Smart Structures and Systems, *Vol. 36, No. 4 (2025) 195-202*
https://doi.org/10.12989/sss.2025.36.4.195

Post-seismic structural damage segmentation using YOLOv8-seg model

Omid Yazdanpanah*[1], Ensieh Ali Bakhshi [2a] and Minwoo Chang [3b]

[1] *Hybrid Structural Testing Center (Hystec), Myongji University, Republic of Korea*
[2] *Industry & Academia Cooperation Foundation, Myongji University, Republic of Korea*
[3] *Department of Civil and Environmental Engineering, Myongji University, Republic of Korea*

(*Received June 26, 2025, Revised November 1, 2025, Accepted November 3, 2025*)

Abstract. A customized YOLOv8-seg architecture is hired in this study to automatically detect and segment post-earthquake damage, such as cracks, spalling, reinforcement exposure, crushing, buckling, and structural failure, that appears on bridge piers tested using slow and fast cyclic tests, shaking table tests, and real-time hybrid simulations. Using a hybrid loss function, the YOLOv8-seg model processes 32×32 and 256×256-pixel image patches, extracted from 124 large RGB images, for cracks and other seismic damage categories, respectively. Training is conducted on the image patches and their corresponding labeled annotations, distinguishing between seismic damage and background (non-damage) pixels. The model is trained with a batch size of 16, utilizing the Adamax optimizer, an exponential learning rate scheduler, and weight decay techniques to improve training stability and performance. The results demonstrate that the generated mask patches closely resemble the actual damage patterns.

Keywords: bridge piers; YOLOv8-seg architecture; multicategory seismic visible damage; real-time pixel-level detection

1. Introduction

Following an earthquake, the rapid and reliable assessment of structural damage is vital for public safety, rehabilitation planning, and resource allocation. Timely evaluations are critical to preventing further harm, facilitating efficient rescue operations, and prioritizing structural repairs. While initial inspections primarily aim to ensure immediate safety, they often provide only superficial evaluations due to time pressures, limited access to damaged sites, and logistical constraints. As a result, comprehensive damage assessments are essential to guide effective retrofitting and reconstruction of reinforced concrete (RC) structures such as buildings and bridges. Engineers require accurate, efficient, and standardized methods to assess structural integrity and prioritize interventions (Maeda *et al*. 2017, KHBD 2015, Yeow *et al*. 2022).

Traditional structural health monitoring (SHM) systems rely on contact-based sensors, including linear variable differential transformers (LVDTs), strain gauges, and accelerometers. These devices offer valuable high-frequency and real-time data but are typically labor-intensive to install, limited in spatial coverage, and require continuous maintenance and calibration. Additionally, they are often impractical in post-disaster environments where speed, accessibility, and scalability are crucial. In contrast, non-contact methods, particularly those based on computer vision (CV), have emerged as a promising alternative. These systems offer wider spatial coverage, simplified deployment, lower maintenance, and the ability to passively monitor structures without physical interaction (Hamidia *et al*. 2024, Sajedi and Liang 2021, Kim *et al*. 2022).

Post-earthquake structural evaluations have traditionally depended on visual inspections by certified engineers. While effective in identifying major damage, this approach is time-consuming, subjective, and constrained by the availability of qualified personnel. In large-scale seismic events, such limitations can delay assessments and hinder emergency response efforts. Recent advances in computer vision and deep learning have shown significant promise in automating structural damage detection from images and video, enabling rapid, objective, and scalable assessments (Narazaki *et al*. 2021, Miao *et al*. 2021). These methods can not only accelerate the evaluation process but also improve consistency and reproducibility. However, several challenges remain, including the need for large annotated datasets, the difficulty of detecting fine-grained or overlapping damage features, and the limited generalizability of current models to varied real-world conditions (Zhang *et al*. 2024, Xia *et al*. 2025, Zou *et al*. 2022, Dong *et al*. 2024).

To address these gaps, this study employs a customized YOLOv8-seg architecture to detect and segment various forms of post-earthquake damage, such as cracks, spalling, reinforcement exposure, crushing, buckling, and structural failure, on RC bridge piers subjected to controlled laboratory conditions including slow and fast cyclic loading, shaking table tests, and real-time hybrid simulations (see Fig. 1). Unlike conventional object detectors, YOLOv8-seg integrates object detection and

*Corresponding author, Ph.D., Research Professor,
E-mail: omidyazdanpanah@mju.ac.kr
[a] M.Sc. Researcher
[b] Ph.D., Associate Professor

ISSN: 1738-1584 (Print), 1738-1991 (Online)

(a) Fast and slow cyclic tests

(a) Shaking table test

(b) RTHS

Fig. 1 Bridge pier experiments under different loading protocols

semantic segmentation into a unified, real-time framework. This allows for both the localization and pixel-wise delineation of structural damage, enhancing interpretability and resolution. Its ability to detect subtle and overlapping damage types at multiple scales makes it particularly well-suited for post-earthquake scenarios. By leveraging this architecture, the present research aims to enhance the accuracy, efficiency, and real-world deployability of automated damage assessment tools, contributing to more resilient infrastructure management and faster post-disaster recovery for RC bridge piers.

