# EMYNet-BDD: EfficientViTB Meets Yolov8 in the Encoder–Decoder Architecture for Building Damage Detection Using Postevent Remote Sensing Images

Masoomeh Gomroki, *Member, IEEE*, Mahdi Hasanlou, *Member, IEEE*, Jocelyn Chanussot, *Fellow, IEEE*, and Danfeng Hong, *Senior Member, IEEE*

***Abstract*—Natural disasters commonly occur in all regions around the world and cause huge financial and human losses. One of the main effects of earthquakes and floods is the destruction of buildings. Photogrammetric and remote sensing (RS) data track changes and detect damages in these events. Considering the evolution in deep learning (DL) techniques, the possibility of accurate information extraction from the RS-based data is increased. DL methods effectively show the damaged regions for decision making and immediate actions for crisis management. The present study is based only on postevent RS images, which apply an encoder–decoder network composed of pretrained EfficientViTB and Yolov8 network blocks as encoder path and the modified-Unet blocks as decoder path for building damage detection (BDD). Compared with methods that use only one network in their encoder path, the presented method achieves better results. To investigate the performance of the proposed method, three datasets affected by different natural disasters are considered. The first dataset is the satellite images of the 2023 Turkey earthquake, the second dataset is associated with the satellite images of the 2023 Morocco earthquake, and the third dataset contains the satellite images of the 2023 Libya flood. The proposed method ultimately reaches the overall accuracy of 97.62%, 98.63%, and 96.43% and the kappa coefficient of 0.86, 0.85, and 0.84 for the first, second, and third dataset, respectively, which shows the proper performance of the proposed method for BDD.**

***Index Terms*—Building damage detection (BDD), EfficientViTB, encoder–decoder network, Libya flood, Morocco earthquake, Turkey earthquake, Yolov8.**

## I. Introduction

IN 2023, natural disasters, such as the Turkey earthquake (6 February), the Maui wildfire (8 August), hurricane Idalia (29 August), the Morocco earthquake (9 September), and the Libya flood (12 September), happened in the world and caused huge human and financial losses. Availability of the maps of the damaged regions for rapid rescue and recovery is one of the most important challenges during these natural disasters [1]. Developments in remote sensing (RS) technology have currently made it possible to collect diverse and extensive data from the regions affected by natural disasters. Nonetheless, a basic challenge is finding methods to extract effective information within a reasonable time accurately [1], [2], [3]. The RS data, such as the synthetic-aperture radar (SAR), the light detection and ranging (LiDAR) system, and the unmanned aerial vehicles (UAVs), as well as the satellite images and multimodal data are used for damage assessment of natural disasters, e.g., earthquakes, floods, and hurricanes.

The SAR data can be captured in all weather conditions, even in the presence of clouds. Although such data seem to be suitable for real-time analysis, its interpretations are difficult and influenced easily by speckle noise [4]. Ferrentino et al. [5] investigated the damage of the August 2016 Central Italy earthquake using the dual-polarization SAR data captured by Sentinel-1. In their proposed method, a convolutional single polarization feature was initially analyzed based on the correlation between the polarized images of pre- and postevent. Then, the same analysis was performed through the cross-polarized channel, which shows that the simultaneous use of the copolarized and cross-polarized channels reaches better results. Finally, a coherent dual-feature method was suggested based on the interchannel coherence (ICC) technique. They indicated that the ICC method results are superior to those techniques that utilized only two SAR images for earthquake damage detection. Li et al. [6] studied the damage of the L'Aquila earthquake in 2009, Italy, utilizing the SAR images of Envisat. Using basic components of the improved principal component analysis method and multitexturing technique, they achieved favorable results for building damage detection (BDD). ElGharbawi and Zarzoura [7] attempted to assess the damages caused by the massive Beirut explosion in Lebanon using the SAR data and the correlation technique. In their proposed method, a spatial-phase frequency filter was applied to improve the correlation estimates, which enhanced the damage detection results by 44%.

The LiDAR system can be similarly utilized to collect the data in all weather conditions and the presence of clouds. As well, it provides the necessary information about ground elevation,

Manuscript received 21 April 2024; revised 26 June 2024; accepted 9 July 2024. Date of publication 11 July 2024; date of current version 26 July 2024. *(Corresponding author: Mahdi Hasanlou.)*

Masoomeh Gomroki and Mahdi Hasanlou are with the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran (e-mail: masoomeh.gomroki@ut.ac.ir; hasanlou@ut.ac.ir).

Jocelyn Chanussot is with the CNRS, Grenoble INP, GIPSA-Lab, University Grenoble Alpes, 38000 Grenoble, France, and also with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: jocelyn.chanussot@grenoble-inp.fr).

Danfeng Hong is with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: hongdf@aircas.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3427017

which allows a more precise BDD. However, capturing LiDAR data seems to be time-consuming and it cannot be immediately provided after events [8]. Zhou et al. [1] investigated Sandy superstorm damage using the pre- and postevent LiDAR data. In their proposed method, building clusters were first extracted by the density-based method. Then, a new cluster-matching method was proposed to match the buildings before and after the event. Finally, some features, such as the roof area, orientation, and shape of the buildings, were considered to extract the damaged buildings. Janalipour and Mohammadzadeh [9] exploited the postearthquake LiDAR data to assess the damages of buildings. For this purpose, three spectral features viz., Haralick feature, Gabor filter, and Laws' mask were used, and then the buildings were extracted with a fuzzy system. Axel and van Aardt [10] suggested a supervised technique for LiDAR-based data segmentation for BDD. In their method, local surface feature was first extracted and then the damaged buildings were estimated using rooftop inclination.

Satellite images are more available than other RS data so they are employed more than other sources for BDD purposes. Unlike the LiDAR and SAR data, such images are much easier to interpret and process, so they are mostly used to detect damages [8], [11]. Merlin and Jiji [12] used pre- and postevent satellite images for BDD. In their proposed method, the buildings were extracted by thresholding and color invariance and then building changes were detected by image differencing. Afterward, the spectral and spatial features were considered for the feature-ranking method and, finally, the classification was executed using the mean ratio feature selection. Ji et al. [13] made use of the visual geometry group (VGG) network to detect the damage of the 2010 Haiti earthquake with the pre- and postevent satellite images. They also study the effects of data augmentation and freezing on some layers of the given network in the final results. Li et al. [14] analyzed the damage of Hurricane Sandy impacting the Caribbean in 2012 using postsatellite images. For this purpose, an encoder–decoder network based on the VGG16 network was utilized. Different data augmentation techniques, such as mirroring, rotation of the Gaussian blur, and the Gaussian noise, were considered to get better results. Chen et al. [15] proposed an unsupervised change detection object-based framework for large-scale areas. Zhou et al. [16] presented a context aggregation network for change detection. They used various benchmarks to show the performance of their methods. Hang et al. [17] introduced an ambiguity-aware network to address the ambiguous regions, where pseudochange occurred or real changes were corrupted. Ji et al. [18] used satellite images before and after the 2010 Haiti earthquake for BDD. In their proposed method, a combination of random forest (RF) algorithm and convolutional neural network (CNN) was employed and their findings revealed that the performance of CNN together with RF was better than the RF algorithm alone. Hang et al. [19] implement a CNN-based network for segmentation by using high-resolution RS. Wu et al. [20] benefited from the pre- and postevent satellite images on the xBD dataset to detect the damaged buildings based on different U-Net architectures with the attention mechanism. Shen et al. [21] used the same dataset for BDD based on the dual-branch U-Net architecture. Zheng et al. [22] investigated the damage of buildings within the xBD dataset. They executed an object-based method in combination with a DL-based network. Employing both pre- and postevent data of the xBD dataset for damage detection, Chen et al. [23] examined the encoder–decoder and transformer-based networks. First, nonlocal deep features were extracted from both data using the transformer-based encoder. Then, the fuse module integrated these features and ultimately the dual-task decoder combined the hierarchical features to map the damage in buildings. Virtriana et al. [24] used the satellite images of the Anak Krakatau Volcano tsunami in Indonesia for BDD and their proposed method consisted of the RF classification algorithm with different predictors.

Multimodal data are also sometimes exploited to detect damage and improve the results. Such data need registration and a high processing time [8]. Gupta and Shah [25] used both the UAV data and satellite images for BDD. Their proposed network was a combination of the ResNet-50 and the Astrous Spatial Pyramid Pooling, which could extract multiscale features and simultaneously obtain the types of damage by segmenting the buildings. Chen et al. [26] used the advantages of structural relationship graph convolutional autoencoder for multimodal change detection. This architecture also is useful for emergency cases, such as disaster management. Also, Chen et al. [27] proposed a framework based on a Fourier-domain graph for unsupervised multimodal change detection. Li et al. [28] combined the LiDAR data and satellite images to detect damaged buildings. In their developed method, the 3-D model of the buildings was first reconstructed with pre-event data and then the damages were detected by the rooftop patch-oriented 3-D estimation. Moreover, Adriano et al. [29] quickly assessed damage by fusing the Sentinel-1, Sentinel-2, and ALOS-2 satellite data.

Most afore-mentioned methods have generally benefited from both pre- and postevent data for BDD but the pre-event ones have not sometimes been available. Furthermore, extracting the changes from two-time data typically encountered some challenges, such as registration error, atmospheric conditions, and noise [11], [30], [31]. Taking into account the main challenges facing each type of data, satellite images are used in this study for BDD, thanks to their easy access and interpretation. In the present study, a new architecture, namely, an encoder–decoder architecture containing the EfficientViTB and the Yolov8 network blocks in its encoder path, is proposed, which uses only the postevent data for BDD (that is, EMYNet-BDD). The EfficientViTB is a cutting-edge recognition model and the Yolov8 is one of the recent detection models with proper performance. A combination of these two models can extract efficient features from input data and overcome the task of BDD. This architecture outperforms the methods that employed a single network in their encoder path. Both the EfficientViTB and the Yolov8 networks are also pretrained and their blocks are utilized in the encoder path, so the semitransfer learning technique [32] is implemented. To show the proper performance of this network, it is used for BDD in the 2023 Turkey earthquake, the 2023 Morocco earthquake, and the 2023 Libya flood.

The main contributions of this study are as follows.

