Computer Vision based Automated Timber Structural Defect Detection Framework

Afia Rasool[1], Qipei Mei[2], Rafiq Ahmad[1,*]

[1]Smart & Sustainable Manufacturing Systems Laboratory (SMART Lab), Department of Mechanical Engineering, University of Alberta, Edmonton, Canada
[2]Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada
arasool@ualberta.ca, qipei.mei@ualberta.ca, rafiq.ahmad@ualberta.ca

Abstract –

In the wood construction industry, timber structural defect detection is usually considered a pre-manufacturing inspection step done manually. To address this issue, the proposed study discusses the timber structural defect detection method based on YOLOv8 variants. The evaluation matrices used are precision, recall, mAP.5, and mAP.5-.95, and the results indicate stable convergence and consistent accuracy on the complex dataset instances. This research contributes to the automation of timber defect detection for precise and robust manufacturing of timber structures. The proposed method further improves resource utilization and contributes towards eliminating waste in the residential construction industry.

Keywords –

Timber; Defect Detection; Mass-Timber Manufacturing; YOLOv8; Construction Industry; Computer Vision

1 Introduction

Timber, a finished wood product, is a major building material for residential and commercial buildings worldwide. Timber exhibits unique structural capacity and fineness, making it suitable for panels, trusses, ridges, beams, and staircases. Like every other wood type, natural timber exhibits multiple defects due to genetic and growth factors, processing phase, and environmental and storage conditions [1]. Defective timber can harm the overall building integrity and durability and must be excluded in final processing. Surface defects like knots and burls originate from growth conditions and can easily be identified at an early stage of timber handling. Contrarily, twisting, warping, and splits often occur during the manufacturing and drying phases of finished stud handling, and many construction workshops use experts' eye and appropriate angles for manual stud quality inspection and anomaly identification.

Therefore, a low-cost, fast solution is needed to identify complex structural irregularities like twists and bends [2-5]. Machine vision and deep learning solutions are popular for object detection, which utilize CNNs to mimic human brain learning behaviors [6-8]. Some research proposes generalized novel algorithms utilizing different layering of CNNs [6,9], whereas other researchers focus on utilizing and fine-tuning the proposed novel and hybrid algorithms to better tune them for specific tasks, environments, and datasets [10-12]. A recent study on defect detection across variant timber types has combined image preprocessing techniques with CNN-based architecture to build a vision-enabled defect detection platform [13]. Another similar approach was utilized for precise sawing and grading automation in timber processing [14]. A method called EMINet uses cross-fusion modules for the surface defect detection of sawn timber [15]. This algorithm is designed to learn and infer multi-scale features with greater speed and accuracy. Xi et al. have utilized SPD-Convolution, SiAFF-PA Net, and multi-attention mechanisms in their model to refine the localization and classification of timber defects [16]. Furthermore, different types of stained wood grain defects have been discussed in literature by incorporating CNN architecture to boost the quality inspection and grading mechanism defects [17].

Image detection algorithms can be divided into two categories: two-stage detection and one-stage detection. R-CNN (Region-based Convolutional Neural Network) [18], Fast R-CNN [19], and Faster R-CNN [20] are popular two-stage models comprising region proposal blocks and classification blocks. Whereas the one-stage object detection has a unified region proposal generation, and classification mechanisms are proposed in YOLO (You Only Look Once) [21-24] and SSD (Single Shot Multibox Detector) [25]. Regarding timber defect detection, the YOLO series has been enhanced in several ways to boost detection accuracy and specified classes. BDCS-YOLO is an enhanced method that showed a 12.3 percent improvement in timber defect detection and sawing for resource optimizations [13]. Another study

has used the SimAm module and Ghost convolution to improve YOLOv5, reducing the computational costs for wood surface defects [26]. Another study has focused on efficiently detecting cracks in wood surfaces by proposing ICDW-YOLO algorithm [27].

