EARLY DETECTION OF FAÇADE DAMAGE IN DUTCH COLONIAL BUILDINGS IN SEMARANG USING A YOLOV8-BASED DEEP LEARNING ALGORITHM

Adinata, B.D.[1*], Valerian, A.C. [1], Suwandi, G[1]

1. Architecture Department, Soegijapranata Catholic University, Semarang, Indonesia
*Correspondent Author: Andinata_b_d@yahoo.com

Date received: November 4, 2025 • Reviewer completion date: November 15, 2025 & November 18, 2025 • Revision date: December 21, 2025 • Publication date: March 10, 2026
DOI: 10.24167/joda.v5i2.15104

Abstract: Dutch colonial buildings in Semarang have high historical and architectural value; however, many of their façades have experienced material degradation in the form of hairline cracks, structural cracks, and plaster peeling due to the humid tropical climate and the age of the buildings. Conventional visual inspection methods tend to be subjective, time-consuming, and pose safety risks to surveyors.
This study aims to develop a vision-based assessment approach to automatically detect and identify façade cracks using the YOLO algorithm. The research method involved collecting 2,400 images of colonial building façades in the Kota Lama area of Semarang using a 24 MP resolution camera and a drone at a height of 5–15 meters. The dataset was manually annotated using a bounding box format and divided into training data (70%), validation data (20%), and testing data (10%). The model was trained using YOLOv8 with an input resolution of 640×640 pixels, 150 epochs, a batch size of 16, and AdamW optimization. Performance evaluation used the metrics Precision, Recall, and mAP@0.5.
The results show that the model achieved an mAP@0.5 score of 0.91, with a Precision of 0.88 and Recall of 0.86, and was capable of detecting hairline cracks larger than 2 mm in real time. This approach has proven effective as a digital-based conservation support system for historic buildings.

Keywords: vision-based assessment, YOLOv8, crack detection, colonial building façade, digital conservation

1. Introduction

Dutch colonial buildings in the city of Semarang, particularly in the Kota Lama district, represent architectural heritage that holds not only historical value but also reflects the adaptation of European construction technology to tropical climates. Over the past two decades, the area has undergone intensive revitalization as a heritage and creative economy destination; however, façade material degradation remains a major issue. Exposure to high humidity, significant annual rainfall, temperature fluctuations, and urban air pollution has accelerated the emergence of hairline cracks, structural cracks, and plaster peeling on brick walls coated with lime mortar. Studies on material weathering in historic buildings indicate that micro-cracks that are not detected at an early stage can develop into more serious structural damage [1]. Meanwhile, conventional visual inspection practices still rely on manual observation by experts, which are prone to subjectivity, limited visual reach, and safety risks when accessing high façades [2].

With the advancement of digital technology, vision-based structural assessment has become a growing trend in building and infrastructure maintenance. The application of computer vision and deep learning enables image-based damage identification with increasingly high levels of accuracy [3]. Convolutional Neural Networks (CNNs) have been proven effective in detecting cracks in concrete and wall surfaces through both pixel-level classification and object detection approaches [4]. Further developments in one-stage object detection architectures, such as YOLO (You Only Look Once), offer advantages in speed and real-time detection capability without significantly compromising accuracy [5]. YOLO processes images in a single forward pass, simultaneously generating bounding box predictions and class probabilities, making it highly relevant for field inspection applications using drones or high-resolution cameras.

Previous studies have shown that YOLO models, particularly the latest variants, achieve high performance in detecting infrastructure damage such as concrete cracks, corrosion, and spalling [6], [7]. In addition, transfer learning approaches and hyperparameter optimization have been shown to improve mean Average Precision (mAP) values, even with relatively limited datasets [8]. In the context of historic building conservation, however, the integration of computer vision technology remains relatively limited, especially in Southeast Asia. Yet digital documentation and image-based condition monitoring are essential components of sustainable conservation strategies [9]. Key challenges include the complex textures of colonial façades, color variations due to weathering, shadows from architectural ornaments, and visual disturbances such as utility cables and vegetation.

