forests

MDPI

Review

A Comparative Literature Review of Machine Learning and Image Processing Techniques Used for Scaling and Grading of Wood Logs

Yohann Jacob Sandvik *,†, Cecilia Marie Futsæther, Kristian Hovde Liland and Oliver Tomic

Faculty of Science and Technology, Norwegian University of Life Sciences, 1432 Ås, Norway; cecilia.futsaether@nmbu.no (C.M.F.); kristian.liland@nmbu.no (K.H.L.); oliver.tomic@nmbu.no (O.T.)

* Correspondence: yohann.jacob.sandvik@nmbu.no

† Current address: Department of Data Science, Faculty of Science and Technology, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway.

Abstract: This literature review assesses the efficacy of image-processing techniques and machine-learning models in computer vision for wood log grading and scaling. Four searches were conducted in four scientific databases, yielding a total of 1288 results, which were narrowed down to 33 relevant studies. The studies were categorized according to their goals, including log end grading, log side grading, individual log scaling, log pile scaling, and log segmentation. The studies were compared based on the input used, choice of model, model performance, and level of autonomy. This review found a preference for images over point cloud representations for logs and an increase in camera use over laser scanners. It identified three primary model types: classical image-processing algorithms, deep learning models, and other machine learning models. However, comparing performance across studies proved challenging due to varying goals and metrics. Deep learning models showed better performance in the log pile scaling and log segmentation goal categories. Cameras were found to have become more popular over time compared to laser scanners, possibly due to stereovision cameras taking over for laser scanners for sampling point cloud datasets. Classical image-processing algorithms were consistently used, deep learning models gained prominence in 2018, and other machine learning models were used in studies published between 2010 and 2018.

Keywords: log scaling; log grading; wood science; computer vision; artificial intelligence; deep learning

check for updates

Citation: Sandvik, Y.J.; Futsæther, C.M.; Liland, K.H.; Tomic, O. A Comparative Literature Review of Machine Learning and Image Processing Techniques Used for Scaling and Grading of Wood Logs. *Forests* **2024**, *15*, 1243. https://doi.org/10.3390/f15071243

Academic Editor: Barbara Ozarska

Received: 28 May 2024
Revised: 3 July 2024
Accepted: 11 July 2024
Published: 17 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Wood as a raw material has a long history in many industries. Wood boards are one of the primary materials used to build residential houses, furniture, and tools. Logs that are not suitable for the production of boards are used to produce pulp in paper manufacturing, charcoal, and chemicals. The global logging and forestry market had an estimated value of $1.08 trillion in 2023 [1], and wood and wood materials are significant exports for forest-rich countries such as Norway, Finland, and Lithuania [2]. Across all logging companies, there is a demand for methods that can rapidly determine the scale, volume, and quality of individual logs and bundles of logs. Manual scaling and grading of logs is often very time-consuming and labor-intensive, but it is the only option if more advanced methods are not available.

A common goal in the field of computer vision is to automate tasks performed by the human vision system using computers. Typical tasks within computer vision are the detection, classification, and tracking of objects in images and videos. Common for many tasks in computer vision is that although it is easy to state what the goal is, it is difficult to state an explicit relation between the input and the goal. Take the detection of wood logs in images as an example; it is easy for a human to point out a log in an image, but it is

BID Group Technologies, Ltd.
EX-2020, Page 1

harder to define an explicit set of rules for a computer that can be used to capture all logs in every image. This type of problem is well-suited for the application of machine learning. Machine learning is a subset of the field of artificial intelligence that is characterized by not requiring explicit rules or guidelines for how a task should be solved. Instead, machine learning models are given input data coupled with a specific goal, then they learn how to attain that goal. Given the exponential increase in computing power over the last decade and the number of image datasets available, many tasks in computer vision have been found to be well-suited for a machine learning approach.

The objective of this work is to investigate recent research conducted on the application of computer vision for the scaling and grading of wood logs. This has led to the following research questions that have prompted this work:

1. How successful have image processing techniques using external imaging been when applied to the grading and scaling of wood logs?
2. How successful has machine learning been when applied in the context of computer vision used for the grading and scaling of wood logs?

2. Method

Three central points decide the content and direction of a literature review: the search terms used to find studies, the choice of article databases in which searches are performed, and how search results are screened to determine the content that will make up the data for the literature review. The search terms and databases used have been decided based on the research questions outlined in Section 1, the Introduction.

2.1. Search Phrases

Since the field of log scaling and grading is relatively narrow, fairly general search terms were used, and more time was invested into screening the search results to ensure that as few relevant studies as possible were missed. Each row of Table 1 makes up a search term that was entered into a database. Four operators were used to construct the searches: "OR", "AND", "*", and "". The search "wood OR lumber" will match with all studies that contain the term "wood" and all studies that contain the term "lumber". In Table 1, the symbol $\vee$ is used to indicate the "OR" operation. The search "wood AND lumber" will only match studies that contain both the terms "wood" and "lumber". In Table 1, the symbol $\wedge$ is used to indicate the AND-operation. To make a search match with an exact phrase and not single words individually, the "" operator is used. The term "machine learning" will only match studies that contain the words "machine" and "learning" together in the phrase "machine learning". It is also useful to match studies that contain variations of a search word. This is achieved by using the "*" operator. The term "*wood" will not only match studies that contain the word "wood" but also with articles that contain the word "hardwood", "softwood", "sapwood", etc. The symbols (S_i) in the first column of Table 1 give the IDs designated to each search. The elements in the Title terms column are terms that are present in the titles of individual studies. The elements in the General terms are terms that are present somewhere in the full text of the studies.

Table 1. The four search terms submitted to the databases to find studies.

ID	Title Terms	General Terms
S_1	*lumber $\vee$ *timber $\vee$ * wood	"machine learning"
S_2	*lumber $\vee$ *timber $\vee$ *wood	"computer vision"
S_3	log $\vee$ logs	(*lumber $\vee$ *timber $\vee$ *wood) $\wedge$ "machine learning"
S_4	log $\vee$ logs	(*lumber $\vee$ *timber $\vee$ *wood) $\wedge$ "computer vision"

S_1 and S_2 were meant to capture all studies containing the words "wood", "timber", or "lumber" in the title and those that used methods within the fields of "machine learning" or "computer vision". These two searches were not combined ("machine learning" $\vee$

"computer vision") because the search tools did not allow for combining multiple exact phrases in a single search. It was found that some relevant studies only contained the word "log" in the title and not "wood", "timber", or "lumber", which is why S_3 and S_4 were added. S_3 and S_4 match studies containing the words "log" or "logs" in the title; the words "wood", "timber", or "lumber"; and "machine learning" or "computer vision" in the full text.

2.2. Databases

There were two criteria set for the potential databases to be used in the literature review. First, they had to contain content relevant to the literature review, and second, they had to have a search engine where it was possible to search for title-specific and full-text-specific search terms using the four operators outlined earlier, and they had to allow for bulk export of the search results. Table 2 shows the different databases used in this study and the number of studies matched with the four searches. The columns show the number of results returned by the specific searches. In total, there were 1288 studies from the four searches in the aforementioned databases for scientific studies, but based on the screening and selection process, only 33 of them were included in the final literature study. All the searches were performed using the individual databases on 1 October 2023.

Table 2. The number of search results returned from the four searches (S_1, S_2, S_3, and S_4) that were submitted to the four individual databases.

Database	S_1	S_2	S_3	S_4	Sum
IEEE Xplore	86	83	5	14	188
Scopus	271	186	6	17	480
SpringerLink	157	96	23	22	298
Web of Science	196	89	30	7	322
Sum	710	453	64	60	1288

2.3. Inclusion/Exclusion Criteria

Some criteria were set for the literature included in this study. The formal criteria were that the studies were written in English and were accepted to a conference or published in a peer-reviewed scientific journal. The requirements set for the content of the literature are outlined in the list below.

- Log-centric analysis:It was required that the included studies were focused on whole wood logs, regardless of species. This means that all studies centered on the analysis of wood boards, veneer, and processed products of wood logs were excluded.
- Target alignment:The target or goal of the included studies must align with the tasks of scaling or grading wood logs. If the goal was not explicitly scaling or grading, it had to contribute to log scaling or grading.
- Input data:Since this review is specific to computer vision methods, only studies using images or point-cloud data were included. The scope was further restricted to exclude studies using cross-sectional imaging of the logs, such as X-ray, ultrasound, and other computer tomography methods.

Figure 1 shows how many studies were removed during the screening steps. In the first step, a general check for non-relevant matches was performed. Since the search terms used were fairly broad and the databases included some of the same journals, duplicates were expected. A total of 374 duplicates were found. Nine results were not articles, conference studies, or proceedings, but were instead chapters in books or notices of retraction and were hence excluded. A total of 103 non-relevant term matches were found and excluded. Non-relevant matches in this context meant that matches were not made with the search words as intended. The majority of the non-relevant matches were results containing the words "Hollywood" or "Bollywood" in the title. In the first content-filtering step, illustrated by the second green box in Figure 1, one can see that a total of 454 search results were

excluded. A total of 273 studies were centered around processed logs, meaning timber boards, wood veneer, wood furniture, etc. A total of 138 studies focused on chemical compounds extracted from wood, such as "Agarwood oil", or the biological properties of wood materials. A total of 43 studies focused on the social science aspects of the timber trade. The second step of the content-filtering process was to ensure that the goals of the studies were aligned with the goal of the literature review; i.e., that the studies were related to the scaling or grading of wood logs. From the third green box in Figure 1, one can see that a total of 269 were excluded. For 81 studies, the target was centered on ecology; 66 studies had a target focused on the state of manufacturing equipment or optimizing operations at sawmills; and 122 studies were centered on classifying the species of a particular log or wood sample. The third and final step of the content filtering entailed separating the research that did not use imaging methods or used cross-sectional imaging methods. A total of 45 studies were excluded in this step, of which 32 studies used cross-sectional images for log grading and 13 studies did not use imaging data in their analysis. After screening, 33 studies remained and were included in the review.

