ENHANCING SEISMIC STRUCTURAL DAMAGE ASSESSMENT OF LOW-TO-MEDIUM RISE REINFORCED CONCRETE FRAMED BUILDINGS USING ARTIFICIAL INTELLIGENCE

KAVEESH GUWANINDU ABEYSURIYA[1], MIHAELA ANCA CIUPALA[1],
SEYED ALI GHORASHI[1] & ALPER ILKI[2]
[1]School of Architecture, Computing and Engineering, University of East London, UK
[2]Faculty of Civil Engineering, Istanbul Technical University, Turkey

ABSTRACT

This paper has introduced an artificial intelligence (AI) integrated method for automating the assessment of seismic structural damage in reinforced concrete (RC) buildings, curtailing the need for conventional, time-intensive on-site visual inspections. A deep learning-based damage assessment model has been developed using pre-trained convolutional neural networks to identify damage-indicators, such as cracks, spalling and crushing from images, and subsequently to predict two crucial local element structural failure modes in low-to-medium rise RC framed buildings: shear and flexural failure. The incorporation of local element structural failure modes within this damage assessment model has been aligned with current damage assessment guidelines, facilitating a transition from simply assessing the level of structural damage to providing more actionable insights for structural integrity evaluations and retrofitting decisions. To develop a high-quality model and tackle key challenges in adopting AI in earthquake/structural engineering domain, particularly the scarcity and imbalance of image datasets, this paper has employed transfer learning, data augmentation and synthetic data generation techniques. These techniques have significantly improved model performance and generalisability, ensuring robust and reliable predictions. The proposed model has achieved a uniform score of 0.91 (91%) in accuracy, precision, recall and F1-score without overfitting, showcasing its reliability for real-world implementation. This research marks a significant step forward in AI-integrated seismic structural damage assessment, providing a rapid, accurate, and scalable method to enhance structural integrity evaluations and urban resilience.

Keywords: seismic damage, reinforced concrete buildings, damage assessment, artificial intelligence, convolutional neural networks, data augmentation, synthetic data generation, transfer learning.

1 INTRODUCTION

Seismic activities worldwide cause varying levels of damage to buildings, ranging from partial to complete collapses, resulting in loss of life and substantial economic consequences [1]. Therefore, rapidly and accurately assessing the damages in buildings after a seismic activity is important for ensuring the safety, functionality and longevity of the structures in seismic-prone regions [2]. These assessments are typically carried out by qualified engineers as per damage assessment guidelines and design codes, to enable accurate identification of structural damage levels [3]. Qualitative methods, such as visual inspections, field assessments, and image-based damage detection, remain the most common methods to assess damages in reinforced concrete (RC) buildings, which rely on identifying damage indicators, such as cracks, spalling, crushing, buckling of steel reinforcement and residual deformation [4]. These insights provide useful information on the building's structural integrity, which forms the basis for targeted retrofit and repair strategies, reducing risks to occupants and buildings while expediting recovery from the disaster and helping communities return to normality [5]. However, these manual approaches rely on subjective interpretations that can result in errors or omissions, and the hazardous conditions encountered during on-site

WIT Transactions on The Built Environment, Vol 217, © 2025 WIT Press
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/ERES250031

inspections further compromise the reliability of the assessment. This limitation highlights the urgent need for a rapid and accurate damage assessment methods, which can be executed remotely, particularly in disaster-prone regions with frequent seismic events and numerous substandard buildings. In response, artificial intelligence (AI) techniques, particularly deep learning (DL) techniques, offer promising potential to automate the qualitative seismic structural damage assessment process using image-based data. AI-driven approaches cannot only address the inherent limitations of manual on-site inspections but also pave the way for more reliable and rapid damage assessments. Fig. 1 shows schematics of two potential avenues for adopting AI in seismic structural damage assessments: an image-based, qualitative approach and a vibration-based, quantitative approach. This paper has focused on qualitative assessments using image-based data, as it remains the most prevalent and feasible practice in the field, as opposed to quantitative assessments using vibration-based data, which is costly in implementation in many disaster-prone economies [4].