1) Introduce a new encoder–decoder architecture that includes the EfficientViTB and the Yolov8 network blocks as encoders and the modified-Unet blocks as decoder paths to

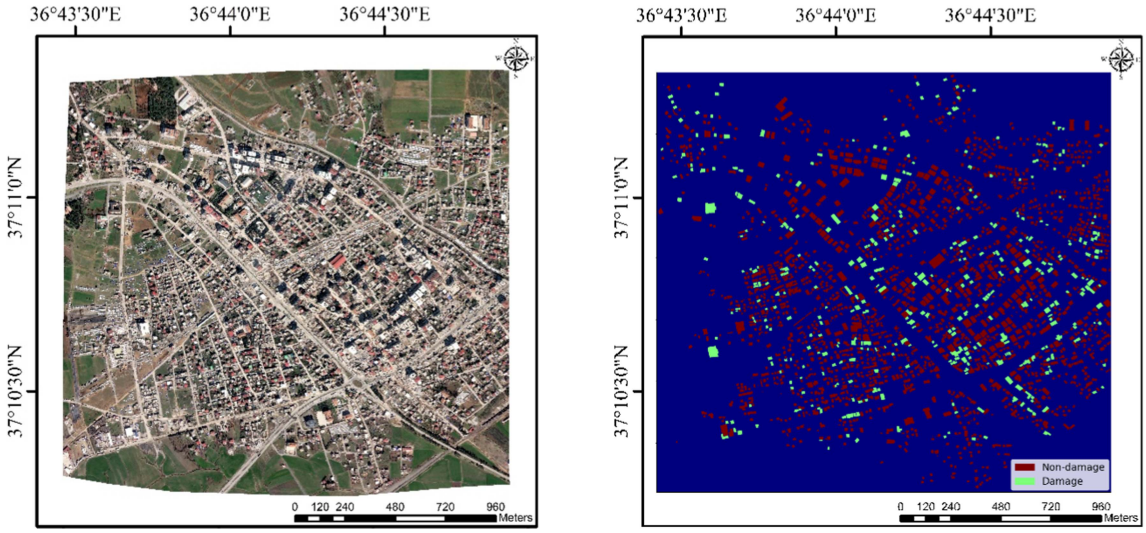



Fig. 1. Turkey earthquake dataset. (a) RGB postimage. (b) GT.

TABLE I
INFORMATION ON THE DATASETS USED IN THIS STUDY

| Datasets | Satellite | Spectral Bands | Special Resolution (m) | Image size | Acquisition Date | Damage Building (polygons) | Nondamage building (polygons) |
|---|---|---|---|---|---|---|---|
| Turkey Earthquake | WorldViewII | R, G, B | 0.5 | 6784×7040 | 11th February 2023 | 400 | 1833 |
| Morocco Earthquake | WorldViewII | R, G, B | 0.5 | 2688×2816 | 10th September 2023 | 44 | 107 |
| Libya Flood | GEO1 | R, G, B | 0.46 | 5248×4352 | 13th September 2023 | 215 | 351 |

improve the overall accuracy (OA) and kappa coefficient (KC).
2) Use only postevent RS images without any auxiliary data.
3) Apply the proposed network to various natural disasters, such as earthquakes and floods.
4) Get the advantages of a semitransfer learning technique in the encoder path (the EfficientViTB and the Yolov8 pretrained by ImageNet and MS-COCO, respectively).

The rest of this article is organized as follows. The data are introduced in Section II and the proposed method is explained in Section III. In Section IV, the experimental results and comparisons are provided. The discussion is presented in Section V. Finally, Section VI concludes this article.

## II. MATERIALS AND DATASETS

### A. 2023 Turkey Earthquake

On 6 February 2023, at 4:17 A.M. (1:17 UTC), an earthquake of 7.8 (on the Richter scale) hit southern and central Turkey as well as southwestern Syria. Fig. 1 illustrates the WorldView II image of the city of Nurdağı after the earthquake and the corresponding ground truth (GT), which shows that the city was densely built and undergoes postearthquake damage. The polygons related to the damaged and nondamaged buildings are drawn in the global mapper and then layered with different numbers to generate the GT. The number of polygons of the nondamaged and damaged buildings is, thus, equal to 1833 and 400, respectively. To the train–test split, 60% and 40% of the total data are considered as the train–validation and test data, respectively. Additional information regarding these data is outlined in Table I.

### B. 2023 Morocco Earthquake

On 8 September 2023, at 22:11 UTC, a 6.8 magnitude earthquake struck the High Atlas Mountain, 71 km southwest of Morocco. Fig. 2 displays the WorldViewII image of the Talat N'Yaaqoub region and the corresponding GT. To generate the GT, we operate like the previous dataset. In this dataset, there are 107 nondamaged polygons of buildings and 44 damaged ones. After training the network with the 2023 Turkey earthquake dataset, the Morocco earthquake dataset is utilized for prediction purposes. Additional information about this dataset is given in Table I.

### C. 2023 Libya Floods

On 11 September 2023, heavy rain led to the collapse of two dams in Libya, and about 30 million $m^3$ of water flowed into the city of Derna, which caused much damage and destroyed many buildings. Fig. 3 shows the GeoEye-1 image of a part of the city of Derna and the corresponding GT in the floodway area. To get the GT, the same procedure is fulfilled similar to

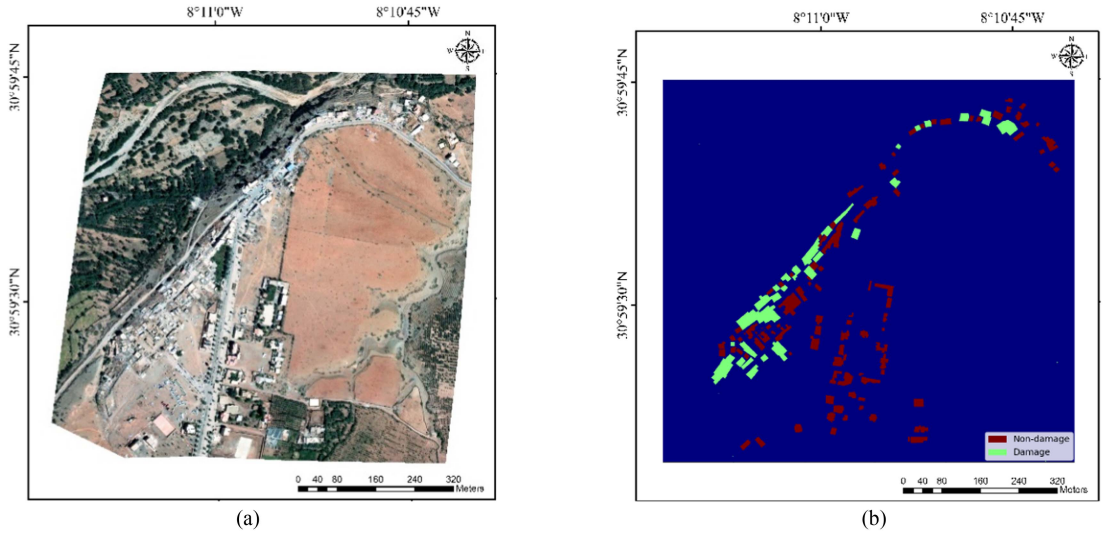



Fig. 2. Morocco earthquake dataset. (a) RGB postimage. (b) GT.

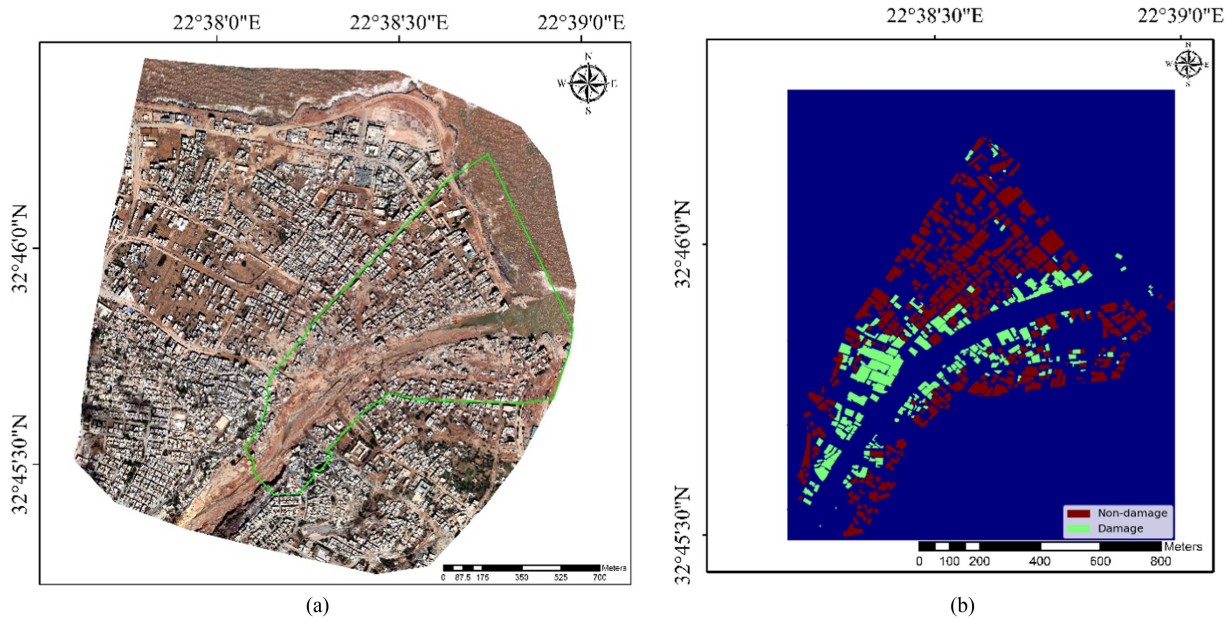



Fig. 3. Libya flood dataset. (a) RGB postimage of the floodway showed by green polygon. (b) GT.

the two previous datasets. The number of polygons of buildings damaged by the floods is 215 and the nondamaged ones are 351. To the train–test split, 60% and 40% of the total data are considered as the train–validation and test data, respectively. Additional information about these data is provided in Table I. All the data related to the present study have been made available to the public by Maxar Technologies Inc.[1]

## III. Proposed Method

The proposed method for BDD consists of three main steps:

1) data preprocessing;
2) training the encoder–decoder network containing the EfficienViTB and the Yolov8 blocks as encoder path and the modified-Unet blocks as decoder path for BDD (EMYNet-BDD);
3) prediction by the trained network and evaluate its accuracy.