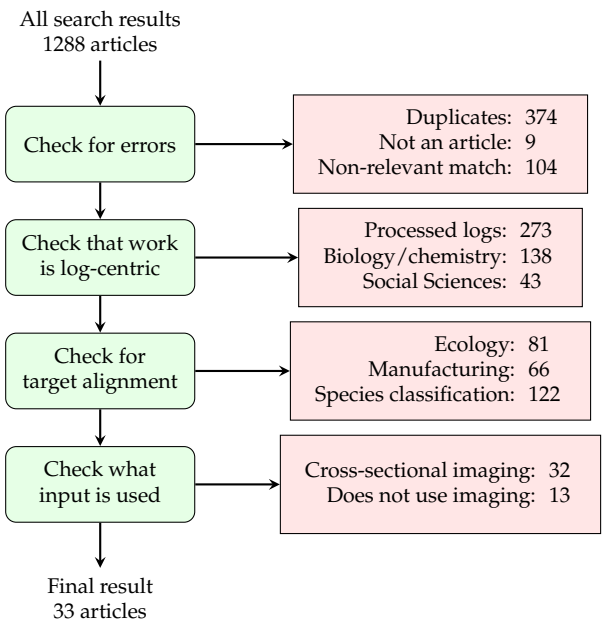

Figure 1. Flow diagram to illustrate screening method. The green boxes with the rounded corners on the left represent the steps in the screening process, and the red boxes with straight corners on the right represent the number of studies excluded, together with the reason for exclusion.

3. Results

In this section, the studies are outlined and compared in terms of the goals, input data used, and choice of algorithm. Figures are provided that illustrate the similarities and differences between these studies.

3.1. Input and Goals

Table 3 shows which input data and which goals were used among the 33 studies included in this review. There are two categories of input data: images and point-clouds. The Image category refers to arrays where the individual elements contain measurements of light intensity and/or frequency, and the point-cloud category refers to arrays where the elements contain measurements of distance. As seen in Table 3, five overarching goals were found. "Log scaling" refers to the estimation of the dimensions or volume of wood logs, whereas "log grading" refers to an estimation of the quality of the log and often determines which types of wood products are worth extracting from the log. The goal category "log end grading" includes studies that attempt to determine the quality of logs or detect defects in logs based on images or scans of the log ends. The goal category "log side grading"

includes studies that attempt to determine the quality of logs or detect defects in logs based on images or scans of the sides of logs. The goal category "individual log scaling" includes studies that attempt to compute the volume or dimensions of individual logs using images, while the studies assigned to the "log pile scaling" goal category attempt to compute the volume of piles of logs or the distribution of diameters among the logs in a pile of logs. The final goal category, "log segmentation" contains the studies that had the end goal of segmenting log ends in images.

Table 3. Type of input data used and the goals of the 33 studies included in this review.

Goal/Input	Images	Point Clouds	Number of Studies
Log end grading	[3–6]		4
Log side grading	[7]	[8–12]	6
Individual log scaling	[13–15]		3
Log pile scaling	[16–23]	[24]	9
Log segmentation	[25–35]		11
Number of studies	27	6	33

This division of studies into goal categories was chosen in such a way as to yield the best grounds for comparison of the performance the different methods to attain similar targets. It was entirely coincidental that the goal division resulted in such a clear division of the input data. Table 3 illustrates that four of the five studies that used point-cloud data as inputs had the goal of grading log sides, and the final study using point-cloud data had the goal of scaling log piles. The majority of the studies used images as the input data, which is to be expected as cameras in general are cheaper and come in a greater variety than sensors that record point-cloud data. Almost all the studies implemented secondary models to benchmark their main model, but for brevity, the comparisons made between different studies will focus on the main model of each study.

3.2. Log End Grading

Grading logs by assessing the ends alone is useful, because logs are often stacked in piles such that the ends are the only visible part. Much information about the log's quality can also be deduced from a cross-sectional view of the log, such as wood density [5]. The four studies with goals in the category log end grading had specific goals that were quite different, as illustrated in Figure 2. Carratù et al. [3] detected rot in logs, Cao and Li [4] had the goal of detecting cracks in logs, Du et al. [5] extracted information about the annual ring distribution, and Decelle et al. [6] aimed to locate the pith region in log ends. Images were the only type of data used for log end grading, which is logical, considering that piths, annual rings, and rot are challenging to detect using point-cloud data. One might be able to detect large cracks in the log end, but depending on the resolution of the scanner, smaller cracks would be easier to locate using images. The model developed by Du et al. [5] was influenced by the saw mark disturbances present in the images. Sawing marks would likely be even more prevalent in point-cloud data, which is another possible reason that it was not found to be used in studies of log end grading.

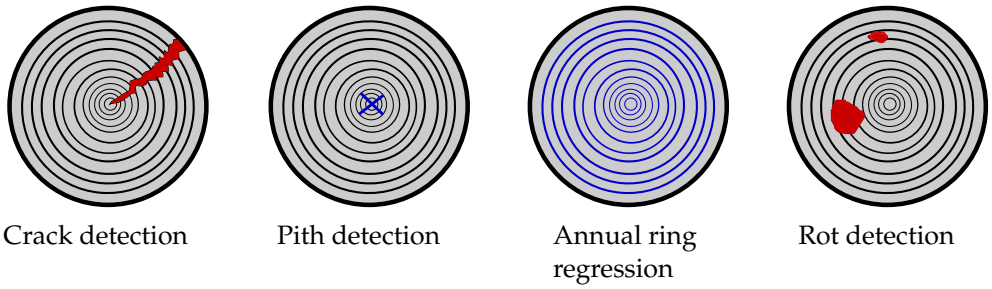

Figure 2. Log end grading: An illustration of the specific goals of the studies within the goal category, including studies that attempt to grade logs using images or scans of the log ends. The red color indicates the detection of defects, and the blue color indicates the estimation of a desired attribute.

In all the studies in this goal category, the images were converted to grayscale, which is an indication that only the structural information of the images was of importance and that a fairly low spectrum of color frequencies was necessary to extract that information. Cao and Li [4] used image graying and histogram equalization to denoise images and used global thresholding with a user-determined threshold to generate a binary mask of cracks followed by morphological operations to further remove noise. Finally, the detected cracks that had a length-to-width ratio smaller than a user-defined threshold were removed, and the remaining cracks left in the binary bitmap were counted as cracks. The model attained an average crack detection accuracy of 84%. The model developed by Decelle et al. [6] computes the local image gradients of the images of individual log ends and accumulates the gradients using ant colony optimization to locate the pith. When evaluated on two different datasets, the model attained a mean distance to the ground truth pith location of 2.34 mm and 2.26 mm, respectively. Du et al. [5] used pith detection as a subroutine of their model but went one step further and extracted information about the log's annual rings. Specifically, they determined the number of annual rings, the width of the annual rings, and the average distance of the 15th annual ring from the pith and from the outside. Du et al. [5] only used the value channel of the image when represented in a hue, saturation, and value (HSV) format. The total variation algorithm was used to denoise the image, followed by the Hough transform, and local peaks were accumulated to locate the pith. Finally, circle fitting was used to locate the annual rings, measure the distance between the rings, and estimate the average distance to the 15th ring. The method attained a relative RMSE of 21.13% for estimating the average distance of the 15th ring from the center and 24.55% for estimating the average distance of the 15th ring from the outside. Carratù et al. [3] partitioned an image of the end of a transport truck into a set of image cut-outs of individual log ends, which were fed into a self-developed convolutional neural network that classified logs with a level of rot above a certain threshold as "rotten". The logs were then categorized as suitable (1) or not suitable (0). At the task of binary classification, the network attained an F1-score of 0.89.

Considering that none of the end goals of the studies were the same, it is not meaningful to compare the performance of the different studies. It would have been interesting if Du et al. [5] had reported the accuracy of their pith location estimator, as that would have given more grounds for comparison with Decelle et al. [6]. The two methods used for pith detection are also somewhat similar: Decelle et al. [6] used image gradients to represent annual rings, which were accumulated to find the pith, and Du et al. [5] used accumulated peaks to locate the pith. Carratù et al. [3], Du et al. [5], Decelle et al. [6] all had fully autonomous models on deployment, meaning that none of their methods relied on operator interference when they were set to estimate their respective goals. Cao and Li [4] on the other hand relied on the operator for setting the appropriate threshold for the crack-identification algorithm, making it less autonomous. Carratù et al. [3] is the only study in this goal category that used images of the entire log piles as input and had a preprocessing step in its approach of extracting cut-outs of singular log-ends. Cao and Li [4], Du et al. [5], Decelle et al. [6] all relied on image cut-outs of singular log ends being extracted before the respective models could be applied. This makes the deep learning model developed by Carratù et al. [3] the most autonomous model in this goal category.

3.3. Log Side Grading

While much information can be extracted from a single cross-sectional image of a log end, log ends reveal little about the distribution of knots in the log. This can affect the quality of the boards or the yield of boards that could be extracted from the log. To detect these types of defects, one has to inspect and grade the log sides, which was done in the studies in this goal category, as illustrated in Figure 3.

