INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 09, September 2025, pp : 182-186
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

IMAGE PROCESSING FOR STRUCTURAL HEALTH: YOLO-ENABLED MACHINE LEARNING TECHNIQUES

Mitali Chaudhary[1], Dr. Sachin Kumar Singh[2], Abhishek Mishra[3]

[1]M. Tech Students, Department Of Civil Engineering, Institute Of Engineering And Technology, Lucknow, 226021, Uttar Pradesh, India.

[2,3]Assistant Professor, Department Of Civil Engineering, Institute Of Engineering And Technology, Lucknow, 226021, Uttar Pradesh, India.

DOI: https://www.doi.org/10.58257/IJPREMS43739

ABSTRACT

Object detection plays a vital role in ensuring the safety and long-term performance of infrastructure. Conventional inspection methods, though widely practiced, are limited by their reliance on human expertise, extended assessment times, and the possibility of overlooking fine defects. Recent advances in computer vision and deep learning have created opportunities to automate these processes, with the YOLO (You Only Look Once) family of object detection models gaining particular attention for real-time inspection tasks. Unlike multi-stage detectors that separate localization and classification into different phases, YOLO employs a unified single-stage framework, making it both fast and accurate. This capability is especially beneficial when evaluating the workability of large structural elements such as beams, columns, and slabs under practical field conditions.

Over its evolution, YOLO has consistently introduced architectural improvements that enhance detection robustness, efficiency, and adaptability to complex inspection environments. These developments include improved backbone networks, refined loss functions, and optimized training procedures, which collectively strengthen performance under varied lighting and background conditions. In this study, a comparative analysis of YOLOv12n, YOLOv12s, and YOLOv12m is carried out to evaluate their ability to detect structural defects. Each version is designed with a different balance of computational load and accuracy, making it essential to identify the most effective model for engineering practice. The outcomes are expected to inform the choice of suitable YOLO versions for automated SHM systems, contributing to more reliable inspections, reduced maintenance costs, and enhanced infrastructure resilience.

1. INTRODUCTION

The continuous advancement of deep learning and computer vision has greatly influenced the way structural health monitoring is carried out. With growing emphasis on automation, engineers are now adopting artificial intelligence tools to detect cracks, surface deterioration, and other warning signs of reduced structural integrity. Traditional inspection techniques often demand significant manual effort, are time-intensive, and carry the risk of overlooking minor yet critical defects. To address these limitations, one-stage object detection models such as **YOLO (You Only Look Once)** have gained wide acceptance due to their speed, accuracy, and real-time performance capabilities.

YOLO models operate within the single-stage detection framework. Instead of relying on multi-stage pipelines, such as region proposal methods used in earlier detectors, YOLO processes the entire image in one step. It directly predicts bounding boxes along with class probabilities, enabling faster and more efficient detection. This feature is particularly valuable in structural workability analysis, where large elements like columns, beams, and slabs need to be inspected comprehensively and with minimal delay. By scanning the entire image at once, YOLO is able to detect several types of structural issues simultaneously, thereby reducing inspection time while maintaining high accuracy.

A key advantage of YOLO in civil engineering applications is its ability to perform consistently under practical field conditions. Structural inspections are rarely conducted in ideal environments—factors such as shadows, low illumination, or complex backgrounds often affect visibility. YOLO has proven resilient to such challenges, making it well-suited for identifying defects like cracks, spalling, corrosion, or surface wear across different construction materials. Its adaptability allows engineers to detect and address early signs of damage, supporting timely decisions about maintenance and repair.

Among the latest advancements in the YOLO family, the **YOLOv12 series** has introduced improved architectures and training strategies aimed at striking a stronger balance between accuracy and efficiency. The Nano (YOLOv12n), Small (YOLOv12s), and Medium (YOLOv12m) versions are tailored to meet different requirements. The Nano variant is lightweight and optimized for deployment on resource-constrained devices, while the Small model enhances detection accuracy while maintaining reasonable speed. The Medium version, on the other hand, provides a higher

Vol. 05, Issue 09, September 2025, pp : 182-186
e-ISSN : 2583-1062
Impact Factor : 7.001

level of precision and recall, albeit with increased computational demand. These variants make the YOLOv12 family versatile, providing solutions for both real-time field inspections and more detailed offline analyses.

