Coarse-to-Fine Seismic Assessment Based on Computer Vision

YANG XU, QIANGQIANG ZHANG and HUI LI

ABSTRACT

Recently, remote sensing satellites, unmanned aerial vehicles (UAVs), and smartphones have been extensively utilized in non-contact post-earthquake inspection at different scales with cutting-edge computer vision and machine learning techniques. This study establishes a computer-vision-based coarse-to-fine seismic assessment framework to localize dense buildings in urban areas, classify collapsed and non-collapsed states, recognize multi-type surface damage on structural components, and evaluate seismic performances. First, a Transformer-CNN deep learning architecture is designed for semantic segmentation of dense buildings and binary classification of collapsed states using large-scale remote sensing satellite images. It consists of a Swin Transformer encoder, multi-stage feature fusion module, and UPerNet decoder to extract global correlations and local features of dense buildings synchronously. Then, a multi-task learning strategy is proposed to simultaneously recognize multi-type structural components (column, beam, wall), seismic damage (concrete crack, spalling, and rebar exposure), and multi-level damage states (minor, moderate, major) using medium-scale UAV images. It contains a CNN-based encoder-decoder backbone with skip-connection modules and multi-head segmentation subnetworks for different tasks. The geometric consistency loss of split line, area, and curvature is further designed to refine the semantic segmentation of local details, increase boundary smoothness, and suppress inner voids. Finally, a machine learning neural network is established to quantify the seismic damage index of structural components using damage-related parameters (lengths, areas, and numbers of concrete crack, spalling, and rebar exposure) and design-related parameters (axial compression ratio, shear span ratio, and volumetric stirrup ratio) as inputs. A seismic damage indicator with an explicit bound of [0,1] can be obtained to reflect the nonlinear accumulation of seismic damage. The effectiveness and applicability under real-world post-earthquake scenarios have been validated by the 2017 Mexico City Earthquake M7.1, 2008 Beichuan Earthquake M8.0, 2010 Yushu Earthquake M6.9, 2015 Nepal Earthquake M7.8, 2016 Ecuador Earthquake M7.8, and 2016 Meinong Earthquake M6.7.

Yang XU[a*], Qiangqiang ZHANG[b], Hui LI[a]
[a]School of Civil Engineering, Harbin Institute of Technology, Harbin, 150090, China
[b]School of Civil Engineering and Mechanics, Lanzhou University, Lanzhou 730000, China (*Corresponding Author Email: xyce@hit.edu.cn)

1 INTRODUCTION

Earthquakes may significantly impact the safety of city buildings, and most of the casualties and economic losses are closely related to seismic disasters. Therefore, precise damage localization and rapid condition assessment of post-earthquake buildings in urban areas are critical for emergency responses and rescue decisions. Recently, the rapid development of machine learning and computer vision techniques profoundly promoted the evolution of earthquake engineering, integrating with remote sensing, UAVs, and robot techniques [1-11]. Images acquired by different platforms have unique advantages and characteristics. The satellite remote sensing images can quickly obtain the large-scale general location of building groups, but the only visible information of building roofs affects the evaluation accuracy; UAV and smartphone images can convey more precise details on building facades and local components, but the inspection range is limited. This study establishes a multi-scale damage recognition, localization, and assessment framework for post-earthquake buildings using large-scale satellite images, median-scale UAV images, and small-scale smartphone images.

2 METHODOLOGY

2.1 Overall Framework

Figure 1 shows the overall schematic of coarse-to-fine seismic assessment based on computer vision, including large-scale dense building localization and semantic segmentation using remote sensing images, medium-scale recognition of multi-type structural components, seismic damage, and deterioration states using UAV images, and fine-scale evaluation of seismic damage index based on experimental data and images.

2.2 Large-scale Coarse Assessment by Remote Sensing Satellite Images

To accomplish the fast localization and preliminary evaluation of post-earthquake buildings in city areas, a two-stage building location and damage assessment method is developed considering the spatial distribution, small size, and imbalanced numbers of dense buildings [12]. As shown in Figure 2, it includes a modified You Only Look Once (YOLOv4) object detection module and a support vector machine (SVM) based classification module. The multi-scale features of densely distributed buildings are extracted and fused, and three statistics (e.g., the angular second moment, dissimilarity, and inverse difference moment) are further discovered based on the gray-level co-occurrence matrix as the texture features to distinguish damage intensities of buildings.