In the preprocessing step, each original images are partitioned into patches with the dimensions of 128 × 128, so 2915, 462, and 1394 patches are obtained from the Turkey earthquake, Morocco earthquake, and Libya flood, respectively. In Turkey's earthquake and Libya flood, 60% of patches are considered as the train–validation data and the remaining 40% are used for the test data. All patches of the Morocco earthquake are considered as the test data. After the "train–test split," the

[1][Online]. Available: http://www.maxar.com/open-data

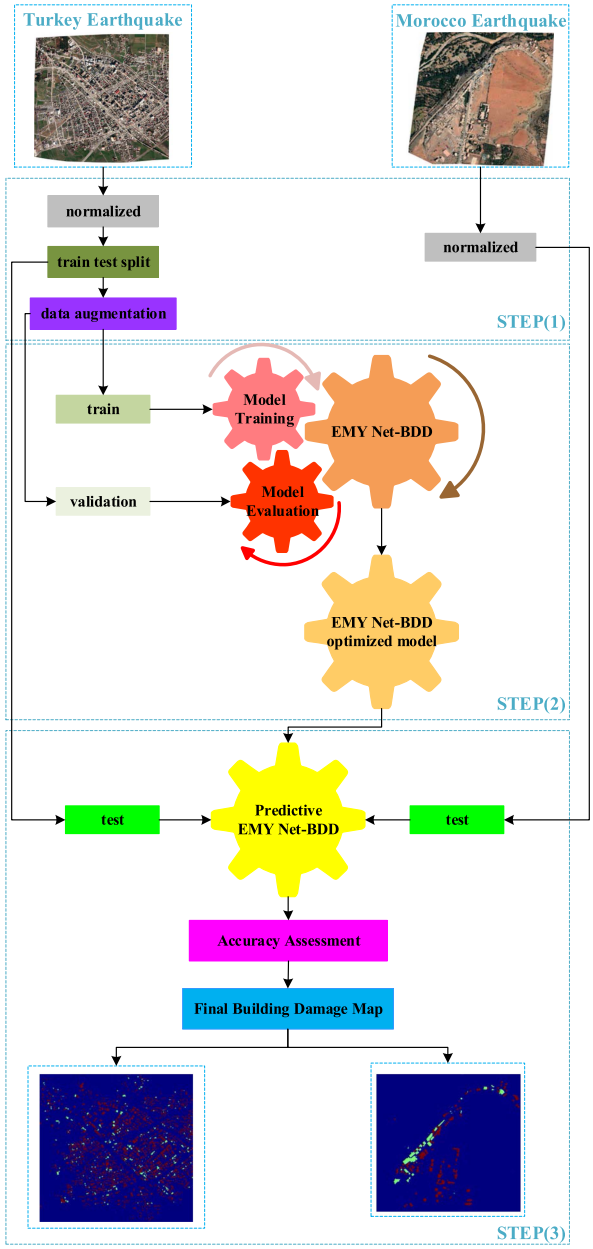



Fig. 4. Flowchart of the proposed method for BDD.

augmentation (rotation) is applied to the train–validation data, so the train–validation data are raised from 1749 to 3498 in the Turkey earthquake and from 836 to 3344 in the Libya flood. In the second step, the training and validation data are imported to the EMYNet-BDD network. The network is subsequently trained with the training data and evaluated through validation and, finally, the optimized model is attained. In the third step, the EMYNet-BDD network predicts the damaged buildings, the network accuracy is evaluated, and the map of the damaged buildings is generated. Fig. 4 depicts the proposed method flowchart. The EMYNet-BDD network initially is trained by Turkey training data and when the network is optimized, then it is used to predict the Turkey and Morocco earthquake building damage map. For the Libya flood, similarly, the EMYNet-BDD network is trained by Libya flood training data and the test data are predicted by this optimized network.

### A. *EfficientViTB Architecture Blocks*

The EfficientViTB network as a new family of the vision transformer models was introduced by Cai et al. [33], which is used for efficient high-resolution dense prediction. Its main core is a multiscale linear attention module that enables the network to have a global field and multiscale learning with hardware-efficient operation [33]. The multiscale linear attention module substitutes the rectified linear unit (ReLU) instead of softmax attention [33]. However, the ReLU linear attention has weaknesses in local information extraction and multiscale learning, so it is improved by convolutions and multiscale linear attention module [33]. In the ReLU linear attention, the similarity function is expressed as follows:

$$\text{Sim}\ (Q,K) = \text{ReLU}\,(Q)\,\text{ReLU}(K)^T. \tag{1}$$

Considering (1), the ReLU attention is then written according to the following equation:

$$O_i = \sum_{j=1}^{N} \frac{\text{ReLU}\,(Q_i)\,\text{ReLU}(K_j)^T}{\sum_{j=1}^{N}\text{ReLU}\,(Q_i)\,\text{ReLU}(K_j)^T},$$
$$V_j = \frac{\sum_{j=1}^{N}\left(\text{ReLU}\,(Q_i)\,\text{ReLU}(K_j)^T\right)V_j}{\text{ReLU}\,(Q_i)\sum_{i=1}^{N}\text{ReLU}(K_j)^T}. \tag{2}$$

Equation (2) can be rewritten to reduce the memory and computational complexity from quadratic to linear, as shown in the following equation:

$$\begin{aligned} O_i &= \frac{\sum_{j=1}^{N}\left[\text{ReLU}\,(Q_i)\,\text{ReLU}(K_j)^T\right]V_j}{\text{ReLU}\,(Q_i)\sum_{j=1}^{N}\text{ReLU}(K_j)^T} \\ &= \frac{\sum_{j=1}^{N}\text{ReLU}\,(Q_i)\left[\left(\text{ReLU}(K_j)^T V_j\right)\right]}{\text{ReLU}\,(Q_i)\sum_{j=1}^{N}\text{ReLU}(K_j)^T} \\ &= \frac{\text{ReLU}\,(Q_i)\left(\sum_{j=1}^{N}\text{ReLU}(K_j)^T V_j\right)}{\text{ReLU}\,(Q_i)\left(\sum_{j=1}^{N}\text{ReLU}(K_j)^T\right)}. \end{aligned} \tag{3}$$

Here, the terms $\left(\sum_{j=1}^{N}\text{ReLU}(K_j)^T V_j\right) \in \mathbb{R}^{d\times d}$ and $\left(\sum_{j=1}^{N}\text{ReLU}(K_j)^T\right) \in \mathbb{R}^{d\times 1}$ are independent from $i$ and they only need to be computed once so that they can be reused for each query, thereby computing $O_i$ has memory and computational complexity of $\mathcal{O}(N)$. Therefore, ReLU linear attention becomes more efficient on hardware.

Fig. 5 shows the EfficientViTB and multiscale linear attention blocks, which contain a multiscale linear module and a feedforward neural network with depthwise convolution (FFN+DWConv). The multiscale linear attention accordingly extracts context information, while the FFN+DWConv extracts local information. In the multiscale linear attention, a multiscale token is generated upon obtaining the $Q$, $K$, and $V$ tokens through the linear projection layer by aggregating the nearby

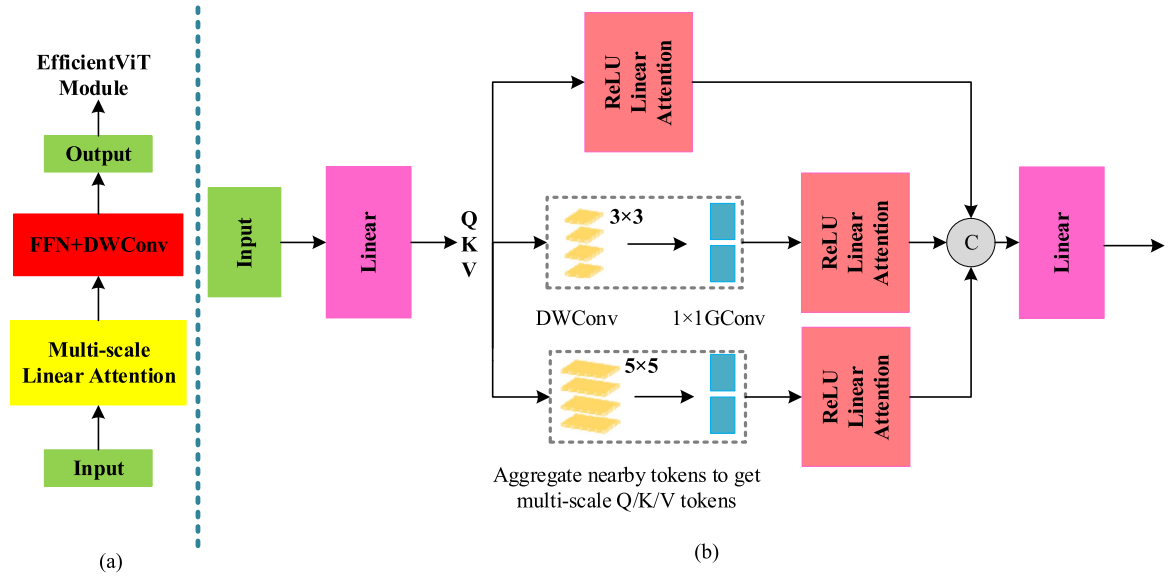



Fig. 5. (a) EfficientViTs building block. (b) Multiscale linear attention [33].

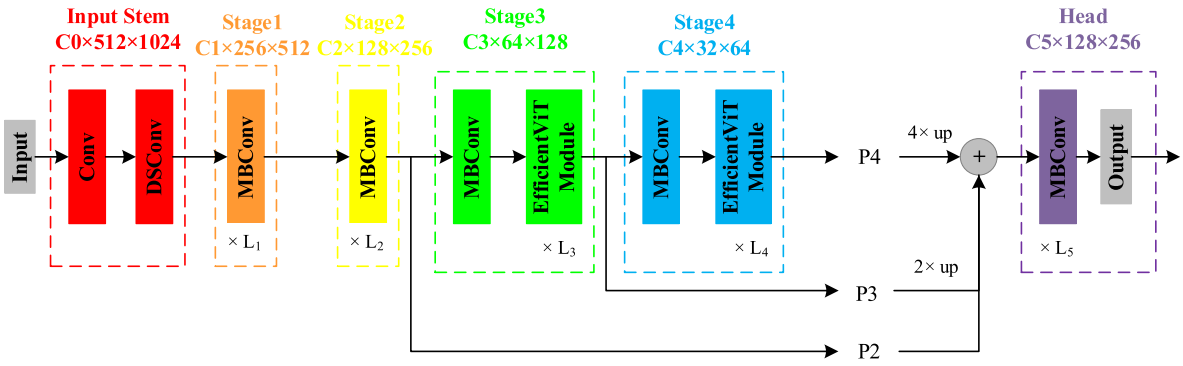



Fig. 6. Architecture of EfficientViTB [33].

tokens via the lightweight small-kernel convolution. The ReLU linear attention is also applied to the multiscale tokens and then the outputs are concatenated and presented to the final linear projection layer to merge the features [33].