2. Damage state definition

Seismic damage evaluation guidelines classify structural damage into five distinct levels based on visual indicators and their correlation with the lateral force-displacement response (see Fig. 2), which is used to assess structural performance (Maeda *et al*. 2017). Damage State I is characterized by fine cracks less than 0.2 mm in width, occurring without reinforcement yielding and maintaining essentially linear elastic behavior. Damage State II, involving cracks between 0.2 mm and 1.0 mm, marks the onset of yielding. Damage State III includes wider cracks (1.0–2.0 mm) and minor cover concrete loss, with structural behavior transitioning into the nonlinear hardening phase. More severe deterioration is observed in Damage State IV, defined by cracks exceeding 2.0 mm, significant loss of cover concrete, and exposed reinforcement. While lateral load capacity may be reduced, gravity load support is typically maintained. The most critical, Damage State V, involves severe structural degradation such as reinforcement buckling, concrete crushing, vertical deformation, and a substantial loss of lateral load-carrying capacity (Maeda *et al*. 2017). These defined damage states are essential for determining appropriate repair and retrofitting strategies. In this study, visible seismic damage is classified into five categories aligned with these damage states: Class 1 (Cracks), Class 2 (Spalling), Class 3 (Reinforcement Exposure), Class 4 (Concrete Crushing), and Class 5 (Buckling/Failure). Undamaged areas are assigned to Class 0, serving as the background. These categories serve as the basis for supervised learning in the segmentation model.

3. Image preprocessing and class imbalance handling

Preprocessing plays a foundational role in preparing image data for semantic segmentation using deep learning. It not only improves training efficiency and model stability but also ensures more reliable and generalizable predictions. In this study, several preprocessing techniques are adopted to optimize the performance of the segmentation network. Pixel values are first normalized to a [0,1] range to ensure consistency across inputs. Data augmentation methods, such

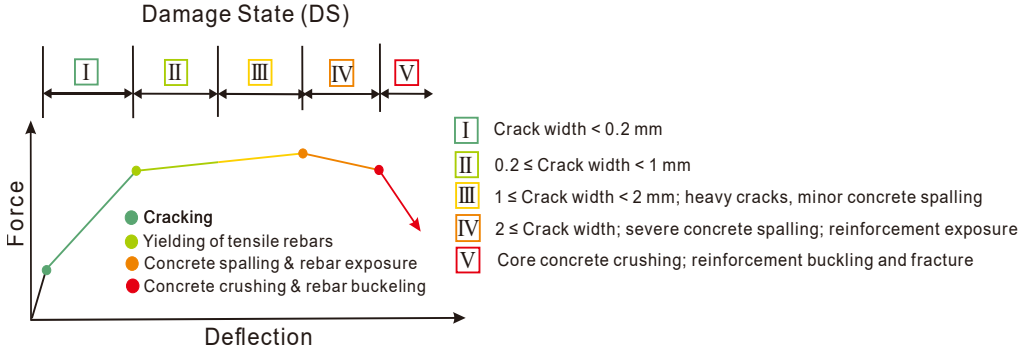


Fig. 2 Bridging visual seismic evidence and structural damage state in RC components

as rotation, flipping, and blurring, are employed to increase variability and prevent overfitting. Image geometry is corrected for distortion and perspective to enable precise semantic segmentation of damage types. Furthermore, to handle large image sizes and avoid out-of-memory issues, a patch-based approach is adopted. The YOLOv8-seg model is first developed for a multicategory seismic damage segmentation, which segments classes 0, 2, 3, 4, and 5 using 256×256-pixel patches, and then for crack segmentation, which detects classes 0 and 1 using finer 32×32-pixel patches. Both models are trained with a batch size of 16. For the multicategory model, a total of 1723 image patches are used, with 20% (344 patches) reserved as unseen test data. The remaining 1379 patches are split into 70% training (965 patches) and 30% validation (414 patches).