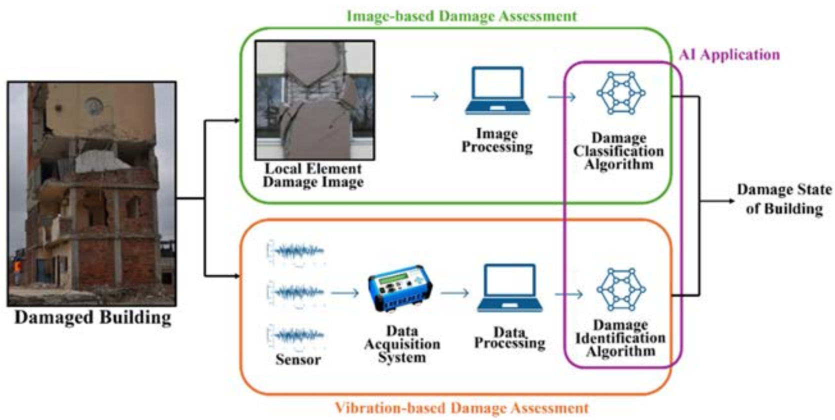

Figure 1: Seismic structural damage assessment approaches adopting AI.

Since the mid-2010s, advances in machine learning (ML)/deep learning (DL), computing power and data availability have rapidly expanded AI's role in structural damage assessment [6]. DL, particularly multi-layer artificial neural network algorithms, now underpins sophisticated research on crack detection, damage classification and failure mode prediction [7]. Early image-based crack detection employed edge detection algorithms (e.g. Canny and Sobel), which later gave way to ML approaches such as support vector machines [8]. Taking advantage of convolutional neural networks (CNN), researchers achieved higher accuracy in identifying cracks, spalling and related damage [9]. Transfer learning (TL) techniques further refined CNN-based detection by leveraging pre-trained models to overcome limitations in datasets within structural/earthquake engineering domain [10]. However, relatively small size of datasets presents challenges for ML/DL-based classification, that the TL strategies (e.g. fine-tuning and feature extraction) help to address [11]. A very few studies demonstrated the efficacy of customised CNN-based models derived from pre-trained CNN models, achieving high accuracy in real-world testing scenarios [12]. Accurate labelling remains crucial, as mislabelled data can degrade model performance. Local element failure mode prediction identifies whether a structural component fails in shear, flexure or a combination of both. Deep neural network-based models have successfully predicted shear wall failure modes based on geometry, reinforcement and material properties [13].

Based on the review of literature above, the following key challenges faced by researchers who are using AI for seismic structural damage assessment using image-based data have been addressed in this paper:

- The image-based datasets acquired in structural/earthquake engineering domains are smaller than datasets in other disciplines. This data scarcity poses a challenge for accurately training, testing and validating DL-based models as they rely on large number of data to ensure model does not overfit and is able to generalise to new and unseen data.
- The lack of open access image-based datasets that depict local element failure modes of building structures, such as shear failure and flexural failure.
- The lack of comparative studies on the performance of the available pre-trained CNN models, together with TL techniques in seismic structural damage assessment using image-based data.

Therefore, this paper focusses on enhancing seismic structural damage assessment of low-to-medium-rise RC framed buildings through a DL-based damage assessment model utilising image-based data. The proposed model automatically detects damage-indicative features, such as cracks, spalling, and crushing, to predict local failure modes in structural elements. To achieve this, a suitable pre-trained CNN has been selected by evaluating relevant pre-trained CNN models, followed by hyperparameter tuning for optimal alignment with the target dataset of this research. The target dataset used for training and testing this model includes images of RC columns, RC beams, and masonry infill walls depicting shear and flexural failure modes. To enhance model performance and generalisability, transfer learning, data augmentation, and synthetic data generation have been employed.

2 METHODOLOGY

To develop a seismic structural damage assessment model that predicts local element failure modes of RC framed buildings utilising CNNs, this paper adopts a five-step methodology. The model's classification criteria have been defined according to local element failure modes, for which a database of damage images has been developed. Data augmentation and synthetic data generation techniques have subsequently been employed to increase dataset size and achieve greater balance. Transfer learning and model regularisation techniques have then been incorporated, and finally, performance assessment criteria have been established to ensure comprehensive evaluation of the developed model.