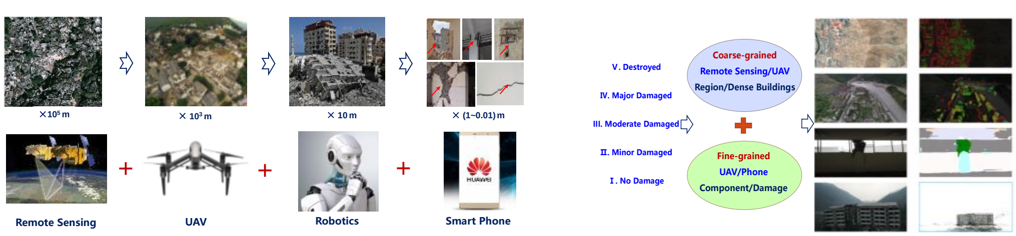


Figure 1. Schematic of computer-vision-based coarse-to-fine seismic assessment

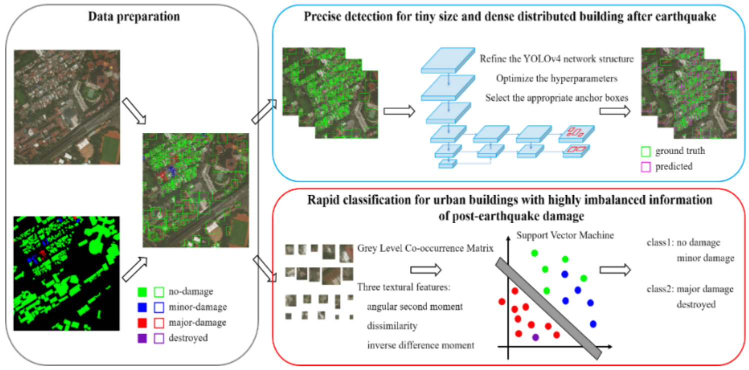


Figure 2. Two-stage building localization and collapse classification based on YOLOv4 and SVM [12]

Furthermore, a novel improved Swin Transformer is proposed to segment dense urban buildings at pixel level by remote sensing images with complex backgrounds, as shown in Figure 3 [13]. The original Swin Transformer is utilized as the backbone of the encoder, and a convolutional block attention module is inserted at patch embedding and patch merging stages. Hierarchical feature maps are fused together and then fed into UPerNet decoder. It enables the model to learn the separability of building damage states and location semantics and improve multi-class segmentation accuracy.

2.3 Medium-scale Moderate Assessment by Multi-task Learning of UAV Images

Because potential features and inherent dependencies of seismic damages on structures are supposed to exist among different post-earthquake inspection tasks, a multi-task learning approach is proposed to simultaneously accomplish the semantic segmentation of seven-type structural components, three-type seismic damages, and four-type deterioration states [14]. The proposed method contains a backbone network and a multi-head task-specific recognition network, as shown in Figure 4. The backbone network follows a CNN-based encoder-decoder structure with skip-connection modules and is designed to extract multi-level features based on the physical properties and structural mechanics of post-earthquake RC structures. The multi-head task-specific recognition network consists of three individual self-attention pipelines, each of which utilizes the extracted multi-level features from the backbone network as the mutual guidance for the corresponding individual segmentation task of structural component, seismic damage, and deterioration state, respectively.