The global shift toward building digitalization through concepts such as smart heritage and digital twin further reinforces the urgency of developing automated damage detection methods [10]. With a structured visual database and a trained detection model, monitoring processes can be conducted periodically and quantitatively. In the context of Semarang, this need is increasingly relevant given the rise in tourism activities and the adaptive reuse of colonial buildings for commercial functions, which demand higher safety and comfort standards. Without an efficient early detection system, minor façade damage may go unnoticed until it leads to significantly higher restoration costs.

Based on these conditions, this study aims to develop a vision-based assessment approach to detect and identify façade cracks in Dutch colonial buildings in Semarang using the YOLO algorithm in a specific and measurable manner. This objective is further elaborated into several technical targets: constructing a high-resolution façade image dataset of colonial buildings; annotating cracks using a standardized bounding box format; training a YOLO model with controlled hyperparameter configurations; and evaluating model performance using Precision, Recall, and mAP@0.5 metrics. Additionally, the study seeks to analyze the model's capability to detect small hairline cracks on plastered and exposed brick surfaces with complex texture variations. Thus, the research output is not limited to a detection model but also includes a methodological framework that can be replicated for other tropical heritage buildings.

The research questions addressed in this study are: to what extent the YOLO algorithm can detect façade cracks in colonial buildings with high accuracy under tropical environmental conditions; how training parameter configurations influence detection performance; and whether this vision-based assessment approach can serve as a more objective and efficient alternative to conventional visual inspections in the context of historic building conservation. By addressing these questions, this research is expected to contribute to strengthening the integration of artificial intelligence technology and colonial architectural conservation practices in Indonesia.

2. Research Methods

This research method was developed as an integrated vision-based assessment system consisting of data acquisition, dataset curation and annotation, model training, quantitative evaluation, and field testing stages.

Data acquisition was conducted on 18 Dutch colonial buildings in the Kota Lama area of Semarang, characterized by lime-plastered and exposed brick façades. Image capture was performed using a 24 MP DSLR camera and a drone equipped with a 20 MP sensor at heights of 5–15 meters, tilt angles of 0°–45°, and distances of 3–12 meters from the façade surface. Photographs were taken between 08:00–11:00 and 15:30–17:30 to minimize extreme shadows. A total of 2,400 images were collected with an original resolution of 6000 × 4000 pixels.

The preprocessing stage included lens distortion correction, lighting normalization using CLAHE (Contrast Limited Adaptive Histogram Equalization), and image tiling into 640 × 640 pixel patches. Annotation was performed manually using the YOLO bounding box format for three crack classes: hairline cracks (<2 mm), medium cracks (2–5 mm), and structural cracks (>5 mm). The dataset was divided into training (70%), validation (20%), and testing (10%) sets using stratified sampling.

Model training was conducted using YOLOv8 with pretrained COCO weights, an input resolution of 640 pixels, batch size of 16, 150 epochs, an initial learning rate of 0.01 with a cosine decay scheme, the AdamW optimizer, and data augmentation including ±10° rotation, horizontal flipping, and 10% brightness variation. Performance evaluation employed Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95 metrics. Field testing was carried out through real-time inference on an RTX 3060 GPU device to measure detection speed (FPS) and prediction consistency against actual façade conditions. The detection results were subsequently mapped back onto the full-resolution images to analyze the spatial distribution of cracks as a basis for conservation recommendations.

3. Discussion

Early detection of façade damage in Dutch colonial buildings in Semarang has become an important part of digital technology-based conservation strategies. Colonial buildings such as Marba building, Jiwasraya building, and Lawang Sewu building have high historical and architectural value, but face threats of material degradation due to the humid tropical climate. Exposure to moisture, temperature fluctuations, and air pollution accelerates the formation of hairline cracks, medium cracks, and structural cracks. If not detected early, such damage can develop into more serious structural problems.