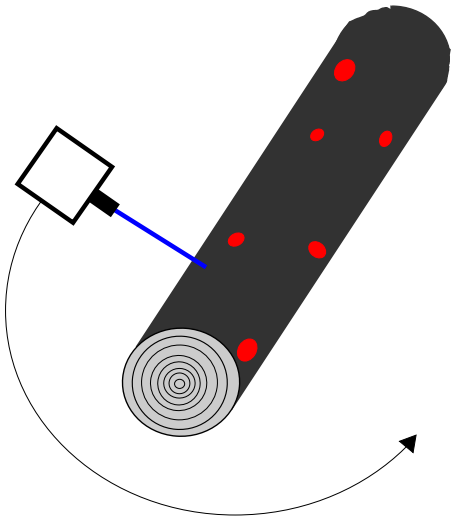

Figure 3. Log-side Grading: An illustration of the set-up and goal of the studies within the goal category, which includes studies that attempt to grade logs using images or scans of the log sides. The red color indicates defects, and the box emitting the blue line illustrates the laser scanner, which rotates around the log to measure the point cloud data.

Four out of the five studies in the log side grading goal category collect the point cloud data used for grading the log using a laser scanner. Laser scanners are a common choice for grading the sides of logs, as the majority of studies in this goal category perform some form of defect detection. Lee et al. [7] is the only study that used images as inputs, where they applied edge detection algorithms to extract depth information from images. Therefore, one can make the argument that some form of 3D scanner would be more appropriate for their application. However, considering that the work of Lee et al. [7] was first presented in 1991, one should assume that 3D scanners were not as readily available at that time. Khazem et al. [12] is the only study included in this review that used cross-sectional imaging. They used an X-ray machine to capture cross-sectional images of logs at varying intervals. The work of Khazem et al. [12] is an exception to the rule of not including works using cross-sectional imaging because after the model was trained, the final model was intended to be applied to radial scans of logs without requiring cross-sectional imaging.

The studies in this goal category can be partitioned into two groups based on their goals. Lee et al. [7], Thomas and Mili [8], Nguyen et al. [9], Thomas et al. [10] aim to detect defects on the log surface, whereas Zolotarev et al. [11], Khazem et al. [12] also detect surface defects, but they used the detected defects to infer the internal knot structure of the log as well. The work performed by Thomas and Mili [8] appears to be a continuation of the work performed by Thomas et al. [10], so only the work performed by Thomas and Mili [8] will be referred to in this section. The models implemented by Thomas and Mili [8] and Nguyen et al. [9] are very similar, although the work of Nguyen et al. [9] was presented almost a decade after the work of Thomas and Mili [8] was published. Both studies used point cloud data of the edges of the logs in polar coordinate format. In both studies, the polar coordinates were "unrolled" into a height map in Cartesian coordinate format. Thomas and Mili [8] fit circles to the radial point clouds of individual log slices and unrolled the polar coordinates relative to the center of the fitted circle. Nguyen et al. [9] evaluated multiple log slices simultaneously and estimated the center line of said slices using cubic spline interpolation. The polar coordinates were then unrolled with reference to the estimated center line. Nguyen et al. [9] performed an additional conversion from the Cartesian height map to cylindrical coordinates. Both studies estimated a reference distance from the center of the log to the edge that represented the expected distance. In the final step, both studies used a form of automatic thresholding to detect large surface rises and depressions, which were classified as defects.

Lee et al. [7] captured images of partitions of the log sides while the log was rotated 10 degrees between each scan. An unspecified edge detection algorithm was used to generate a binary mask of the log edge, which was filtered by an algorithm called the skeleton thinning algorithm. The filtering algorithm aimed to fit multiple small straight lines to the edge of the log. Potential knots were then identified by searching for long diagonal lines among the detected edges. Finally, knots were then distinguished from bark based on their color and texture. The model of Zolotarev et al. [11], which is initially quite similar to the work of Nguyen et al. [9], consists of five steps: point-cloud filtering and center line estimation, log surface height map generation, knot segmentation, volumetric reconstruction of knots, and virtual sawing of the log. The DBSCAN clustering algorithm was used to filter the point-cloud data, removing artifacts in the radial direction. Circle fitting and cubic spline interpolation were used to estimate the pith location at individual cross-sections and interpolate them into a center line, which is used as a reference when "unrolling" the radial point-cloud data into a log surface height map. Laplacian of the Gaussian filter was used to segment the knots, and the internal structure of the knots was estimated using a biological knot-property model. The height map with estimated knot structures was then converted back to the 3D Cartesian coordinate representation of the log, where virtual sawing was performed by inserting planes into the 3D model of the log, creating intensity maps that represented virtual board faces. The pixel intensities corresponded to the probability that there was a knot in that location of the virtual board. Khazem et al. [12] trained a deep learning model to predict the internal knot structure of the log given the external point cloud coordinates. Khazem et al. [12] developed a mixed neural network using LSTM cells for the recurrent layers as well as convolutional and fully connected layers. The network was intended to take only a radial scan of the surface of the log as input and generate a segmentation mask of the internal knot structure of the log.

Lee et al. [7], Nguyen et al. [9] did not present any quantitative evaluation metrics, so they cannot be compared to the other studies in terms of performance. The intensity map produced by the model developed by Zolotarev et al. [11] could be compared to the segmentation bitmaps produced by Khazem et al. [12]. However, Zolotarev et al. [11] chose to evaluate the performance based on the correlation between the generated intensity map values and the probability of a knot being present in the real log. The model developed by Zolotarev et al. [11] attains a Pearson correlation coefficient of 0.98 and a Spearman correlation coefficient of 1. While this metric illustrates the effectiveness of the model, it is somewhat convolved and difficult to understand. Khazem et al. [12] used multiple metrics to evaluate the segmentation performance of the model they developed, where the dice similarity coefficient (similar to F1-score) is the most relevant for comparing the study with other works in this goal category. Khazem et al. [12] attained dice similarity coefficients of 0.74 and 0.70 when the model was applied to datasets of fir logs and spruce logs, respectively. What is missing in the performance metrics used by Khazem et al. [12] is some measure of the detection rate, as that would have given better grounds for comparison with the other studies in this goal category. Thomas and Mili [8] split defects into two classes: those that were expected to be detected and those that were not expected to be detected. Of the 59 defects the model was expected to find, it detected 47, and of the 103 defects the model was not expected to find, the model detected 11. This yields an overall detection rate of 35.8%, but when limiting the classification task to only the most visible defects, the detection rate was improved to 79.7%. In terms of autonomy, all the models discussed within this goal category are intended to run without requiring operator interference.

3.4. Individual Log Scaling

Log scaling includes measuring the log length, diameter, or volume, as shown in Figure 4. The three studies in the category of individual log scaling all had the goal of estimating the dimensions of individual logs using images captured from stereo cameras or an equivalent setup. Kruglov and Chiryshev [13] detected and tracked the logs in videos

and calculated their dimensions and volumes in real-time. Kalmari et al. [14] measured the length of logs in images taken from the harvester head. Yang et al. [15] created a 3D reconstruction of logs using a dual-camera setup.

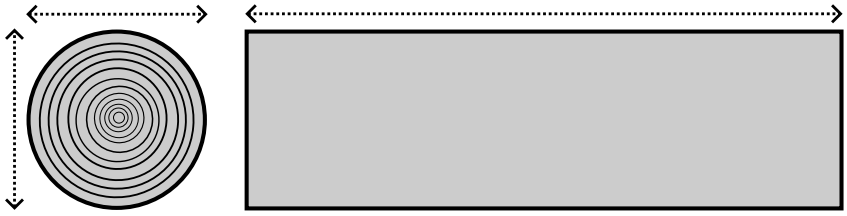

Figure 4. Individual Log Scaling: An illustration of the goal of the studies within this category, which includes studies that attempt to estimate the dimensions or volume of individual logs.

All the studies in this category used some form of stereographic imaging techniques. Kruglov and Chiryshev [13] employed video frames from a stereo camera system to detect, track, and scale logs on a conveyor belt. Kalmari et al. [14] used images captured by a stereovision camera mounted on a harvester head to measure the length of logs. Yang et al. [15] used synchronized cameras to create 3D reconstructions of logs on a conveyor belt.

The structures of the models developed in these studies share some similarities. All three studies involved an initial stage of feature extraction or object segmentation. Kruglov and Chiryshev [13] relied on the background of the images being static while the logs moved through the video frames. They generated a stochastic pixel model of the background using multivariate normal distribution when logs were not present, and they used this to remove the background when logs were present in the frames to generate a segmentation mask. Kruglov and Chiryshev [13] then used a key point detection algorithm to recognize specific points on the logs that were tracked over multiple frames using optical flow. The final stage of the detection was the combination of the two synchronized segmentation masks. This was done by minimizing the sum of Euclidean distances between two synchronized frames. Kruglov and Chiryshev [13] approximated the individual logs by a set of cylinders, and the volume of the logs was estimated by adding the volume of the individual cylinders. The method was tested in a laboratory, but no quantitative metrics were presented on the performance of the volume measurements. Kalmari et al. [14] also utilized the Harris detector and optical flow to track a log across video frames. However, instead of tracking logs on a conveyor belt, the logs were tracked as they were passed through a harvester head. Random sample consensus (RANSAC) was used to remove false feature matches to improve the estimate of the log's motion. The model developed by Kalmari et al. [14] was tested on seven logs and attained a mean absolute error of 2.9 mm and a mean absolute relative error of 0.09% when estimating log length.