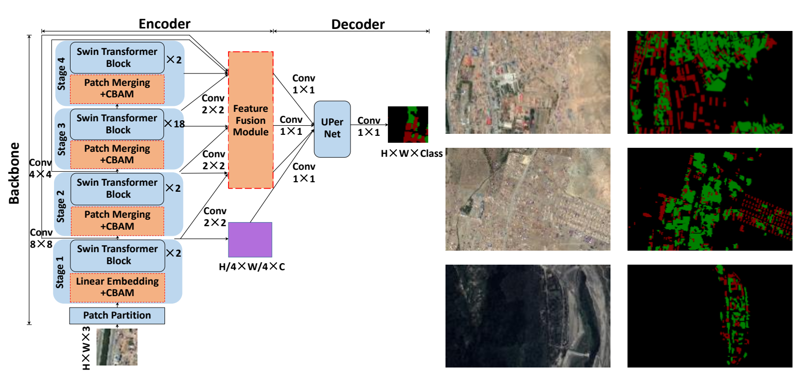


Figure 3. Improved Swin Transformer for dense building segmentation and state classification [13]

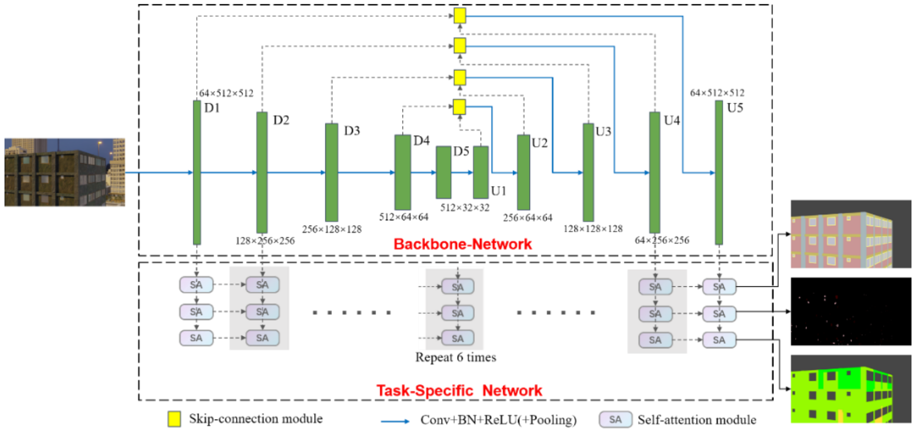


Figure 4. Multi-task learning semantic segmentation of structural component, damage, and state [14]

To train the multi-task learning model, a synthetical loss function is accordingly designed with real-time adaptive coefficients to balance the multi-task losses and focus on the most unstable fluctuating tasks:

$$L=\sum_{i=1}^{K}\lambda_i L_i\,, L_i=-\frac{1}{N}\sum_{j=1}^{N} y_j log\left(p_j\right), \lambda_i^{t+1}=\frac{K exp\left(\frac{\omega_i^t}{T}\right)}{\sum_{k=1}^{K} exp\left(\frac{\omega_k^t}{T}\right)}, \omega_i^t=\frac{L_i^t}{L_i^{t-1}}(t\geq 2) \quad (1)$$

where L denotes the overall loss, λ_i denotes the coefficient of ith task-specific loss, K denotes the number of tasks, L_i denotes the cross-entropy loss for training the ith task, N denotes the pixel number of the input image, y_j denotes the label of the jth pixel, and p_j denotes the positive classification probability by a softmax function. t denotes the current iteration, ω_i denotes the change rate of current and previous losses for the ith task loss, and T is the scaling hyperparameter avoiding the exponential explosion. The real-time proportion of each task loss is controlled by λ_i^{t+1}.

2.4 Boundary Refinement by Geometric Consistency

A novel loss function named geometric consistency enhanced (GCE) loss is designed considering the geometrical constraints of split line length, curvature, and area to focus on local boundaries and improve the segmentation details [15-17]:

$$\begin{aligned}
&L_{GCE}=\alpha_1 L_{length}+\alpha_2 L_{curvature}+\alpha_3 L_{area}+L_{CE}, \alpha_1, \alpha_2, \alpha_3 \geq 0\\
&L_{length}=\left|\left[\sum_{h,w} L_x^v(h,w)+\sum_{h,w} L_y^v(h,w)\right]-\left[\sum_{h,w} L_x^u(h,w)+\sum_{h,w} L_y^u(h,w)\right]\right|\\
&L_{curvature}=\left|\sum_{h,w} C^v(h,w)-\sum_{h=1,w} C^u(h,w)\right|, L_{area}=\left|\sum_{h,w} A^v(h,w)-\sum_{h,w} A^u(h,w)\right|\\
&L_{CE}=-\sum_{h,w} A^u(h,w)\cdot\log\left[\sum_{h,w} A^v(h,w)\right]
\end{aligned} \quad (2)$$