The hypergeometric distribution (Rice 2007), characterized by the total sample size (N), the number of samples in category k (N_k), and the batch size (N_b = 16), is used to calculate the probability of selecting each damage category within a batch (see Eq. (1)), thereby promoting balanced representation (see Table 1).

$$P(X=i)=\frac{\binom{N_k}{i}\binom{N-N_k}{N_b-i}}{\binom{N}{N_b}} \tag{1}$$

Due to the limited availability of image patches for Class 3 (exposure category) in both the training and validation sets, Table 1 illustrates the probability of having no more than two samples from this class within a batch. Remarkably, there is a 97.586% likelihood that each batch contains more than this threshold, underscoring the severity of the imbalance.

Class imbalance at the pixel level poses a significant challenge in segmentation tasks. When dominant classes overshadow underrepresented ones, models tend to prioritize majority features, often neglecting rare but critical damage types. This leads to biased learning, suboptimal loss convergence, and misleading evaluation outcomes. To counteract this, balancing strategies are essential, not only to improve the detection of minority classes but also to enhance overall model performance and stability.

In this work, we assign higher importance to underrepresented classes by applying weighted loss terms in the multicategory segmentation network. These weights, calculated inversely to the pixel distribution of each class as shown in Eq. (2), help ensure balanced learning across all categories. Table 2 highlights the disproportionate representation of Class 3, which receives a correspondingly higher weight during training.

To further mitigate class imbalance, we adopt a hybrid loss function (Eq. (3)) that integrates Dice (Zhou *et al*. 2018), Focal (Lin *et al*. 2017), Jaccard (Wang *et al*. 2023), and Tversky losses (Salehi *et al*. 2017), see Eqs. (4)-(7), respectively. This composite approach is designed to strengthen the network's ability to detect minority-class features, enhance generalization on unseen data, and maintain robust optimization throughout training. The same methodology is also applied to the crack segmentation network to ensure consistency across tasks.

$$w_i = {(1/p_i)}\Big/{\sum_{j=1}^{K}(1/p_j)} \tag{2}$$

$$L_{Total} = L_{Dice} + L_{Focal} + L_{Jaccard} + L_{Tversky} \tag{3}$$

where

$$L_{Dice} = 1 - \frac{2.\sum_i \ y_{true}^i \cdot y_{pred}^i + \varepsilon_1}{\sum_i \ y_{true}^i + \sum_i \ y_{pred}^i + \varepsilon_1}, \qquad \varepsilon_1 = 1 \tag{4}$$

$$\begin{aligned} L_{Focal} &= -\alpha_t.(1-p_t)^{\gamma}.\log(p_t), && \gamma = 3 \\ \alpha_t &= y_{true}.\alpha + (1-y_{true}).(1-\alpha), && \alpha = 0.25 \\ p_t &= y_{true}.y_{pred}.w + (1-y_{true}).(1-y_{pred}.w) \end{aligned} \tag{5}$$

Table 1 Number of samples with each damage category and the corresponding probability for multicategory segmentation network

	Class 0	Class 2	Class 3	Class 4	Class 5	Total
Training data	560	820	370	605	582	965
Validation data	225	366	173	285	258	414
Probability for training data	2.102×10^{-4}	1.567×10^{-10}	2.414×10^{-2}	4.508×10^{-5}	1.017×10^{-4}	-