Similar to Kruglov and Chiryshev [13], Yang et al. [15] also attempted to scale logs on a conveyor belt using a stereovision setup, but they also generated a point cloud representation of the log without the background. In this setup, two cameras were placed at opposite ends of a conveyor belt capturing images of opposite ends of a log. Two synchronized frames were warped and rotated such that the log was aligned with the x-axis of both images. The common x-axis after rectification is referred to as the epipolar line. Individual partitions of pixels, referred to as "blocks", were then matched based on their pixel characteristics. For the block-matching, they used a window of 7×7 pixels and searched along the epipolar line that minimized the sum of absolute differences algorithm. The two cameras were only able to see one side of the log at a time, so to obtain a point cloud model of the entire log, it was rotated in batches of 10 degrees to compute new point cloud coordinates, which were then added to the 3D reconstruction of the log. Their model was tested on three logs of different shapes, and their model output was compared with the output of a laser scanner. The 3D reconstruction estimated by their model was found to coincide well with the output of the laser scanner, but no quantitative performance metrics were given.

In terms of performance and autonomy, Kalmari et al. [14] provided quantitative performance metrics, reporting a mean absolute error and mean absolute relative error for their model. In contrast, Kruglov and Chiryshev [13], Yang et al. [15] did not provide quantitative results regarding the accuracy of their model estimates. This variation in reporting makes it difficult to compare these studies in terms of model performance. All the models presented in this goal category are intended to run relatively autonomously. The model developed by Kalmari et al. [14] could even complement one of the other two models such that log scaling could be performed during harvesting and at the sawmill.

3.5. Log Pile Scaling

Since logs are often stored in piles, it is useful to obtain an estimate of the aggregate volume of an entire pile. This was the goal of the studies included in this goal category. Figure 5 illustrates the two specific goals encountered in this category: the volume estimation of a log pile and the estimation of the distribution of diameters among the logs that make up the log pile.

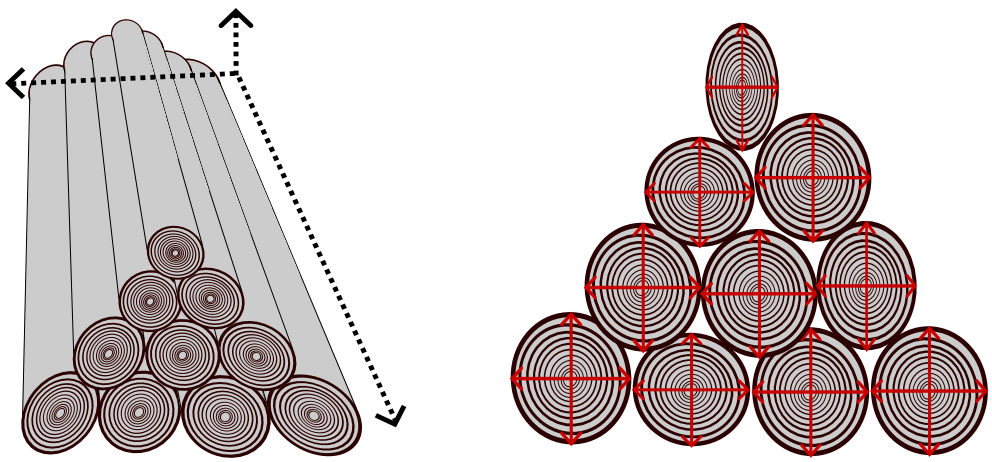

Figure 5. Log pile scaling: An illustration of the goal of the studies within this category, which includes studies that attempt to estimate the volume of log piles or the distribution of the diameter of the logs that make up the log piles.

In terms of input data, all the studies in this goal category utilized images, except for Martí et al. [24], which used LiDaR to collect point cloud representations of the log end side of a log pile. Herbon et al. [17], Kruglov et al. [18], Correia et al. [19] used images of log piles taken from different angles to create a 3D reconstruction of the pile, similar to the methods used by Yang et al. [15]. Galsgaard et al. [16], Li et al. [20], Carratù et al. [21], Zheng et al. [22], Carratù et al. [23] used images captured only from the log end side of a log pile. Zheng et al. [22] used a binocular camera that gave depth measurements of the objects in the image, which were used to convert the relative pixel measurements to physical measurements. Galsgaard et al. [16], Li et al. [20], Carratù et al. [21,23] used cameras that captured 2D images but relied on detecting reference objects in the image of known sizes, which were then used to convert relative pixel measurements to physical measurements.

Table 4 shows the different types of models used by the studies in this goal category. There are an equal number of studies that apply deep learning models and classical image processing algorithms and a single study that uses a clustering algorithm. It should be noted that all the studies that used artificial neural networks were published in 2023, and the studies using classical image processing techniques were published between 1991 and 2021. Li et al. [20] developed an instance segmentation model based on combining an object detection model and semantic segmentation model run in parallel. The outputs from the object detection and semantic segmentation are combined using a "metric learning

paradigm" to produce the final log end instances. Li et al. [20] developed a custom loss function that combines the loss in the detection and segmentation branches such that the instance segmentation model is end-to-end trainable. To obtain the real-world dimensions of the logs, they used the inner diameter of the rear wheel of the loading trucks as a reference object of known size, which yields a scaling factor. The two studies (Carratù et al. [21,23]) are written by many of the same authors, and they have almost identical approaches. The goal of both studies was to detect the logs in the pile and measure their diameters. Carratù et al. [23] used YOLOv4 to detect log ends on the backs of loading trucks, marked them with bounding boxes, and used direct linear transformation (DLT) to convert the pixel measurements to real measurements. The method developed by Carratù et al. [23] relied on operators manually marking the corners of two yellow triangles in each image to act as a reference object of known size. The DLT algorithm then used the goal points to calibrate the camera and create a "homography plane" of the truck rear-end, which represented both the distance to and orientation of the truck rear-end with respect to the camera. The length of the longest side of each log end bounding box was then used to estimate the log end diameters. Carratù et al. [21] seems to be a continuation of the work performed by Carratù et al. [23], where they used a newer object detection model, YOLOv5s, and incorporated the detection of the reference object into the deep learning model. Carratù et al. [21] also replaced the two yellow triangles with a checkered square as the reference object. Zheng et al. [22] used a customized version of Mask R-CNN to perform instance segmentation of the log ends on loading trucks to estimate the wood volume of the entire vehicle. Individual log cutouts were fitted with ellipses using the least squares method. Since a binocular camera was used, Zheng et al. [22] can estimate the depth at different pixels. Estimating the diameter of the individual log end masks then becomes a matter of estimating the distance from the camera to the log end by matching the coordinates of the center of the fitted ellipse curve to the corresponding coordinate in the depth image and using the distance to scale the pixel diameter to real measurements. Assuming that the length of the truck was known, the length of the logs and hence their volume could be estimated given their distance from the camera.

Table 4. The choice of models in studies in the log pile scaling goal category.

Model	Studies
Clustering algorithm	[16]
Classical image processing	[17–19,24]
Artificial neural network	[20–23]

Herbon et al. [17], Kruglov et al. [18], Correia et al. [19], Martí et al. [24] all used classical image processing models. In the setup in Martí et al. [24], the log piles being scanned were always placed in a rectangular support frame and were measured from a fixed range of distances. Martí et al. [24] used depth–threshold filtering to separate the point cloud that corresponds to the log ends from the point cloud that corresponds to the support frame and the rest of the background. To perform the segmentation, Martí et al. [24] tested two different algorithms: a region-growing segmentation algorithm and a circle fitting algorithm. The region-growing algorithm classified point cloud coordinates into regions based on their coordinate value and the value of the neighboring coordinate values, and the circle fitting algorithm is given a minimum and maximum diameter and finds circles in the image within the valid diameter domain. In Kruglov et al. [18], a model that allows for input of one or two images of the same log pile from different angles is validated. It used a combination of segmentation and clustering for detecting, segmenting, and scaling the log piles. The fast radial symmetry algorithm was used to detect the log ends in images, and a combination of the Stoer–Wagner algorithm and the Watershed method was used to segment the log ends in the images. If multiple images were given as inputs, both images were segmented, and the minimum Euclidean distance was used to match the log end segmentations from both images and compute the physical measurements of the

log ends. The logs in the pile were assumed to have more or less the same length when computing the volume of the full pile. Correia et al. [19] segmented log piles on the back of loading trucks. Correia et al. [19] made use of images captured from the side and the end of each log pile. Additionally, they relied on the images always being captured with the same background and at the same distance to filter out the background and convert from relative measurements to physical measurements. They made use of image gradients, spatial-average filtering, and a region-growing algorithm to segment the pile in both images. The difference in brightness between the solid wood pixels and the empty space between the logs was used to estimate the portion of the segmented pile that was solid wood. Herbon et al. [17] developed a model that utilizes multiple machine learning techniques. They used a quadratic filtering technique to approximate the log ends in images with circles. The K-nearest-neighbors estimator was used to estimate the contour that enveloped the entire pile. A random sample consensus-based plane fitting method and principal component analysis were used to fit a plane to the surface of the log pile with the log ends and orient the pile according to a Cartesian coordinate system. The volume of the individual logs was estimated by multiplying the circular area of each individually segmented log with the average length of the logs in the pile.

Galsgaard et al. [16] used the circular Hough transform (CHT) and local circularity measures (LCM) to detect the image region with the highest density of circular shapes. The detected area corresponds to an approximate detection of the log pile in the image. The detected area was regarded as foreground when the graph cut segmentation was initiated. The "blobs", detected in the graph-cut segmentation step, that had a diameter below a certain threshold were discarded as false detections, while the remaining blobs were subjected to a series of morphological operations to refine the circular shape of the segmented logs. To convert the relative measurements in the segmentation mask to physical measurements, Galsgaard et al. [16] made use of short blue rods placed on the log ends as reference objects of a known scale. It was also assumed that the average length of the logs that made up the pile was known, such that the volume of each log could be calculated as the segmented log end area multiplied by the average log length.

