

A Fast and Robust Artificial Intelligence Technique for Wood Knot Detection

Dercilio Junior Verly Lopes,^{a,*} Gabrielly dos Santos Bobadilha,^a and Karl Michael Grebner^b

This study reports the feasibility of using deep convolutional neural networks (CNN), for automatically detecting knots on the surface of wood with high speed and accuracy. A limited dataset of 921 images were photographed in different contexts and divided into 80:20 ratio for training and validation, respectively. The “You only look once” (YoloV3) CNN-based architecture was adopted for training the neural network. The Adam gradient descent optimizer algorithm was used to iteratively minimize the generalized intersection-over-union loss function. Knots on the surface of wood were manually annotated. Images and annotations were analyzed by a stack of convolutional and fully connected layers with skipped connections. After training, model checkpoint was created and inferences on the validation set were made. The quality of results was assessed by several metrics: precision, recall, F1-score, average precision, and precision x recall curve. Results indicated that YoloV3 provided knot detection time of approximately 0.0102 s per knot with a relatively low false positive and false negative ratios. Precision, recall, f1-score metrics reached 0.77, 0.79, and 0.78, respectively. The average precision was 80%. With an adequate number of images, it is possible to improve this tool for use within sawmills in the forms of both workstation and mobile device applications.

Keywords: Object-detection; Bounding box; Knots; Boards; Fast; YoloV3

Contact information: a: Department of Sustainable Bioproducts/Forest and Wildlife Research Center (FWRC), Mississippi State University, Starkville, MS - 39762-9820 - USA; b: Department of Mechanical Engineering, Mississippi State University, Starkville, MS – 39762-9820.

* Corresponding author: dvl23@msstate.edu

INTRODUCTION

According to the Food and Agriculture Organization of the United Nation - FAO (FAO 2020), the United States is the second largest producer of sawn wood in the world. Traditionally, lumber grading is done by trained and experienced human graders. In fact, sawmills are increasingly requiring intimate knowledge of the condition of raw material to effectively manage their inventory and improve business performance (Rudakov 2018). Over the years, automated sorting machines have been implemented to expedite the process in order to increase production and accuracy (Kontzer 2019).

In general, wood is a central material for furniture, fine-working, and civil construction. In the first two cases, visual appearance is important. Therefore, wood defects negatively affect the aesthetic appearance of the final products. In the latter case, defects play a critical role in degrading the mechanical properties of wood, including detrimental effects on modulus of rupture (MOR) and modulus of elasticity (MOE) (Zhong *et al.* 2012; Rocha *et al.* 2018).

According to Nasir and Cool (2018), most research efforts in sawing have been concentrated in advancements of primary and secondary wood processing with band and circular saws. Nasir and Cool (2018) have advocated for new prediction techniques during processing for increasing optimization and overall sawmill yield. Prediction, monitoring, and controlling artificial intelligence techniques have the capability of helping the wood machining field transition to industry 4.0.

In a study conducted by Nasir and Cool (2020), vibration signals combined with self-organizing maps (SOM) were fed into the adaptive neuro-fuzzy inference system (ANFIS) and multi-layer perceptron (MLP) for cutting power and waviness prediction by using a circular saw in climb cutting. Both methods obtained nearly perfection in prediction of average cutting power and surface waviness for the testing set. Furthermore, a vibration signal could be successfully used for online monitoring of cutting power and surface waviness.

Several investigations have been conducted on machine-vision-enabled defect detection. Schmoldt *et al.* (1997) investigated using a multi-layer perceptron neural network to identify and locate internal log defects. In addition, techniques such as Gabor or wavelet can also be used to scan products and report defects (Lampinen *et al.* 1998; Cetiner *et al.* 2016). However, expensive equipment and several steps are necessary for defect classification and detection. Moreover, these techniques are not fast, nor are they accurate enough to meet the required speed in sawmill processing.

Machine-learning (ML) studies for defect detection have been developed by Gu *et al.* (2010), Mahram *et al.* (2012), and Urbonas *et al.* (2019). These authors used ML to classify defects on wood surfaces using support vector machine (SVM) and feature extraction with gray level co-occurrence, local binary pattern (LBP), principal component analysis (PCA), and linear discriminant analysis (LDA) techniques.