The EfficientViTB network architecture is illustrated in Fig. 6, which is composed of a backbone/head encoder–decoder. The EfficientViT module is, thus, imported into stages 3 and 4 of the backbone. The features from stages 2, 3, and 4 are then given to the head component. In this architecture, the addition is used to integrate the desired features [33]. The head design is also simple, which includes several MBConv blocks and the output layer [33].

### *B. Yolov8 Architecture Blocks*

In the field of computer vision, the Yolo networks have so far had many accomplishments. The Yolov8 introduced by the Ultralytics on 10 January 2023 has higher detection accuracy and speed than previous Yolo family networks [34]. The Yolov8, as depicted in Fig. 7, comprises three parts, namely, the backbone, the neck, and the head [34], [35].

The Yolov8 network makes use of CSPDarknet-53 [36] as the backbone. The input data are downsampled five times to obtain five different feature pyramid levels, as shown in B1–B5 of Fig. 7 In the Yolov8 network, the CSP module in the original version of the backbone is replaced by the C2f module. The central battery system (CBS) module acts like a convolution operator on the input data and then it leads to batch normalization. This module activates the stream information using the sigmoid linear unit to show the output [34]. According to Fig. 7, the backbone ultimately utilizes the spatial pyramid pooling-fast (SPPF) module to pool the input feature maps into a fixed-size one to adapt the output size [34].

As depicted in Fig. 7, the Yolov8 network is designed in the neck part based on the path aggregation network-feature pyramid network (PAN-FPN) structure. In contrast to the Yolov7 [37] and the Yolov5 networks, the Yolov8 eliminates the convolution operator after upsampling in the PAN structure, which has the main network performance and achieves a lightweight model with higher speed [34]. Moreover, the Yolov8 network uses the decoupled head structure for detection purposes. This structure employs two separate branches for object classification and

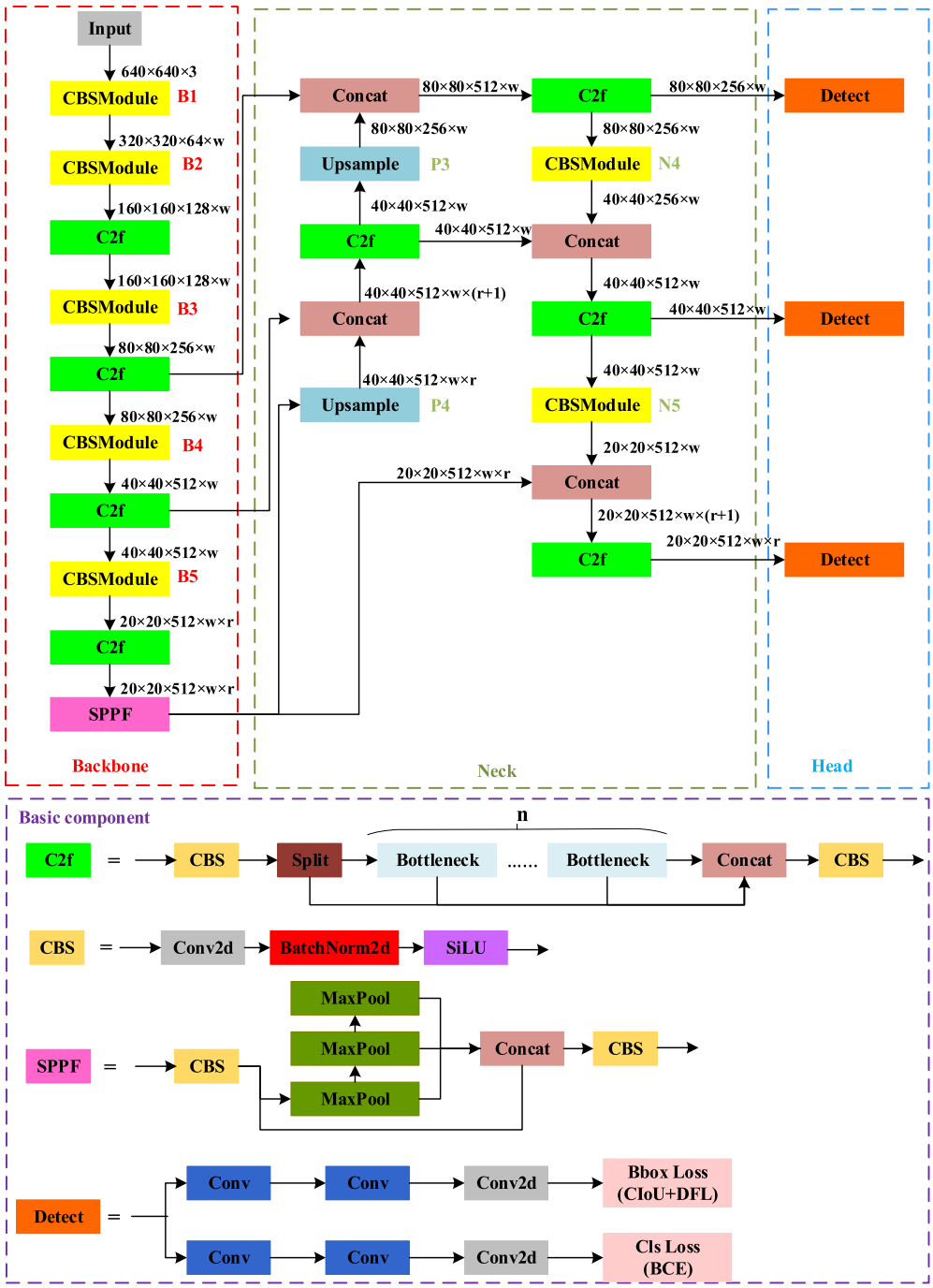
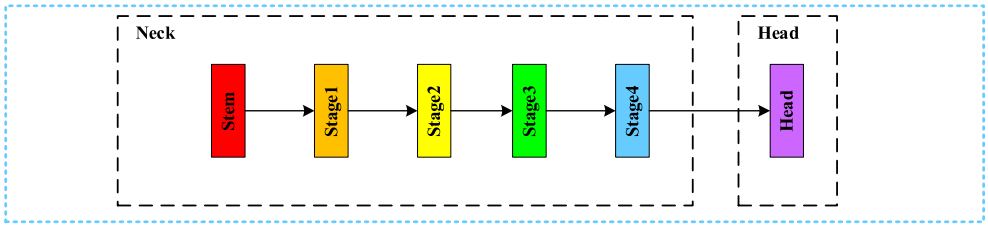



Fig. 7. Network architecture diagram of the Yolov8 consists of three parts: backbone, neck, and head. These three parts include four basic components of C2f, CBS, SPPF, and Detect [34].

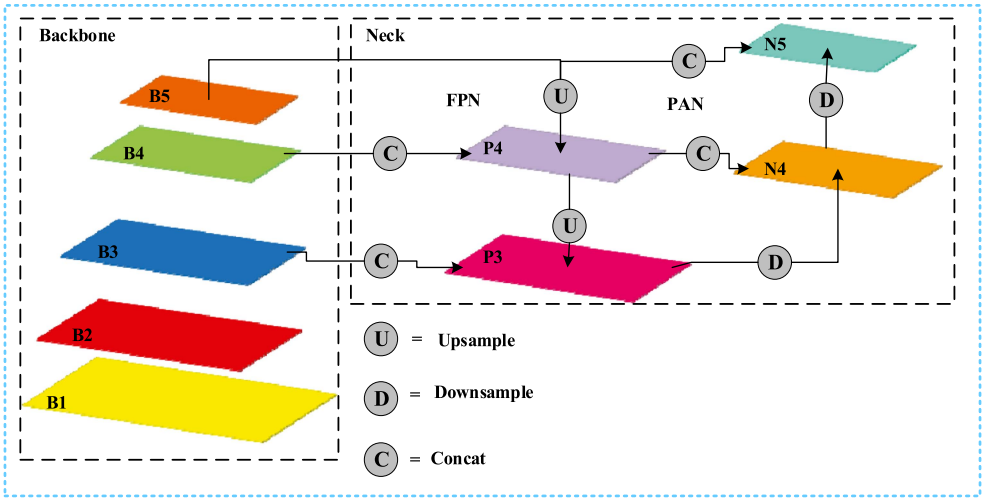
predicted bounding box regression and two different loss functions for both tasks [34], [35].

### C. *EfficientViTB Meets the Yolov8 (EMYNet-BDD) Framework Training*

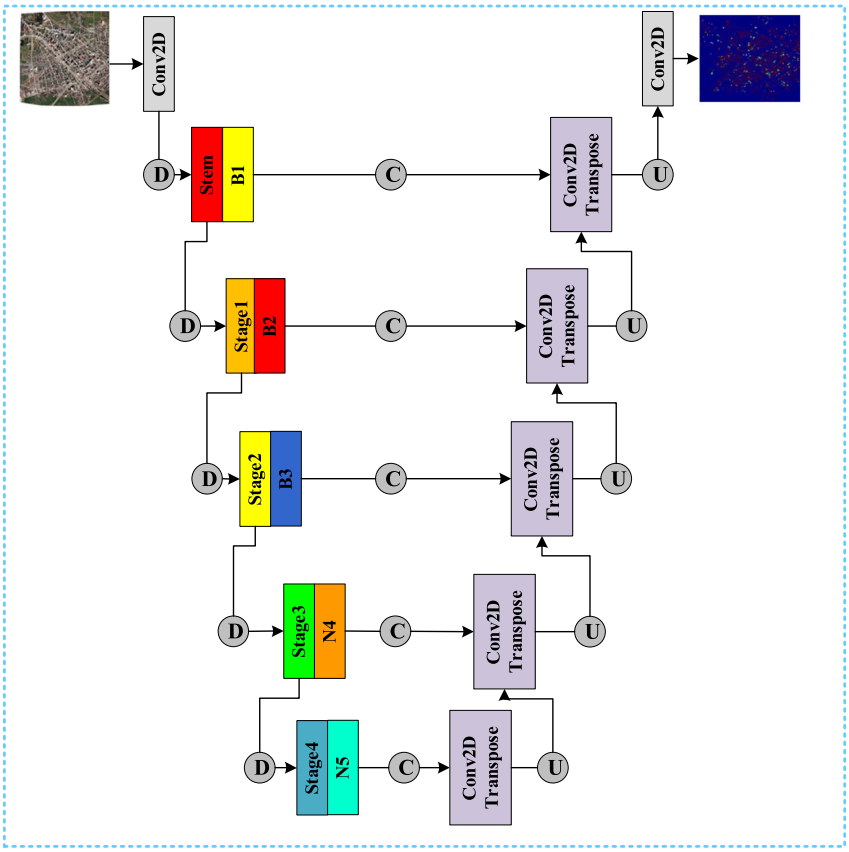
As mentioned earlier, the EMYNet-BDD in the encoder path is made up of the combination of the EfficientViTB and the Yolov8 networks, pretrained by the ImageNet and the MS-COCO datasets, respectively. Both networks are used in the encoder path, so the semitransfer learning technique [32] is implemented. In the decoder path, the improved-Unet blocks [38] are utilized. Compared with methods in which only one network is used, the EMYNet-BDD architecture uses two networks for the encoder path, which efficiently extracts features from the input data and reaches higher accuracy for BDD. As shown in Fig. 8, the EfficientViTB network includes one stem and four stages and the Yolov8 network has five backbone blocks (B1–B5) and two neck blocks (N4 and N5). The EfficientViTB and the Yolov8 network blocks form the encoder path in five




(a)




(b)




(c)

Fig. 8. Architecture of EMYNet-BDD. (a) Structure of the EfficientViTB network, which contains one stem and four stages (represented with a different color). (b) Schematic representation of PAN-FPN structure in the Yolov8 (each block of neck and backbone represented by a different color) [35]. (c) Structure of the proposed EMYNet-BDD network using the EfficientViTB and the Yolov8 as encoder and modified Unet as decoder. The encoder concatenates with the decoder at five different resolutions.

