in Figure 1, since the quality classes are basically the same for every log, but the features (e.g. color) that determine the actual quality are dependent of the type of tree. Using the classification results, a type specific $price_{t_i}$ and a quality depended factor q_i in combination with the volume of the associated log, the value of the pile can be calculated as shown in Equation 6.

$$v_{pile} = \sum_{i=1}^{logs} v_i \cdot q_i \cdot price_{t_i} \tag{6}$$

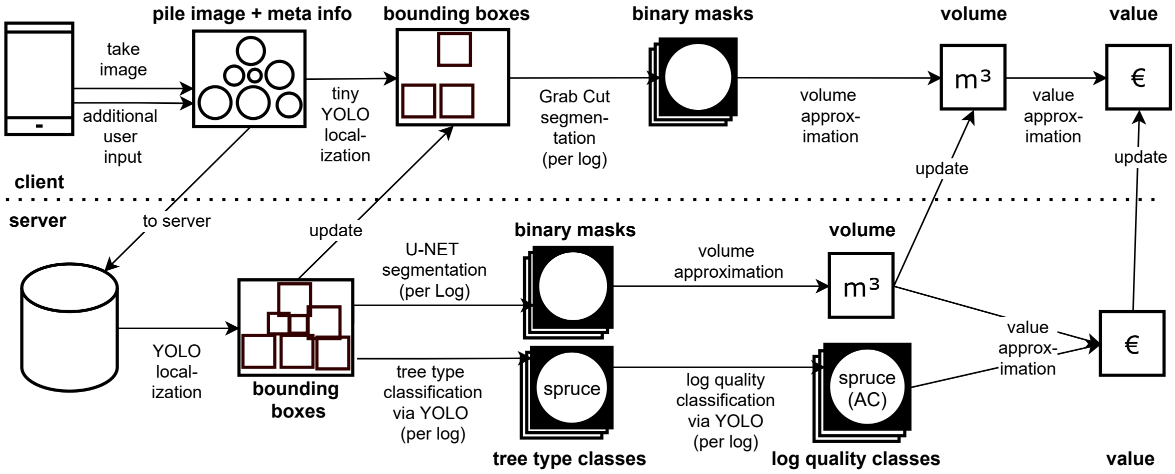

Figure 2: The overall process of determining the value of a wooden pile based on the volume and optional the class of individual logs, consisting of a client-side for a first user preview and the server-side for a more accurate calculation.

5 MATERIAL

To the best of the authors' knowledge, there is neither a wooden pile nor a log data set publicly available, that is suitable for the evaluation of the proposed methodologies. There was the HAWKwood database, which is not available anymore (Herbon, 2014). For this reason, a data set of 440 pile images from the project partners' archives is created with a resolution between 640×480 pixels and 4032×3024 pixels. The images are distributed according to the seasons, with 42 images in spring, 125 in summer, 97 in autumn, 133 in winter and 43 remaining images, which cannot be associated with the recording time, because of a missing timestamp. This distribution is shown in Table 1 and leads to different visual conditions of the logs related with the seasonal weather events, like moist wood or partially snow or mud covered logs.

Table 1: The distribution of the data set according to the seasons.

Spring	Summer	Autumn	Winter	Unknown
42	125	97	133	43
9.5%	28.4%	22%	30.2%	9.8%

The data set is manually labelled with bounding boxes, classes and binary masks of the individual logs. The pile images contain in total 18521 and on average 42 individual logs, with a minimum amount of 1 log, a maximal amount of 395 logs and a standard deviation of 63.95. Currently, 2243 spruces, 15 firs, 346 douglas firs, 75 pines and 1135 larchs are manually classified, as shown in Table 2. These logs are mostly also classified according to their quality, as listed in Table 3. In addition to that, cross-sections are segmented for 166 individual logs.

Table 2: The distribution of manually classified logs and their types relative to the 18521 source logs.

Type of Tree	# Logs	Relative Distribution
Spruce	2243	12.1%
Fir	15	0.1%
Douglas Fir	346	1.9%
Pine	75	0.4%
Larch	1135	6.1%

Table 3: The distribution of the log quality per type based on the manually classified logs shown in Table 2.