Besides multiple pieces of research exhibiting good performance on surface-level defects such as minor cracks, knots, and stains, it is clear from the literature study that no significant work has focused on utilizing computer vision to inspect physical defects such as bends, twists, and missing of wood parts within timber studs. This research is focused on generating industrial solutions for the precise localization of defective parts using the YOLOv8 algorithm. To automate the precise detection of structural defects in timber studs, this research begins with the data collection and generation of the timber stud structural defect detection dataset, specifically focusing on twists and bends. The developed dataset is used for training and prediction using the YOLOv8 image detection method. An extensive study of evaluation matrices is performed on multiple variants of YOLOv8 (YOLOv8nano, YOLOv8small, YOLOv8medium, YOLOv8large, YOLOv8x_large) to balance the overall system accuracies and computational cost [2]. Moreover, due to the complexity of the task nature, a prediction study is discussed, emphasizing the importance of initial setup settings during data collection in the overall success of deep learning training.

The precise and easy timber defect detection system can minimize the wasting of complete studs and maximize material utilization, potentially reducing the additional tree harvest. Furthermore, prior screening of timber quality can reduce repair overengineering resulting in durable construction. These factors contribute towards decarbonizing the construction industry and helping easily adopt natural timber as a cheap, sustainable material, lessening carbon emissions. This research can also directly advance industrial practices of automating timber inspection using computer vision and robotics.

The overall flow of the study is shown in Figure 1, and the key steps involve dataset preparation, explained in Section 2; model training on YOLOv8 variants detailed in Sections 3, 4, and 5; and model prediction results are presented in Section 6, followed by conclusion and discussion of future insights.

2 Dataset Preparation

Timber studs are used in the manufacturing of wood buildings as wall panels, roof trusses, floor units, and supporting ridges, and they vary in dimensions a lot. An average stud can be 2-4 meters long, 2-4 inches wide, and 4-8 inches in depth and have a face side (the wider and flatter surface) and an edge side (the narrower surface). They can vary hugely depending on the grade and species, moisture contents, load requirements, and code compliances. The changes in shapes and structural integrity of these timber studs, called structural defects, can occur due to drying, external stress, and improper seasoning. Bend (curvature along the length of stud), twist (spiral warping), bow (curve along the face side), cup (curve along the width), and warp (an amalgamation of bends, twists, and bows) are categorized as structural defects.

Table 1 shows the dataset preparation steps and their details. As the defect can form at any angle or side, we have used 3 camera heights (2 feet, 6 feet, and 18 feet) to

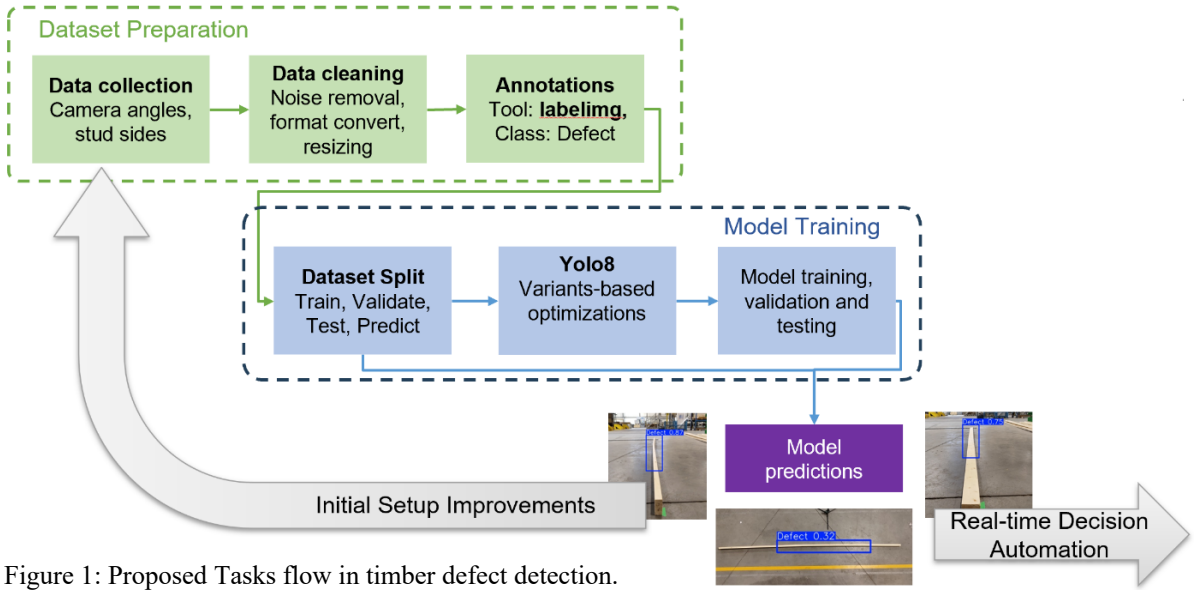


Figure 1: Proposed Tasks flow in timber defect detection.