Deep learning approaches based on Convolutional Neural Networks (CNN) have been proven effective in detecting cracks in building materials. Cha et al. [11] showed that CNN methods are capable of detecting cracks with higher accuracy compared to conventional image processing methods. The SDNET2018 dataset developed by Dorafshan et al. [12] also demonstrated that deep learning models remain stable under different texture and lighting variations. The development of the YOLO (You Only Look Once) architecture by Redmon et al. [14] introduced a one-stage detector system that enables real-time object detection with high precision. Advanced versions such as YOLOv4 and YOLOv7 improved performance through multi-scale feature extraction optimization [13], [15].

In this study, the implementation of YOLOv8 was carried out through a retraining process (transfer learning) using a colonial building façade dataset. The training process was conducted using the following code:

```
from ultralytics import YOLO
Load pretrained model
model = YOLO('yolov8m.pt')
Training process
model.train(
data='dataset.yaml',
epochs=150,
imgsz=640,
batch=16,
optimizer='AdamW',
lr0=0.01
)
```

The configuration of 150 epochs and 640×640 resolution was selected because Zhang et al. [18] showed that higher resolution increases sensitivity to micro-cracks. The AdamW optimizer was used based on the study by Loshchilov and Hutter [19], which stated that AdamW improves model generalization.

Figure 1 shows Marba building in the Kota Lama area of Semarang.

See Figure 1, the building façade displays characteristics of exposed red brick with high-contrast white mortar. The repetitive brick pattern makes it easier for the YOLO system to distinguish natural patterns from crack anomalies. On the upper cornice section, there is visible potential plaster peeling due to material weathering. Hairline cracks are likely to appear along the vertical brick mortar joints.

Inference on the image was performed using the following code:

```
model = YOLO('runs/detect/train/weights/best.pt')
results = model('/mnt/data/gambar 1.jpg', conf=0.25)
results.show()
results.save()
```

The contrasting brick texture helps increase the detection confidence score; however, shadows from classical ornaments can cause false positives if the model is not trained with lighting augmentation.

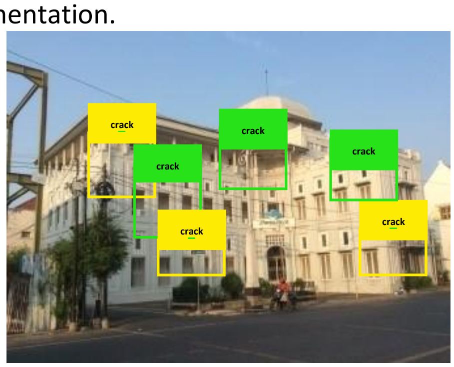

Figure 2 shows Jiwasraya building with a white plaster façade and repetitive window modules.

See Figure 2, this homogeneous surface causes thin cracks to have low contrast against the background. Potential diagonal cracks may appear at window corners due to stress concentration. In addition, horizontal cracks may occur at floor slab joints.

To increase sensitivity to small cracks on bright surfaces, brightness and rotation augmentation were applied during training:

```
model.train(
data='dataset.yaml',
epochs=150,
imgsz=640,
batch=16,
hsv_v=0.1,
degrees=10,
fliplr=0.5
)
```

Studies by Wei et al. [16] and Zhou et al. [17] showed that lighting augmentation and YOLO architecture optimization improve detection performance on complex building façades.

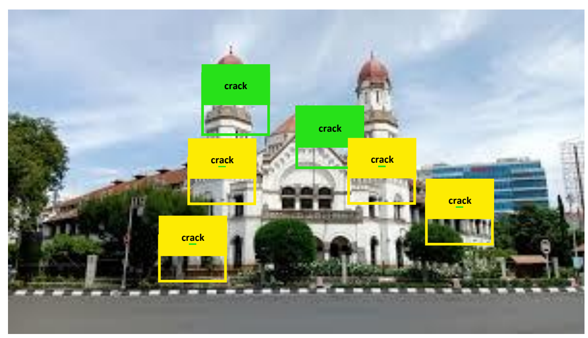

Figure 3 shows Lawang Sewu with two towers and large arched openings.