Type of Tree	AC	BR	CX	K
Spruce	1464	214	13	543
	65.3%	9.5%	0.6%	24.2%
Fir	11	0	4	0
	73.3%	0%	26.7%	0%
Douglas Fir	270	55	2	18
	78%	15.9%	0.6%	5.2%
Pine	74	1	0	0
	98.7%	1.3%	0%	0%
Larch	886	76	35	52
	78.1%	6.7%	3.1%	4.6%

6 PRELIMINARY RESULTS

The following preliminary results are based on a subset of the mentioned dataset, that was not considered for the training of the proposed models. This subset is used, because of the situation that no comparable dataset is publicly available to our knowledge. Using a tiny, scaled YOLOv4 (Wang et al., 2020) model on the client-side results in a real-time capable log detection accuracy of 82.9% with 23 MB storage usage

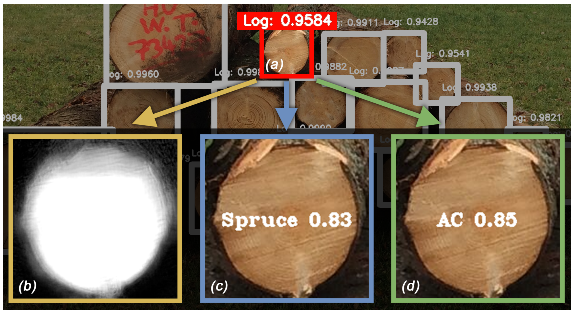

Figure 3: Sample of a wooden pile highlighting the results for a single log with (a) its localization, (b) its segmentation and (c) its tree type as well as (d) its quality classification, together with the detection confidences.

only. On the server, a classic, scaled YOLOv4 model results in an accuracy of 94.1% with storage requirements of 250 MB (Praschl et al., 2021). Using the resulting bounding boxes, an U-NET (Ronneberger et al., 2015) model for the segmentation and multiple YOLOv4 classification models are used in parallel. A detailed evaluation of these networks is still pending, but first results (compare Figure 3) show a mean average precision of 95% for the type model, 96% for a spruce quality model and 89% for the segmentation model. The quality model shows storage requirements of 250 MB and the segmentation model 330 MB.

7 CONCLUSION AND OUTLOOK

The preliminary results show that state-of-the-art computer vision algorithms for localization, segmentation and classification can be combined to a multiprocess analysis approach for wooden piles and for this push the digitization of the forestry industry. The separation of the problem into localization with subsequent segmentation / classification allows to boost the overall quality of results and further facilitates a client / server infrastructure where preview results can be provided on common smartphones in real-time while further analysis and higher accuracy are asynchronously performed on a server. Especially, the consideration of the proposed ontology, allows improving the classification accuracy and makes the process more robust. In future the extension of the training data set, especially with logs of underrepresented wood and quality types, as well as evaluations of classification and segmentation models are planned.

ACKNOWLEDGEMENTS

Our thanks to the province of Upper Austria for facilitating the project Woodmaster+ with the easy2innovate funding program.

The authors thank *Philipp Auersperg-Castell* and *Brigitte Forster-Heinlein* for their contributions to the research, as well as *Ulrich* and *Luis Hainberger* for providing the labeled test data set and for the valuable input.

REFERENCES

Auersperg-Castell, P. (2018). Photooptische holzpoltervermessung mittels haar-kaskaden. Bachelor's thesis, University of Passau, Germany.

Cao, J., Liang, H., Lin, X., Tu, W., and Zhang, Y. (2017). Potential of near-infrared spectroscopy to detect defects on the surface of solid wood boards. *BioResources*, 12(1):19–28.

Carpentier, M., Giguere, P., and Gaudreault, J. (2018). Tree species identification from bark images using convolutional neural networks. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1075–1081. IEEE.

Chen, W., Chen, S., Guo, H., and Ni, X. (2020). Welding flame detection based on color recognition and progressive probabilistic hough transform. *Concurrency and Computation: Practice and Experience*, 32(19):e5815.

Chiryshev, Y. V. and Atamanova, A. S. (2016). Automatic wood log detection based on random decision forests learning algorithm and histogram of oriented gradients. In *Proceedings of the 3rd International Young Scientists Conference on Information Technologies, Telecommunications and Control Systems*, pages 7–12.

Decelle, R. and Jalilian, E. (2020). Neural networks for cross-section segmentation in raw images of log ends. In *2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS)*, pages 131–137.

Fiel, S. and Sablatnig, R. (2010). *Automated identification of tree species from images of the bark, leaves or needles*. na.

Gunawan, P., Prakasa, E., Sugiarto, B., Wardoyo, R., Rianto, Y., Damayanti, R., Dewi, L. M., et al. (2018). 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)*, pages 23–27. IEEE.

Gutzeit, E. and Voskamp, J. (2012). Automatic segmentation of wood logs by combining detection and segmentation. In Bebis, G., Boyle, R., Parvin, B., Koracin, D., Fowlkes, C., Wang, S., Choi, M.-H., Mantler, S., Schulze, J., Acevedo, D., Mueller, K., and Papka, M., editors, *Advances in Visual Computing*, pages 252–261, Berlin, Heidelberg. Springer Berlin Heidelberg.

Herbon, C. (2014). The hawkwood database. *arXiv preprint arXiv:1410.4393*.

Herbon, C., Tönnies, K., and Stock, B. (2014). Detection and segmentation of clustered objects by using iterative classification, segmentation, and gaussian mixture models and application to wood log detection. In *German Conference on Pattern Recognition*, pages 354–364. Springer.

