ISBN: 978-981-94-1156-6
2024 the 14th International Workshop on Computer Science and Engineering (WCSE 2024)
doi: 10.18178/wcse.2024.06.005

Using YoloV8 Book Defect Detection Model with MIRNet LLI

John Paul Tomas[1+], Katrice Asher Albano[2], Jill Samantha Alday[3], Matthew Raphael Corbe[4] and Gian Nicole Pangan [5]

[12345]School of Information Technology, Mapúa University, Philippines

Abstract. Books serve as fundamental repositories of knowledge. In an age where digitization of texts is rampant, the tangible versions of these tomes are at risk of degradation due to environmental factors, mishandling, and the natural decay of materials. The integrity of a book, both in terms of its legibility and visual appeal, can be compromised by various forms of damage. It is crucial for institutions like museums and libraries to preserve these concrete sources of knowledge. Traditional methods for identifying damages in books have relied on manual checks and rudimentary image processing. However, the inefficiencies of these methods have highlighted the urgent need for more sophisticated, automated systems for the detection of book defects. This paper introduces a novel approach that utilizes Roboflow and YOLOv8, an advanced object detection model, in conjunction with MIRNet for Low-Light Image (LLI) enhancement to detect and classify various book defects. Our proposed system aims to significantly reduce human effort and increase the accuracy of defect detection. Through extensive experiments and evaluations, we demonstrate the effectiveness of our model in identifying a range of defects such as torn pages, water damage, and broken bindings under various lighting conditions. The integration of MIRNet LLI ensures that the defects are accurately detected even in low-light images, which is common in the storage environments of many historical books. The outcomes suggest that our model not only streamlines the preservation process but also provides a scalable solution for libraries and archives worldwide.

Keywords: Computing methodologies, Artificial intelligence, Computer vision, Object detection, Computer graphics, Image processing

1. Introduction

Books are one of the central carriers of knowledge. Though there are a great deal of digitalized books nowadays, physical books risk deteriorating with age, improper handling, and fading ink. Book defects affect the quality of a book's readability and aesthetics. Preserving the physical vessels of information and their contents becomes vital in museums and libraries. Existing methods for detecting book imperfections include manual inspection and basic image processing techniques. The necessity for automated and efficient book defect detection solutions arises from the challenges and limitations encountered by existing procedures. Manual inspection is time-consuming, labor-intensive, and susceptible to human error. Basic image processing methods, however, require more accuracy and robustness for detecting precise book defects. Thus, there is a demand for modern technological solutions to automate the process and contribute accurate defect detection. Object detection techniques in machine learning bears great capabilities for addressing this need. Through training models on large datasets, machine learning algorithms can learn to identify and classify varying kinds of physical flaws in books with high accuracy. The research gap undertaken by this work is the application of Roboflow, YoloV8, and MIRNET LLI. Roboflow is a cloud platform for training and deploying machine learning models, providing a suitable and accessible framework for the development of the book defect detection model. A state-of-the art object detection algorithm, YoloV8, is known to be efficient and effective in detecting objects in complex displays. By integrating these technologies, this approach can potentially enhance the productivity and scalability of book defect detection, contributing to the advancement of preservation efforts and the accessibility of high-quality books. The rest of the paper discusses a review of related literature, methodologies, results and discussion, conclusion, and recommendations.

+ Corresponding author. Tel.: + 639178695257
E-mail address: jpqtomas@mapua.edu.ph (J.P.Q. Tomas).

2. Review of Related Literature

2.1. YoloV8

The YOLO (You Only Look Once) algorithm series has been instrumental in the field of object detection, with YOLOv8 emerging as a significant advancement in this lineage. It encapsulates the culmination of progressive enhancements in speed, accuracy, and efficiency, achieved through sophisticated neural network architectures and advanced training methodologies. The algorithm's ability to process images in real-time while maintaining high accuracy makes it an invaluable tool for a variety of detection tasks [5]

In the specific area of defect detection, YOLOv8 has demonstrated transformative potential. Its application has been explored across different industries, including infrastructure maintenance and manufacturing quality control. The algorithm's versatility is evident in its successful deployment in detecting road surface defects, identifying irregularities on wind turbine blades, and pinpointing imperfections in miniature electronic components. These advancements underscore the adaptability of YOLOv8 to diverse defect detection scenarios, proving it to be a robust solution for automated inspection systems [6].