TABLE II
INFORMATION FORMULAE FOR ACCURACY ASSESSMENT METRICS

| Evaluation metrics | Formula |
|---|---|
| OA | $\frac{TP+TN}{TP+FN+TN+FP}$ |
| KC | $\frac{2\times(TP\times TN-FN\times FP)}{(TP+FP)\times(FP+TN)+(TP+FN)\times(FN+TN)}$ |
| Precision | $\frac{TP}{TP+FP}$ |
| IOU | $\frac{TP}{TP+FP+FN}$ |
| $F1$-score | $\frac{2\times TP}{(2\times TP)+FP+FN}$ |

TABLE III
PARAMETERS USED FOR TRAINING EMYNET-BDD

| Parameters | Value |
|---|---|
| Pretrained (EfficientViTB) | ImageNet dataset |
| Pretrained (Yolov8) | MS COCO dataset |
| $\gamma$ (Focal loss hyperparameters) | 2 |
| $\alpha_i$(Focal loss hyperparameters) | 0.25 |
| Number of epochs | 80 |
| optimizer | Adam |
| Learning rate | $1\times10^{-4}$ |
| Learning decay | $1\times10^{-6}$ |

levels of the pyramid feature and five different sizes are then concatenated with the Conv2D transpose blocks to produce the EMYNet-BDD.

### D. Accuracy Assessment

As shown in Table II, five metrics are introduced in this study to evaluate the performance of the proposed method. TP denotes the true positive (number of labels predicted to be damaged that are actually damaged), FP indicates the false positive (number of labels predicted to be damaged that are actually nondamaged), FN denotes the false negative (number of labels predicted to be nondamaged that are actually damaged), and TN indicates the true negative (number of labels predicted to be nondamaged that are actually nondamaged).

## IV. EXPERIMENTAL RESULTS

The results and the performance of the proposed method and the comparison with other methods are presented in this section. This section includes four parts, namely, experimental parameter settings, accuracy assessment, comparison of experimental results, and ablation study.

### A. Experimental Parameter Settings

All DL networks in this study are implemented using TensorFlow 2.10.0 and Python 3.8.8 through a system with an NVIDIA GeForce GTX 1050Ti GPU, a 3.5 GHz CPU, and 32 GB RAM. The data are partitioned into patches with the dimensions of $128 \times 128$ and the focal loss function is employed. Considering the multilabel building damage data patches, multiple focal loss functions are utilized according to the following equation [39]:

$$L_{\text{focal loss}} = -\sum_{i=1}^{c} \alpha_i (1-y_i)^{\gamma} t_i \log(y_i) \tag{4}$$

in which $y_i$ represents the predicted probability distribution, $t_i$ denotes the true probability distribution, $C$ is the number of classes, and $\gamma$ and $\alpha_i$ show the hyperparameters of the focal loss function. The performance of Adam and SGD optimizers was computed and Adam had a proper result. The learning rate of the Adam optimizer tested for values of $\{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$ and $10^{-4}$ had better performance. The values of 0.25 and class weights were assumed for $\alpha$ so that 0.25 was adequate. Also, the default value 2 for $\gamma$ had appropriate results. Table III lists other parameters for the EMYNet-BDD network training.

### B. Comparison of Experimental Results

The performance of the EMYNet-BDD network is compared with other DL methods in Table IV. With the limited computer capacity, versions of networks that can be executed in the available system are taken into account. In addition, the dimensions of the input images are considered $128 \times 128$.

According to Table IV, the EMYNet-BDD network can reach an OA of 97.63% and a KC of 0.86 for the 2023 Turkey earthquake dataset, indicating its better performance than other networks. In terms of other parameters, such as precision, intersection over union (IoU), and the $F1$-score, the proposed network reaches 99.93%, 98.99%, and 98.81%, respectively. Compared with the Yolox network, which is utilized commonly in RS-based tasks, the Yolov8 network shows better performance in the proposed method. Although the TinyViT network is lightweight and has fewer parameters due to its structure, its performance is not as proper as the proposed method because it cannot handle the complexities of BDD. Moreover, the performance of EMYNet-BDD is compared with the ResidualUnet, the VGG16, and the modified DamFormer that were used in the previous studies of BDD. The residual blocks are used in the ChangeOS [22] network and the DamFormer [23] is based on the SegFormer [40] network blocks, both of these two networks have the Siamese architecture and have proper performance in BDD. Also, the VGG16 was applied for BDD methods with only postevent data [13], [40]. It should be considered that the parameters and architecture of networks that are applied for comparison are different from the original ones.

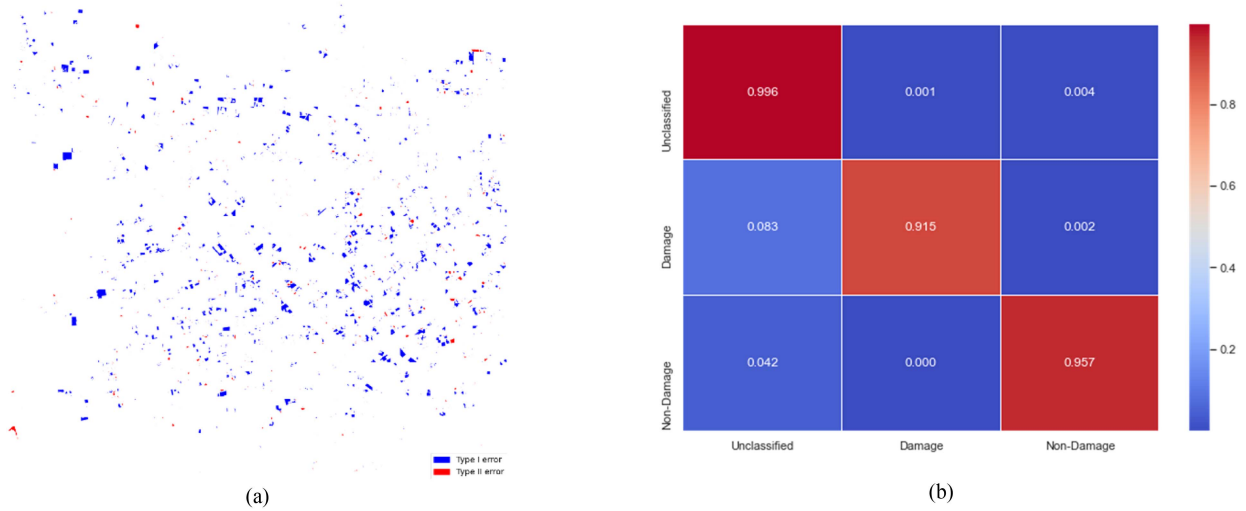
To shed light on the performance of the proposed method for BDD in various disasters, it is applied to the 2023 Morocco earthquake and the 2023 Libya flood datasets. The EMYNet-BDD network is applied to the 2023 Morocco earthquake dataset after training on the 2023 Turkey earthquake. According to Table IV, the OA and KC of this network are, respectively, 98.63% and 0.85 for the 2023 Morocco earthquake dataset, which shows its proper performance.

As floods are among the most common natural disasters across the world and generally cause damage to buildings, the EMYNet-BDD network is utilized to detect the damages after the 2023 Libya flood. According to Table IV, this network achieves the OA of 96.43% and the KC of 0.84.

TABLE IV
QUANTITATIVE EVALUATION OF THE RESULTS

| Dataset | Method | OA (%) | KC | Precision (%) | IOU (%) | F1-score (%) | Parameters (millions) |
|---|---|---|---|---|---|---|---|
| Turkey earthquake | EffformerV2 | 95.87 | 0.81 | 99.66 | 98.74 | 99.36 | 26.1 |
| | EfficientNet | 96.00 | 0.81 | 99.63 | 98.73 | 99.36 | 8.8 |
| | TinyViT | 96.16 | 0.82 | 99.73 | 98.83 | 98.84 | 5.3 |
| | Yolox | 95.52 | 0.79 | 99.90 | 98.71 | 99.35 | 4.8 |
| | Residual Unet | 74.03 | 0.19 | 78.90 | 72.32 | 86.40 | 4.7 |
| | VGG16 | 73.05 | 0.18 | 79.43 | 71.45 | 84.66 | 15.3 |
| | Modified DamFormer | 95.34 | 0.79 | 99.61 | 98.66 | 99.31 | 3.7 |
| | EMYNet-BDD | **97.62** | **0.86** | **99.93** | **98.99** | **99.49** | 21.1 |
| Morocco earthquake | EMYNet-BDD | 98.63 | 0.85 | 99.59 | 98.64 | 98.81 | 21.1 |
| Libya flood | EMYNet-BDD | 96.43 | 0.84 | 99.59 | 98.71 | 98.81 | 21.1 |

The bold values are the best performance of the proposed method.




(a) (b)

Fig. 9. Turkey earthquake results. (a) Type-I and -II errors. (b) Confusion matrix.

As an example, Fig. 9 shows the type-I and -II errors of the Turkey earthquake in addition to its related confusion matrix. As only postevent data are exploited in our study, completely damaged buildings are challenging to detect as a building even with human vision. The confusion matrix shows proper performance in detecting pixels of each class.

### C. Ablation Study

In the ablation study, the existence and nonexistence of some parts of the proposed method and their effects on the results are discussed. Since the proposed method is based on an encoder path with two networks, the encoder–decoder architecture is supposed in two models, S#1 and S#2. The S#1 model is designed without the Yolov8 network and only the EfficientViTB network is considered. The S#2 model is designed without the EfficienViTB network and the Yolov8 network is merely used. The results of both models and the main EMYNet-BDD network are presented in Table V.