Janai, J., Güney, F., Behl, A., Geiger, A., et al. (2020). Computer vision for autonomous vehicles: Problems, datasets and state of the art. *Foundations and Trends® in Computer Graphics and Vision*, 12(1–3):1–308.

Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., and Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2:100033.

Khalid, M., Lee, E. L. Y., Yusof, R., and Nadaraj, M. (2008). Design of an intelligent wood species recognition system. *International Journal of Simulation System, Science and Technology*, 9(3):9–19.

Knyaz, V., Visilter, Y., and Zheltov, S. (2004). Photogrammetric techniques for measurements in woodworking industry. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Proceedings, V. XXXIII. Part B5/2*, pages 42–47.

Kruglov, A. V. (2016). Development of the rounded objects automatic detection method for the log deck volume measurement. In *First International Workshop on Pattern Recognition*, volume 10011, page 1001104. International Society for Optics and Photonics.

Kryl, M., Danys, L., Jaros, R., Martinek, R., Kodytek, P., and Bilik, P. (2020). Wood recognition and quality imaging inspection systems. *Journal of Sensors*, 2020.

Mu, H., Qi, D., Zhang, M., and Yu, L. (2008). Image edge detection of wood defects based on multi-fractal analysis. In *2008 IEEE International Conference on Automation and Logistics*, pages 1232–1237. IEEE.

Nasirzadeh, M., Khazael, A. A., and bin Khalid, M. (2010). Woods recognition system based on local binary pattern. In *2010 2nd International Conference on Computational Intelligence, Communication Systems and Networks*, pages 308–313. IEEE.

Paula Filho, P., Oliveira, L. S., Britto, A. d. S., and Sabourin, R. (2010). Forest species recognition using color-based features. In *2010 20th International Conference on Pattern Recognition*, pages 4178–4181. IEEE.

Paula Filho, P. L., Oliveira, L. S., Nisgoski, S., and Britto, A. S. (2014). Forest species recognition using macroscopic images. *Machine Vision and Applications*, 25(4):1019–1031.

Praschl, C., Auersperg-Castell, P., Brigitte, F.-H., and Zwettler, G. A. (2021). Multi-resolution localization of individual logs in wooden piles utilizing yolo with tiling on client/server architectures. *Proceedings of 33rd European Modeling & Simulation Symposium*.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation.

Schraml, R. and Uhl, A. (2014). Similarity based cross-section segmentation in rough log end images. In Iliadis, L., Maglogiannis, I., and Papadopoulos, H., editors, *Artificial Intelligence Applications and Innovations*, pages 614–623, Berlin, Heidelberg. Springer Berlin Heidelberg.

Seng, L. K. and Guniawan, T. (2018). An experimental study on the use of visual texture for wood identification using a novel convolutional neural network layer. In *2018 8th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, pages 156–159. IEEE.

Shustrov, D. (2018). Species identification of wooden material using convolutional neural networks. Master's thesis, Lappeenranta University of Technology School of Engineering Science.

Tou, J. Y., Lau, P. Y., and Tay, Y. H. (2007). Computer vision-based wood recognition system. In *Proceedings of International workshop on advanced image technology*, pages 197–202. Citeseer.

Tou, J. Y., Tay, Y. H., and Lau, P. Y. (2009). Rotational invariant wood species recognition through wood species verification. In *2009 First Asian Conference on Intelligent Information and Database Systems*, pages 115–120. IEEE.

Urbonas, A., Raudonis, V., Maskeliūnas, R., and Damaševičius, R. (2019). Automated identification of wood veneer surface defects using faster region-based convolutional neural network with data augmentation and transfer learning. *Applied Sciences*, 9(22):4898.

Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Mer, J., Buchwitz, M., and Wellbrock, W. (2019). Deep learning for industrial computer vision quality control in the printing industry 4.0. *Sensors*, 19(18).

Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y. M. (2020). Scaled-yolov4: Scaling cross stage partial network. *arXiv preprint arXiv:2011.08036*.

Xu, S., Wang, J., Shou, W., Ngo, T., Sadick, A.-M., and Wang, X. (2020). Computer vision techniques in construction: a critical review. *Archives of Computational Methods in Engineering*, pages 1–15.

Yadav, A. R., Dewal, M., Anand, R., and Gupta, S. (2013). Classification of hardwood species using ann classifier. In *2013 fourth national conference on computer vision, pattern recognition, image processing and graphics (NCVPRIPG)*, pages 1–5. IEEE.

Yusof, R., Khalid, M., and M. Khairuddin, A. S. (2013). Application of kernel-genetic algorithm as nonlinear feature selection in tropical wood species recognition system. *Computers and Electronics in Agriculture*, 93:68–77.