The foundation of YOLOv8's success in defect detection lies in the quality of dataset training. Robust datasets that encompass a wide array of defect types, environmental conditions, and object variations are crucial for training the algorithm. This comprehensive training process equips YOLOv8 with the ability to generalize effectively across different real-world scenarios, maintaining high accuracy levels in identifying defects. The extensive nature of these datasets ensures that the algorithm is well-prepared to handle the complexities of practical defect detection [7].

Empirical evidence from experimental results further solidifies YOLOv8's standing in defect detection. Comparative studies highlight the model's superior performance metrics, such as precision and recall, over its predecessors and competing algorithms. These metrics are critical for evaluating the effectiveness of YOLOv8 in practical applications, and they consistently indicate that YOLOv8-based models are more adept at identifying defects accurately and reliably [6][7].

The integration of YOLOv8 into ongoing defect detection research has catalyzed further innovations within the field. Researchers have tailored YOLOv8's architecture to enhance its performance for specific detection tasks. For instance, integrating components like BiFPN and SimAM has led to even more precise detection capabilities. This continuous refinement and adaptation of YOLOv8 not only pushes the boundaries of current automated inspection systems but also paves the way for future breakthroughs in object detection technology [8].

2.2. YoloV7

The YOLOv7 algorithm marks a significant milestone in the evolution of object detection models. Its architecture is designed to optimize both speed and accuracy, a critical consideration for real-time applications. This is achieved by employing advanced techniques such as batch normalization, residual connections, and multi-scale predictions, which enhance the model's ability to generalize across different object sizes and shapes [9]. The innovations in YOLOv7 not only refine its detection capabilities but also streamline its training process, allowing it to rapidly adapt to new datasets and environments. This makes YOLOv7 a versatile tool for researchers and practitioners who require a reliable and efficient object detection model [10].

The YOLOv7 framework has shown substantial improvements in processing speed without compromising the accuracy of detection. This is particularly evident in complex detection scenarios where the model must discern between closely similar objects or defects. Its ability to do so with minimal false positives is a testament to the robustness of the underlying algorithms and the effectiveness of the training methodologies used [11].

In the realm +of defect detection, YOLOv8 stands out as a robust solution capable of identifying a range of anomalies with high precision. Its application extends across various industries, from detecting imperfections in product manufacturing to identifying structural damage in engineering applications. The model's enhanced feature extraction capabilities allow it to discern subtle differences between normal and defective items, which is crucial for maintaining quality control and ensuring the safety of products and structures [7].

Defect detection has been revolutionized by deep learning models like YOLOv7, which offer unprecedented accuracy in identifying anomalies within various materials and products. The application of YOLOv7 in defect detection tasks across different industries has demonstrated its capability to detect fine-grained irregularities that would otherwise be overlooked by human inspectors or less sophisticated automated systems [12]. The model's proficiency in defect detection is particularly valuable in fields such as manufacturing, where it can identify defects in real-time, thus preventing the distribution of flawed products and ensuring high-quality standards [13].

The training of deep learning models like YOLOv7 requires a well-annotated and diverse dataset to ensure the model's efficacy in accurately detecting objects or defects. The creation of such datasets involves meticulous labeling of images to provide the model with clear examples of both normal and defective items. Research has shown that the quality of these datasets is as important as their size; a smaller but well-curated dataset can be more beneficial than a larger, poorly labeled one [14]. Moreover, augmenting datasets with synthetic data generated through various techniques can enhance the model's ability to generalize and perform well on unseen data [15].

Experimental results are crucial for validating the performance of object detection models like YOLOv7. Research has consistently shown that YOLOv7 outperforms its predecessors across standard benchmarks, exhibiting higher precision and recall rates [16]. These experiments typically involve challenging datasets that include a wide range of objects, lighting conditions, and occlusions, thereby providing a comprehensive assessment of the model's capabilities. The results from these experiments are invaluable for researchers looking to apply YOLOv7 to new domains or optimize its performance for specific applications [17].