2.1 Model classification criteria

The existing damage assessment guidelines generally assess the level of damages in structural elements [14]. This paper aims to develop a model that predicts the local element failure modes of low-to-medium rise RC framed buildings. The failure modes depicted in Fig. 2 form the basis for the model's classification criteria, although this study specifically focuses on shear and flexural failure modes in RC columns, RC beams, and masonry infill walls.

2.2 Database development

In current damage assessment practices, structural element damage is primarily documented through photographs obtained during on-site or remote building inspections. These image collections capture crucial indicators of physical damage, such as crushed regions, distinct

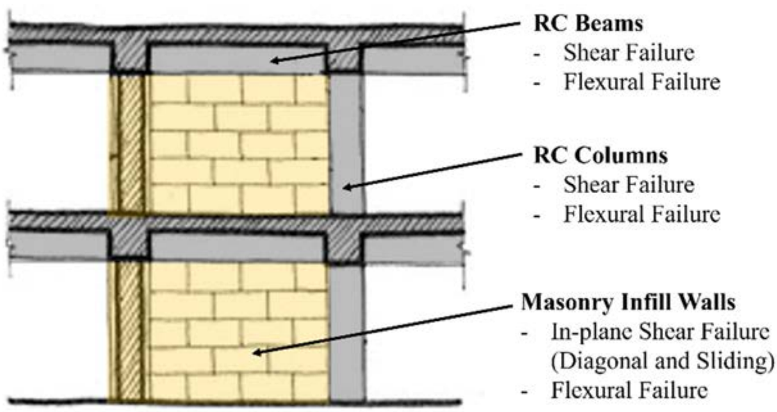

Figure 2: Structural elements and failure modes applicable to low-to-medium rise RC framed buildings.

Figure 3: Sample images of shear failure and flexural failure of RC column, RC beam and masonry infill walls included in the database developed in this paper.

collections capture crucial indicators of physical damage, such as crushed regions, distinct crack patterns and measurements of crack width and length, which enables precise characterisation of observed damage. The database has been developed by following a multi-step process including web scraping, data segregation, data cleaning, quality checks, database structuring, versioning, and documentation. Each image is ultimately classified according to the type of RC structural element and the failure mode observed in the damaged component, constituting the training data for the damage assessment model. As illustrated in Fig. 3, the dataset comprises shear failure and flexural failure in RC columns, RC beams, and masonry infill walls, thereby establishing the model's classification criteria.

2.3 Data augmentation and synthetic data generation

Table 1 presents an overview of the database developed for this study. Notably, the datasets for RC beams and masonry infill walls are comparatively smaller and exhibit imbalance relative to those for RC columns. Such imbalances, along with limited dataset sizes, can yield overfitted models that are unable to generalise on new and unseen data. To address these challenges and enhance variability of the dataset, data augmentation (Fig. 4) and synthetic data generation (Fig. 5) techniques have been employed to generate new images, thereby simulating a wider range of real-world conditions.

Table 1: Overview of the database developed in this paper.

Element type	**Local element failure mode**		**Total**
	Shear failure	**Flexural failure**	
RC columns	1,048	858	1,906
RC beams	152	100	252
Masonry infill walls	880	327	1,207

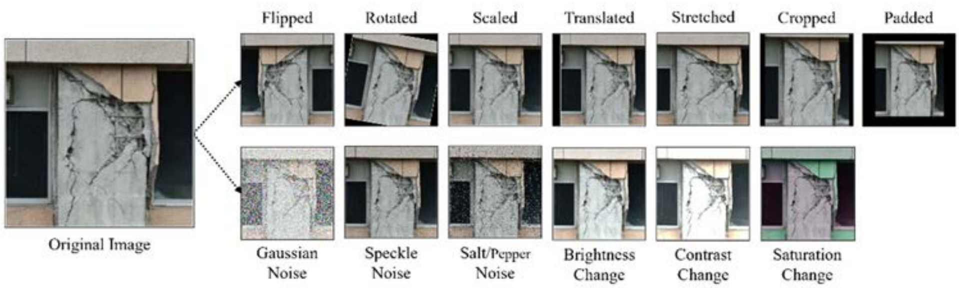

Figure 4: Images generated using synthetic data generation techniques.