Table 5 shows the performances recorded by the different studies in the goal category. Martí et al. [24] did not specify any quantitative metrics related to the scaling of the log pile, whereas Correia et al. [19], Carratù et al. [23] specified the performance of their methods with slight caveats. Correia et al. [19] reported that the volume estimates of their model differed less than 5% from the manual measurement estimates for 90% of the loads it was tested on. Carratù et al. [23] report that the errors in the diameter estimates of their model followed a normal distribution centered at zero, 80% of the measurements had errors in the domain $[-4\text{ cm}, 4\text{ cm}]$, and 50% of the measurements had errors in the domain $[-2\text{ cm}, 2\text{ cm}]$. These reported error measurements give an idea of the performance of their models, but they are not easy to directly compare with the other models in this category. Among the studies that have the specific goal of estimating log pile volume, the clustering model developed in Galsgaard et al. [16] has the worst performance. The three models that used classical image processing techniques, Herbon et al. [17], Kruglov et al. [18], Correia et al. [19], seem to have approximately equal performances, depending on which of the average deviations reported by Herbon et al. [17] one takes into account. The deep learning model developed by Zheng et al. [22] has the highest performance at estimating the wood volume of log piles, with an average relative error of 0.646%. For the studies with the specific goal of estimating the diameter distribution of the logs, we see that the model developed by Li et al. [20] has a higher performance than the model developed by Carratù et al. [21].

All the models presented in the studies in this goal category are intended to run fully autonomously, except the object detection model by Carratù et al. [23], which required operators to manually select the reference objects in the images for each measurement. However, this seems to have been improved in the second iteration of the model presented by Carratù et al. [21], where detection of the reference object is included in the deep learning model.

Table 5. The performance of models among the studies in the log pile scaling goal category. The table only shows the performance metrics related to the estimation of the log pile volume and diameter distribution. The studies in the upper part of the table estimate log pile volume, and the studies in the lower part estimate the diameter distribution of the logs.

Study	Performance	Metric
[16]	14.22%	Average relative error
[17]	−6.3%, 1.4%	Average deviation from ground truth on two datasets
[18]	5.14%	Average relative error
[19]	Not explicitly specified	
[22]	0.646%, 1.68%	Average and max relative error
[24]	Not specified	
[20]	−4.62%	Average relative error
[21]	8.50%	Coefficient of variance
[23]	Not explicitly specified	

3.6. Log Segmentation

Segmentation of logs in images is an important step in the process of log scaling or log grading using computer vision. However, additional information is required to convert the relative sizes detected in images to absolute sizes. The motivations for developing segmentation models are the same as those for developing automatic log scaling and log grading models; namely, to reduce the labor costs and inaccuracies associated with manual log scaling. Many studies have segmentation as a subroutine of their method, but the studies included in this goal category all have segmentation of logs as their final goal.

In contrast to some of the other goal categories, the goals of all the studies included in log segmentation were quite similar. All of them used images as inputs and had the aim of producing a segmentation mask separating the logs from the backgrounds as output. The slight variations in the goals are illustrated in Figure 6. The studies that performed semantic segmentation of log ends produced a binary bitmap where individual pixels are regarded as "log end" or "not log end". The studies that performed instance segmentation of the log ends produced bitmaps where the log ends were separated from the background. The individual instances of log ends were separated as well, such that they could be analyzed separately. The study that performed instance segmentation of entire logs yielded bitmaps where the pixels of the entire logs were separated from the background, and the individual log instances were held separate. The majority of the studies in this goal category used images of log piles, taken from the log end side as input, with three notable exceptions: Fortin et al. [32] worked with images of entire logs taken from random angles, as their goal was instance segmentation of entire logs, and Schraml and Uhl [31], Decelle and Jalilian [33] used small images of individual log ends as the input.

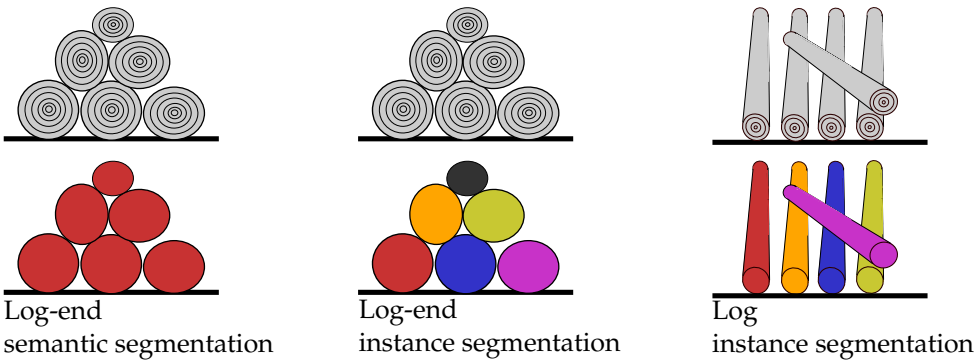

Figure 6. Log Segmentation: An illustration of the specific goals of the studies within the goal category, which includes studies that attempt to semantically segment log ends, studies that perform instance segmentation of log ends, and a study that performs instance segmentation of whole logs.

Table 6 shows the different types of models encountered in this goal category. The models used by the studies in this goal category mainly fall within two categories, namely clustering algorithms or artificial neural networks. The only exception is Chiryshev et al. [25], which used histograms of oriented gradients (HOG) for feature extraction and the Random Forest algorithm to classify individual pixels as "log face" or "not log face". Graph-cut segmentation is the most-used clustering algorithm among the studies within the log segmentation goal category, although it is implemented in different ways. Gutzeit and Voskamp [26] first used Haar cascades to perform a type of object detection of the log ends in the images, which were represented as circles. Graph-cut segmentation was then used to segment the image into the foreground, background, and unknown pixels. The Graph-cut segmentation was initiated, with the circular areas detected with Haar cascades as the foreground, and the bitmap was refined by capturing the pixels that were not included in the original circular object detection. Finally, the object detection results and semantic segmentation results were combined to yield an instance segmentation of the log ends. Gutzeit et al. [27,30] are studies written by the same authors and have a very similar approach. The algorithms developed by Gutzeit et al. [30] rely on the assumption that the log piles are always located in the center of the image. They initiated the graph-cut segmentation with portions of the bottom and top of the images as the background and a portion of the image center as the foreground. The yellow component of the image in RGB format and the value component of the image in hue, saturation, and value format (HSV) were extracted from the image and were used to adjust the initial weights of the graph-cut segmentation algorithm when it was applied to log end segmentation. Herbon et al. [28], Schraml and Uhl [31] are two studies that perform semantic segmentation of log ends in images using a clustering algorithm other than graph-cut segmentation. Herbon et al. [28] have designed an iterative model consisting of an initial stage of pixel classification using local binary patterns (LBP) in combination with HOG. The initial classifiers output a binary bitmap, which is passed on to the iterative pipeline of Gaussian mixture model (GMM) clustering, thresholding, watershed transform, distance transform, and another set of LBP and HOG classifications. The iterative pipeline is repeated until the set of detected objects does not change. Schraml and Uhl [31] worked with images of individual log ends and aimed to segment only the pixels corresponding to the log end. Schraml and Uhl [31] developed a three-stage region-growing algorithm that used two fast computable texture features to describe pixel blocks in images and the earth movers distance to measure the distance between neighboring pixel blocks. The three stages consisted of cluster initialization, where a fixed number of clusters are initialized that are equidistant from the image center; the growing procedure, where clusters are grown using fixed thresholds for the distance to neighboring pixels; and cluster merging and trimming, where the clusters are merged to form one continuous bitmap for the log end and erroneous pixels are removed using ellipse fitting.

Table 6. The choice of models in the studies that are part of the log segmentation goal category.

Model	Studies
Random forest classifier	[25]
Graph-cut segmentation	[26,27,30]
Other clustering algorithm	[28,31]
Artificial neural network	[29,32–35]

Five studies applied artificial neural networks in their segmentation models. Samdangdech and Phiphobmongkol [29] created a three-stage algorithm that detects the log pile using the single shot multi-box detector object detection network, segments the log ends in the detected pile using a fully convolutional VGG-16 network, and finally separates the individual log ends using the "connected component labeling" function that is part of the OpenCV library. Fortin et al. [32] compared the Mask R-CNN, Rotated Mask R-CNN, and Mask2Former neural networks at the task of instance segmentation of whole logs in images. The model that attained the highest mean average precision (mAP) was the visual

transformer-based network Mask2Former. Decelle and Jalilian [33] also compared a set of different networks, but at the task of semantic segmentation of log ends in images. They compared U-net Mask R-CNN, RefineNet, and SegNet. RefineNet attained the highest dice score. Praschl et al. [34] used a two-stage algorithm for instance segmentation of log ends in images. The YOLOv4 object detection network is used to detect the individual log ends in images and extract the image cut-outs from the computed bounding boxes. The individual cut-outs were then processed individually by a U-net segmentation network that detects the individual pixels in each cut-out that correspond to the log ends. Zheng et al. [35] used a modified version of the YOLOACT network to perform instance segmentation of the log ends in images. YOLOACT itself is a modified version of the object detection network YOLO, which is made for instance segmentation.