Integrating YOLOv7 into defect detection research represents a significant leap forward in the pursuit of automated quality assurance systems. The model's architecture, designed for real-time processing, allows for its deployment in dynamic environments where rapid detection and response are crucial. For instance, in the field of book defect detection, YOLOv7's precision enables it to identify issues such as misprints, tears, or warping that can occur during production [18].

Researchers have leveraged YOLOv7's capabilities to develop specialized systems tailored for specific defect detection tasks. For example, in the case of transmission line insulator defects, YOLOv7's adaptability has been harnessed to create a system that can operate under various environmental conditions while maintaining high detection rates [17]. This flexibility is partly due to YOLOv7's ability to learn from diverse data sources, including those that mimic real-world imperfections.

Moreover, integrating YOLOv7 into defect detection workflows often involves fine-tuning the model on specialized datasets. This process ensures that the model can recognize the subtle nuances between defective and non-defective items within specific contexts. The integration process also typically includes rigorous validation phases where the model's predictions are compared against ground truth data to iteratively improve its accuracy [10]. The successful integration of YOLOv7 into defect detection research not only underscores the model's technical prowess but also highlights its practical utility in automating tasks that require high precision and reliability. The ongoing developments in this area suggest a future where such advanced AI-driven systems will become commonplace in ensuring quality across various manufacturing and production industries.

3. Methodology

3.1. Data Collection

Fig. 1: Example of Images Captured from a Single Book

The researchers photographed books from personal collections, with a minimum of 5 pictures each taken per book as shown in Figure 1. These were taken from the following angles: (1) Top view of the front cover (2) Top view of the back (3) Top View of the Spine (4) 45 angle of the book wherein the front cover and foreedges are visible and (5) Pages of the Book. Since there is variation in how defects manifest even within the same book, the scans can be 1 or more, especially for books with more defects. Images were photographed with a white background and overhead lighting.

3.2. Dataset

The final dataset is composed of 372 training images, 107 validation images and 52 test images resulting in 531 images in a 70-20-10 training-test-valid split. These were uploaded to the Roboflow database and manually annotated with the following classification label that were derived based on commonly used terminology for book defects found in book selling platforms such as AbeBooks [3].

1. **Chip-** used to describe where small pieces are missing from the edges of the boards or where fraying has occurred on a dust jacket or the edge of a paperback.
2. **Spot-** spots in paper due to age, otherwise known as foxing in book selling terminology.
3. **Crease-** folds in paper or cover that are visible.
4. **Tear-** holes or missing pieces larger than chip, is not necessarily restricted to the edges of the book or its dust jacket (ex. torn pages).
5. **Stain-** dirt due to liquid damage, faded text etc.
6. **Nick-** scratches or indented marks on the cover or paper, may not be as visible as annotations or mark (ex. scratch marks due to fingernails or coin leaving dents in a book cover).
7. **Mark-** stain due to pen/pentel/paint (ex. Remaindered copy)
8. **Annotations-** highlighted or underlined sections or handwriting, commonly found in used textbooks.
9. **Barcode-**sticker that is used for merchandise identification, usually includes pricing.
10. **Label-** library or owner identification stickers.

The distinction for each of the labels chosen was based on several factors. Although writing and highlight in pages can look visually distinct, they were lumped into one category (Annotations) since these are usually grouped together in this category and have a similar impact on the book's perceived quality. While Mark is also a result of writing implements, it was decided it should have its own distinct copy due to its less intrusive appearance (usually only found in the foredge) and is considered a separate indicator for a book's perceived quality (Remaindered books with marks are often considered a different class from Used Copies with annotations). Library labels and barcodes also have different categories even though these are both stickers since in book grading, library copies are identified as a separate category, and it is therefore useful to have a model that can distinguish the differences.