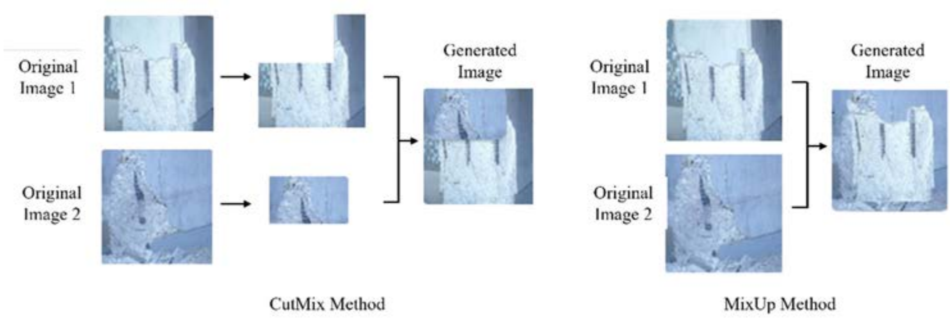

Figure 5: Images generated using data augmentation techniques.

2.4 Model development

AI consists of computational techniques designed to simulate an array of cognitive functions representative of a human brain, such as reasoning, classification, appraisal, and determination of optimal courses of action. The AI technique employed in the damage assessment model developed in this paper is convolutional neural networks (CNN), which is a specialised DL model specifically designed to process structured data on a grid (e.g. images of damaged RC elements). CNNs are highly effective at recognising spatial hierarchies in data by using convolutional layers that apply filters to the input, detecting fundamental features, such as edges, colours and textures. These layers progressively combine these simpler features to identify more complex patterns, such as shapes or objects in the images. As illustrated in Fig. 6, CNN typically consists of three key types of layers: convolutional layers, pooling layers, and fully connected layers. In convolutional layers, small filters move across the input data, creating feature maps that capture specific patterns in different regions. Pooling layers then reduce the dimensionality of these feature maps, improving the model's computational efficiency and helping to prevent overfitting. The fully connected layers integrate all the learned features to produce predictions or classifications [15]. In the context of this paper, the model input consists of images depicting damaged structural elements, while the output class indicates the corresponding local element failure mode: shear failure or flexural failure. Other failure modes can be incorporated in future research to further expand the scope of this damage assessment model.

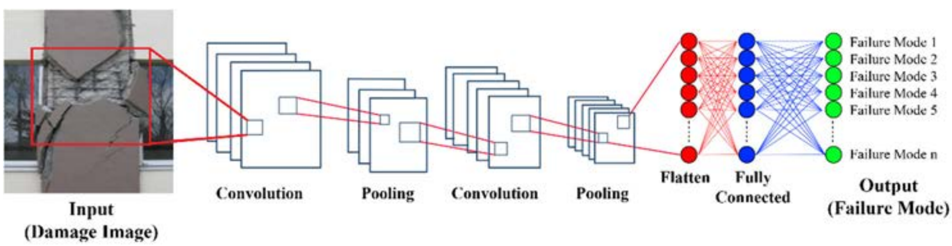

Figure 6: Convolutional neural network (CNN) architecture.

Pre-trained CNN architectures have shown superior capabilities in image-based damage assessment tasks. The most suitable pre-trained CNN-based architecture for this research has been selected by evaluating the performance of pre-trained CNN architectures renowned for image-based prediction tasks, namely VGGNet (VGG16 and VGG19) [16], GoogleNet (Inception) [17], Xception [18], MobileNet [19], DenseNet (DenseNet121, DenseNet169, DenseNet201) [20], ResNet (ResNet50, ResNet101 and ResNet152) [21], EfficientNet-B0 [22] and AlexNet [23], with the target dataset. Since these models have initially been trained on datasets from different domains, assessing their ability to extract meaningful features from the target datasets relevant to the domain of this paper is crucial, hence the need of the comparative experiment. Next, transfer learning techniques have been applied to the most applicable pre-trained CNN model in terms of performance to the target dataset. As illustrated in Fig. 7, two types of transfer learning techniques have been adopted in this paper: feature extraction and fine tuning. In feature extraction, the pre-trained model's layers remain unchanged, allowing it to automatically capture generalised patterns from the new dataset. On the other hand, fine-tuning modifies specific layers, usually the last ones, to optimise the model's parameters for improved adaptation to the target dataset.