More recently, state-of-the-art classification, detection, and segmentation of images and videos with important theoretical and practical achievements have been achieved by leveraging deep convolutional neural networks (CNN) (Gu *et al.* 2017). Inspired by the breakthrough of CNNs in object-detection, the present work investigates using one of the latest deep learning CNN approaches to perform real-time object detection of wood surface defects. In the wood science field, image data are scarce and are expensive and time-consuming to acquire. Knots on a wood surface vary in size, location, and type. According to Cao *et al.* (2018), knots can be classified as encased or intergrown and further extended to sound or unsound classifications.

The overall goal of this work is to advance the wood science field by demonstrating that artificial intelligence can be used for defect detection. This is done by employing a real-time object-detection algorithm called You Only Look Once (YoloV3) (Redmon and Farhadi 2018). The hypothesis to be evaluated in this work is: compared to common methods that are slow and expensive, convolutional neural networks are capable of classifying, and locating wood surface knots automatically, accurately, fast, and reliably.

EXPERIMENTAL

Materials

Usually, in sawmill processing, line bucking and merchandising play a critical role in optimization. Those steps are responsible for separating defective boards to increase clear cuttings and help in the sorting process. To that end, the authors created a small image

dataset that included both Southern yellow pine (*Pinus* spp.) defect and defect-free boards of varied sizes. Images were taken by the Department of Sustainable Bioproducts at Mississippi State University by using a high-definition commodity camera. The surface of the boards also was photographed in several different contexts in order to improve model robustness (Fig. 1). A total of 921 high-definition images were obtained.



Fig. 1. Southern yellow pine boards with knots in two different contexts, namely, in the field (left) and as an isolated image with white background (right).

Methods

Data annotation

Image annotation is a crucial step for object detection. The main objective of annotation is to label the position and defect class (type of defect) on the boards. For this stage, the work was carried out with open-source software written in Python 3.6.7 called *labelImg* (Tzutalin 2020). With this algorithm and knowledge, it was possible to select, locate, and label areas on the image that contained knots. As previously mentioned, images with no defects were not annotated but were included in the neural network training. The labeling algorithm allows exporting location of knots in .txt format, which is the format needed by YoloV3. The .txt format contained the defect class and pixel coordinates of the knots in the image. The labeling specifications included: a) one row per defect object, b) each row consisted of the data: *class*, *x_center*, *y_center*, *width*, *height* of a box that enclosed the defect; c) box coordinates were normalized between 0 – 1; and d) class number was zero-indexed, *i.e.*, started from 0 (zero).

Detection model

The Yolo (You only look once) CNN-based object detection algorithm is an architecture designed for real-time image processing. The first release of the algorithm was made by Redmon *et al.* (2015) in which the authors framed object detection as a regression problem for identifying spatially separated bounding boxes and associated class probabilities. The second Yolo release was made by Redmon and Farhadi (2016) that included a series of improvements, namely batch normalization, high resolution classifier, convolutional layers with bounding boxes, dimension clusters, direct location prediction, fine-grained features, and multi-scale training. The YoloV3 architecture was released by Redmon and Farhadi (2018) with a feature extractor called Darknet-53 that uses 53 convolutional layers and adds the idea of skipped connection from ResNet architecture (He *et al.* 2015). Darknet-53 is much more powerful than Darknet-19 but is still more efficient than ResNet-101 or ResNet-152 backbones (Redmon and Farhadi 2016). It also accepts

images with different sizes. YOLOv3 uses multi-scale prediction and multiple scale feature maps, which leads to better accuracy for target detection. Figure 2 shows the structure of the architecture.