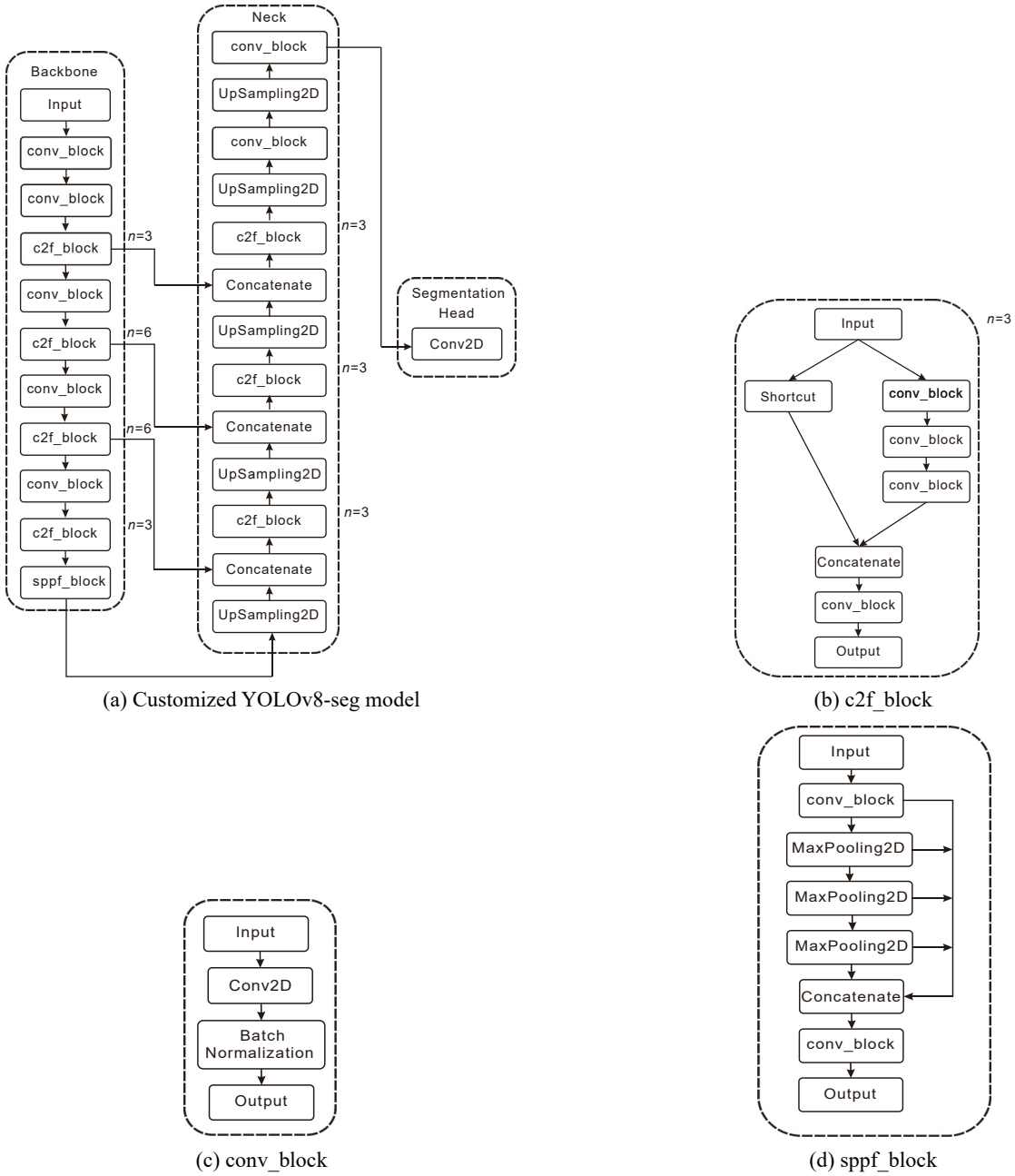
Table 2 Proportion of pixels in each category and the corresponding normalized class weights

	Class 0	Class 2	Class 3	Class 4	Class 5
Training data	0.2867	0.4188	0.0308	0.1738	0.0898
Normalized class weights (w)	0.0631	0.0433	0.5881	0.1042	0.2013

$$L_{Jaccard} = \left(1 - \frac{\sum_i \left|y^i_{true}.y^i_{pred}.w\right| + \varepsilon_2}{\sum_i \left|y^i_{true}\right| + \sum_i \left|y^i_{pred}.w\right| - \sum_i \left|y^i_{true}.y^i_{pred}.w\right| + \varepsilon_2}\right) \cdot \varepsilon_2, \qquad \varepsilon_2 = 100 \tag{6}$$

$$L_{Tversky} = 1 - \frac{\sum_i y^i_{true}.y^i_{pred}.w}{\sum_i y^i_{true}.y^i_{pred}.w + \alpha_1.\sum_i \left(1 - y^i_{true}\right).y^i_{pred}.w + \beta.\sum_i y^i_{true}\left(1 - y^i_{pred}.w\right)}, \quad \alpha_1 = 0.3,\ \beta = 0.7 \tag{7}$$

The proposed hybrid loss function integrates four complementary components to address class imbalance and improve segmentation robustness. Dice loss leverages the overlap between predicted (y_{pred}) and ground truth (y_{true}) masks, with a smoothing factor $\varepsilon_1 = 1$ to ensure numerical stability, enhancing sensitivity to small structures. Focal loss introduces a focusing parameter $\gamma = 3$ and a weighting factor $\alpha = 0.25$ to down-weight easy examples and emphasize hard-to-classify pixels, effectively mitigating class imbalance. Jaccard loss (IoU) incorporates a weighting term w and a smoothing term $\varepsilon_2 = 100$ to stabilize training while improving boundary accuracy by penalizing mismatched regions. Tversky loss generalizes Dice loss through asymmetric penalties controlled by $\alpha_1 = 0.3$ and $\beta = 0.7$, enabling finer control over false positives and false negatives. Collectively, these parameters ensure that the