capture the pictures of a stud on both sides (face and edge). A total of 15 studs, with 12 defective and 3 non-defective, are used to generate 400 multi-dimensional images. These images are manually cleaned to remove irrelevant details and are resized to 640 x 640 sizes to be fed to the YOLOv8 models. LabelImg [28] is a tool for ground truth generation for popular image detectors and is used here to annotate the dataset, creating defect class bounding boxes. The images showing no defect are not assigned to the defect class and are passed to the model training as background images. After cleaning and annotations, the final dataset comprising 285 images with defects and 65 background images is split stud-wise for training, validation, testing, and prediction at 65%, 15%, 15%, and 5%, respectively.

The dataset generation and model training for structural defects is tricky, as a defect visible on the face side at a particular angle may not be detectable across the edge and vice versa. The complexity of the task can be seen in Figure 2, which shows some real instances from the dataset generated.

Table 1: Timber Stud Surface Defect Detection Dataset preparation steps.

Phase	Parameter	Value
Initial Setup	Stud sides	2 (face, edge)
	Camera heights	3 (2 feet, 6 feet, 18 feet)
Data collection	Total studs	15
	Defective studs	12
	Non-defective studs	3
	Total images	400 (varying dimensions)
Resizing, cleaning, and annotations	Total images after data cleaning	350 (640 x 640)
	Background images	65
	Annotated with class: Defect	285
Model training	Training (65 %)	10 studs, 234 images
	Validation (15 %)	2 studs, 48 images
	Testing (15 %)	2 studs, 48 images
	Prediction (5%)	1 stud, 20 images

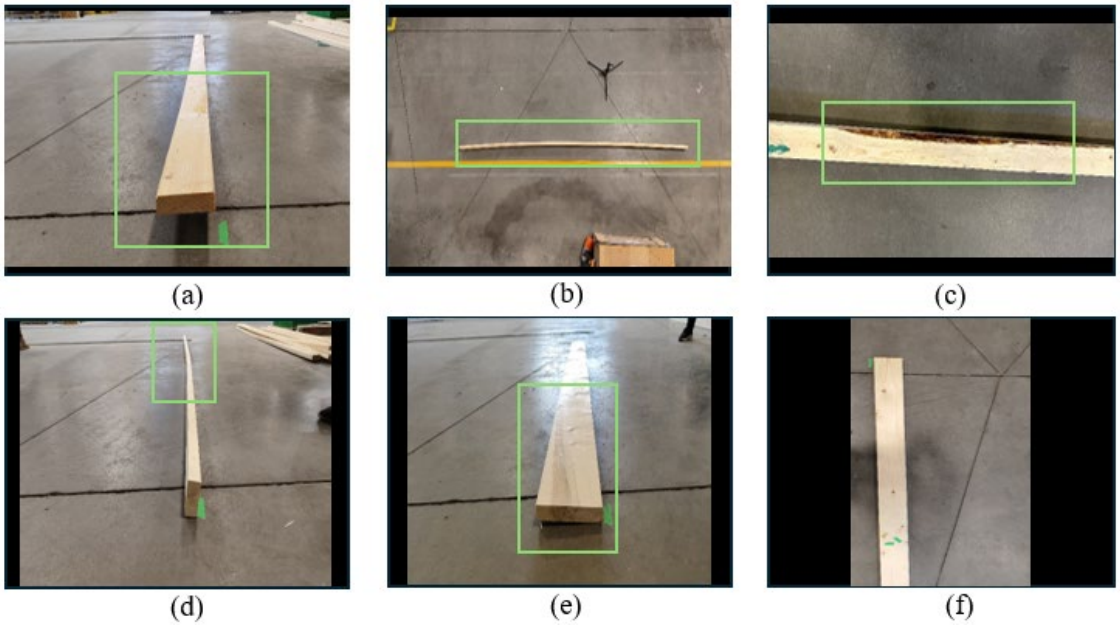


Figure 2: Instances of dataset (a) a bended stud captured from 2 feet height (b) the defect captured from 18 feet height (c) missing wood defect (d) bended and twisted stud from the edge (e) an example of twist and bend (f) background instance showing no defect