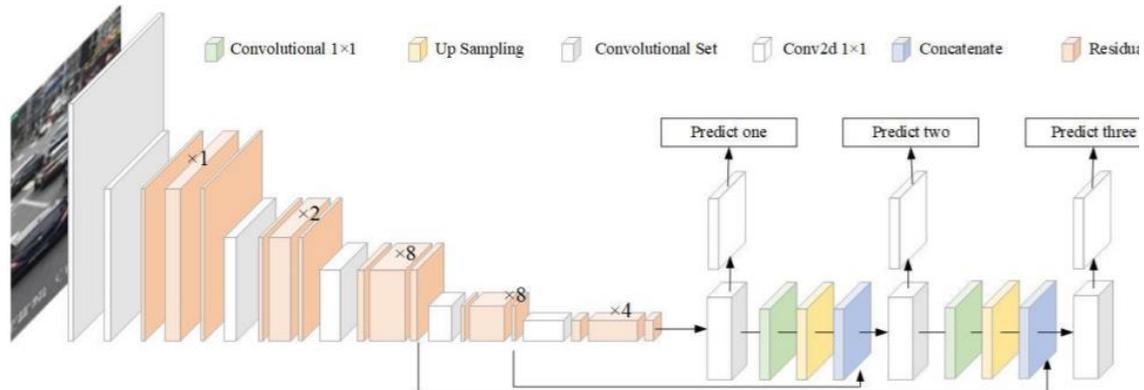


Fig. 2. YoloV3 detailed architecture. Adapted from Redmon and Farhadi (2018) and Mao *et al.* (2019).

By default, each YoloV3 layer has 255 outputs filters: 85 values per anchor (4 box coordinates + 1 object confidence + 80 class confidences, times 3 anchors). The settings to filters were updated, = $[5 + n] * 3$, which resulted in 18 filters. The entire network was trained from scratch; *i.e.*, pre-trained weights were not employed. A standard 608 pixels x 608 pixels input image was used, with standard anchor boxes ([116 x 90, 156 x 198, 373 x 326], [30 x 61, 62 x 45, 59 x 119], [10 x 13, 16 x 30, 33 x 23]) in order to detect large, medium, and small objects in the images, respectively. The object threshold was set to 0.5.

Experimental evaluation

The training was performed on a CentOS 7 Linux computer with an Intel i9-9920X CPU @3.5 GHz accelerated by 4x Nvidia RTX 2080Ti with each GPU having 4,352 CUDA cores and 11Gb of memory. The YoloV3 CNN was implemented in PyTorch 1.5.1 and torchvision 0.6.1. The implementation was derived from Jocher *et al.* (2020). A batch size of 96 was used, with an adaptive momentum estimation with initial learning rate of 0.01 to iteratively optimize the generalized intersection-over-union cost function as described in Rezatofighi *et al.* (2019). The weights and biases were updated through the backpropagation algorithm with a maximum number of batches of 4000 over 1500 epochs. Training required approximately 12 hours on 737 training images. The model was validated on 184 images. Data augmentation was performed on-the-fly by rotating and translating images. The CNN-based object detection quality metrics included precision, recall, F1-score, precision-recall curve, and average precision at 0.5 intersection-over-union (IOU) threshold. The precision x recall curve is a method to evaluate the performance of an object detection. A detector is considered satisfactory if its precision stays high as recall increases, which means that if the threshold varies, the precision and recall will still be high. The metrics are given by Eq. 1, 2, 3, and 4.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1\text{-Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

where TP is the true positive, correct detection (detection with $\text{IOU} \geq \text{threshold}$, FP is the false positive, a wrong detection (detection with $\text{IOU} \leq \text{threshold}$, and FN is the false negative, a ground truth not detected.

The intersection-over-union (IOU) function is the ratio of the area of intersection to the area of union between detected and the ground-truth bounding box(es). Our knot detection system was considered to work satisfactorily if:

$$\text{IOU} = \frac{\text{Area of overlap}}{\text{Area of Union}} > 0.5 \quad (4)$$

Better understanding is provided in Fig. 3.