3.3. Data Preprocessing

Table 1: Class Balance

Class	*No. of Annotations*	*Status*
Spot	**347**	**Overrepresented**
Crease	**302**	
Stain	**276**	

Chip	**260**	
Annotations	**162**	
Barcode	**67**	**Underrepresented**
Label	**62**	**Underrepresented**
Nick	**62**	**Underrepresented**
Tear	**56**	**Underrepresented**
Mark	**48**	**Underrepresented**

All preprocessing steps were natively handled in the Roboflow platform during dataset generation. The images were auto oriented and resized into a 640*640 px for all model variants. Moreover, due to the observed class imbalance and poor performance shown in Table 1, models were further evaluated with variants that utilize label remapping as a preprocessing step. This will result in the reduction of the 10 original classes in an attempt to combine the underrepresented classes due to the lack of diversity that will adequately represent all types of defects. The following remappings were tested and considered, with the following justifications.

1. **Barcode and Label (combined as Label)-, due** to their visually similar appearance and purpose as mentioned in 3.2.
2. **Mark and Nick (combined as Nick)-** combined due to the small nature of their appearances.
3. **Chip and Tear (combined as Chip) -** combined since chips and tears involve pieces of the paper missing or in a separated state, albeit at different levels of severity as per their description.

3.4. Data Augmentation (LLI Enhancement)

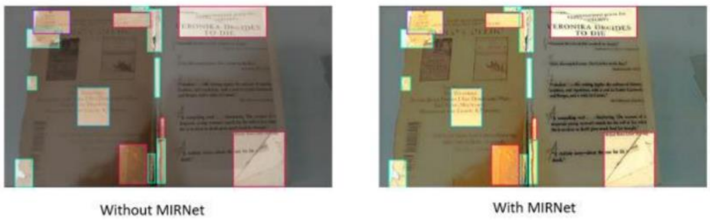

Fig. 2: MIRNet Augmentation Sample

MIRnet is a low light augmentation model developed with the intention of adjusting low light output without sacrificing the original image's feature details, resulting in an effective approach for enhancing the quality of dimly lit captured output. The researchers tested this effectiveness in a previous study involving car type detection, which is why it was considered for this project.

The augmentations using MIRNet was applied to a fully annotated dataset with the aid of Roboflow API as shown in Figure 2. The original dataset with annotations was downloaded from Roboflow and then fed to a custom-made model that automates the process of enhancing the images then uploading it to Roboflow with the same annotations as the original to the dataset.

3.5. Model

The researchers will compare the performance metrics of models trained using Roboflow 3.0 Object Detect and Yolov8, with both models being deployed on the Roboflow dashboard [3]. It also allows for training based on previous checkpoints, which was useful for the purpose of comparing the effect of the applied augmentation and preprocessing steps. The researchers were able to use the free amount of training credits to iteratively train the model to improve based on previous results.

YOLOv8 on the other hand, is a real-time object detection model that efficiently identifies and classifies objects within images or videos [2]. As an evolution of the YOLO architecture, YOLOv8 combines speed and accuracy, dividing images into a grid and predicting bounding boxes and class probabilities for each grid cell making it useful for developing object detection models. Compared to earlier models, YOLOv8 eliminates the need for anchoring boxes which streamlines the training process. It was also chosen specifically since it is the latest version supported by Roboflow for deployment.

For Roboflow Train, the model was less flexible, and the only settings that wre modified were fast for the very first iteration (No Augmentations+ Preprocessing) and accurate for future model. For YoloV8, the model was set to train in 500 epochs with early stopping implemented if no improvement is found for at least 50 epochs. Compared to the Roboflow model, the Yolov8 variant are not trained based on previous iterations. In both cases, the Roboflow platform was useful in generating evaluation metrics by having both models deployed directly on the platform which will result in the generation of mAp, precision and recall scores as well as confusion matrix and vector analysis.