3 YOLOv8 Method

YOLO is a one-stage detection model that combines the predecessor's functionalities into improved architecture in its version series. YOLOv8 [2] is an adaptable and efficient model comprising three main modules. The initial backbone is a variant of CSPDarknet, used for multi-scale feature extraction in the model. The second unit, called the neck, implements the PAN-FPN module for feature aggregation, and finally, the head contributes to object classification and bounding box regression. One powerful feature of YOLOv8 is the real-time prediction generation that supports inspection-based applications for faster, lightweight, and more accurate real-time detections. The trade-off between speed and accuracy is well-managed in YOLOv8, which provides precise detections at the rate of high throughput. Moreover, the refined multi-scale feature fusion of this model makes it best suited for the current case of structural defect detection where the images are captured from different heights, offering multi-sized imperfection instances.

As the available dataset is of small size, the training is run on different variants of YOLOv8 to achieve high performance, considering factors of resource optimization, overfitting avoidance, comparative performance analysis, scalability and adaptation. These variants are YOLOv8 nano, YOLOv8 small, YOLOv8 medium, YOLOv8 large, and YOLOv8 extra-large, exhibiting different training speeds, computational requirements, and number of trainable parameters. From nano to large, these variants differ based on the layer count, sizes, and trainable parameters, and support limited to industry real-time applications.

4 Model Training

After the dataset preparation and splitting of the dataset, the data is passed to YOLOv8 variants for training and validation. Table 2 discusses the training machine specifications. The model is trained with the generated dataset on a machine with Central Processing Unit (CPU) Intel(R) Core(TM) Ultra 9 185H 2.50 GHz, Graphics Processing Unit (GPU) NVIDIA GeForce RTX 4070 Laptop GPU (8,188 MiB VRAM), 64 GB RAM (Random Access Memory), with installed version of CUDA 11.2, python 3.9.20 and Pytorch 2.5.1+cu121.

Table 2: Training machine specifications.

Property	Value
CPU	Intel(R) Core(TM) Ultra 9 185H 2.50 GHz
GPU	NVIDIA GeForce RTX 4070 Laptop GPU (8,188 MiB VRAM)
CUDA version	11.2
Operating System	Windows 11 Home
RAM	64 GB
Python version	3.9.20
PyTorch version	2.5.1+cu121

Table 3: Model Training parameters on YOLOv8

Parameter	Value	Parameter	Value
Image Size	640 x 640	Epochs	250
Batch Size	8	Augmentation	hsv_h=0.015
Learning rate	0.001		hsv_s=0.7
Momentum	0.937		hsv_v=0.4
Weight Decay	0.0005		mosaic=1.0

As the generated timber stud structural defect detection dataset is small and multifaceted, the training was performed on different hyperparameters to choose the best settings suitable for better accuracy, and Table 3 shows the carefully selected set of model training parameters on YOLOv8 variants. These hyperparameters were chosen to balance generalizability and computational efficiency, given the dataset constraints. The input image size is 640 x 640 with a batch size kept at 8 trained for 250 epochs during each model training. The batch size was selected based on the GPU availability to avoid any computational instability. The learning rate is initiated with 0.001, momentum of 0.937, and weight decay of 0.0005 to achieve a smoother convergence. The training was also set to early stopping to save the best-trained parameters for inference. During real-time image capturing for real-time inference, the brightness and hues vary greatly depending on the workshop environment and lighting conditions. To make the training robust of inference time conditions, the data augmentation is kept true during the training, and the augmentation parameters are passed as hue 0.015, saturation 0.7, and value 0.4, and the mosaic set is kept on. These data augmentation settings can avoid the overfitting of smaller datasets, making the model resilient to surrounding conditions by slightly changing the picture effects during training.

5 Evaluation Metrics

Typically, YOLOv8 does not support model testing on a different dataset. Instead, it generates the model predictions based on the trained weights. However, to assist the optimization study and hyperparameter evaluation, this study used a separate model testing mechanism by passing the unseen test images to a separate model evaluation method. This proved effective for understanding the real model accuracies and finetuning accordingly.