where α_1, α_2, and α_3 denote three individual weight coefficients; they are treated as hyperparameters, and appropriate values should be selected for a specific task. v and u denote the predicted and ground-truth masks, respectively. $\sum_{h,w} L_x^v(h,w)$ and $\sum_{h,w} L_x^u(h,w)$,

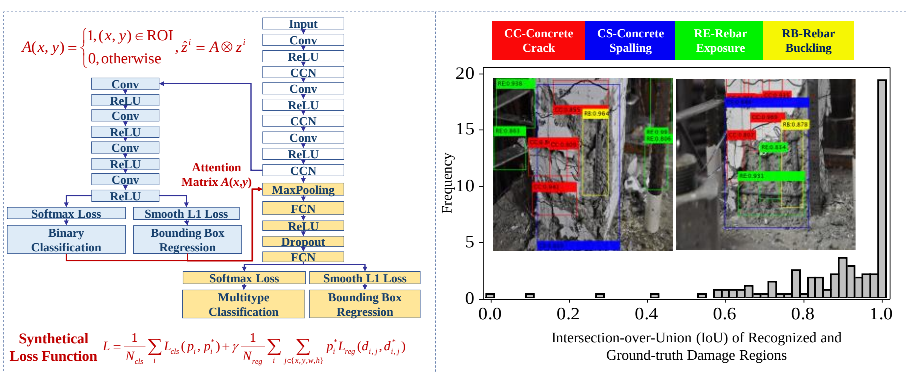
$\sum_{h,w} L_y^v(h,w)$ and $\sum_{h,w} L_y^u(h,w)$ denote the split line lengths in the height and width directions of x and y axes, respectively. $\sum_{h,w} C^v(h,w)$ and $\sum_{h,w} C^u(h,w)$ denote the split line curvatures. $\sum_{h,w} A^v(h,w)$ and $\sum_{h,w} A^u(h,w)$ denote the pixel values, respectively.

2.5 Small-scale Fine Assessment by Hysteresis Machine Learning

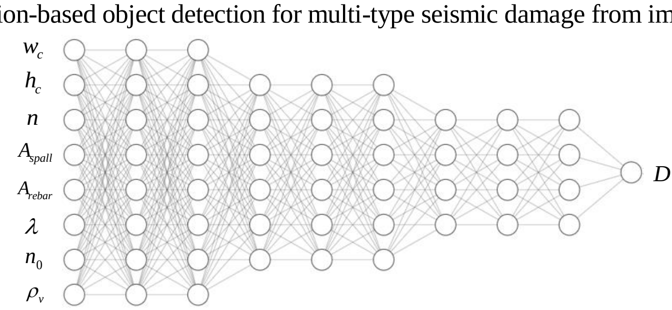
Using machine learning and computer vision techniques can establish classifiers to model the relationship between seismic images and damage levels. However, they are black-box models lacking interpretability. An earthquake engineering knowledge-enhanced machine learning method is established for seismic damage assessment of structural components, as shown in Figure 5 [18,19]. First, a synthetical indicator of seismic damage with an explicit bound of [0,1] is designed based on refined Park-Ang model. Then, a deep neural network is established for the damage index regression. Both damage-related parameters (lengths, areas, and numbers of various damage regions of concrete crack, concrete spalling, and rebar exposure) and design-related parameters (axial compression ratio, shear span ratio, and volumetric stirrup ratio) are used as inputs.

3 RESULTS AND DISCUSSION

The investigated xBD dataset includes 386 optical images of the 2017 Mexico City earthquake with a resolution of 1024 × 1024 pixels. Statistical analysis shows that the average localization accuracy of buildings exceeds 95.7% and that the binary classification accuracy for damage assessment reaches 97.1% [12]. Some representative results of dense building localization and collapse classification are shown in Figure 6.