4. Results and Discussion

Table 2: Results of the Model

Model	*Preprocessing*	*Augmentation*	*mAp*	*Precision*	*Recall*	*F1 Score*
Roboflow 3.0 Object Detection	Original Labels	w/o Mirnet	0.10	0.39	0.09	0.15
		with MIRNet	0.11	0.20	0.12	0.15
	Remapped Labels	w/o MIRNet	0.09	0.17	0.12	0.14
		with MIRNet	0.11	0.18	0.13	0.15
YoloV8	Original Labels	w/o Mirnet	0.07	0.41	0.13	0.19
		with MIRNet	0.07	0.41	0.13	0.19
	Remapped Labels	w/o MIRNet	0.07	0.41	0.13	0.19
		with MIRNet	0.07	0.41	0.13	0.19

4.1. Impact of MIRnet LLI Data Augmentation

The impact of MIRNet data augmentation on the performance of the Roboflow 3.0 Object Detection and YOLOv8 models is relatively limited based on the provided results in Table 2. For the Roboflow 3.0 Object Detection model, using MIRNet augmentation shows a slight increase in the mean Average Precision (mAP) from 0.10 to 0.11 when applied with Original Labels preprocessing, while other metrics such as Precision, Recall, and F1 Score do not exhibit significant improvements. Similarly, when Remapped Labels preprocessing is used, MIRNet augmentation leads to similar slight improvements in mAP but minimal changes in other performance metrics. However, for the YOLOv8 model, the impact of MIRNet augmentation appears to be negligible, with performance metrics remaining largely unchanged regardless of whether the augmentation is applied or not. Overall, while MIRNet augmentation may offer some marginal benefits for the Roboflow 3.0 Object Detection model, its impact on other metrics and the YOLOv8 model is minimal based on the provided data.

4.2. Effects of Data Preprocessing

The overall effect of data preprocessing on the performance of object detection models, as depicted in Table 2, appears to be marginal. Despite variations in the choice of preprocessing techniques and the inclusion of Mirnet, the changes in performance metrics such as mean Average Precision (mAP), precision, recall, and F1 score are relatively minor across different model configurations. These preprocessing steps, natively handled in the Roboflow platform during dataset generation, included autoorientation and resizing of images to a standardized format of 640*640 pixels for all model variants. Additionally, due to observed class imbalance and poor performance shown in Table 1, models were further evaluated with variants that utilized label remapping as a preprocessing step. This involved reducing the 10 original classes to combine underrepresented classes to adequately represent all types of defects. Specifically, remapping were tested and considered based on visual similarities and characteristics of the defects, with Barcode and Label combined as Label due to their visually similar appearance and purpose, Mark and Nick combined as Nick due to the small nature of their appearances, and Chip and Tear combined as Chip since chips and tears involve pieces of paper missing or in a separated state, albeit at different levels of severity. However, despite these preprocessing efforts, the observed differences in performance metrics across model configurations were not substantial, suggesting that the impact of data preprocessing on model performance in this context may be limited. Further analysis and experimentation may be warranted to fully understand the implications of these preprocessing techniques on the effectiveness of object detection models.

4.3. Model Comparison (YoloV8 and Roboflow)

Comparing the performance of the Roboflow 3.0 Object Detection and YOLOv8 models reveals several key differences. When examining the mean Average Precision (mAP), Roboflow 3.0 Object Detection demonstrates a slight improvement from 0.10 to 0.11 with MIRNet augmentation, irrespective of whether Original Labels or Remapped Labels preprocessing is used. In contrast, YOLOv8 maintains a consistent mAP of 0.07 regardless of the augmentation or preprocessing method employed. Precision varies between the models; while Roboflow 3.0 Object Detection experiences a decrease from 0.39 to 0.20 for Original Labels and a slight change from 0.17 to 0.18 for Remapped Labels with MIRNet augmentation, YOLOv8 maintains a steady precision of 0.41 across all scenarios. In terms of recall, Roboflow 3.0 Object Detection sees marginal improvements from 0.09 to 0.12 for Original Labels and from 0.12 to 0.13 for Remapped Labels with MIRNet augmentation, while YOLOv8 maintains a constant recall of 0.13 regardless of augmentation or preprocessing. Finally, both models exhibit consistent F1 Scores, with Roboflow 3.0 Object Detection showing minimal changes and YOLOv8 maintaining a stable F1 Score of 0.19. Despite the slight improvements observed in some metrics for Roboflow 3.0 Object Detection with MIRNet augmentation, YOLOv8 displays a more consistent performance across all scenarios.