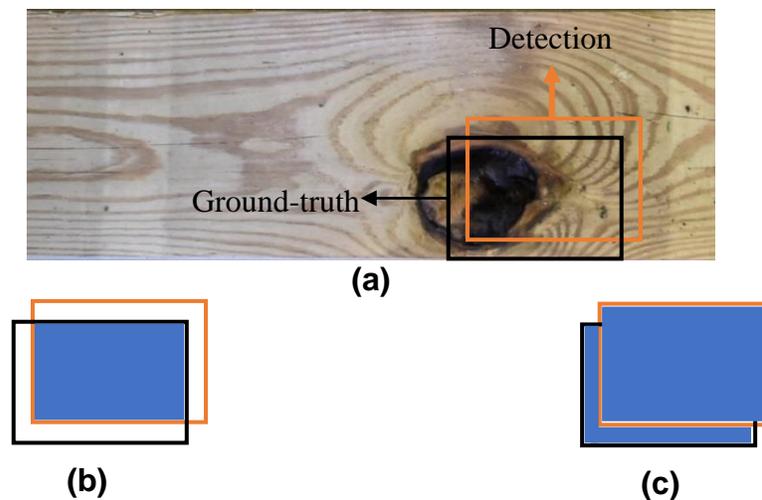


Fig. 3. Intersection-over-union for knots detection. (a) Ground-truth bounding box and detected knot bounding box. (b) intersection of boxes, and (c) union of boxes

Three videos were recorded of knotted boards and the YoloV3 model was applied the trained with them. The video files were recorded by commodity smartphones (the authors have no access to sawmill quality scanning equipment). The videos had a frame rate of 30 frames per second, a frame width of 1920 pixels, and frame height of 1080 pixels. Lopes *et al.* (2020a; 2020b).

RESULTS AND DISCUSSION

The generalized intersection-over-union loss function of the object detection algorithm is plotted for 1500 epochs (Fig. 4). The loss function is defined as the difference between the output and target variable overlapping and is a good indicator of the quality of a trained model. The training was done until the cost function remained below 1.0 to ensure model convergence. The loss values are initially very large and slowly decay to a somewhat constant value below 1.0 from epoch 1200 and onward.

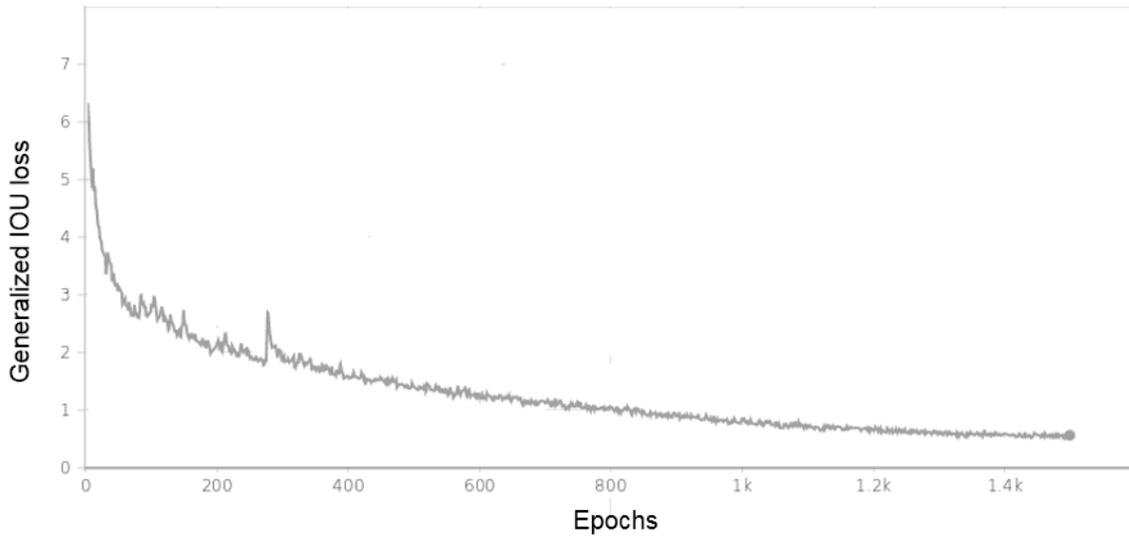


Fig. 4. Loss function for YoloV3 trained on wood surface knots for 1500 epochs

The results of knots detection quality evaluation are presented in Table 1 and Precision x Recall curve can be seen in Fig. 5.