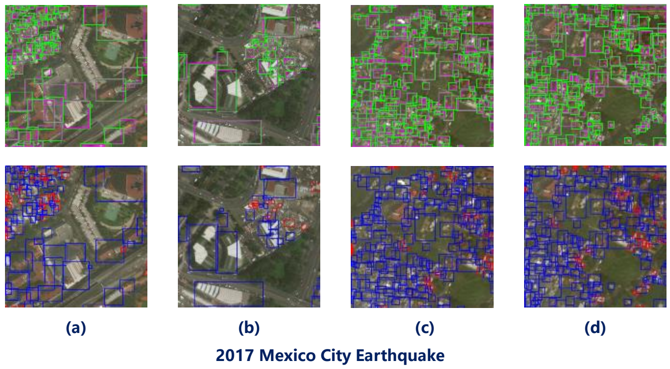


(a) Region-based object detection for multi-type seismic damage from images [18]



(b) Deep neural network for seismic damage index regression [19]

Figure 5. Schematic of damage index quantification using quasi-static experimental data and images



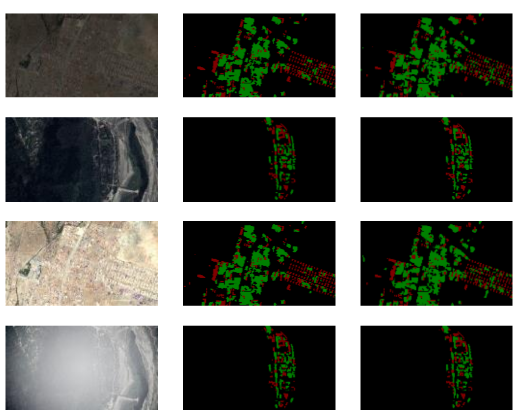
Test Image	No Damage/Minor Damage	Major Damage/Destroyed
(a)	87/115 (75.7%)	7/8 (87.5%)
(b)	26/39 (66.7%)	2/2 (100%)
(c)	194/227 (85.5%)	1/1 (100%)
(d)	188/229 (82.1%)	1/1 (100%)

Figure 6. Representative results of two-stage dense building localization and collapse classification [12]

To further demonstrate the semantic segmentation accuracy on post-earthquake dense buildings, remote sensing seismic images in Yushu City and Beichuan city with various weather disturbances are tested [13]. Figure 7 shows representative test results of large-scale images with a resolution of 4608 × 2560 under different weather disturbances. The results show that the improved Swin Transformer gains higher accuracy and better robustness against light overexposure, darkness, and fog occlusions than the original Swin Transformer.

For the geometric consistency guided medium-scale assessment, 480 original UAV images from Beichuan city are investigated, as shown in Figure 8. The results show that the proposed GCE loss can effectively suppress the false-positive small-region noise by adding constraints of overall geometrical shapes [17].

For the quantification of the seismic damage index, images and experimental data of 124 RC columns during the entire quasi-static experiment process are utilized [19]. The results show that the established regression model of the seismic damage index is unbiased and stable without overfitting. Additional comparative studies are further performed to verify the effectiveness, necessity, robustness, and generalization ability under other real-world post-earthquake scenarios, as shown in Figure 9.



Method	Back-ground IoU	Non-collapsed Building IoU	Collapsed Buidling IoU	mIoU
Original Swin Transformer	94.97%	84.50%	82.21%	87.23%
Swin Transformer+CBAM	95.10%	85.80%	83.56%	87.95%
Swin Transformer + Feature Fusion	95.29%	85.64%	84.15%	88.36%
Improved Swin Transformer (Full Model)	**95.35%**	**86.07%**	**84.16%**	**88.53%**

Figure 7. Representative test results of large-scale images under different weather disturbances using the improved Swin Transformer [13]