Table 7 shows the performance metrics of the best model in each study. Even though the goals for the studies in this category are close to identical, the evaluation metrics used are vastly different. Some studies only reported the performance metrics on the log detection part of their model, some only reported the performance of the pixel-wise segmentation stages of their model, and some studies reported the performance of both. Gutzeit et al. [30], Schraml and Uhl [31] measured performance in terms of pixel-wise false positives and false negatives and pixel-wise absolute error, respectively. These metrics are somewhat comparable, since false positives and false negatives must be a part of the absolute error. The reported metrics indicate that the region-growing algorithm of Schraml and Uhl [31] performed better than the graph-cut segmentation method of Gutzeit et al. [30]. However, their respective goals are somewhat different. Schraml and Uhl [31] had the goal of segmenting single log ends in images of individual log ends, whereas Gutzeit et al. [30] aimed to segment all the log ends in images of log piles. The latter was a more complex task than the former. Gutzeit and Voskamp [26], Gutzeit et al. [27], Samdangdech and Phiphobmongkol [29], Decelle and Jalilian [33], Praschl et al. [34] all reported the pixel-wise segmentation in terms of F1-score, which makes it possible to compare them. Gutzeit and Voskamp [26], Gutzeit et al. [27] used graph-cut segmentation-based models and attained F1-scores of 0.91 and 0.901, respectively, for pixel-wise segmentation. The deep learning-based methods presented by Samdangdech and Phiphobmongkol [29], Decelle and Jalilian [33], Praschl et al. [34] attained F1-scores of 0.970, 0.976, and 0.944, respectively, for pixel-wise segmentation, which is 0.03 to 0.07 higher than the clustering-based methods. The RefineNet model of Decelle and Jalilian [33] performed approximately equally well as the SSD model of Samdangdech and Phiphobmongkol [29] in terms of F1-score. The RefineNet and SSD networks outperformed the YOLO model of Praschl et al. [34] in terms of the segmentation F1-score. The remaining studies all evaluated their models in terms of the log detection rate or log detection error rate. Chiryshev et al. [25], Herbon et al. [28] both evaluated their model's performance with the recall and false positive rate. They reported equal recall scores, but Chiryshev et al. [25] reported a false positive rate that was three times higher than Herbon et al. [28]. Chiryshev et al. [25] also reported the F1-score for their log detection rate, which was nearly identical to the log detection rate F1-score reported by Samdangdech and Phiphobmongkol [29] of 0.990. Fortin et al. [32], Zheng et al. [35] are the only two studies that reported performance in terms of mean average precision at the intersection over union (IoU) threshold 50 (mAP_{50}). Although it may seem as though the model of Zheng et al. [35] outperforms the model of Fortin et al. [32] in terms of mAP_{50}, it should be noted that Zheng et al. [35] had the goal of segmenting the log ends on loading trucks in images with fairly homogeneous backgrounds, and Fortin et al. [32] had the goal of segmenting whole logs in images with heterogeneous backgrounds. Hence, the task of Fortin et al. [32] was more complex, and their performance cannot be directly compared to that of the model developed by Zheng et al. [35].

Concerning autonomy, all the studies developed models that are intended to work without any direct operator intervention. That being said, the type of input data the models use indicates what level of indirect intervention they require. The model developed by Fortin et al. [32] was the only model that was intended to segment entire logs independent

of the orientation and background conditions. The remaining models require the logs to be stacked in a pile with the log ends facing the camera. This would make the model developed by Fortin et al. [32] the most autonomous. Among the log end segmentation studies, there are different levels of autonomy. The studies that used clustering methods in their model require the log piles to be centered in the image and the background to be homogeneous above and below the pile. The studies that used deep neural networks do not specify any of these requirements explicitly, so if the data used to train said networks included a variety of background conditions, such requirements could be mitigated. The least-autonomous model in this category is the one developed by Herbon et al. [28], as it uses small rectangular images of singular log ends as the input. This means that the step of extracting the cut-out images of singular log ends either has to be performed manually or by another model.

Table 7. Performance among the studies in the log segmentation goal category. The table only shows performance metrics related to segmentation and log detection.

Study	Performance	Metric
[30]	5.35%, 2.38%	Pixel-wise relative false positives and false negatives.
[31]	5.10%,	Pixel-wise relative absolute error.
[26]	0.910, 0.151	Pixel-wise F1-score and log detection error.
[27]	0.901	Pixel-wise F1-score.[1]
[33]	0.976, 98.7%,	Pixel-wise F1-score and accuracy.
[34]	0.944%, 93.4%	Pixel-wise F1-score [2] and accuracy.
[29]	0.990, 0.970	F1-score for log end detection and pixel-wise segmentation.
[28]	99.3%, 0.360%	Recall and false positive rate for log end detection.
[25]	0.986%, 99.3%, 1.10%	F1-score recall and false positive rate for log end detection.
[32]	0.575,	mAP_{50}.
[35]	0.791	mAP_{50}.

[1] Calculated from true positives, false positives, and false negatives reported by Gutzeit et al. [27]. [2] Calculated from true positives, false positives, and false negatives reported by Praschl et al. [34].

4. Discussion

4.1. Input Data

Figure 7 shows the number of studies reviewed in Section 3 that used point clouds and images as inputs. Between 1991 and 2023, 27 studies used images as model input for their log scaling or grading models. Point cloud data were used as model input in six studies between 2006 and 2023.

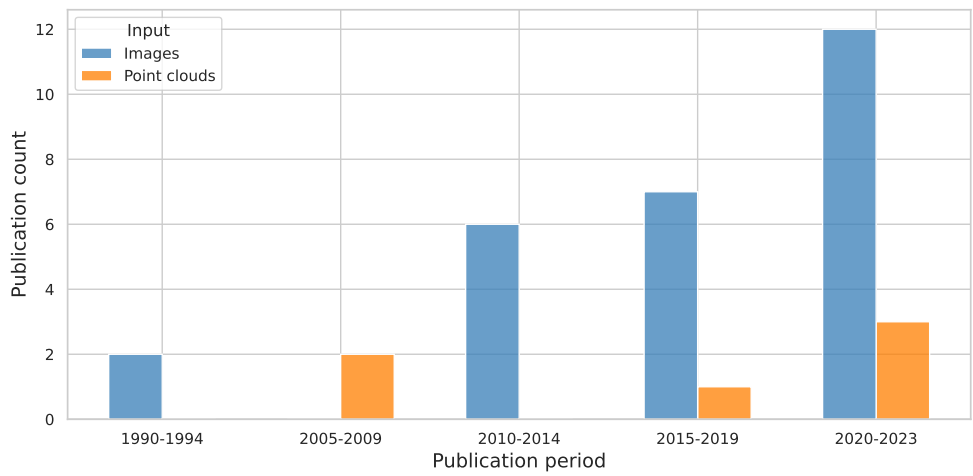

Figure 7. Bar plot showing how many studies were published in five-year intervals using images or point clouds as input data. The period of 1995–1999 is omitted from the plot because none of the studies included in this review were published during that time.

Studies with point clouds as input used laser scanners to sample the point cloud datasets. From Figure 7, it is clear that the number of publications on the topic handled in this review has increased over time, and that the use of cameras has increased more than the use of laser scanners. A possible reason for this may be that cameras have become more affordable than laser scanners, and that stereovision cameras can provide both images and point cloud data as input to the model. All the studies in the individual log scaling goal category [13–15] and four of the nine studies in the log pile scaling goal category [17,18,22] used images from different angles and stereovision techniques to generate point cloud representations of the logs. This supports the argument that stereo cameras can replace laser scanners in some scenarios where a point cloud representation of logs is involved. However, five of the six studies in the grade log sides goal category [8–12] used laser scanners instead of stereo cameras to generate point cloud representations of the logs, three of which were published after 2015 [9,11,12]. This indicates that there are instances where laser scanners are preferred over cameras to generate point cloud representations of logs. The advantage of using laser scanners over stereo cameras for generating point clouds is that laser scanners sample the point cloud representation directly, whereas stereo cameras require computationally expensive algorithms to generate point cloud representations, making them less computationally efficient.

4.2. Goal Categories

In the Results section (Section 3), the studies were also partitioned into goal categories to facilitate comparisons between models designed for similar goals. There were four studies in the log end grading category, six studies in the log side grading category, three studies in the individual log scaling category, nine studies in the log pile scaling category, and eleven studies in the log segmentation category. Based on Figure 8 we can see certain trends in the choices of goals among the studies over time. The goal category of log side grading has been present at an approximately constant rate, except for the period of 2010–2014, where it was not present. Studies in the log segmentation goal category started being published during the period of 2010–2014 and have been published regularly since. Studies with the goal of log pile scaling were published once in 1991 and became more frequent from 2015 to 2023, where three of the publications were made in 2023 alone. Studies have been published spuriously in the other goal categories, without a clear trend.

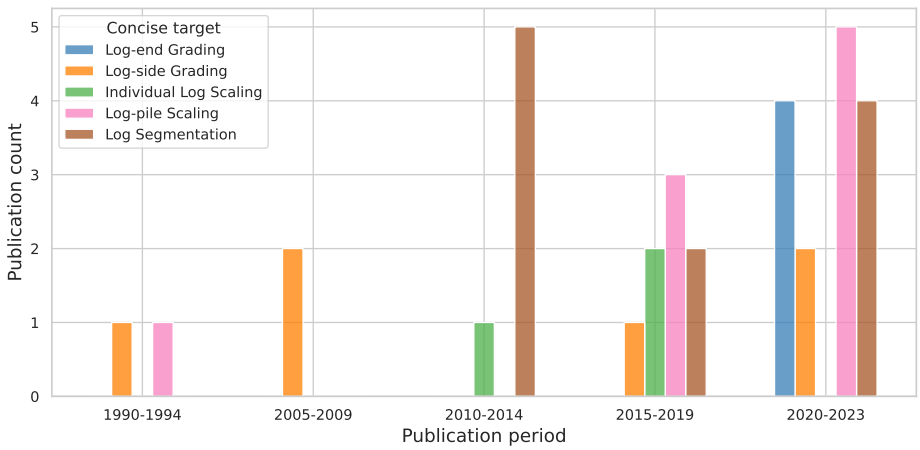

Figure 8. Bar plot showing how many studies were published in five-year intervals within the five goal categories: log end grading, log side grading, individual log scaling, log pile scaling, and log segmentation. The period of 1995–1999 is omitted, as none of the studies included in this review were published during that time.