TABLE V
ABLATION STUDY OF THE PROPOSED METHOD FOR THE TURKEY EARTHQUAKE DATASET

| Method | OA (%) | KC | Precision (%) | IOU (%) | F1-score (%) | Parameters (millions) |
|---|---|---|---|---|---|---|
| S#1 | 96.45 | 0.80 | 99.67 | 98.73 | 99.37 | 14.1 |
| S#2 | 96.93 | 0.81 | 99.89 | 98.75 | 99.38 | 10.1 |
| proposed method | 97.62 | 0.86 | 99.93 | 98.99 | 99.49 | 21.1 |

As outlined in Table V, the overall performance of the given method decreases when any of the networks are eliminated.

## V. DISCUSSION

Quantitative and qualitative evaluation, as well as visual analysis investigated the performance of the proposed method for BDD. Fig. 10 displays the proposed method results in comparison with other methods in Table IV. Accordingly, the damaged buildings are in green and those that are nondamaged are in crimson.

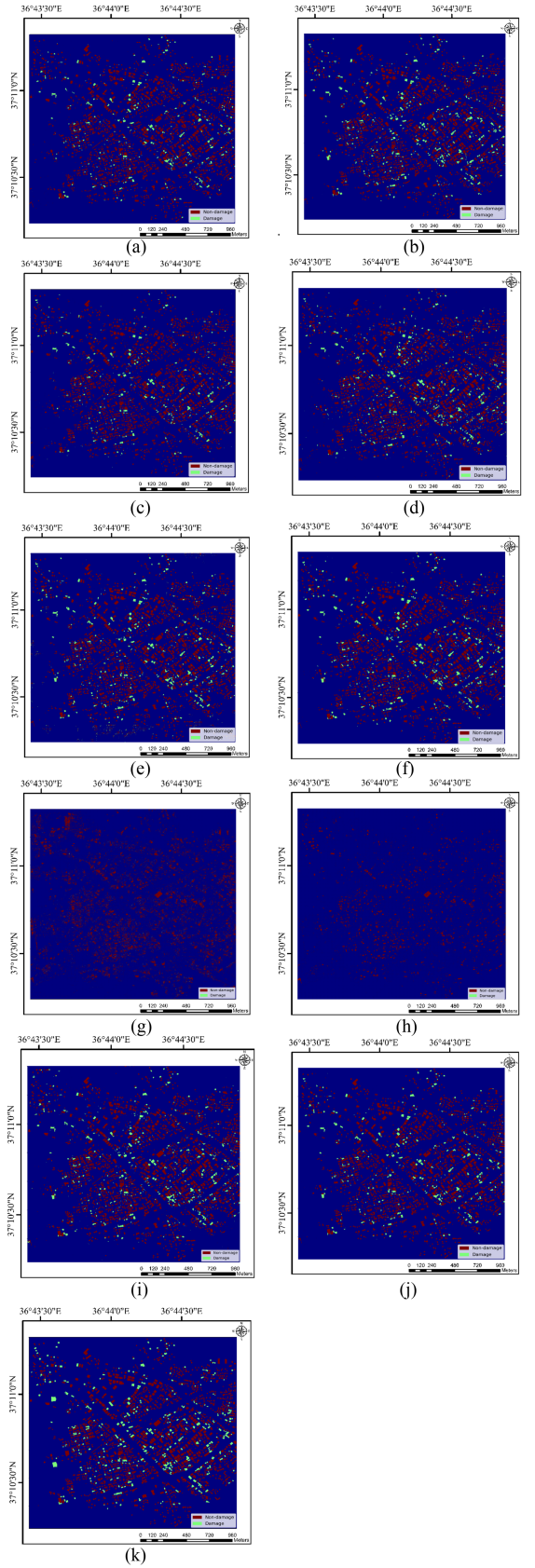



Fig. 10. Comparison between the proposed method and other networks. (a) EffformerV2. (b) TinyViT. (c) EfficientnetV2. (d) Yolox. (e) Yolov8. (f) EfficientViTB. (g) Residual Unet. (h) VGG16. (i) Modified DamFormer. (j) Proposed method. (k) GT.

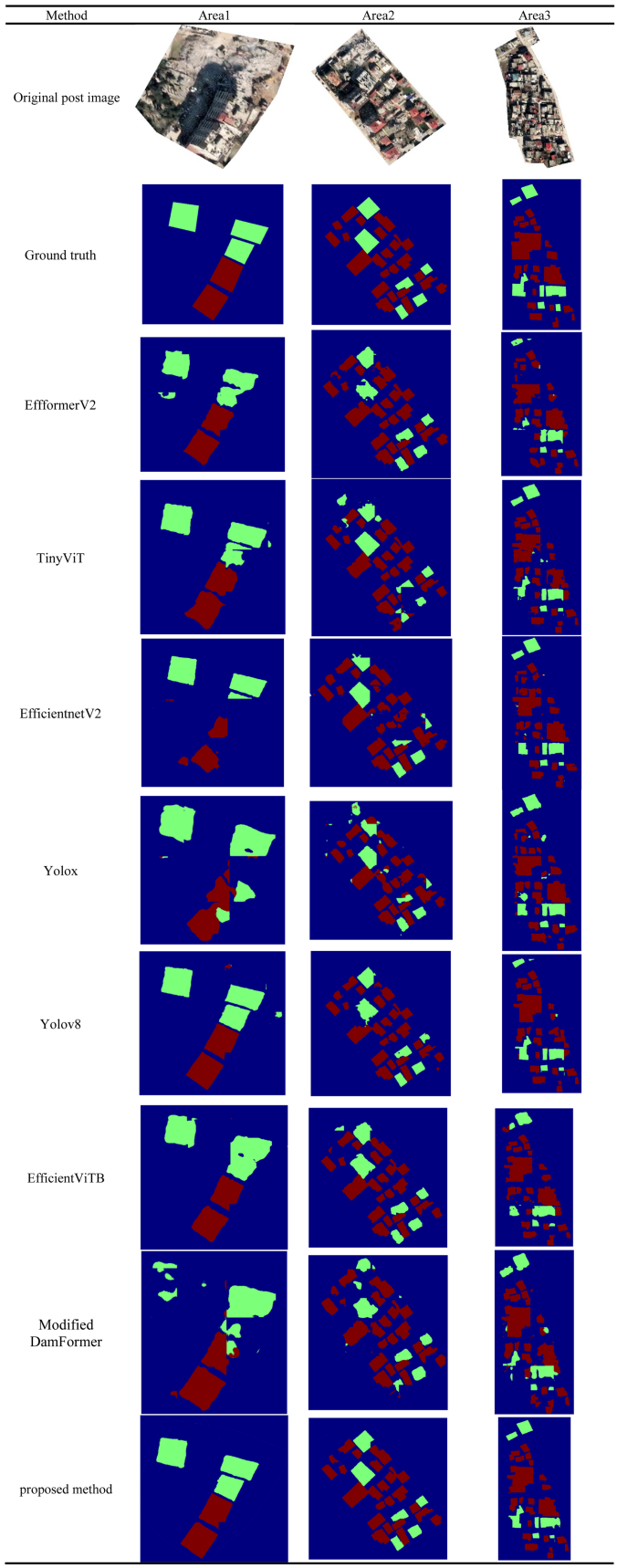



Fig. 11. Comparison between the result of the proposed method and other networks.

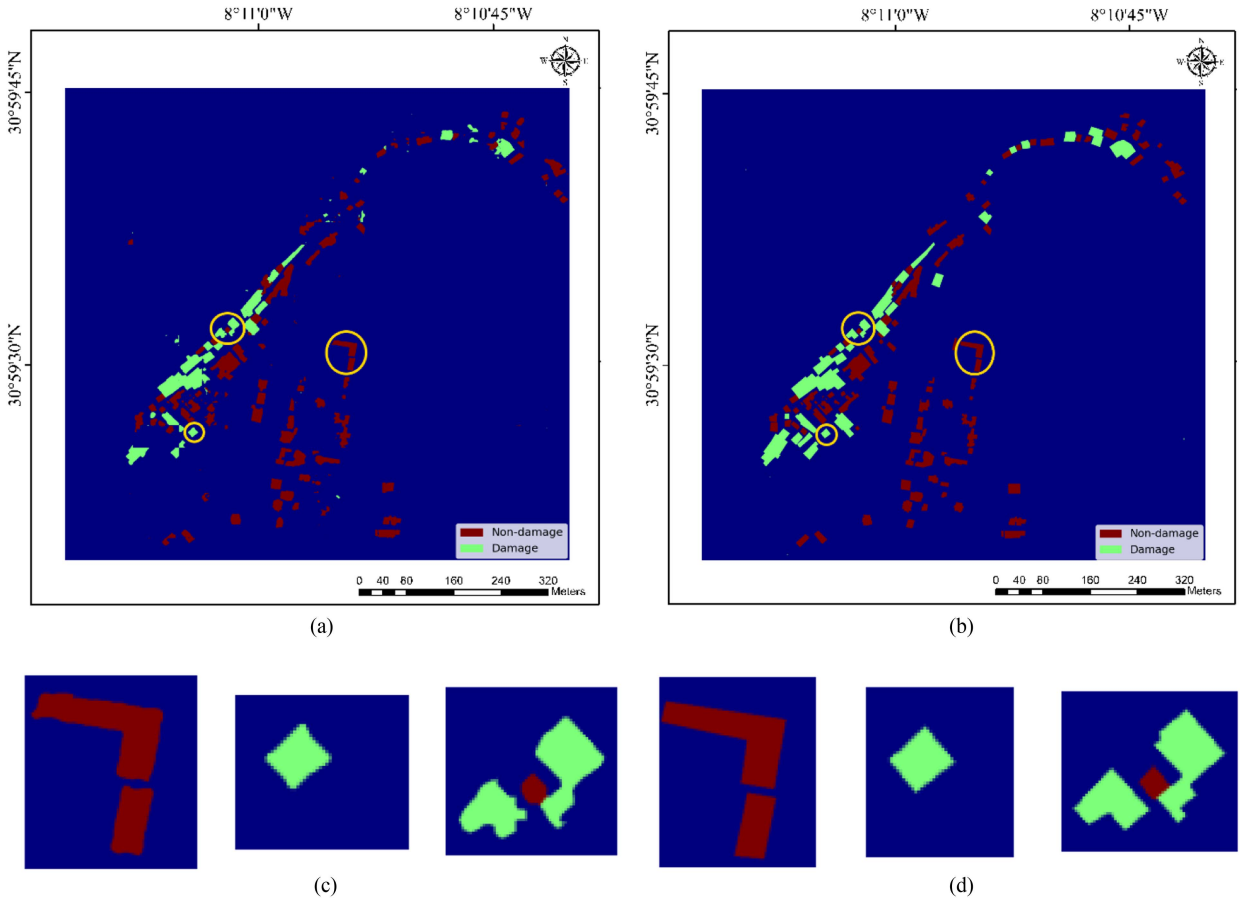



Fig. 12. Proposed method results for the Morocco earthquake dataset. (a) Proposed method. (b) GT. (c) Three example areas with different shape markets with yellow circles. (d) Corresponding GT of three examples.