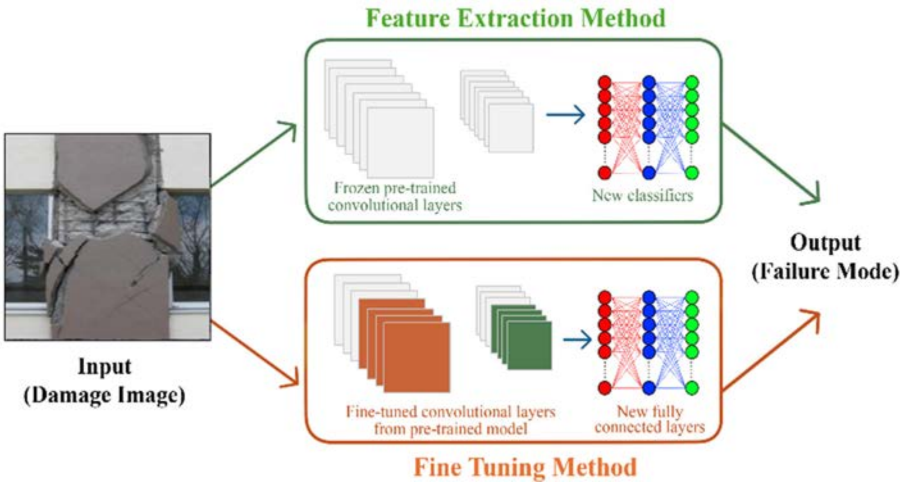

Figure 7: Transfer learning techniques employed in this paper.

2.5 Model evaluation

The performance of CNN-based models has been evaluated using a confusion matrix [24]. This matrix consists of four key elements: true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). Accuracy (eqn (1)) is calculated as the percentage of correctly classified instances out of the total cases. Precision (eqn (2)) represents the percentage of correctly predicted positive cases among all predicted positives. Recall (eqn (3)) measures the percentage of actual positive cases correctly identified by the model. The F1 score (eqn (4)) is the harmonic mean of precision and recall, expressed as a percentage, providing a balanced measure of performance [25].

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}, \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}, \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}, \tag{3}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}. \tag{4}$$

To check if the model overfits, the difference between the training accuracy (TRA) and validation accuracy (VLA) is checked as per eqn (5) below:

$$\text{If difference \%} = \text{TRA} - \text{VLA}; \quad \begin{cases} \leq 5\%; \text{ no overfitting;} \\ \text{between } 5-10\%; \text{ mild overfitting} \\ > 10\%; \text{ severe overfitting.} \end{cases} \tag{5}$$

Similarly, the ability of the model to generalise to new and unseen data has been checked by comparing the gap between the validation and testing accuracy (eqn (6)) and loss (eqn (7)).

$$\text{Accuracy gap} = \text{VLA} - \text{TSA} \ \leq 3\% \text{ well generalisable,} \tag{6}$$

$$\text{Loss gap} = \text{VLL} - \text{TSL} \ \leq 0.1\% \text{ well generalisable.} \tag{7}$$

3 RESULTS

3.1 Dataset

Table 2 presents the final dataset (total of 6,314 images), which has been enhanced using data augmentation and synthetic data generation techniques to ensure sufficient size and balance for effective model training, validation, and testing. This dataset was divided into training, validation, and testing sets using a 70:10:20 ratio.

Table 2: Overview of the dataset used for model training, validation and testing.

Element type	**Local element failure mode**		**Total**
	Shear failure	**Flexural failure**	
RC columns	1,048	1,058	2,106
RC beams	1,049	1,059	2,108
Masonry infill walls	1,063	1,037	2,100

3.2 Model performance

A comparative experiment has been conducted to determine the most suitable pre-trained CNN model specifically for the target dataset of this paper by evaluating the performance of various pre-trained CNN models previously used within this research domain. MobileNet pre-trained CNN model that has demonstrated the best performance in this comparative experiment, with a uniform score of 0.88 (88%) for accuracy, precision, recall and F1 score, with the shortest run-time, completing the task in 43 minutes has been selected as the baseline model. Feature extraction and fine-tuning transfer learning techniques have been applied to the baseline MobileNet model, and subsequently the regularisation parameters of the model, such as loss function regularisation factor, dropout rate, and learning rate and number of epochs have been adjusted. The final model's confusion matrix is presented in Fig. 8. As indicated in Table 3, this model has achieved uniform scores of 0.91 (91%) for accuracy, precision, recall, and F1 score.