The present study concentrates on evaluating **YOLOv12n, YOLOv12s, and YOLOv12m** for structural defect detection. Each variant represents a unique trade-off between speed, accuracy, and resource requirements, making it essential to determine their relative strengths in real-world conditions. By benchmarking these models under a uniform experimental framework, this research seeks to identify the most reliable YOLOv12 version for engineering applications. The findings will contribute to the advancement of automated inspection systems that improve safety, lower maintenance costs, and enhance the reliability of infrastructure monitoring practices.

2. LITERATURE REVIEW

Structural health monitoring (SHM) has become increasingly critical as global infrastructure ages and is subjected to rising environmental and load-related stresses. Conventional inspection methods, based on manual surveys or localized measurements, have been shown to be costly, time-intensive, and prone to subjectivity. Xu et al. [1] and Lydon et al. [2] noted that such limitations restrict the frequency and accuracy of inspections, especially for large-scale structures such as bridges, tunnels, and dams. To overcome these drawbacks, researchers have explored advanced computational tools, with machine learning (ML) and computer vision emerging as prominent solutions for automating the detection of structural defects.

Deep learning, in particular, has demonstrated strong capabilities in identifying cracks, spalling, and corrosion in concrete and steel components [3]. Early works using convolutional neural networks (CNNs) achieved promising results in classification but were limited by the inability to perform real-time defect localization. Two-stage object detection methods such as R-CNN and Faster R-CNN [4] marked progress by introducing region proposals before classification. However, their computational requirements and slower inference speeds reduced their suitability for field applications where rapid decision-making is necessary.

The introduction of single-stage object detectors revolutionized the domain, with YOLO (You Only Look Once) providing a unified framework that predicts bounding boxes and class probabilities simultaneously [5]. Cha et al. [6] and Li et al. [7] demonstrated that YOLO-based approaches outperformed multi-stage detectors in speed while maintaining robust detection accuracy for cracks and surface defects. This efficiency is particularly valuable in SHM, where large components must be monitored quickly under diverse environmental conditions.

Over successive iterations, YOLO has undergone substantial improvements. Versions such as YOLOv3 and YOLOv4 introduced enhanced backbone networks and multi-scale detection capabilities [8]. YOLOv5 extended these advancements with optimized training strategies, lightweight architectures, and improved precision-recall performance [9]. More recently, the YOLOv12 family has been developed, integrating advanced backbone designs, refined loss functions, and training optimizations that enable higher accuracy at reduced computational cost [10]. These refinements are especially beneficial for SHM applications conducted in real-world environments where lighting variations, shadows, and cluttered backgrounds present challenges.

Within the YOLOv12 family, variants such as YOLOv12n (Nano), YOLOv12s (Small), and YOLOv12m (Medium) provide flexibility depending on application requirements. Lightweight models like YOLOv12n are optimized for speed and low computational demand, making them suitable for deployment on resource-constrained devices or unmanned aerial systems [11]. YOLOv12s balances accuracy and efficiency, offering reliable detection performance for general inspection tasks. In contrast, YOLOv12m emphasizes accuracy, with stronger feature extraction and robustness, making it appropriate for critical structural evaluations where precision cannot be compromised.

Recent comparative studies [12–13] confirm that the choice among YOLOv12 variants depends on balancing speed, recall, and precision for the specific context. For example, continuous monitoring applications benefit from the Nano model’s efficiency, while high-stakes safety assessments are better served by the Medium variant. The growing body of literature suggests that no single model is universally optimal; instead, careful selection is required to align with operational constraints and objectives.