(a) Customized YOLOv8-seg model

(b) c2f_block

(c) conv_block

(d) sppf_block

Fig. 3 Proposed YOLOv8-seg model architecture

(a) Loss (b) mIoU

Fig. 4 Hybrid loss and mean IoU (mIoU) variations observed during assigned epochs

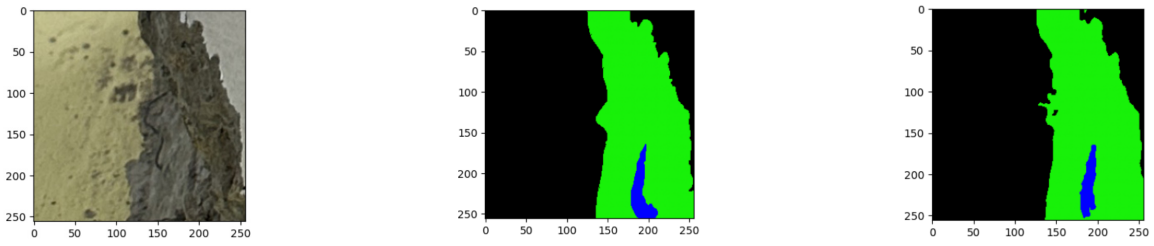
hybrid loss captures diverse error characteristics, balancing overlap accuracy, class weighting, and hard-sample sensitivity, to achieve stable optimization, improved feature discrimination, and enhanced model generalization.

4. Proposed YOLOv8-seg model for multicategory pixel-level segmentation

YOLOv8-seg (Zhang *et al*. 2024, Xia *et al*. 2025, Zou *et al*. 2022, Dong *et al*. 2024) is an advanced variant of the YOLO (You Only Look Once) object detection model (Redmon *et al*. 2016), optimized for semantic segmentation, assigning a class label to each pixel in an image. Unlike standard object detection, which provides bounding boxes, segmentation requires spatially detailed outputs, making architectural tweaks necessary.

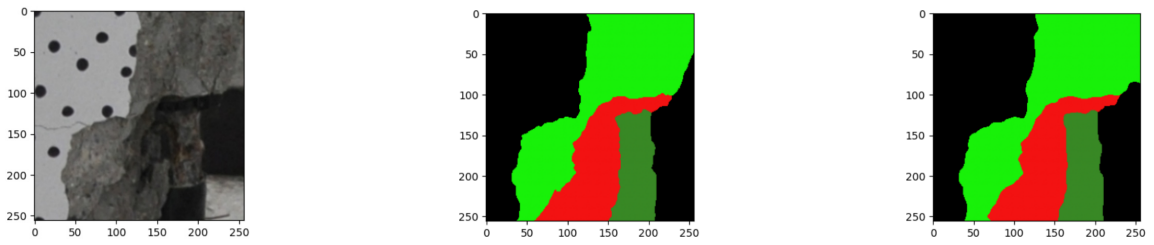
The proposed architecture in Fig. 3 is a customized YOLOv8-seg model designed for efficient and accurate semantic segmentation of seismic damage on bridge pier surfaces. It is composed of a backbone, neck, and segmentation head. The backbone extracts hierarchical features using alternating conv_block and c2f_block modules and ends with a sppf_block to capture multiscale spatial context. The neck aggregates these features through multiple UpSampling2D, c2f_block, and Concatenate operations, forming a feature pyramid that enhances spatial resolution and contextual understanding. Finally, the segmentation head, consisting of a single Conv2D layer, produces dense pixel-wise class predictions. Supporting modules such as the conv_block, c2f_block, and sppf_block are designed for lightweight computation and rich feature extraction, enabling the model to balance accuracy and speed effectively for real-time segmentation tasks.