Table 1. Knots Detection Metrics

# of images	# of knots	Avg. # of knots per image	# of knots detect	# of not knots detected	P.	R.	F1-score	AP
185	427	2.3	351	46	0.77	0.79	0.78	0.8

= Number; Avg = Average; P. = Precision; R. = Recall; AP = Average precision

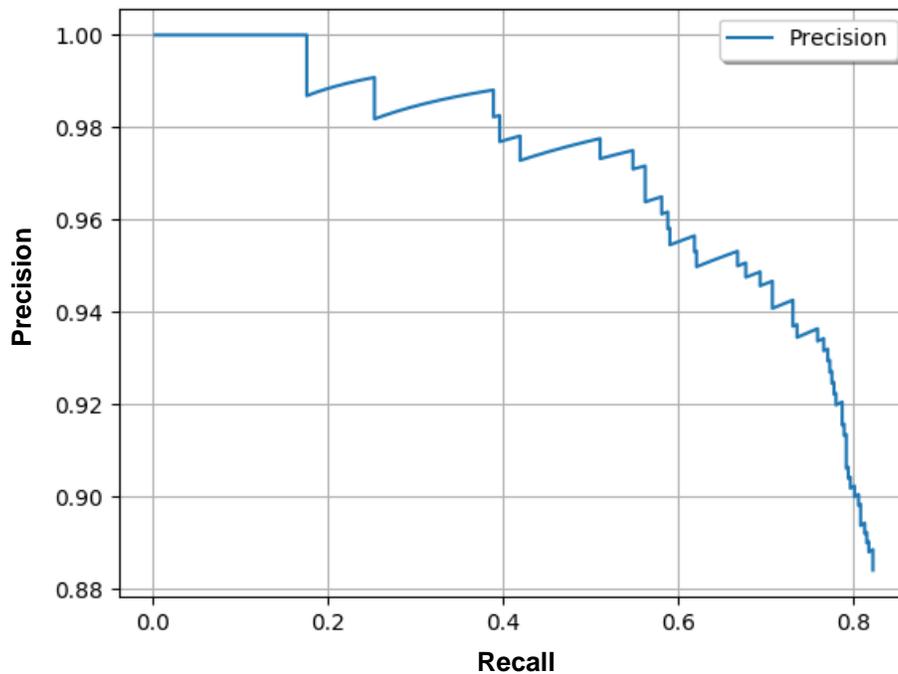


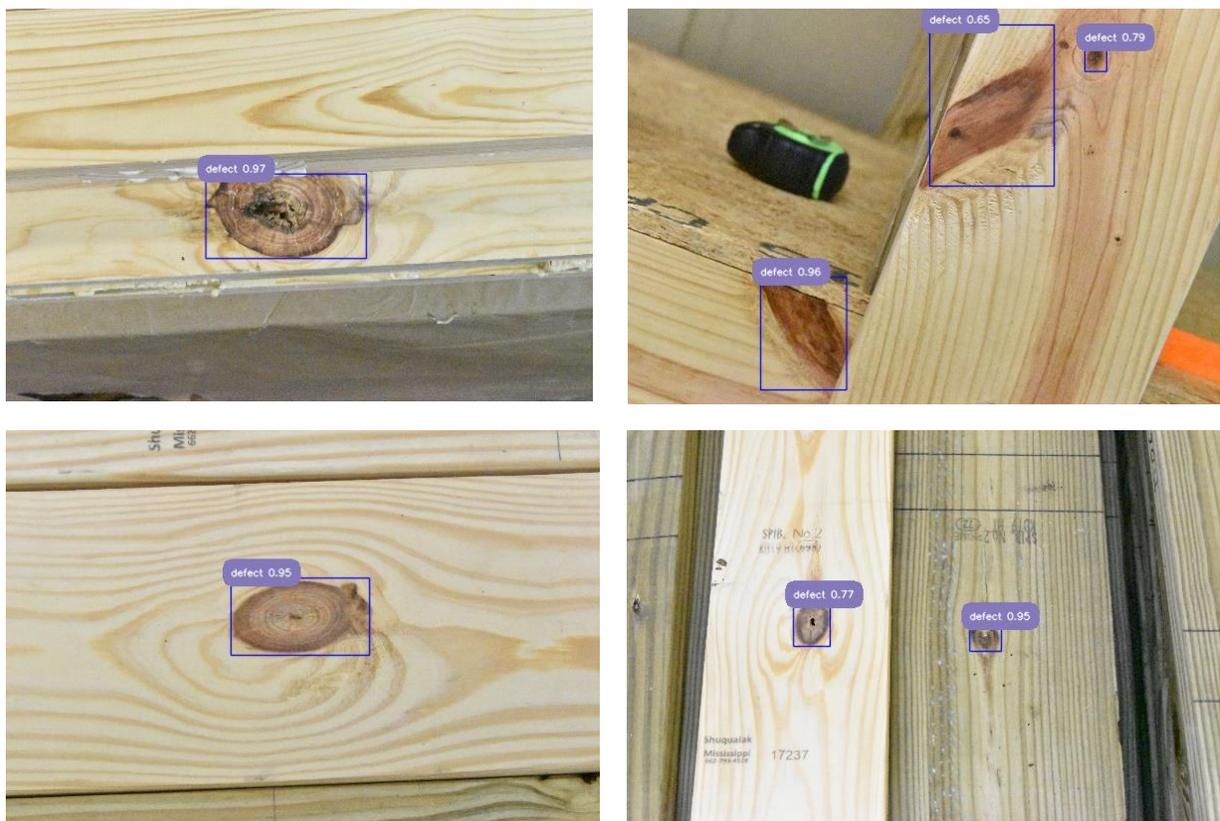
Fig. 5. Precision x recall curve

The quality metrics indicated acceptable results, since the number of false positives and false negatives were relatively low. In general, the YoloV3 model recognized and accurately detected small, medium, and large knots sizes. In some cases, the algorithm did not correctly identify knots on the surface that were easily observed by the human eye. This may be likely explained by the low number of training images used in the CNN training. Even though the dataset was augmented, commonly used object detectors use more than 3,500 training images (Lin *et al.* 2015). In comparison, the present dataset had less than 1,000 images.

Wane was not labeled for this study. Usually, wane is found lengthwise on a board. It was observed that the ground-truth bounding box spanned almost the whole width of the board and on the length of the board. In other words, bounding boxes drawn for wane included large portions of clear wood. A better approach based on semantic segmentation is being developed to optimize this issue so that it is possible to label the image pixel-wise, which would preserve the precise shape of each defect and could include both knots and wane for detection.