Table 4: Variants of YOLOv8 and their specifications

YOLOV8 Variants	#Layers	#Params	Val Precision (%)	Test Precision (%)	Val Recall (%)	Test Recall (%)
YOLO8nano	225	3,011,043	79.5	62.8	63.0	72.3
YOLO8sml	225	11,135,987	74.1	60.9	67.4	74.5
YOLO8med	295	25,856,899	93.0	74.8	60.9	72.3
YOLO8lg	365	43,630,611	85.7	79.8	60.9	67.1
YOLO8x_lrg	365	68,153,571	84.1	66.3	57.4	62.8

The evaluation metrics used for the inspection of model training and testing are precision (P), recall (R), and mean average precision (mAP) of the predicted bounding box of defects. Precision counts the correct predictions from all prediction instances, whereas recall counts the correct predictions out of all actual predictions. The formula for precision and recall is given in Equation (1) and Equation (2).

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

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

In Equation (1), *P* indicates precision, whereas *TP* and *FP* donate true positives and false positives, respectively. In Equation (2), *R* represents recall, and *TP* and *FN* present true positives and false negatives.

Mean average precision (mAP) considers the classification and location accuracy, and mAP.50 and mAP.5:.95 are reported here. mAP.50 reports the average precision of the model when the intersection over the union (IOU) threshold is 50% or up between predicted and ground truth bounding boxes. mAP@.50:.95 is another metric that computes the average precision for IoU thresholds ranging from 0.50 to 0.95 in 0.05-step increments. This metric is more critical than mAP@.50 since it penalizes projections far from ground truths and displays the model's ability to precisely locate objects. A high mAP@.50:.95 value indicates greater performance in localization and classification.

6 Discussion of Results

As indicated in Table 4, the number of layers and trainable parameters are increasing from YOLOv8nano to larger models. The validation and test results indicate varying accuracies of these trained models on the timber defect dataset. The validation and test precision for YOLOv8 nano, YOLOv8 small (sml), YOLOv8 medium (med), YOLOv8 large (lrg), and YOLOv8 extra-large (x_lrg) are 79.5% and 62.8%, 74.1% and 60.9%, 93.0% and 74.8%, 85.7% and 79.8%, 84.1% and 66.3% accordingly. The validation and test recall for YOLOv8 nano, YOLOv8 small, YOLOv8 medium, YOLOv8 large, and YOLOv8 extra-large are 63.0% and 72.3%, 67.4% and 74.5%, 60.9% and 72.3%, 60.9% and 67.1%, 57.4% and 62.8% accordingly. These results indicate that YOLOv8 medium and YOLOv8 large are best suitable for the current detection task, as the precision is high for these models, which means the model is detecting true instances and recall is comparatively better, demonstrating fewer missed detections.

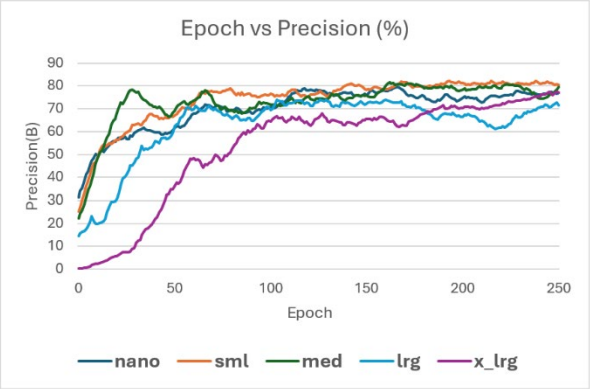
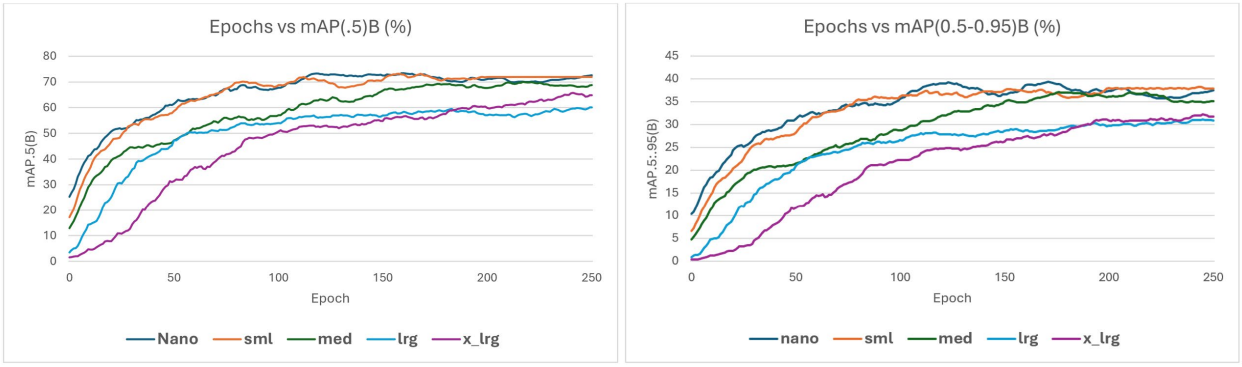