The proposed YOLOv8-seg architecture integrates carefully tuned hyperparameters across its modular design to optimize semantic segmentation performance. The conv_block modules use increasing filter sizes (64–512) and a mix of 3×3 and 1×1 kernels, combined with Batch Normalization and ReLU activation function for stable feature extraction and downsampling. The c2f_block leverages partial fusion with n = 3 or 6 repetitions,



(a) RGB image patch (b) RGB label patch (c) Predicted mask patch

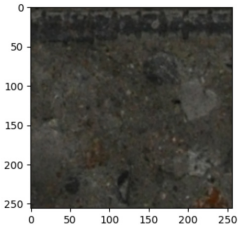
per-class IoU					mIoU
Class 0	Class 2	Class 3	Class 4	Class 5	0.8719
0.9846	0.9541	0.6771	-	-	

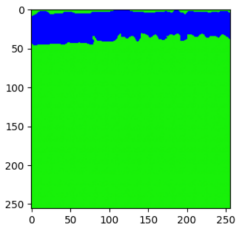


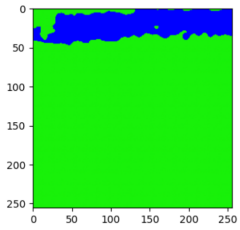
(d) RGB image patch (e) RGB label patch (f) Predicted mask patch

per-class IoU					mIoU
Class 0	Class 2	Class 3	Class 4	Class 5	0.9727
0.9816	0.9700	-	0.9654	0.9739	

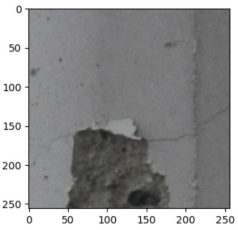
Fig. 5 Patched image, corresponding label, and predicted mask including multicategory damage

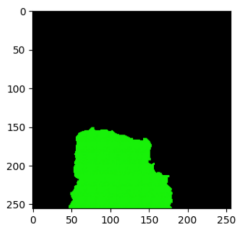

(g) RGB image patch

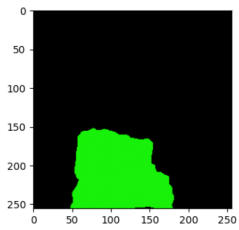

(h) RGB label patch


(i) Predicted mask patch

per-class IoU					mIoU
Class 0	Class 2	Class 3	Class 4	Class 5	0.8897
-	0.9690	0.8104	-	-	


(j) RGB image patch


(k) RGB label patch


(l) Predicted mask patch

per-class IoU					mIoU
Class 0	Class 2	Class 3	Class 4	Class 5	0.9891
0.9964	0.9818	-	-	-	

Fig. 5 Continued

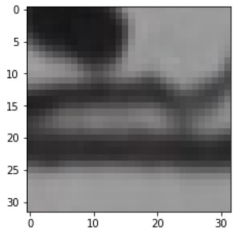
depending on depth, enhancing gradient flow and efficiency. The sppf_block uses three successive 5×5 MaxPooling layers with a stride of 1 to capture multiscale spatial context before fusion through concatenation and a 1×1 convolution. In the neck, UpSampling2D operations restore spatial resolution while skip connections and c2f_blocks refine feature maps. Finally, a 1×1 Conv2D layer with softmax activation serves as the segmentation head, producing per-pixel class predictions for accurate and real-time semantic segmentation. Fig. 4 tracks the changes in L_{Total} and mean Intersection over Union (mIoU) over 100 epochs, demonstrating that the validation loss and mIoU closely follow the trends of the training loss and mIoU. The mIoU scores for the training, validation, and unseen datasets are 0.9546, 0.9042, and 0.9008, respectively.

Fig. 5 presents the semantic segmentation results for seismic damage on unseen bridge piers, evaluated using per-class IoU. The analysis is based on randomly extracted RGB image patches, their corresponding ground truth labels, and predicted mask outputs. As observed in Fig. 5(a)-(c), the model achieves IoU scores of 0.9846 for Class 0, 0.9541 for Class 2, and 0.6771 for Class 3, while Classes 4 and 5 are absent in this sample. The overall mIoU for this instance is 0.8719. For Figs. 5(d)-(f), the model results in a mIoU of 0.9727, with high per-class IoU scores: 0.9816 for Class 0, 0.9700 for Class 2, 0.9654 for Class 4, and 0.9739 for Class 5. Figs. 5(i) and (l) offer additional evidence of the segmentation model's performance.

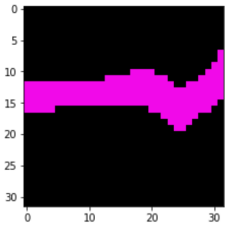
Crack segmentation results are presented in Fig. 6. As illustrated, the proposed model can identify crack pixels with a high level of accuracy, achieving mIoU scores of0.9021 and 0.9247 for the image patches in Figs. 6(a) and d(d), respectively. Moreover, the model effectively distinguishes crack regions from non-crack areas such as gridlines and background occlusions, demonstrating strong discriminative capability. Figs. 6(i) and 6(l) further substantiate the effectiveness of the segmentation model. For the training, validation, and unseen datasets, the mIoU scores are 0.8036, 0.8019, and 0.8003, respectively.