Fig. 9 illustrates the training/validation accuracy curve and loss curve of the model, where the details for overfitting and generalisability check have been extracted to Table 3. These results confirm that the model developed in this paper has been successful in predicting the failure modes of RC columns, RC beams and masonry infill walls accurately, without overfitting, while ensuring superior generalisability to new and unseen data.

4 CONCLUSION

This paper has made notable progress in adopting deep learning techniques to assess seismic structural damages in low-to-medium rise reinforced concrete framed buildings and subsequently predict the failure modes of reinforced concrete columns, beams and masonry

www.witpress.com, ISSN 1743-3509 (on-line)

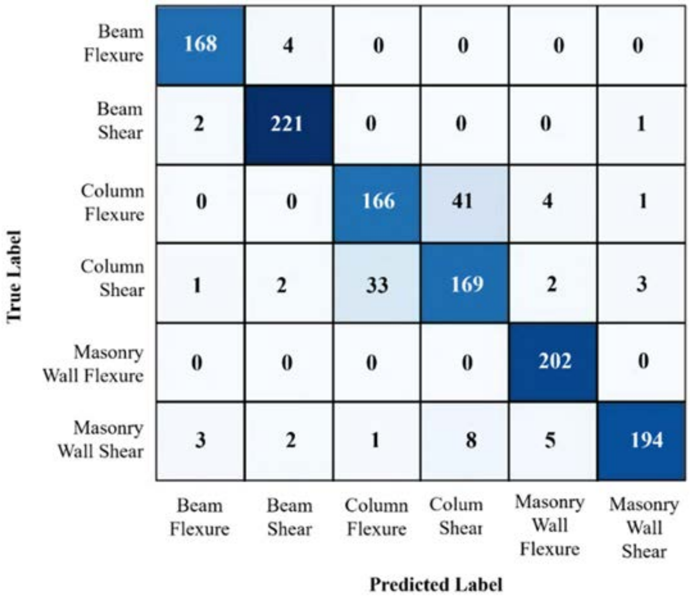

Figure 8: Confusion matrix of the damage assessment model developed in this paper.

Table 3: Model performance results.

Criteria	Matrix/Check	Score	Limit
Performance check	Accuracy	0.9084	N/A
	Precision	0.9085	N/A
	Recall	0.9084	N/A
	F1 score	0.9080	N/A
Overfitting check	Training accuracy and validation accuracy difference %	0.25%	<5%
Generalisation check	Fine-tuning validation accuracy and overall model test accuracy gap	0.08%	<3%
	Fine-tuning validation loss and overall model test loss gap	0.00%	0.01%

infill walls. By utilising pre-trained convolutional neural networks, a deep learning technique along with transfer learning, data augmentation and synthetic data generation, this paper has produced a damage assessment model that can efficiently predict local element structural failure modes using image-based data. The model developed in this paper has achieved uniform scores of 0.91 (91%) for accuracy, precision, recall, and F1 score, the performance indicators computed using confusion matrix, the standard method to evaluate the performance of deep learning-based models. Further model performance checks conducted has ensured that the model does not overfit and demonstrates superior generalisability to new and unseen data. This research development has contributed to a transferable body of knowledge that bridges cutting-edge AI techniques with conventional earthquake engineering, providing a scalable and practical approach to strengthening urban resilience and disaster preparedness worldwide.