As shown in Fig. 10, the number of damaged buildings is lower than the nondamaged ones in the 2023 Turkey earthquake dataset, so it is unbalanced. The proposed method can overcome this problem and separate the damaged buildings from the nondamaged ones. Furthermore, one of the challenges in this method is the lack of pre-event data, so it is not easy to detect the boundary of the buildings that are completely damaged, but the proposed method copes with it to a large extent and achieves favorable results for BDD. The EffformerV2 and the TinyViT methods [see Fig. 10(a) and (b)] show overdetect damaged buildings and they have noise in some places. In addition, the EfficientnetV2 network [see Fig. 10(c)] fails to detect the buildings in some places and the results of this method are lower than those in the GT. Moreover, the Yolox network [see Fig. 10(d)] fails to detect the boundary of the buildings correctly, especially in places with a high density of buildings and also it misses some damaged buildings. The Yolov8 and the EfficientViTB networks [see Fig. 10(e) and (f)], as investigated in the ablation study, cannot detect the building boundaries in some places. The Residual Unet and VGG16 [see Fig. 10(g) and (h)] cannot detect the damaged building. The Modified DamFormer [see Fig. 10(i)] fails to detect some building boundaries.

To better examine the performance of the proposed method, three areas from the 2023 Turkey earthquake dataset with different complexities are considered in Fig. 11. The first area has a sparse urban region with towers that are completely damaged. The proposed network can accurately distinguish the damaged buildings from the nondamaged ones. The Yolox and the modified DamFormer networks wrongly classified some nondamaged towers into damaged classes and failed to detect building boundaries accurately. The EfficientnetV2 network also misses one damaged building and cannot reconstruct its boundary. The EfficientViTB network considers two damaged towers as one building and could not reconstruct their boundaries correctly. Nonetheless, the proposed network accurately distinguishes the damaged buildings from the nondamaged ones and reconstructs the boundary of buildings accurately.

The second area has a very dense urban region that contains buildings of different shapes near each other. The Yolox, the TinyViT, and the modified DamFormer networks miss some damaged buildings, and the boundaries of the nondamaged ones are not properly detected and some buildings are detected wrongly or have noise. The Yolov8 network fails to recognize some building boundaries correctly. Although the EfficientViTB network provides relatively good results, it excessively detects some buildings that are not damaged. The EfficientnetV2 network misses some damaged buildings and cannot reconstruct the boundaries of the nondamaged ones correctly. It also performs poorly in boundary reconstruction. On the other hand, the proposed network can cope with the complexities of this

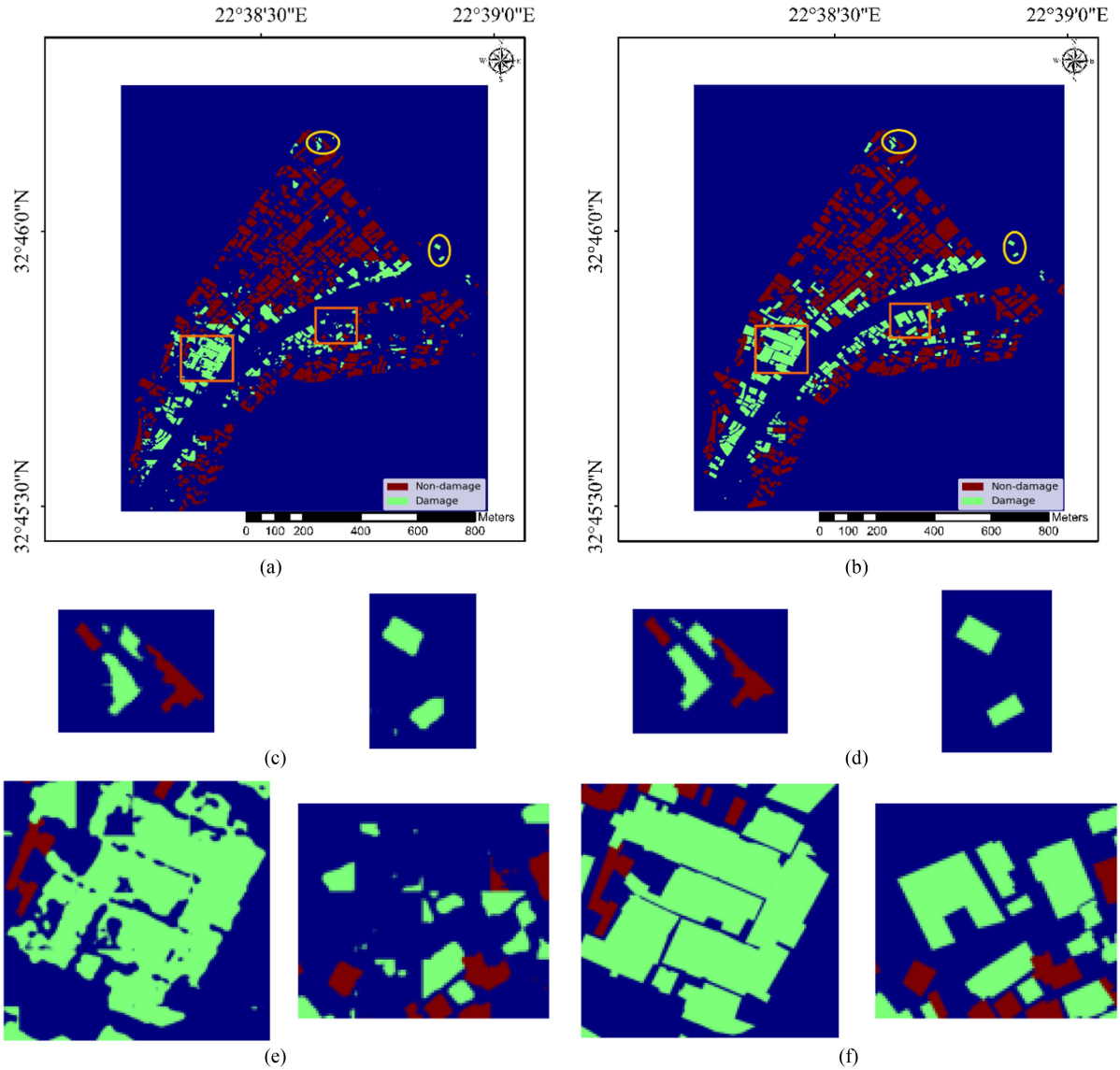



Fig. 13. Proposed method results for Libya flood dataset. (a) Proposed method. (b) GT. (c) Two example areas with different shapes and small damage market with yellow ellipsoid. (d) Corresponding GT of two examples. (e) Two examples of highly damaged areas marked with orange rectangle. (f) Corresponding GT of two examples.

area and distinguish the damaged and nondamaged buildings accurately.

The third area also has a very dense urban region and buildings with complex boundaries that are difficult to distinguish. The EffformerV2 network misses some damaged buildings and cannot recognize the complex boundaries. The EfficientnetV2 network detects noises in some places and does not detect the boundaries correctly. Also, the EfficientViTB and the modified DamFormer networks wrongly detect some places as the damaged ones, although they are not so. Additionally, it cannot reconstruct the boundaries properly. The TinyViT network detects some noise areas, but it performs relatively better than other methods. The Yolov8 network is better at reconstructing the boundary of buildings as compared with other methods, but it has errors in the detection of some places and misses one of the damaged buildings. The Yolox network accordingly cannot cope with the complexity of the boundaries and does not detect them correctly. It also has errors in distinguishing the damaged and nondamaged buildings. The proposed method can handle the density of the buildings and the complexity of their boundaries in this area.

The second dataset used in this study is the 2023 Morocco earthquake. The proposed network is applied to it after training on the 2023 Turkey earthquake dataset. According to Fig. 12, the network performs properly in this dataset and correctly distinguishes the damaged and nondamaged buildings. In Fig. 12, three areas are considered as the examples and the corresponding GT [see Fig. 12(c) and (d)]. These areas show that the proposed network can correctly distinguish the damaged buildings from the nondamaged ones and reconstruct their boundaries.

The third dataset exploited in this study is the 2023 Libya flood. The results of this dataset examined by the proposed

network and its corresponding GT are displayed in Fig. 13. The proposed network can reach a relatively suitable accuracy to detect damaged buildings. Two examples of the small buildings that were damaged and detected by the proposed method are marked with a yellow ellipsoid. However, the number and density of the buildings damaged due to flood, marked with orange rectangles, are abundant in some areas, so the proposed method cannot detect them accurately.

## VI. Conclusion

A new encoder–decoder architecture based on the EfficientViTB and the Yolov8 networks is presented in this study for BDD. Both networks are pretrained and used in the encoder path, which provides better performance compared with the methods in which only one network is used in their encoder path. Three datasets affected by different natural disasters are taken into account in this study. The first dataset is related to the WorldView II images of the 2023 Turkey earthquake, the second dataset is associated with the WorldView II images of the 2023 Morocco earthquake, and the third dataset contains the GeoEye-I images of the 2023 Libya flood. One of the advantages of this study is only the postevent data that are used, compared with the methods that utilize both pre- and postevent data. The proposed method also reaches the OA of 97.62%, 98.63%, and 96.43% for the three datasets, respectively. It also has proper performance in distinguishing the damaged and nondamaged buildings.

According to the high performance of this method, it is suggested to apply it to BDD after other natural disasters, such as hurricanes and fires. Moreover, it is suggested to utilize the proposed method for BDD on other types of RS data.

## Data Availability

Data will be made available on request.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

## Acknowledgment

The authors would like to thank Maxar Technologies, Inc., for providing satellite images.

## References

[1] Z. Zhou, J. Gong, and X. Hu, "Community-scale multi-level post-hurricane damage assessment of residential buildings using multi-temporal airborne LiDAR data," *Autom. Construction*, vol. 98, pp. 30–45, 2019.

[2] D. Hong et al., "SpectralGPT: Spectral remote sensing foundation model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 8, pp. 5227–5244, Aug. 2024.