5. Conclusions

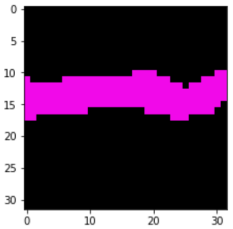
This study introduces a computer vision-based approach utilizing a customized YOLOv8-seg model for pixel-level, multicategory segmentation of seismic damage in RC bridge piers. A comprehensive and diverse image database is assembled, incorporating data from cyclic loading tests, shaking table experiments, and RTHS, offering a more realistic representation of seismic damage than traditional cyclic tests alone. The annotation process categorizes damage into five classes: crack, spalling, exposure, crushing, buckling, and failure. To address the significant data imbalance, the training strategy incorporates data augmentation, class weighting, and a hybrid loss function, applied at both the sample and pixel levels. The multicategory segmentation model (excluding cracks) is trained on 256×256 image patches, while the dedicated model for crack segmentation uses 32×32 image patches. This patch-based approach enables efficient processing of high-resolution images while avoiding GPU memory limitations. Unlike prior studies that rely on simplified cyclic test protocols, which may misrepresent true seismic demands, this research emphasizes the use of RTHS-generated imagery for more accurate damage representation. Furthermore, critical structural response indicators,



(a) RGB image patch

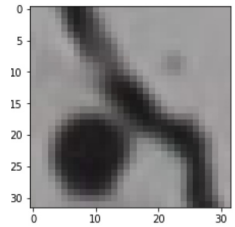


(b) RGB label patch

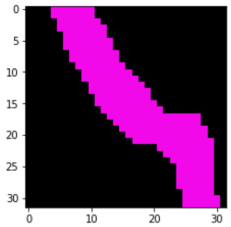


(c) Predicted mask patch

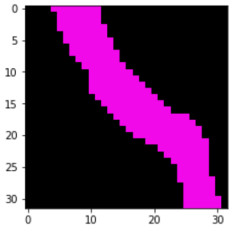
per-class IoU		mIoU
Class 0: 0.9640	Class 1: 0.8402	0.9021



(d) RGB image patch

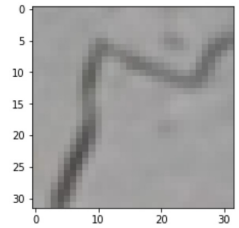


(e) RGB label patch

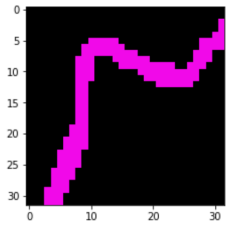


(f) Predicted mask patch

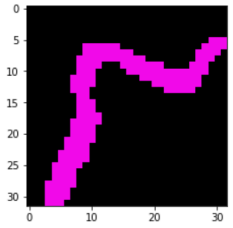
per-class IoU		mIoU
Class 0: 0.9600	Class 1: 0.8893	0.9247



(g) RGB image patch

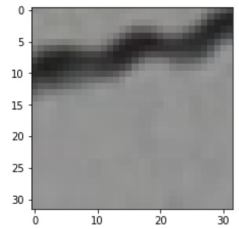


(h) RGB label patch

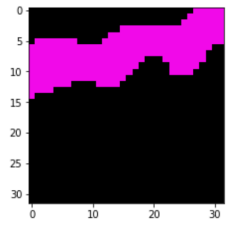


(i) Predicted mask patch

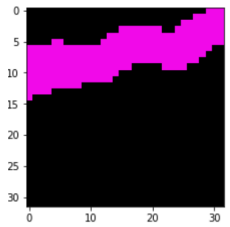
per-class IoU		mIoU
Class 0: 0.9423	Class 1: 0.7330	0.8377



(j) RGB image patch



(k) RGB label patch



(l) Predicted mask patch

per-class IoU		mIoU
Class 0: 0.9604	Class 1: 0.8710	0.9157

Fig. 6 Image patches and the corresponding labels and predicted mask patches including crack category damage

such as residual crack width and lateral force-drift hysteresis, benefit directly from the improved segmentation accuracy. By integrating deep learning with realistic seismic testing data, this work enhances the precision of damage detection and contributes to more reliable post-earthquake assessment and resilience planning for RC bridge infrastructure.

Acknowledgments

The authors gratefully acknowledge financial support from the Creative Challenge Research project (RS-2023-00241807) and the Brain Pool program (RS-2023-00264271), both funded by the Ministry of Science and ICT through the National Research Foundation of Korea.