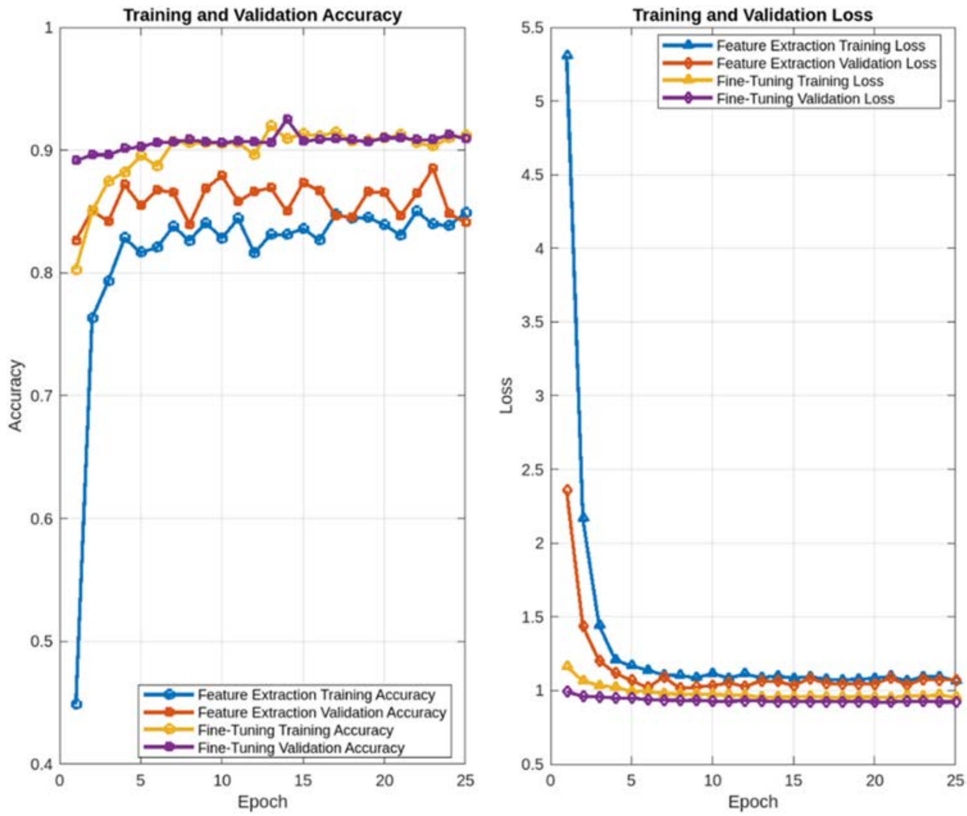

Figure 9: Training/validation accuracy and loss curves of the damage assessment model developed in this paper.

ACKNOWLEDGEMENTS

The financial contribution of the University of East London, UK, and the 2024 Earthquake Engineering Field Investigation Team (EEFIT) Research Grant Scheme, the Institution of Structural Engineers, UK, is greatly appreciated.

REFERENCES

[1] Kazama, M. & Noda, T., Damage statistics (Summary of the 2011 off the Pacific Coast of Tohoku Earthquake damage). *Soils Found.*, **52**(5), pp. 780–792, 2012.

[2] Arslan, M.H. & Korkmaz, H.H., What is to be learned from damage and failure of reinforced concrete structures during recent earthquakes in Turkey? *Eng. Fail. Anal.*, **14**(1), pp. 1–22, 2007. https://doi.org/10.1016/j.engfailanal.2006.01.003.

[3] Gatti, M., Structural health monitoring of an operational bridge: A case study. *Eng. Struct.*, **195**, pp. 200–209, 2019. https://doi.org/10.1016/j.engstruct.2019.05.102.

[4] Rosenblatt, W., Laursen, P., McDaniel, C. & Archer, G., Structural damage detection through forced vibration testing. *16th World Conference on Earthquake (16WCEE)*, Santiago Chile, 2017.