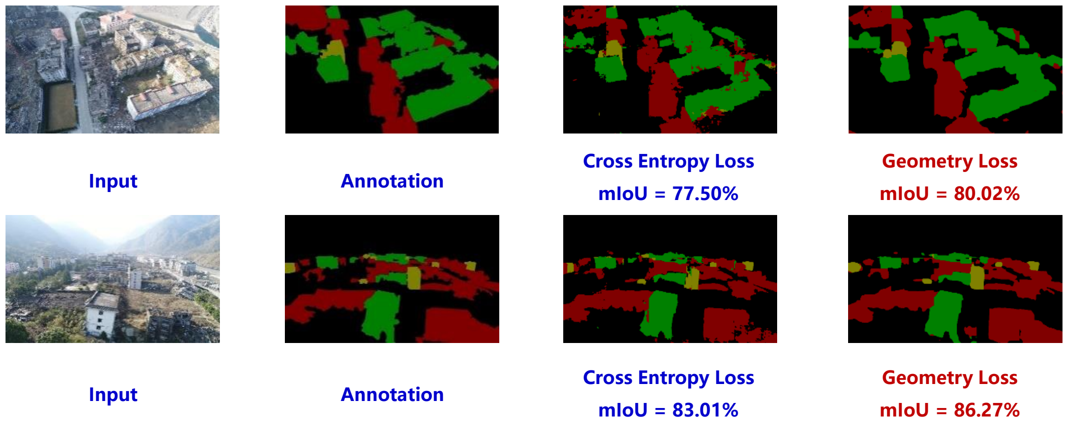
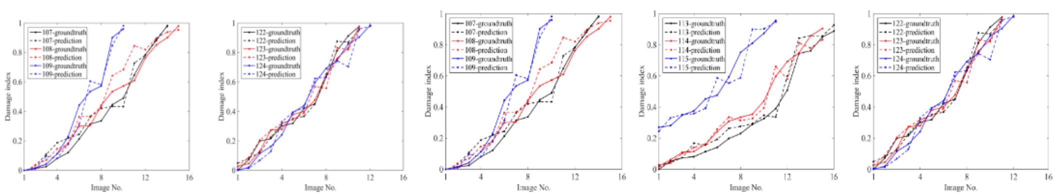


Figure 8. Representative results of geometric consistency enhanced building segmentation [17]



Stage	Precision	Recall	F1-score
1	73.68%	82.35%	77.78%
2	93.22%	82.09%	87.30%
3	87.50%	98.24%	92.56%
4	96.72%	93.65%	95.16%
5	93.33%	95.45%	94.38%

Figure 9. Representative results of machine learning based seismic damage index assessment [19]

4 CONCLUSIONS

This study establishes a computer-vision-based coarse-to-fine seismic assessment framework to localize dense buildings in urban areas, classify collapsed and non-collapsed states, recognize multi-type surface damage on structural components, and evaluate seismic performances. A series of Transformer, CNN, and NN-based deep learning models are designed for the localization, classification, and quantification of dense buildings, deterioration states, and damage index using large-scale remote sensing satellite images, medium-scale UAV images, near-field surface images, and quasi-static experimental data.

ACKNOWLEDGMENT

Financial support for this study was provided by the National Natural Science Foundation of China [Grant Nos. 51921006, 52192661, and 52008138], National Key R&D Program of China [Grant Nos. 2021YFF0501003 and 2019YFC1511102], China Postdoctoral Science Foundation [Grant Nos. BX20190102 and 2019M661286], Heilongjiang Province Postdoctoral Science Foundation [Grant Nos. LBH-TZ2016 and LBH-Z19064], and Heilongjiang Province Natural Science Foundation [Grant No. LH2022E070].