The YoloV3 model, on average, takes about 5 seconds to analyze all 184 validation images and 427 knots present. In comparison to previous works, Cavalin *et al.* (2006) used multi-layer perceptron and support vector machines to detect wood defects, but that approach did not account for how fast their classifier was or if it was possible to implement it in real-time. Schmoltdt *et al.* (1997) used CT scans to classify logs defects and reported that it took 25s for analysis of a single 256 pixels x 256 pixels CT slice. With advanced technology, particularly, powerful GPU processing, the present methodology decreased knots detection speed by a factor of 5.

Figure 6 shows several examples of the present automated knot detection.





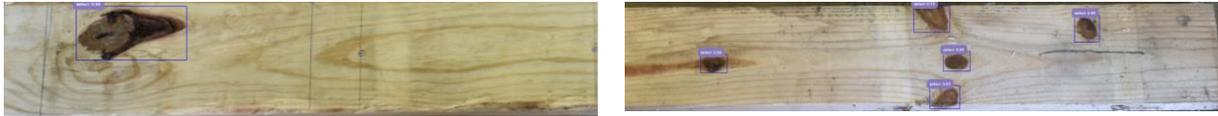


Fig. 6. Examples of knots detection using our implementation of YoloV3

The authors plan to continue developing machine-learning approaches for supporting the wood science field. To that end, the plan is to label defects pixel-wise so that many types of defects can be analyzed. For future research, it is expected that semantic segmentation will provide a sufficient foundation for accurate and fast defect detection, identification of type of defect, and accurate estimation of defect surface area.