[5] Rossetto, T. & Elnashai, A., Derivation of vulnerability functions for European-type RC structures based on observational data. *Eng. Struct.*, **25**(10), pp. 1241–1263, 2003. https://doi.org/10.1016/S0141-0296(03)00060-9.
[6] Falcone, R., Lima, C. & Martinelli, E., Soft computing techniques in structural and earthquake engineering: A literature review. *Eng. Struct.*, **207**, 110269, 2020. https://doi.org/10.1016/j.engstruct.2020.110269.
[7] Bao, Y., Liu, D. & Li, H., A mechanics-informed neural network method for structural modal identification. *Mech. Syst. Signal Process*, **216**, 2024. https://doi.org/10.1016/j.ymssp.2024.111458.
[8] Jahanshahi, M.R. & Masri, S.F., A new methodology for non-contact accurate crack width measurement through photogrammetry for automated structural safety evaluation. *Smart Mater. Struct.*, **22**(3), p. 35019, 2013. https://doi.org/10.1088/0964-1726/22/3/035019.
[9] Cha, Y.J., Choi, W. & Büyüköztürk, O., Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, **32**(5), pp. 361–378, 2017. https://doi.org/10.1111/mice.12263.
[10] Wang, N., Zhao, X., Zou, Z., Zhao, P. & Qi, F., Autonomous damage segmentation and measurement of glazed tiles in historic buildings via deep learning. *Computer-Aided Civil and Infrastructure Engineering*, **35**(3), pp. 277–291, 2020. https://doi.org/10.1111/mice.12488.
[11] Ogunjinmi, P.D., Park, S.S., Kim, B. & Lee, D.E., Rapid post-earthquake structural damage assessment using convolutional neural networks and transfer learning. *Sensors*, **22**(9), 2022. https://doi.org/10.3390/s22093471.
[12] Yilmaz, M., Dogan, G., Arslan, M.H. & Ilki, A., Categorization of post-earthquake damages in RC structural elements with deep learning approach. *Journal of Earthquake Engineering*, **28**(9), pp. 2620–2651, 2024. https://doi.org/10.1080/13632469.2024.2302033.
[13] Barkhordari, M.S. & Massone, L.M., Efficiency of deep neural networks for reinforced concrete shear walls failure mode detection. *27th International Computer Conference, Computer Society of Iran (CSICC)*, pp. 1–4, 2022. https://doi.org/10.1109/CSICC55295.2022.9780477.
[14] Ilki, A., Halici, O.F., Comert, M. & Demir, C., The modified post-earthquake damage assessment methodology for TCIP (TCIP-DAM-2020). *Advances in Assessment and Modeling of Earthquake Loss*, eds S. Akkar, A. Ilki, C. Goksu & M. Erdik, Springer, pp. 85–107, 2021. https://doi.org/10.1007/978-3-030-68813-4_5.
[15] Krizhevsky, A., Sutskever, I. & Hinton, G.E., ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, eds F. Pereira, C.J. Burges, L. Bottou & K.Q. Weinberger, Curran Associates, Inc., 2012. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.
[16] Simonyan, K. & Zisserman, A., Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations (ICLR 2015)*, Computational and Biological Learning Society, pp. 1–14, 2015.
[17] Szegedy, C. et al., Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, 2014. https://api.semanticscholar.org/CorpusID:206592484.
[18] Chollet, F., Xception: Deep learning with depthwise separable convolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807, 2016. https://api.semanticscholar.org/CorpusID:2375110.

[19] Howard, A.G. et al., MobileNets: Efficient convolutional neural networks for mobile vision applications. *ArXiv*, abs/1704.04861, 2017. https://api.semanticscholar.org/CorpusID:12670695.

[20] Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K.Q., Densely connected convolutional networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2261–2269, 2017. https://doi.org/10.1109/CVPR.2017.243.

[21] He, K., Zhang, X., Ren, S. & Sun, J., Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2015. https://api.semanticscholar.org/CorpusID:206594692.

[22] Tan, M. & Le, Q.V., EfficientNet: Rethinking model scaling for convolutional neural networks. *ArXiv*, abs/1905.11946, 2019. https://api.semanticscholar.org/CorpusID:167217261.

[23] Krizhevsky, A., Sutskever, I. & Hinton, G.E., ImageNet classification with deep convolutional neural networks. *Commun ACM*, **60**(6), pp. 84–90, 2017. https://doi.org/10.1145/3065386.

[24] Sokolova, M. & Lapalme, G., A systematic analysis of performance measures for classification tasks. *Inf. Process Manag*., **45**(4), pp. 427–437, 2009. https://doi.org/10.1016/j.ipm.2009.03.002.

[25] Powers, D., Evaluation: From precision, recall and F-factor to ROC, informedness, markedness and correlation. *Mach. Learn. Technol*., **2**, 2008.