4.3. Model Usage

To visualize the general trends in model usage, each study was categorized into one of the three model categories and was aggregated based on its publication year. The models

developed in the studies were regarded as deep learning models if they applied artificial neural networks, other machine learning models if they applied a machine learning model other than artificial neural networks, or classical image processing if they used classical image processing algorithms. Figure 9 shows the trend in model usage over the years. Classical image processing models were overall the most populous model category over the entire period, as they were applied in 15 publications between 1991 and 2021. Other machine learning models were used in seven publications between 2010 and 2018. Since 2018, it seems that the use of deep learning models has taken over for machine learning models in the context of scaling and grading logs using external imaging. Deep learning models were used in 11 studies from 2018 to 2023, where 6 of them were published in 2023 alone. If this trend continues, one would expect deep learning models to become more popular in the future to scale and grade logs using computer vision.

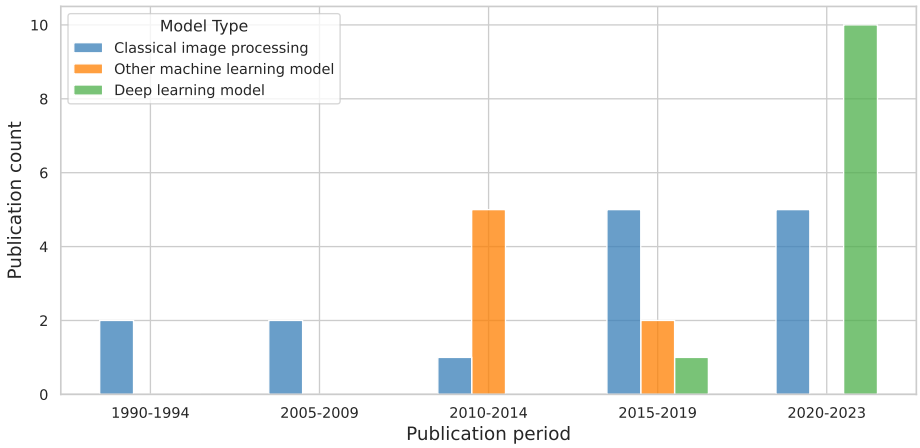

Figure 9. Bar plot showing how many studies were published during different five-year intervals applying deep learning models, other machine learning models, or classical image processing. The period of 1995–1999 is omitted from the plot because none of the studies included in this review were published during that itme.

The use of different model types within the specific goal categories, described in detail in Section 3, is summarized in Figure 10. Classical image processing models were used fairly often in all goal categories except for the log segmentation category. Other machine learning models are used almost exclusively in the log segmentation category, except for one study in the log pile scaling category. Deep learning models are used mostly in the log segmentation and log pile scaling categories but are used in two studies in the log end grading and log side grading categories as well. By considering both Figures 8 and 10, one could reason that part of the reason that the log pile scaling and log segmentation goal categories have been the most common in recent times is that they are the two goal categories that deep learning models perform well.

The log pile scaling and log segmentation goal categories were the only categories where it was meaningful to compare models based on performance. In the other goal categories, the specific targets of the studies were too different, quantitative performance metrics were not reported, or the quantitative performance metrics reported were not similar enough. In the log pile scaling goal category, it was found that the deep learning models outperformed both the classical image processing algorithms and the other machine learning algorithms. The model developed by Zheng et al. [22] performed best at estimating the wood volume of log piles, and the model developed by Li et al. [20] performed best at estimating the diameter distribution of the log pile. In the log segmentation goal category, it was not possible to specify which model had the highest performance since they used such disparate evaluation metrics. In general, the deep learning models also outperformed other machine learning models and classical image processing algorithms in this log segmentation goal category. There was no clear distinction in performance between the

machine-learning models and the classical image-processing models based on the results reported in the studies.

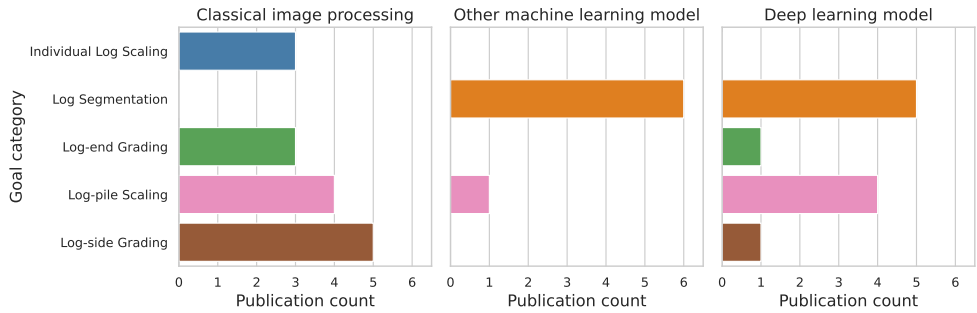

Figure 10. Three bar plots showing how often deep learning, machine learning, or image processing models were used in the different goal categories.

4.4. Adoption and Future Perspectives

It has been established that deep learning models have been increasingly used in the studies reviewed in this paper. This is in line with the general trend in computer vision research, where deep learning models have been shown to outperform other machine learning models and classical image processing algorithms in many tasks. Since the launch of ChatGPT in 2022, large language models have been widely adopted in many industries and are being used in the most imaginative ways to automate tasks and processes. This increase in the usage of large deep learning models has led to a subsequent increase in demand for computational capacity, efficient AI-dedicated hardware, and off-the-shelf software solutions for people without a background in AI. The improvement of deep learning models in natural language processing can lead to further development of deep learning models within computer vision. It is, however, expected that the increase in the availability of dedicated hardware and off-the-shelf software solutions will make it easier and cheaper to apply deep learning models for computer vision tasks such as grading and scaling of logs.

For machine learning and computer vision technology to be adopted in the logging industry, there has to be an economic incentive. Almost all the studies included in this review state that reducing the labor hours required for the scaling and grading of logs is a strong motivating factor for their research. Companies such as Italy-based Microtec (https://microtec.eu/) (accessed on 28 February 2024) and Danish Dralle (https://www.dralle.dk/) (accessed on 28 February 2024) have developed commercial products used for log scaling and grading, and many of their products are already computer vision-based. Both of these observations further strengthen the argument that there is an economic incentive to apply more advanced computer vision models to automate the process of log scaling and grading.

5. Conclusions

A comparative literature review is a tool used in academia to acquire an overview of the current state-of-the-art methods within a specific field. A drawback in the approach of a review of academic literature in the field of log scaling and grading is that much of the research and development is performed in the private sector and is not published as scientific literature. Companies such as the aforementioned Microtec and Dralle most likely perform some research on computer vision, since they develop commercial products used for log scaling and grading. However, their work cannot be included here, as their methods are not described in any conference studies or peer-reviewed journal entries.

Since this review has excluded studies that use internal cross-sectional imaging, it is biased against them. This affects the statements made regarding the trends in input data given in Section 4.1. It was stated that the use of cameras has increased more than the use of laser scanners for collecting data about logs, and that this might be due to the rising

popularity of stereovision systems. Internal cross-sectional images are another input type that is popular in log-grading scenarios, but since it was excluded from this review, one cannot exclude the possibility that internal cross-sectional imaging techniques may have replaced laser scanners for tasks related to log grading.

This literature review has assessed the efficacy of image-processing techniques and machine-learning models in the context of the grading and scaling of wood logs. Four searches were conducted using four scientific databases, yielding a total of 1288 results, which were narrowed down to 33 relevant studies. The studies were categorized according to their goals into log end grading, log side grading, individual log scaling, log pile scaling, and log segmentation. The studies were compared based on the input used, choice of model, model performance, and level of autonomy. The models encountered were subdivided into three types: classical image-processing algorithms, deep learning models, and other machine learning models. It was rarely possible to compare the performances of the models developed in the individual studies, even within the same goal category. This was because the specific goals or the metrics that were used to evaluate the models were too different to be compared. In the log pile scaling goal category, the model developed by Zheng et al. [22] attained the highest performance at estimating the wood volume of log piles, and the model developed by Li et al. [20] performed best at estimating the diameter distribution of the log pile.

In general, deep learning models performed better than classical image-processing techniques and other machine learning models in the log pile scaling and log segmentation goal categories. The trends in the input data showed that cameras have become more popular over time compared to laser scanners. It was theorized that part of the reason for this could be that stereovision systems are taking over for laser scanners for sampling point cloud datasets. Trends in model types showed that classical image processing algorithms were present to some degree at all time intervals and in all goal categories except for log segmentation. Deep learning models started being used in studies in 2018 and were the most common between 2020 and 2023. They were used among the studies in every goal category except for individual log scaling. If the current trend continues, one would expect deep learning models to become more popular in the future. Other machine learning models were used in studies published between 2010 and 2018, and these studies mostly fell into the log segmentation goal category. The log pile scaling and log segmentation goal categories were the two most populous goal categories overall. However, these two goal categories might be favored among studies because deep learning models have become more popular in recent times, and deep learning models excel at both log pile scaling and log segmentation.