5. Conclusion

In this study, the researchers investigated the impact of MIRNet LLI (Low-Light Image) data augmentation on two object detection models: Roboflow 3.0 Object Detection and YOLOv8. The methodology encompassed comprehensive data collection, resulting in a dataset of 372 training images, 107 validation images, and 52 test images, which were annotated with classification labels for various book defects. Preprocessing steps, including auto-orientation and resizing, were conducted to standardize the dataset. Additionally, due to class imbalance, label remapping was employed to combine underrepresented classes, aiming to improve model performance. MIRNet augmentation was then applied to enhance the dataset's quality, particularly in low-light conditions, using a custom-made model. Subsequently, the researchers compared the performance metrics of the two models, focusing on mean Average Precision (mAP), precision, recall, and F1 score. The results highlighted subtle improvements in the Roboflow 3.0 Object Detection model's performance metrics with MIRNet augmentation, particularly in mAP and recall. However, the impact on the YOLOv8 model appeared negligible, with consistent metrics observed regardless of augmentation. Further analysis revealed marginal changes in performance metrics across different preprocessing techniques and model configurations, indicating limited effects of data preprocessing on model effectiveness. Despite variations in preprocessing and augmentation, the differences in performance between the two models were notable. While Roboflow 3.0 Object Detection showed slight improvements with MIRNet augmentation, YOLOv8 demonstrated more consistent performance across all scenarios. While MIRNet augmentation may offer marginal benefits for enhancing the Roboflow 3.0 Object Detection model's performance, its impact on other metrics and the YOLOv8 model is minimal. The study underscores the importance of rigorous experimentation and comparison to understand the effectiveness of different augmentation and preprocessing techniques in improving object detection model performance. Further research may delve deeper into exploring alternative augmentation methods and their implications for enhancing model robustness and accuracy in diverse environmental conditions.

6. Recommendations

As observed in the overall poor performance of all models, the study has more limitations that could be amended in future work. Model performance is highly dependent on dataset quality and diversity to adequately identify the visual differences between the different types of defects, and as evidenced by the effect of class remapping cannot be easily amended by reducing the number of classes. Dataset should be strategically diversified by providing more samples that exhibit the underrepresented classes. Future work could also entail introducing more samples with variations in lighting quality to better assess the impact of the MIRNet LLI. Another avenue to consider is to try applying the LLI augmentation before annotation, since it was also intended to make visual features more pronounced and is therefore useful for highlighting less visible defects like nick and creases. With augmentations applied, it may be easier to label defects and result in an overall improvement in ground truth annotation quality. With all these improvements incorporated and with the training of a model with good enough performance, the the object detection model could be further refined into an advanced computer vision project which involves automatically assigning a book rating quality (e.g Ex-Library Copy, Like New, Good) based on the identified defects. This would be helpful in the standardization and automation of the upload process in used book reselling platforms where this type of standardized terminology is used.

7. Acknowledgement

The authors would like to express their deep gratitude to the School of Information Technology at Mapua University, Makati, Philippines, for providing the resources and environment conducive to carrying out this research. Special thanks are extended to the faculty members for their invaluable guidance and support throughout this project. We also acknowledge the contributions of the participants who generously shared their personal book collections for the data collection phase of our study. Their willingness to assist played a crucial role in the development of our book defect detection model. Our sincere appreciation goes to the Roboflow platform for facilitating the training and deployment of our machine learning models. The platform's robust features and user-friendly interface were instrumental in the successful completion of our experiments. We are grateful to the authors and creators of the YOLOv8 and MIRNet LLI technologies for making their work accessible, which significantly enhanced the capabilities of our object detection model. Lastly, we thank our peers and colleagues who provided feedback and insights that greatly improved the quality of our research. Their constructive criticism and encouragement were vital in refining our methodology and analysis. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The work presented was supported by the dedication and commitment of all those involved.

8. References

[1] Sunanda Perla and Kavitha Dwaram. 2023. Low Light Image Illumination Adjustment Using Fusion of MIRNet and Deep Illumination Curves. 2023. Springer, 620–636.