Figure 3: Training Precision in percentage

The training precision is shown in Figure 3, and it is visible that all models try to converge between 100-200 epochs with a slight increase in precision afterward, but the YOLOv8 large training curve indicates a loss in accuracy in later epochs, indicating that the model is still trying to learn new parameters due to its large size.

The training mAP.5 (B) and mAP.5:.95 (B) for all models in Figures 4 (a) and (b), respectively, indicate a similar pattern as precision(B) where models can achieve between 63% to 73% metrics for mAP.5 and between 35 % to 40% mAP.5-.95. These converged graphs show that models are neither over-fitting nor under-fitting, provided the toughness, small-scaled and complexity of the dataset.

Moreover, the final validation and test mAP.5 (B) and mAP.5:.95 (B) for all models are presented in Figure 5. This comparative graph indicates around 75% and 70% of validation and test mAP.5 (B) and around 45% and 38% achieve validation and test mAP.5:.95 (B). These values



(a) (b)

Figure 4: (a) Training mAP(.5)(B) in percentages (b) Training mAP(0.5-0.95)(B) in percentages

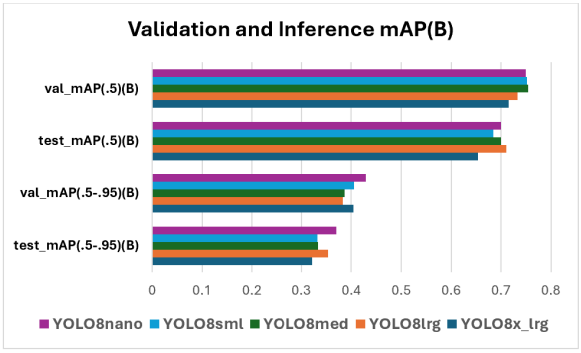
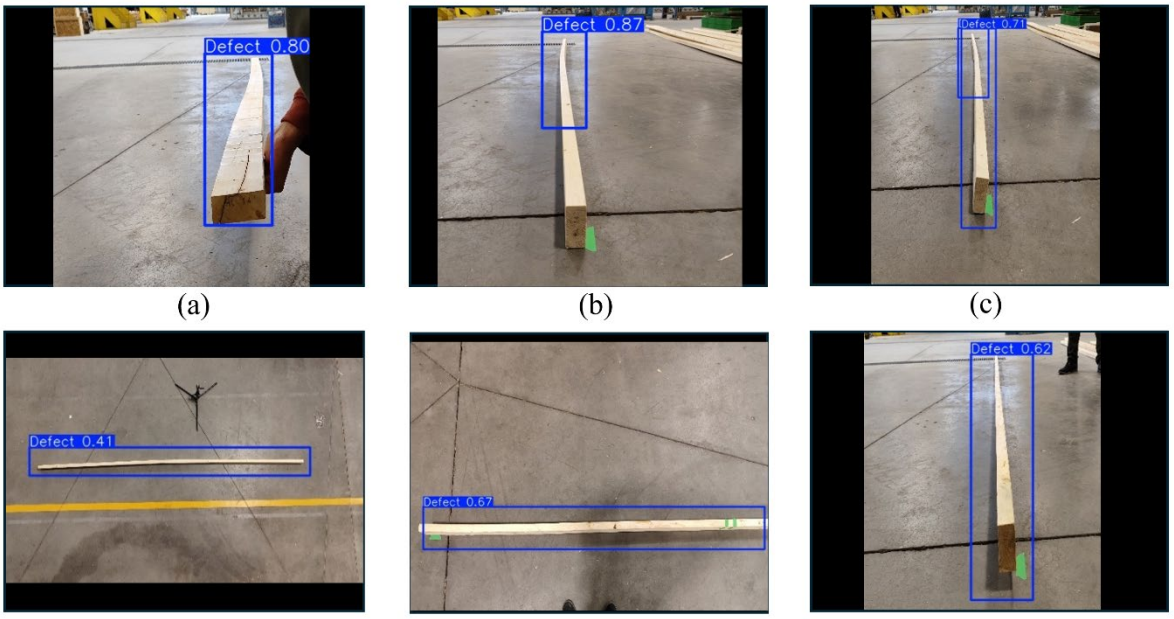