References

Dong, X., Liu, Y. and Dai, J. (2024), "Concrete Surface Crack Detection Algorithm Based on Improved YOLOv8", *Sensors*, **24**(16), p. 5252. https://doi.org/10.3390/s24165252

Hamidia, M., Sheikhi, M., Asjodi, A.H. and Dolatshahi, K.M. (2024), "Computer vision-based quantification of updated stiffness for damaged RC columns after earthquake", *Adv. Eng. Software*, **190**, p. 103597. https://doi.org/10.1016/j.advengsoft.2024.103597

KHBD (2015), Ministry of Land, Infrastructure and Transport,

Korea Highway Bridge Design Code (Limit State Design Method).

Kim, B., Lee, J., Sim, S.H., Cho, S. and Park, B.H. (2022), "Computer vision-based remote displacement monitoring system for in-situ bridge bearings robust to large displacement induced by temperature change", *Smart Struct. Syst., Int. J.*, **30**(5), 521-535. https://doi.org/10.12989/sss.2022.30.5.521

Maeda, M., Nishida, T., Matsukawa, K. and Murakami, M. (2017), "Revision of guideline for postearthquake damage evaluation of reinforced concrete buildings in Japan", In: *The 16th World Conference on Earthquake Engineering*, Santiago, Chile.

Miao, Z., Ji, X., Okazaki, T. and Takahashi, N. (2021), "Pixel-level multicategory detection of visible seismic damage of reinforced concrete components", *Comput.-Aided Civil Infrastr. Eng.*, **36**(5), 620-637. https://doi.org/10.1111/mice.12667

Narazaki, Y., Hoskere, V., Yoshida, K., Spencer, B.F. and Fujino, Y. (2021), "Synthetic environments for vision-based structural condition assessment of Japanese high-speed railway viaducts", *Mech. Syst. Signal Process.*, **160**, p. 107850. https://doi.org/10.1016/j.ymssp.2021.107850

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016), "You only look once: Unified, real-time object detection", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788.

Rice, J.A. (2007), *Mathematical Statistics and Data Analysis*, (Third ed.), Duxbury Press, p. 42.

Lin, T.Y., Goyal, P., Girshick, R., He, K. and Dollár, P. (2017), "Focal loss for dense object detection", In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2980-2988.

Sajedi, S.O. and Liang, X. (2021), "Uncertainty-assisted deep vision structural health monitoring", *Comput.-Aided Civil Infrastr. Eng.*, **36**(2), 126-142. https://doi.org/10.1111/mice.12580

Salehi, S.S.M., Erdogmus, D. and Gholipour, A. (2017), "Tversky loss function for image segmentation using 3D fully convolutional deep networks", In: *International Workshop on Machine Learning in Medical Imaging*, pp. 379-387.

Wang, Z., Ning, X. and Blaschko, M. (2023), "Jaccard metric losses: optimizing the Jaccard index with soft labels", *Adv. Neural Inform. Process. Syst.*, **36**, 75259-75285.

Xia, H., Li, Q., Qin, X., Zhuang, W., Ming, H., Yang, X. and Liu, Y. (2025), "Bridge crack detection algorithm designed based on YOLOv8", *Appl. Soft Comput.*, **171**, p. 112831. https://doi.org/10.1016/j.asoc.2025.112831

Yeow, T.Z., Kusunoki, K., Nakamura, I., Hibino, Y., Fukai, S. and Safi, W.A. (2022), "E-defense shake-table test of a building designed for post-disaster functionality", *J. Earthq. Eng.*, **26**(10), 5153-5174. https://doi.org/10.1080/13632469.2020.1865219

Zhang, C., Chen, X., Liu, P., He, B., Li, W. and Song, T. (2024), "Automated detection and segmentation of tunnel defects and objects using YOLOv8-CM", *Tunnell. Undergr. Space Technol.*, **150**, p. 105857. https://doi.org/10.1016/j.tust.2024.105857

Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N. and Liang, J. (2018), "Unet++: A nested u-net architecture for medical image segmentation", In: *International Workshop on Deep Learning in Medical Image Analysis*, pp. 3-11.

Zou, D., Zhang, M., Bai, Z., Liu, T., Zhou, A., Wang, X., Cui, W. and Zhang, S. (2022), "Multicategory damage detection and safety assessment of post-earthquake reinforced concrete structures using deep learning", *Comput.-Aided Civil Infrastr. Eng.*, **37**(9), 1188-1204. https://doi.org/10.1111/mice.12815