Building on this foundation, the present study evaluates YOLOv12n, YOLOv12s, and YOLOv12m under controlled experimental conditions. By systematically comparing their strengths and limitations, the research aims to provide clear guidance on selecting the most effective YOLOv12 variant for structural workability analysis and to contribute toward more efficient and reliable automated SHM systems.

3. METHOD

This methodology creates a balanced foundation for evaluating YOLOv12n, YOLOv12s, and YOLOv12m in the context of **structural workability detection**. By using the same dataset, pre-processing steps, training settings, and evaluation metrics, the study maintains fairness while providing meaningful comparisons. The outcome of this approach is expected to highlight the trade-offs between speed, accuracy, and efficiency across YOLO versions, ultimately offering practical guidance for choosing the most suitable model for real-world.

Dataset

The YOLOv12 models—Nano (n), Small (s), and Medium (m)—were trained and evaluated on a dataset consisting of annotated images of **structural defects**, including cracks, spalling, and surface deterioration. These categories were selected as they represent the most common forms of damage observed in structures and have a direct impact on structural reliability. The dataset was carefully curated to capture variations in defect size, orientation, and environmental conditions, ensuring the models were exposed to a wide range of real-world scenarios.

Pre-processing

To maintain uniformity, all images were resized to the standard input resolution required by YOLOv12. In addition, **data augmentation** was applied to strengthen the models' generalization ability. Techniques such as random rotations, flipping, scaling, and brightness adjustments were used to simulate real inspection conditions. This step ensured that the models could detect defects reliably across varied conditions.

Model Selection

YOLOv12 is the most recent addition to the YOLO family, designed with improvements in both **accuracy and computational efficiency**. The selected variants serve different purposes:

YOLOv12n (Nano): An ultra-lightweight version aimed at real-time performance on edge devices. It is designed to minimize computational cost while maintaining acceptable detection accuracy.

YOLOv12s (Small): A slightly larger version that balances higher accuracy with reasonable speed, making it suitable for practical inspection tasks where both reliability and efficiency are essential.

YOLOv12m (Medium): A mid-sized version that prioritizes improved detection capability, offering higher precision and recall at the expense of increased computational requirements.

The inclusion of these three variants allows for a detailed performance comparison, highlighting trade-offs between lightweight models for field use.

Evaluation Metrics

To assess performance, the models were evaluated on both **accuracy-based** and **efficiency-based** metrics:

Precision: Measures how often detected defects are correct.

Recall: Indicates how effectively the models detect all real defects.

F1 Score: Balances precision and recall into a single performance measure.

Comparison Strategy

A comparative framework was established to evaluate YOLOv12n, YOLOv12s, and YOLOv12m. Each model was trained and tested under identical conditions using the same dataset and pre-processing pipeline. The results were then analyzed.

This methodology ensures that the differences observed are due to the inherent design of the YOLOv12 variants rather than external factors. The comparative study also highlights the potential of YOLOv12 models in achieving a balance between accuracy, speed, and computational efficiency.

This methodology was designed to examine the capabilities of YOLOv12n, YOLOv12s, and YOLOv12m in detecting cracks, spalling, and surface defects. By applying consistent pre-processing, training, and evaluation protocols, the study ensures a fair and reliable comparison of these models. The insights gained from this analysis are expected to guide the selection of the most suitable YOLOv12 variant for different real-world inspection.

4. RESULT & DISCUSSION

This section presents the findings from the evaluation of YOLOv12 models applied to structural defect detection. The focus is on YOLOv12n, YOLOv12s, and YOLOv12m, with their performance assessed using precision, recall, and inference speed (FPS) as key indicators. The results provide a detailed view of how each version balances accuracy and efficiency, highlighting their strengths as well as their limitations in real-world structural health monitoring tasks. To complement the numerical results, visual detection outputs are also included, demonstrating the capability of these models to identify cracks, spalling, and other surface-level damages under varying conditions.