CONCLUSIONS

1. In this study, a deep learning approach for knot detection was presented based on YoloV3 convolutional neural network architecture. The Darknet-53 was trained from scratch on a custom image dataset. The model proved to be remarkably fast and accurate when identifying small, medium, and large knots. Results indicated an overall detection speed of 0.0102 seconds per knot.
2. A training from scratch strategy was used to train a limited number of images with resulting average precision of 0.8, precision of 0.77, recall of 0.79, and F1-Score of 0.78.
3. Identification and location of knots can increase a sawmill's overall yield. Automatic knot detection using state-of-the-art deep learning methods is a starting point for removing poor quality products from processing lines. With a trained AI model, sawmills can utilize this type of technology with reduced cost. Previously studied methods are accurate; however they are expensive and slow, which represents a crippling bottleneck for real-world applications. To that end, the present approach was 5 times faster than previous literature.

ACKNOWLEDGMENTS

The authors wish to acknowledge the support of U.S. Department of Agriculture (USDA), Research, Education, and Economics (REE), Agriculture Research Service (ARS), Administrative and Financial Management (AFM), Financial Management and Accounting Division (FMAD) Grants and Agreements Management Branch (GAMB), under Agreement No. 58-0204-6-001. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture. This publication is a contribution of the Forest and Wildlife Research Center (FWRC) at Mississippi State University. The authors would also like to thank Dr. Greg W. Burgreen for English proofreading, suggestion and comments.

Data Sharing

Data will be made available upon request to the corresponding author.

REFERENCES CITED

- Cao, Y., Street, J., Mitchell, B., To, F., DuBien, J., Seale, R. D., and Shmulsky, R. (2018). "Effect of knots on horizontal shear strength in Southern yellow pine," *BioResources* 12(2), 4509-4520. DOI:10.15376/biores.13.2.4509-4520.
- Cavalin, P., Oliveira, L. S., Koerich, A. L., Britto, A. S. (2006). "Wood defect detection using grayscale images and an optimized feature set," in: *32nd Annual Conference on IEEE Industrial Electronics*. Paris, France, pp. 3408-3412. DOI:10.1109/OECPM.2006.347618
- Cetiner, I., Var, A. A., and Cetiner, H. (2016). "Classification of knot defect types using wavelets and KNN," *Eletronika IR Eletrotehnika* 22(6), 67-72. DOI:10.5755/j01.eie.22.6.17227.
- Food and Agriculture Organization of the United Nation – FAO. (2020). "Forest product statistics", (<http://www.fao.org/forestry/statistics/80938@180723/en/>), Accessed on July 8, 2020.
- Gu, I. Y. H., Andersson, H., and Vicen, R. (2010). "Wood defect classification based on image analysis and support vector machines," *Wood Sci. Technol* 44, 693-704. DOI: 10.1007/s00226-009-0287-9
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G., Cai, J., and Chen, T. (2017). "Recent advances in convolutional neural networks," *Pattern Recognition* 77(C), 354-377. DOI:10.1016/j.patcog.2017.10.013.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). "Deep residual learning for image recognition," *Computer Science, Computer Vision*. arXiv:1512.03385v1. DOI:10.1109/CVPR.2016.90.
- Jocher, G., guigarfr, perry0418, Ttayu, Veitch-Michaelis, J., Bianconi, G., and IlyaOvodov. (2020). "Ultralytics/yolov3: Rectangular Inference, Conv2d + Batchnorm2d Layer Fusion (Version v6)". Zenodo. DOI:org/10.5281/zenodo.2672652.
- Kontzer, T. 2019. "Going against the grain: How Lucidyne is revolutionizing lumber grading with deep learning," (<https://blogs.nvidia.com/blog/2019/04/18/lucidyne-gradescan-lumber-grading/>), Accessed July 7, 2020.
- Lampinen, J., Smolander, S., and Korhonen, M. (1998). "Wood surface inspection system based on generic visual features," in: *Industrial Applications of Neural Networks*, F. F. Soulié and P. Gallinari (eds.), World Scientific, Singapore, pp. 35-42.
- Lin, T., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., and Dollar, P. (2015). "Microsoft COCO: Common objects in context," *Computer Science, Computer Vision*. arXiv:1405.0312v3.
- Lopes, D. J. V., Bobadilha, G. S., and Grebner, K. M. (2020a). *YOLOV3 on photos of boards with knots*. (<https://drive.google.com/file/d/1Et99YBVBvnbOFsrxhHhg2eVAU3PcBGvA/view?usp=sharing>)
- Lopes, D. J. V., Bobadilha, G. S., and Grebner, K. M. (2020b). *YOLOV3 on videos of boards with knots*.