Figure 5: Validation and Inference results on YOLOv8 variants

demonstrate effective learning and consistently high capabilities, indicating that the models are generalized well on unseen data. Moreover, YOLOv8 medium and large are two modules that consistently achieved superior accuracies and precision, which confirms their ability to perform reliably in real-world scenarios.

Based on the collected results, YOLOv8 medium-trained weights are selected for model inference so that the images from one defective stud are passed to generate predictions, presented in Figure 6. The model generates predictions for the same defective studs differently when different angles of images are provided. The instances captured from 2 feet height generated a bounding box with high accuracy (a,b,c), but the top views 6 feet and 16 feet high indicate a gradual decrease in precision (d,e).



(a) (b) (c)

(d) (e) (f)

Figure 6: The difference in prediction accuracies based on different camera angles.

These observations provide valuable insight into the number of cameras, optimized camera installation location, and images needed for effective real-world predictions using the provided setup and methods.

As the number of defective studs in a working industrial setting is limited, so based on data availability, the number of defective and non-defective studs was adjusted to create a balanced trainable dataset, totalling 15 studs. The smaller datasets are highly prone to overfitting issues, compromising overall model authenticity and generalizability. To tackle this issue, a careful selection of hyperparameters and data augmentation was adopted while training with different modules of YOLOv8. In the future, this research can extend to collect and expand the dataset length to enhance accuracy and attain better results. Further, a performance comparison between other state-of-the-art and latest image detection training models can also be conducted.

7 Conclusion

This research discussed timber stud structural defect detection dataset preparation steps followed by the extensive model training on YOLOv8 variants, specifically focusing on detecting bends, twists, and wraps in timber studs used in wood house construction. The quantitative results indicated a balance between accuracy and computational efficiency, whereas YOLOv8 medium and YOLOv8 large demonstrated the high validation and test precision of 93.0%, and 74.8%, 85.7%, and 79.8%, respectively. The predictions generated provided knowledge about the best device placement and total frames for real-time prediction precision, and the predicted defect bounding boxes can be used for precise automated cutting of defective parts. This research will help reduce carbon emissions through resource utilization optimization and waste mitigation in the construction industry. Moreover, it advances computer vision and robotic automation in the Canadian wood construction industry. The proposed methods can be expanded to accommodate other types of timber defects, expand the dataset, and automate other pre- and post-manufacturing inspection steps.

References

[1] D. Li, Z. Zhang, B. Wang, C. Yang, and L. Deng. Detection method of timber defects based on target detection algorithm. *Measurement (Lond)*, 203:111987, 2022.

[2] Ultralytics. YOLOv8. On-line: https://github.com/ultralytics/ultralytics, Accessed: 25/12/2024.

[3] J. J. Cisneros-Gonzalez, A. Rasool, and R. Ahmad. Digital technologies and robotics in mass-timber manufacturing: a systematic literature review on construction 4.0/5.0. *Construction Robotics*, 8(2): 29, 2024.

[4] A. Rasool, G. R. Satsangee, L. Arickswamy, M. M. Ashfaq, and R. Ahmad. Winter-Safe Slip Prevention Rim for E-Scooter: Design to Production Lifecycle Analysis. In *1st International Conference on Industrial, Manufacturing, and Process Engineering (ICIMP-2024)*, page 88, Basel, Switzerland, 2024.

[5] J. J. Cisneros González, A. Rasool, Y. H. Chui, and R. Ahmad. Robotic cell design and simulation system for cross-laminated timber panel post-construction operations. In *Proceedings of the Transforming Construction with Off-site Methods and Technologies (TCOT) Conference*, pages 1–10, Fredericton, Canada, 2024.