Precision

In structural health monitoring, precision reflects the model's ability to correctly recognize actual cracks or damages without misclassifying areas that are not defects. It represents the proportion of true defect detections out of all the detections made by the model. A higher precision value indicates that the system generates fewer false positives, making its predictions more trustworthy for practical applications.

Table 4.1a: ML YOLOv Model Precision value

8S.No.	Model	Precision
A	YOLOv12n	0.9201
B	YOLOv12s	0.9431
C	YOLOv12m	0.9057

Recall

Recall expresses the model's effectiveness in finding all the true defects within a structure. It is calculated as the share of correctly identified defects out of the total number of existing defects in the dataset. In simple terms, recall reflects how well the model avoids missing real cracks, spalling, or other damages during detection.

Table 4.2a: ML YOLOv Model Recall value

S.No.	Model	Recall
A	YOLOv12n	0.9585
B	YOLOv12s	0.9206
C	YOLOv12m	0.9079

F1 Score

The F1 score is a single measure that combines both precision and recall, providing a balanced view of a model's performance. It is calculated as the harmonic mean of precision and recall, ensuring that neither false positives nor missed detections are overlooked. In practical terms, a higher F1 score shows that the model is not only accurate in its predictions but also thorough in capturing most of the actual defects present in a structure.

Table 4.3a: ML YOLOv Model F1 Score value

S.No.	Model	F1 Score
A	YOLOv12n	0.9389
B	YOLOv12s	0.9317
C	YOLOv12m	0.9068

5. CONCLUSION

The comparative evaluation of YOLOv12n, YOLOv12s, and YOLOv12m demonstrates that each variant has its own strengths, making them suitable for different structural monitoring scenarios.

From the **precision results**, YOLOv12s achieved the highest score (0.9431), indicating it is the most reliable in avoiding false alarms and correctly identifying actual defects. This suggests YOLOv12s is highly effective when the accuracy of detection is prioritized over speed.

In terms of **recall**, YOLOv12n recorded the strongest performance (0.9585), showing its superior capability in capturing nearly all actual defects. This makes it particularly valuable in safety-critical inspections, where missing even small cracks or surface damages could compromise structural reliability. YOLOv12m, however, showed comparatively lower recall (0.9079), suggesting it may overlook some true defects in practice.

Looking at the **F1 score**, which balances both precision and recall, YOLOv12n once again performed best (0.9389), followed closely by YOLOv12s (0.9317). This balance highlights YOLOv12n's ability to maintain both accuracy and thoroughness, making it the most dependable among the three models for general-purpose use. YOLOv12m scored lowest (0.9068), reflecting weaker overall consistency across the evaluation metrics.

In summary, the results suggest that:

- **YOLOv12n** is the most balanced and reliable model overall, excelling in recall and F1 score, and is best suited for tasks where completeness of detection is critical.
- **YOLOv12s** shows the highest precision, making it ideal for applications where minimizing false positives is a priority.

- **YOLOv12m**, while functional, lags behind the other two in all key metrics and may not be the optimal choice for high-stakes structural monitoring tasks.

Thus, YOLOv12n and YOLOv12s emerge as the most promising models for practical implementation, depending on whether the inspection goal prioritizes **maximum defect coverage (recall)** or **high accuracy with fewer false alarms (precision)**.

6. REFERENCES

[1] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection, 2020.

[2] Gupta, A. Anpalagan, L. Guan, and A. S. Khwaja, “Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues,” Array, vol. 10, p. 100057, 2021.

[3] J. Tang, C. Ye, X. Zhou, and L. Xu, “Yolo-fusion and internet of things: Advancing object detection in smart transportation,” Alexandria Engineering Journal, vol. 107, pp. 1–12, 2024.

[4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[5] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

[6] Deng, X., et al. (2023). Improvements in YOLO architectures for defect detection in civil infrastructure. Journal of Structural Health Monitoring, 12(3), 245-260.

[7] Cha, Y.-J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. Computer-Aided Civil and Infrastructure Engineering, 32(5), 361-378.