- (https://drive.google.com/file/d/1irEfT3q3FdDjFvRtv2AnSjxqVRVJFl_y/view?usp=sharing)
- Mahram, A., Shayesteh, M. G., and Jafarpour, S. (2012). "Classification of wood surface defects with hybrid usage of statistical and textural features," in: *Proceedings of the 35th International Conference on Telecommunications and Signal Processing (TSP)*, Prague, Czech Republic, 3-4 July 2012, pp. 749–752.
- Mao, Q., Sun, H., Liu, Y., and Jia, R. (2019). "Mini-YOLOv3: Real-time object detector for embedded applications," in: *IEEE Access*, vol. 7, pp. 133529-133538. DOI: 10.1109/ACCESS.2019.2941547.
- Nasir, V., and Cool, J. (2018). "A review on wood machining: Characterization, optimization, and monitoring of the sawing process," *Wood Material Science & Engineering*, 18 pp. DOI: 10.1080/17480272.2018.1465465.
- Nasir, V., and Cool, J. (2020). "Intelligent wood machining monitoring using vibration signals combined with self-organizing maps for automatic feature selection". *The International Journal of Advanced Manufacturing Technology* 108, 1811-1825.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2015). "You only look once: Unified, real-time object detection," *Computer Science, Computer Vision*. ArXiv:1506.02640.
- Redmon, J., and Farhadi, A. (2016). "YOLO9000: Better, Faster, Stronger," *Computer Science, Computer Vision*. ArXiv:1612.08242. DOI:10.1109/CVPR.2016.91.
- Redmon, J., and Farhadi, A. (2018). "YOLOv3: An incremental improvement," *Computer Science, Computer Vision*. ArXiv:1804.02767.
- Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., and Savarese, S. (2019). "Generalized intersection over union: A metric and a loss for bounding box regression". *Computer Science, Computer vision and Patter Recognition*. arXiv:1902.09630.
- Rocha, M. F. V., Costa, L. R., Costa, L. J., Araujo, A. C. C., Soares, B. C. D., and Hein, P. R. G. (2018). "Wood knots influence the modulus of elasticity and resistance to compression". *Floresta e Ambiente* 25(4), 1-6. DOI: 10.1590/2179-8087.090617.
- Rudakov, N. (2018). *Detection of Mechanical Damages in Sawn Timber Using Convolutional Neural Networks*, Master's Thesis, Programme in Computational Engineering and Technical Physics, Lappeenranta, Finland.
- Schmoldt, D. L., Li, P., and Abbott, A. L. (1997). "Machine vision using artificial neural networks with local 3D neighborhoods," *Computers and Electronics in Agriculture*, 16: 225-271. DOI: 10.1016/S0168-1699(97)00002-1
- Tzotalin. (2020). "LabelImg," (<https://github.com/tzotalin/labelImg>),_Accessed on: June 30 2020.
- Urbanas, A., Raudonis, V., Maskeliunas, R., and Damasevicius, 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. DOI: 10.3390/app9224898
- Zhong, Y., Ren, H. Q., Lou, W. L., and Li, X. Z. (2012). "The effect of knots on bending modulus of elasticity of dimension lumber," *Key Engineering Materials* (517), 677-82.
- Article submitted: August 17, 2020; Peer review completed: October 18, 2020; Revised version received and accepted: October 19, 2020; Published: October 23, 2020.
DOI: 10.15376/biores.15.4.9351-9361