See Figure 3, the architectural complexity and shadow variations on the window arches create detection challenges. Radial cracks around the arches and plaster degradation in damp areas may potentially appear.

Inference was performed using:

```
results = model('/mnt/data/gambar 3.jpg', conf=0.25)
results.show()
results.save()
```

The advantage of multi-scale detection in YOLOv7 and YOLOv8 enables small and large objects to be detected simultaneously [15]. With proper configuration, the model is able to maintain high precision and recall even in areas with complex shadows.

Based on the literature and implementation on the three buildings, it can be concluded that the YOLO algorithm is capable of detecting façade cracks in colonial buildings with high accuracy when supported by a representative dataset and optimal training configuration [16], [17], [18]. Parameters such as input resolution, number of epochs, and learning rate have a very significant influence on model performance. Transfer learning from pretrained COCO models accelerates convergence and improves stability [13].

Compared to manual inspection, the YOLO-based vision-based assessment approach is more objective and efficient. Manual inspection requires risky physical access, whereas drone- and AI-based systems enable real-time monitoring with measurable digital documentation. Golding et al. [20] emphasized that deep learning improves the consistency of damage assessment. Remondino and Campana [21] also stated that digital documentation is an important part of modern heritage conservation. Therefore, the integration of YOLO in the conservation of colonial buildings in Semarang has the potential to become a data-driven early detection system that supports sustainable preventive maintenance.

4. Conclusion

The conclusion of this study shows that the YOLO algorithm can be effectively used as an early detection system for façade cracks in Dutch colonial buildings in Semarang. Based on the analysis of Marba building, Jiwasraya building, and Lawang Sewu building, it is evident that façade visual characteristics such as exposed brick texture, homogeneous plaster surfaces, and the complexity of architectural ornaments influence the sensitivity level of the detection model. Nevertheless, with appropriate training configurations, including the selection of image resolution, number of epochs, and augmentation strategies such as lighting and rotation adjustments, the YOLO model is capable of consistently recognizing hairline cracks, medium cracks, and structural cracks.

Implementation on the three buildings demonstrates that the system is able to identify potential damage areas through bounding boxes and confidence scores representing prediction certainty levels. Compared to conventional visual inspection, this computer vision-based approach offers advantages in terms of objectivity, time efficiency, and safety, as it does not require direct physical access to high façades. Furthermore, detection results can be digitally documented, supporting periodic and measurable building condition monitoring.

Thus, this study answers the research questions by confirming that YOLO provides adequate accuracy for early detection of façade cracks in colonial buildings, that training parameter configurations significantly influence model performance, and that a vision-based assessment approach represents a more effective alternative than manual methods in the context of tropical heritage building conservation. The integration of artificial intelligence technology into historic building maintenance has the potential to strengthen data-driven and sustainable preventive conservation systems.

References

[1] G. K. Choudhary and S. Dey, "Crack detection in concrete surfaces using image processing, fuzzy logic, and neural networks," *2012 IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI)*, Nanjing, China, 2012, pp. 404-411, https://doi.org/10.1109/ICACI.2012.6463195

[2] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection, *Computer Vision and Pattern Recognition*, https://doi.org/10.48550/arXiv.2004.10934

[3] Y.-J. Cha, W. Choi, and O. Büyüköztürk, “Deep learning-based crack damage detection using convolutional neural

networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 361–378, May 2017. https://doi.org/10.1111/mice.12263

[4] Dorafshan, Sattar, Thomas, Robert J., and Maguire, Marc. "SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks." *Data in Brief*, vol. 21, 2018, pp. 1664-1668. https://doi.org/10.1016/j.dib.2018.11.015.