REFERENCES

1. Bao, Y., & Li, H. (2021). Machine learning paradigm for structural health monitoring. Structural Health Monitoring, 20(4), 1353-1372.
2. Xu, Y., Li, S., Zhang, D., Jin, Y., Zhang, F., Li, N., & Li, H. (2018). Identification framework for cracks on a steel structure surface by a restricted Boltzmann machines algorithm based on consumer-grade camera images. Structural Control and Health Monitoring, 25(2), e2075.
3. Xu, Y., Bao, Y., Chen, J., Zuo, W., & Li, H. (2019). Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images. Structural Health Monitoring, 18(3), 653-674.
4. Xu, Y., Fan, Y., Bao, Y., & Li, H. (2023). Task-aware meta-learning paradigm for universal structural damage segmentation using limited images. Engineering Structures, 284, 115917.
5. Xu, Y., Fan Y., & Li, H. (2023). Lightweight semantic segmentation of complex structural damage recognition for actual bridges. Structural Health Monitoring, 14759217221147015: 1-20. https://doi.org/10.1177/14759217221147015.
6. Chen, J., Chen, Z., Xu, Y., & Li, H. (2021). Efficient reliability analysis combining Kriging and subset simulation with two-stage convergence criterion. Reliability Engineering & System Safety, 214, 107737.
7. Xu, Y., Zhu, B., Zhang, Z., & Chen, J. (2022). Hierarchical dynamic Bayesian network-based fatigue crack propagation modeling considering initial defects. Sensors. 2022, 22(18), 6777.
8. Xu, Y., Tian, Y., & Li, H. (2022). Unsupervised deep learning method for bridge condition assessment based on intra- and inter-class probabilistic correlations of quasi-static responses. Structural Health Monitoring-An International Journal, 22(1): 14759217221103016.
9. Xu, Y., Bao, Y., Zhang, Y., & Li, H. (2020). Attribute-based structural damage identification by few-shot meta learning with inter-class knowledge transfer. Structural Health Monitoring-An International Journal, 20(4), 1494-1517.
10. Qiao, W., Zhao, Y., Xu, Y., Lei Y., Wang Y., Yu, S., & Li, H. (2021). Deep learning-based pixel-level rock fragment recognition during tunnel excavation using instance segmentation model. Tunnelling and Underground Space Technology, 115, 104072.
11. Xu, Y., Qian, W., Li, N., & Li, H. (2022). Typical advances of artificial intelligence in civil engineering. Advances in Structural Engineering. 2022, 25(16): 3405-3424.
12. Wang, Y., Cui, L., Zhang, C., Chen, W., Xu, Y., Zhang, Q. (2022). A two-stage seismic damage assessment method for small, dense, and imbalanced buildings in remote sensing images. Remote Sensing, 14(4), 1012.
13. Cui, L., Jing, X., Wang, Y., Huan, Y., Xu, Y., & Zhang, Q. (2023). Improved swin transformer-based semantic segmentation of post-earthquake dense buildings in urban areas using remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 369-385. DOI: 10.1109/JSTARS.2022.3225150.
14. Xu, Y., Qiao, W., Zhao, J., Zhang, Q., & Li, H. (2023). Vision-based multi-level synthetical evaluation of seismic damage for RC structural components: a multi-task learning approach. Earthquake Engineering and Engineering Vibration. https://doi.org/10.1007/s11803-023-2153-4.
15. Wang, Y., Jing, X., Chen, W., Li, H., Xu, Y., & Zhang, Q. (2023). Geometry-informed deep learning-based structural component segmentation of post-earthquake buildings. Mechanical Systems and Signal Processing, 188, 110028. DOI: 10.1016/j.ymssp.2022.110028.
16. Wang, Y., Jing, X., Xu, Y., Cui, L., Zhang, Q., & Li, H. (2023). Geometry-guided semantic segmentation for post-earthquake buildings using optical remote sensing images. Earthquake Engineering & Structural Dynamics, 1-22. https://doi.org/10.1002/eqe.3966.
17. Wang, Y., Jing X., Cui, L., Zhang, C., Xu, Y., Yuan, J., & Zhang, Q. (2023). Geometric consistency enhanced deep convolutional encoder-decoder for urban seismic damage assessment by UAV images. Engineering Structures, 286, 116132.
18. Xu, Y., Wei, S., Bao, Y., & Li, H. (2019). Automatic seismic damage identification of reinforced concrete columns from images by a region-based deep convolutional neural network. Structural Control and Health Monitoring, 26(3), e2313.
19. Xu, Y., Li, Y., Zheng, X., Zheng, X., & Zhang, Q. (2023). Computer-vision and machine-learning-based seismic damage assessment of reinforced concrete structures. Buildings, 2023, 13, 1258. https://doi.org/10.3390/buildings13051258.