Author Contributions: Y.J.S.: conceptualization, methodology, data curation, formal analysis, visualization, writing—original draft preparation, and writing—review and editing. O.T.: conceptualization, methodology, visualization, and writing—review and editing. C.M.F. and K.H.L.: visualization, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: Internal funding was provided by the Norwegian University of Life Sciences for the author's PhD position as part of the Research-driven Innovation Centre SFI SmartForest (project number 309671) financed by the Research Council of Norway.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank Terje Gobakken at the Norwegian University of Life Sciences for productive discussions about this literature review.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Forestry and Logging Global Market Report 2023—Product Image Forestry and Logging Global Market Report 2023. 2023. Available online: https://www.researchandmarkets.com/reports/5781398/forestry-logging-global-market-report (accessed on 14 February 2024).
2. Wood Exports by Country in US$ Thousand 2021. 2023. Available online: https://wits.worldbank.org/CountryProfile/en/Country/WLD/Year/LTST/TradeFlow/Export/Partner/by-country/Product/44-49_Wood (accessed on 14 February 2024).
3. Carratù, M.; Gallo, V.; Liguori, C.; Pietrosanto, A.; O'Nils, M.; Lundgren, J. A CNN-based approach to measure wood quality in timber bundle images. In Proceedings of the 2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Glasgow, UK, 17–20 May 2021; pp. 1–6. [CrossRef]
4. Cao, X.; Li, G. An Effective Method of Wood Crack Trace and Quantity Detection Based on Digital Image Processing Technology. In Proceedings of the 2021 13th International Conference on Machine Learning and Computing, Shenzhen, China, 26 February–1 March 2021; ICMLC 2021; pp. 304–309. [CrossRef]
5. Du, W.; Xi, Y.; Harada, K.; Zhang, Y.; Nagashima, K.; Qiao, Z. Improved hough transform and total variation algorithms for features extraction of wood. *Forests* **2021**, *12*, 466. [CrossRef]
6. Decelle, R.; Ngo, P.; Debled-Rennesson, I.; Mothe, F.; Longuetaud, F. Pith Estimation on Tree Log End Images. In *Proceedings of the Reproducible Research in Pattern Recognition*; Kerautret, B., Colom, M., Krähenbühl, A., Lopresti, D., Monasse, P., Talbot, H., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 101–120.
7. Lee, S.C.; Qian, G.S.; Chen, J.B.; Sun, Y.W.; Hay, D.A. Determining a maximum value yield of a log using an optical log scanner. In Proceedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, Maui, HI, USA, 3–6 June 1991; pp. 747–748.
8. Thomas, L.; Mili, L. A Robust GM-Estimator for the Automated Detection of External Defects on Barked Hardwood Logs and Stems. *IEEE Trans. Signal Process.* **2007**, *55*, 3568–3576. [CrossRef]
9. Nguyen, V.T.; Kerautret, B.; Debled-Rennesson, I.; Colin, F.; Piboule, A.; Constant, T. Algorithms and Implementation for Segmenting Tree Log Surface Defects. In *Proceedings of the Reproducible Research in Pattern Recognition*; Kerautret, B., Colom, M., Monasse, P., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 150–166.
10. Thomas, L.; Mili, L.; Thomas, E.; Shaffer, C. Defect detection on hardwood logs using laser scanning. *Wood Fiber Sci.* **2006**, *38*, 682–695.
11. Zolotarev, F.; Eerola, T.; Lensu, L.; Kälviäinen, H.; Helin, T.; Haario, H.; Kauppi, T.; Heikkinen, J. Modelling internal knot distribution using external log features. *Comput. Electron. Agric.* **2020**, *179*, 105795. [CrossRef]
12. Khazem, S.; Fix, J.; Pradalier, C. Improving Knot Prediction in Wood Logs with Longitudinal Feature Propagation. In *Proceedings of the International Conference on Computer Vision Systems*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 169–180.
13. Kruglov, A.; Chiryshev, Y. *Detection and Tracking of the Objects in Real Time Applied to the Problem of the Log Volume Estimation;* CEUR-WS.org: Aachen, Germany, 2016; Volume 1710, pp. 114–123.
14. Kalmari, J.; Kulovesi, J.; Visala, A. *Measuring Log Length Using Computer Stereo Vision in Harvester Head*; Elsevier: Amsterdam, The Netherlands, 2011; Volume 44, pp. 2226–2229. [CrossRef]
15. Yang, L.; Diago, L.; Hagiwara, I. Real-time 3D reconstruction of wood logs in a production line. *J. Food Sci. Technol.* **2015**, *8*, 49–57. [CrossRef]
16. Galsgaard, B.; Lundtoft, D.; Nikolov, I.; Nasrollahi, K.; Moeslund, T. *Circular Hough Transform and Local Circularity Measure for Weight Estimation of a Graph-Cut Based Wood Stack Measurement*; IEEE: Waikoloa, HI, USA, 2015; pp. 686–693. [CrossRef]
17. Herbon, C.; Tonnies, K.D.; Otte, B.; Stock, B. *Mobile 3D Wood Pile Surveying*; IEEE: Tokyo, Japan, 2015; pp. 422–425. [CrossRef]
18. Kruglov, A.; Shishko, E.; Kozhova, V.; Zavada, S. Testing of the Round Timber Volume Measurement Method. In Proceedings of the 2017 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO), Prague, Czech Republic, 20–22 May 2017; pp. 232–236. [CrossRef]
19. Correia, B.; Davies, R.; Carvalho, F.; Rodrigues, F. *Toros—An Image Processing System for Measuring Consignments of Wood;* Society of Photo-Optical Instrumentation Engineers (SPIE): San Diego, CA, USA, 1991; Volume 1567, pp. 15–24. [CrossRef]
20. Li, H.; Liu, J.; Wang, D. A Fast Instance Segmentation Technique for Log End Faces Based on Metric Learning. *Forests* **2023**, *14*, 795. [CrossRef]
21. Carratù, M.; Gallo, V.; Liguori, C.; Pietrosanto, A.; O'Nils, M.; Lundgren, J. An innovative method for log diameter measurements based on deep learning. In Proceedings of the 2023 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Kuala Lumpur, Malaysia, 22–25 May 2023; pp. 1–6.
22. Zheng, J.; Li, S.; Zhang, S.; Kong, L.; Zhigang, D. Research on Volume Measurement of Logs Based on Embedded Application. *IEEE Access* **2023**, *11*, 19186–19201. [CrossRef]
23. Carratù, M.; Gallo, V.; Liguori, C.; Lundgren, J.; O'Nils, M.; Pietrosanto, A. Vision-Based System for Measuring the Diameter of Wood Logs. *IEEE Open J. Instrum. Meas.* **2023**, *2*, 1–12. [CrossRef]
24. Martí, F.; Forkan, A.R.M.; Jayaraman, P.P.; McCarthy, C.; Ghaderi, H. LogLiDAR: An Internet of Things Solution for Counting and Scaling Logs. In Proceedings of the 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Kassel, Germany, 22–26 March 2021; pp. 413–415. [CrossRef]

25. Chiryshev, Y.V.; Kruglov, A.V.; Atamanova, A.S. Automatic Detection of Round Timber in Digital Images Using Random Decision Forests Algorithm. In Proceedings of the 2018 International Conference on Control and Computer Vision, Singapore, 15–18 June 2018; ICCCV '18; pp. 39–44. [CrossRef]
26. Gutzeit, E.; Voskamp, J. Automatic Segmentation of Wood Logs by Combining Detection and Segmentation. In *Proceedings of the Advances in Visual Computing*; Bebis, G., Boyle, R., Parvin, B., Koracin, D., Fowlkes, C., Wang, S., Choi, M.H., Mantler, S., Schulze, J., Acevedo, D., et al., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 252–261.
27. Gutzeit, E.; Ohl, S.; Voskamp, J.; Kuijper, A.; Urban, B. Automatic Wood Log Segmentation Using Graph Cuts. In *Proceedings of the Computer Vision, Imaging and Computer Graphics: Theory and Applications*; Richard, P., Braz, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; Volume 229, pp. 96–109.
28. Herbon, C.; Tönnies, K.; Stock, B. Detection and Segmentation of Clustered Objects by Using Iterative Classification, Segmentation, and Gaussian Mixture Models and Application to Wood Log Detection. In *Proceedings of the Pattern Recognition*; Jiang, X., Hornegger, J., Koch, R., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 354–364.
29. Samdangdech, N.; Phiphobmongkol, S. *Log-End Cut-Area Detection in Images Taken from Rear End of Eucalyptus Timber Trucks*; IEEE: Nakhon Pathom, Thailand, 2018. [CrossRef]
30. Gutzeit, E.; Ohl, S.; Kuijper, A.; Voskamp, J.; Urban, B. *Setting Graph Cut Weights for Automatic Foreground Extraction in Wood Log Images*; SciTePress: Angers, France, 2010; Volume 2, pp. 60–67.
31. Schraml, R.; Uhl, A. Similarity Based Cross-Section Segmentation in Rough Log End Images. In *Proceedings of the Artificial Intelligence Applications and Innovations*; Iliadis, L., Maglogiannis, I., Papadopoulos, H., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2014; pp. 614–623.
32. Fortin, J.M.; Gamache, O.; Grondin, V.; Pomerleau, F.; Giguère, P. Instance segmentation for autonomous log grasping in forestry operations. In Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 23–27 October 2022; pp. 6064–6071.
33. Decelle, R.; Jalilian, E. Neural networks for cross-section segmentation in raw images of log ends. In Proceedings of the 2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS), Genova, Italy, 9–11 December 2020; pp. 131–137.
34. Praschl, C.; Auersperg-Castell, P.; Forster-Heinlein, B.; Zwettler, G.A. Segmentation and Multi-facet Classification of Individual Logs in Wooden Piles. In *Proceedings of the International Conference on Computer Aided Systems Theory*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 460–467.
35. Zheng, J.; Zheng, J.; Zhang, S.; Yu, H.; Kong, L.; Zhigang, D. Segmentation Method for Whole Vehicle Wood Detection Based on Improved YOLACT Instance Segmentation Model. *IEEE Access* **2023**, *11*, 81434–81448. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.