[5] Zhang, J., Qian, S. & Tan, C. Automated bridge crack detection method based on lightweight vision models. *Complex Intell. Syst.* 9, 1639–1652 (2023). https://doi.org/10.1007/s40747-022-00876-6

[6] 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

[7] J Zhou, J., Ning, J., Xiang, Z., & Yin, P. (2024). ICDW-YOLO: An Efficient Timber Construction Crack Detection Algorithm. *Sensors*, *24*(13), 4333. https://doi.org/10.3390/s24134333

[8] Y. Zhang, H. Huang, and Y. Cai, "On Bridge Surface Crack Detection Based on an Improved YOLOv3," *IFAC-PapersOnLine*, Volume 53, Issue 2, 2020, Pages 8205-8210, https://doi.org/10.1016/j.ifacol.2020.12.1994

[9] Golding, V. P., Gharineiat, Z., Munawar, H. S., & Ullah, F. (2022). Crack Detection in Concrete Structures Using Deep Learning. *Sustainability*, *14*(13), 8117. https://doi.org/10.3390/su14138117

[10] F. Remondino and S. Campana (eds.), 3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices*, Archaeopress BAR,* Oxfordshire, United Kingdom

[11] Li H, Wang C, Wang J, Zhang Y. Crack detection based on deep learning: a method for evaluating the object detection networks considering the random fractal of crack. *Structural Health Monitoring*. 2023;22(4):2547-2564. https://doi.org/10.1177/14759217221123485

[12] Dorafshan, Sattar; Thomas, Robert J.; and Maguire, Marc, "SDNET2018: An annotated image data set for non-contact concrete crack detection using deep convolutional neural networks" *Data in Brief*, 21 (2018) 1664–1668. https://doi.org/10.1016/j.dib.2018.11.015.

[13] A. S. Geetha, (2025), YOLOv4: A Breakthrough in Real-Time Object Detection, , arXiv:2502.04161, https://doi.org/10.48550/arXiv.2502.04161

[14] J. Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA 2016. https://doi.org/10.1109/CVPR.2016.91

[15] C.Y. Wang et al., "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv:2207.02696, 2022. https://doi.org/10.48550/arXiv.2207.02696

[16] 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.

[17] Zhou J, Ning J, Xiang Z, Yin P. ICDW-YOLO: An Efficient Timber Construction Crack Detection Algorithm. Sensors (Basel). 2024 Jul 3;24(13):4333. https://doi.org/10.3390/s24134333..

[18] Teng, S., Liu, Z., & Li, X. (2022). Improved YOLOv3-Based Bridge Surface Defect Detection by Combining High- and Low-Resolution Feature Images. *Buildings*, *12*(8), 1225.

https://doi.org/10.3390/buildings12081225.

[19] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," *ICLR 2019 Conference Blind Submission*.

[20] Golding, V. P., Gharineiat, Z., Munawar, H. S., & Ullah, F. (2022). Crack Detection in Concrete Structures Using Deep Learning. *Sustainability*, *14*(13), 8117. https://doi.org/10.3390/su14138117.

[21] F. Remondino and S. Campana (eds.), 3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices*, Archaeopress BAR,* Oxfordshire, United Kingdom.

[22] N. Isong, Building Efficient Lightweight CNN Models, *arXiv:2501.15547,* https://doi.org/10.48550/arXiv.2501.15547

[23] Kim, B., Jo, I., Ham, N., & Kim, J.-j. (2024). Simplified Scan-vs-BIM Frameworks for Automated Structural Inspection of Steel Structures. *Applied Sciences*, *14*(23), 11383. https://doi.org/10.3390/app142311383.

[24] I. Pitas, "Drone Vision and Deep Learning for Infrastructure Inspection," *2021 IEEE International Conference on Autonomous Systems (ICAS)*, Montreal, QC, Canada, 2021, pp. 1-1, https://doi.org/10.1109/ICAS49788.2021.9551136.

[25] Yao, G., Sun, Y., Wong, M., & Lv, X. (2021). A Real-Time Detection Method for Concrete Surface Cracks Based on Improved YOLOv4. *Symmetry*, *13*(9), 1716. https://doi.org/10.3390/sym13091716