[2] Juan Terven, Diana-Margarita Córdova-Esparza, and Julio-Alejandro Romero-González. 2023. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Machine Learning and Knowledge Extraction 5, 4 (2023), 1680–1716.

[3] A guide to used book conditions. Retrieved February 5, 2024 from https://www.abebooks.com/books/rarebooks/collecting-guide/understanding-rare-books/guide-book-conditions.shtml

[4] Train a Model in Roboflow - Roboflow Docs. Retrieved February 5, 2024 from https://docs.roboflow.com/train/train

[5] Wang, X., Gao, H., Jia, Z., & Li, Z. (2023). BL-YOLOV8: an improved road defect detection model based on YOLOV8. *Sensors*, *23*(20), 8361. https://doi.org/10.3390/s23208361

[6] Hussain, M. (2023). YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. *Machines*, *11*(7), 677. https://doi.org/10.3390/machines11070677

[7] Narlan, R. (2023). Automated Pavement Defect Detection Using YoloV8 Object Detection Algorithm. *proceeding.hpji.or.id*. https://doi.org/10.58674/phpji.v16i1.388

[8] Yao, H., Tan, W., Li, L., & Wu, L. (2023b). WFRE-YOLOV8S: a new type of defect detector for steel surfaces. *Coatings*, *13*(12), 2011. https://doi.org/10.3390/coatings13122011

[9] Wei, G., Wan, F., Zhou, W., Xu, C., Ye, Z., Liu, W., Lei, G., & Xu, L. (2023). BFD-YOLO: A YOLOV7-Based detection method for building façade defects. *Electronics*, *12*(17), 3612. https://doi.org/10.3390/electronics12173612

[10] Liu, B., Wang, H., Cao, Z., Wang, Y., Lü, T., Yang, J., & Zhang, K. (2024). PRC-Light YOLO: an efficient lightweight model for fabric defect detection. *Applied Sciences*, *14*(2), 938. https://doi.org/10.3390/app14020938

[11] Wang, Y., Wang, H., & Xin, Z. (2022). Efficient Detection model of steel strip surface defects based on YOLO-V7. *IEEE Access*, *10*, 133936–133944. https://doi.org/10.1109/access.2022.3230894

[12] Ji, Y., & Di, L. (2023). Textile defect detection based on multi-proportion spatial attention mechanism and channel memory feature fusion network. *Iet Image Processing*. https://doi.org/10.1049/ipr2.12957

[13] Zheng, J., Wu, H., Zhang, H., Wang, Z., & Xu, W. (2022). Insulator-Defect detection algorithm based on improved YOLOV7. *Sensors*, *22*(22), 8801. https://doi.org/10.3390/s22228801

[14] Wang, R., Liang, F., Mou, X., Chen, L., Yu, X., Peng, Z., & Chen, H. (2023). Development of an improved YOLOV7-Based model for detecting defects on strip steel surfaces. *Coatings*, *13*(3), 536. https://doi.org/10.3390/coatings13030536

[15] Fan, X., Tan, J., Zhang, J., Gong, Y., Zhang, H., Huang, C., & Chen, J. (2023). TOFD defect Image recognition based on YOLOV7. *Journal of Physics: Conference Series*, *2560*(1), 012047. https://doi.org/10.1088/1742-6596/2560/1/012047

[16] Cui, W., Li, Z., Duanmu, A., Xue, S., Guo, Y., Ni, C., Zhu, T., & Zhang, Y. (2024). CCG-YOLOV7: A wood defect detection model for small targets using improved YOLOV7. *IEEE Access*, 1. https://doi.org/10.1109/access.2024.3352445

[17] Panigrahy, Satyajit & Karmakar, Subrata. (2022). Research on Transmission Line Insulator Defects Detection using YOLOv7. 384-388. 10.1109/CATCON56237.2022.10077650.

[18] *Steel Strip Quality Assurance with YOLOV7-CSF: a coordinate attention and SIOU fusion approach*. (2023). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/10320359
CRE Press, 1998.