[6] E. A. Holm, P. P. Choi, M. M. Nowell, and S. P. Lynch. "Overview: Computer Vision and Machine Learning for Microstructural Characterization and Analysis. *Metallurgical and Materials Transactions A,* 51(12):5985–5999, 2020.

[7] G. R. Satsangee, H. Al-Musaibeli, and R. Ahmad. A defect detection method based on YOLOv7 for automated remanufacturing. *Applied Sciences*, 14(13):5503, 2024.

[8] H. Z. Imam, H. Al-Musaibeli, Y. Zheng, P. Martinez, and R. Ahmad. Vision-based spatial damage localization method for autonomous robotic laser cladding repair processes. *Robotics and Computer-Integrated Manufacturing*, 80:102452, 2023.

[9] A. Rasool, M. M. Fraz, and S. Javed. Multiscale Unified Network for Simultaneous Segmentation of Nerves and Micro-vessels in Histology Images. In *Proceedings of the 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2)*, pages 1–6, Islamabad, Pakistan, 2021.

[10] R. Wang, Y. Chen, F. Liang, B. Wang, X. Mou, and G. Zhang. BPN-YOLO: A Novel Method for Wood Defect Detection Based on YOLOv7. *Forests*, 15(7):1096, 2024.

[11] W. Meng and Y. Yuan. SGN-YOLO: Detecting Wood Defects with Improved YOLOv5 Based on Semi-Global Network. *Sensors*, 23(21):8705, 2023.

[12] R. Ehtisham, M. A. Khan, M. Sharif, S. L. Fernandes, and G. S. Siddiqui. Computing the characteristics of defects in wooden structures using image processing and CNN. *Automation in Construction*, 158:105211, 2024.

[13] C. Fan, Z. Zhuang, Y. Liu, Y. Yang, H. Zhou, and X. Wang. Bilateral Defect Cutting Strategy for Sawn Timber Based on Artificial Intelligence

Defect Detection Model. *Sensors,* 24(20):6697, 2024.

[14] M. Ji, W. Zhang, X. Diao, G. Wang, and H. Miao. Intelligent Automation Manufacturing for Betula Solid Timber Based on Machine Vision Detection and Optimization Grading System Applied to Building Materials. *Forests*, 14(7):1510, 2023.

[15] Y. Zhu, Z. Xu, Y. Lin, D. Chen, K. Zheng, and Y. Yuan. Surface defect detection of sawn timbers based on efficient multilevel feature integration. *Measurement Science and Technology,* 35(4):045401, 2024.

[16] H. Xi, R. Wang, F. Liang, Y. Chen, G. Zhang, and B. Wang. SiM-YOLO: A Wood Surface Defect Detection Method Based on the Improved YOLOv8. *Coatings*, 14(8):1001, 2024.

[17] J. Zhang, L. Bi, J. Lian and C. Guan. A Single-Trial Event-Related Potential Estimation Based on Independent Component Analysis and Kalman Smoother. *In Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM),* pages 9–12, Auckland, New Zealand, 2018

[18] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. arXiv preprint arXiv:1311.2524, 2014.

[19] R. Girshick. Fast R-CNN. On-line: https://github.com/rbgirshick/, Accessed: 25/12/2024.

[20] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149, 2017.

[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society*, pages 779-788, Las Vegas, USA, 2016.

[22] J. Redmon and A. Farhadi. YOLO9000: Better, faster, stronger. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6517–6525, Hawaii, USA, 2017.

[23] J. Redmon and A. Farhadi. YOLOv3: An Incremental Improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[24] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*, 2020.

[25] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. SSD: Single shot multibox detector. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 21–37, Amsterdam, Netherlands, 2016.

[26] J. Xu, H. Yang, Z. Wan, H. Mou, D. Qi, and S. Han. Wood Surface Defects Detection Based on the Improved YOLOv5-C3Ghost With SimAm Module," *IEEE Access,* 11:105281–105287, 2023.

[27] J. Zhou, J. Ning, Z. Xiang, and P. Yin. ICDW-YOLO: An Efficient Timber Construction Crack Detection Algorithm. *Sensors*, 24(13):4333, 2024.

[28] Tzutalin. LabelImg: Graphical Image Annotation Tool. On-line: https://github.com/tzutalin/labelImg, Accessed: 25/12/2024.