[3] D. Hong et al., "Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks," *Remote Sens. Environ.*, vol. 299, 2023, Art. no. 113856.

[4] P. Ge, H. Gokon, and K. Meguro, "A review on synthetic aperture radar-based building damage assessment in disasters," *Remote Sens. Environ.*, vol. 240, 2020, Art. no. 111693.

[5] E. Ferrentino, F. Nunziata, M. Migliaccio, and A. Vicari, "A sensitivity analysis of dual-polarization features to damage due to the 2016 central-Italy earthquake," *Int. J. Remote Sens.*, vol. 39, no. 20, pp. 6846–6863, 2018.

[6] Q. Li, L. Gong, and J. Zhang, "A correlation change detection method integrating PCA and multi-texture features of SAR image for building damage detection," *Eur. J. Remote Sens.*, vol. 52, no. 1, pp. 435–447, 2019.

[7] T. ElGharbawi and F. Zarzoura, "Damage detection using SAR coherence statistical analysis, application to Beirut, Lebanon," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 1–9, 2021.

[8] S. T. Seydi, M. Hasanlou, J. Chanussot, and P. Ghamisi, "BDD-Net+: A building damage detection framework based on modified coat-net," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 4232–4247, Apr. 2023.

[9] M. Janalipour and A. Mohammadzadeh, "Evaluation of effectiveness of three fuzzy systems and three texture extraction methods for building damage detection from post-event LiDAR data," *Int. J. Digit. Earth*, vol. 11, no. 12, pp. 1241–1268, 2018.

[10] C. Axel and J. van Aardt, "Building damage assessment using airborne LiDAR," *J. Appl. Remote Sens.*, vol. 11, no. 4, 2017, Art. no. 046024.

[11] S. T. Seydi, M. Hasanlou, J. Chanussot, and P. Ghamisi, "Leveraging involution and convolution in an explainable building damage detection framework," *Eur. J. Remote Sens.*, vol. 56, 2023, Art. no. 2252166.

[12] G. S. Merlin and G. W. Jiji, "Building damage detection of the 2004 Nagapattinam, India, Tsunami using the texture and spectral features from IKONOS images," *J. Indian Soc. Remote Sens.*, vol. 47, no. 1, pp. 13–24, 2019.

[13] M. Ji, L. Liu, R. Zhang, and M. F. Buchroithner, "Discrimination of earthquake-induced building destruction from space using a pretrained CNN model," *Appl. Sci.*, vol. 10, no. 2, 2020, Art. no. 602.

[14] Y. Li, W. Hu, H. Dong, and X. Zhang, "Building damage detection from post-event aerial imagery using single shot multibox detector," *Appl. Sci.*, vol. 9, no. 6, 2019, Art. no. 1128.

[15] H. Chen, J. Song, C. Wu, B. Du, and N. Yokoya, "Exchange means change: An unsupervised single-temporal change detection framework based on intra- and inter-image patch exchange," *ISPRS J. Photogramm. Remote Sens.*, vol. 206, pp. 87–105, 2023.

[16] F. Zhou, C. Xu, R. Hang, R. Zhang, and Q. Liu, "Mining joint intraimage and interimage context for remote sensing change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, May 2023, Art. no. 4403712, Accessed on: Jun. 17, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10123995/

[17] R. Hang, S. Xu, P. Yuan, and Q. Liu, "AANet: An ambiguity-aware network for remote-sensing image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, Feb. 2024, Art. no. 5612911, Accessed on: Jun. 17, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10454002/

[18] M. Ji, L. Liu, R. Du, and M. F. Buchroithner, "A comparative study of texture and convolutional neural network features for detecting collapsed buildings after earthquakes using pre- and post-event satellite imagery," *Remote Sens.*, vol. 11, no. 10, 2019, Art. no. 1202.

[19] R. Hang, P. Yang, F. Zhou, and Q. Liu, "Multiscale progressive segmentation network for high-resolution remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Sep. 2022, Art. no. 5412012.

[20] C. Wu et al., "Building damage detection using U-Net with attention mechanism from pre-and post-disaster remote sensing datasets," *Remote Sens.*, vol. 13, no. 5, 2021, Art. no. 905.

[21] Y. Shen et al., "BDAnet: Multiscale convolutional neural network with cross-directional attention for building damage assessment from satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, May 2021, Art. no. 5402114.

[22] Z. Zheng, Y. Zhong, J. Wang, A. Ma, and L. Zhang, "Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters," *Remote Sens. Environ.*, vol. 265, 2021, Art. no. 112636.

[23] H. Chen, E. Nemni, S. Vallecorsa, X. Li, C. Wu, and L. Bromley, "Dual-tasks Siamese transformer framework for building damage assessment," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 1600–1603.

[24] R. Virtriana et al., "Machine learning remote sensing using the random forest classifier to detect the building damage caused by the Anak Krakatau Volcano tsunami," *Geomatics, Natural Hazards Risk*, vol. 14, no. 1, pp. 28–51, 2023.

[25] R. Gupta and M. Shah, "RescueNet: Joint building segmentation and damage assessment from satellite imagery," in *Proc. 25th Int. Conf. Pattern Recognit.*, 2021, pp. 4405–4411.

[26] H. Chen, N. Yokoya, C. Wu, and B. Du, "Unsupervised multimodal change detection based on structural relationship graph representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Dec. 2022, Art. no. 5635318.
[27] H. Chen, N. Yokoya, and M. Chini, "Fourier domain structural relationship analysis for unsupervised multimodal change detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 198, pp. 99–114, 2023.
[28] M. Li et al., "Post-earthquake assessment of building damage degree using LiDAR data and imagery," *Sci. China Ser. E, Technol. Sci.*, vol. 51, pp. 133–143, 2008.
[29] B. Adriano, J. Xia, G. Baier, N. Yokoya, and S. Koshimura, "Multi-source data fusion based on ensemble learning for rapid building damage mapping during the 2018 Sulawesi earthquake and tsunami in Palu, Indonesia," *Remote Sens.*, vol. 11, no. 7, 2019, Art. no. 886.
[30] E. Khankeshizadeh et al., "A novel weighted ensemble transferred U-Net based model (WETUM) for postearthquake building damage assessment from UAV data: A comparison of deep learning- and machine learning-based approaches," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, Jan. 2024, Art. no. 4701317.
[31] M. Janalipour and A. Mohammadzadeh, "A novel and automatic framework for producing building damage map using post-event LiDAR data," *Int. J. Disaster Risk Reduction*, vol. 39, 2019, Art. no. 101238.
[32] M. Gomroki, M. Hasanlou, and P. Reinartz, "STCD-EffV2T Unet: Semi transfer learning EfficientNetV2 T-Unet network for urban/land cover change detection using Sentinel-2 satellite images," *Remote Sens.*, vol. 15, no. 5, 2023, Art. no. 1232.
[33] H. Cai, J. Li, M. Hu, C. Gan, and S. Han, "EfficientViT: Multi-scale linear attention for high-resolution dense prediction," 2024. [Online]. Available: https://arxiv.org/abs/2205.14756
[34] G. Wang, Y. Chen, P. An, H. Hong, J. Hu, and T. Huang, "UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios," *Sensors*, vol. 23, no. 16, 2023, Art. no. 7190.
[35] X. Wang, H. Gao, Z. Jia, and Z. Li, "BL-YOLOv8: An improved road defect detection model based on YOLOv8," *Sensors*, vol. 23, no. 20, 2023, Art. no. 8361.
[36] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," 2018. [Online]. Available: https://arxiv.org/abs/1804.02767
[37] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," 2022. [Online]. Available: https://arxiv.org/abs/2207.02696
[38] M. Gomroki, M. Hasanlou, and J. Chanussot, "Automatic 3D multiple building change detection model based on encoder-decoder network using highly unbalanced remote sensing datasets," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 10311–10325, Oct. 2023.
[39] W. Liu, L. Chen, and Y. Chen, "Age classification using convolutional neural networks with the multi-class focal loss," in *Proc. IOP Conf. Ser., Mater. Sci. Eng.*, 2018, Art. no. 012043.
[40] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "SegFormer: Simple and efficient design for semantic segmentation with transformers," in *Proc. 35th Conf. Neural Inf. Process. Syst.*, 2021, vol. 34, pp. 12077–12090.
[41] A. Calantropio, F. Chiabrando, M. Codastefano, and E. Bourke, "Deep learning for automatic building damage assessment: Application in post-disaster scenarios using UAV data," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 1, pp. 113–120, 2021.

**Masoomeh Gomroki** (Member, IEEE) received the B.Eng. degree in geomatics engineering and the M.Eng. degree in photogrammetry from the University of Isfahan, Isfahan, Iran, in 2012 and 2015, respectively, and the Ph.D. degree in photogrammetry and remote sensing from the University of Tehran, Tehran, Iran, in 2024.

Her research interests include the application of deep learning in remote sensing, change detection, hazard detection, medical image processing, language modeling, and LiDAR-based applications for forest analysis.

**Mahdi Hasanlou** (Member, IEEE) received the B.Sc. degree in surveying and geomatics engineering, the M.Sc. degree in remote sensing, and the Ph.D. degree in remote sensing from the University of Tehran, Tehran, Iran, in 2003, 2006, and 2013, respectively.

Since 2013, he has been an Associate Professor with the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, where he is the Head of the Remote Sensing Laboratory. He is currently the Dean of the School of Surveying and Geospatial Engineering. His research interests include change detection, damage detection, hyperspectral, thermal, optical, and SAR remote sensing for urban and agro-environmental applications.

**Jocelyn Chanussot** (Fellow, IEEE) received the M.Sc. degree from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from the Universit de Savoie, Annecy, France, in 1998, both in electrical engineering.

Since 1999, he has been with Grenoble INP, where he is currently a Professor of signal and image processing. He was a Visiting Scholar with Stanford University, Stanford, CA, USA, KTH, Stockholm, Sweden, and NUS, Singapore. Since 2013, he has been an Adjunct Professor with the University of Iceland, Reykjavik, Iceland, and Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. From 2015 to 2017, he was a Visiting Professor with the University of California, Los Angeles, Los Angeles, CA. His research interests include image analysis, hyperspectral remote sensing, data fusion, machine learning, and artificial intelligence.

Dr. Chanussot is the AXA Chair in remote sensing of Aerospace Information Research Institute, Chinese Academy of Sciences. He was the Founding President of the IEEE Geoscience and Remote Sensing French Chapter from 2007 to 2010, which received the 2010 IEEE GRS-S Chapter Excellence Award. He was a recipient of multiple outstanding paper awards. He was the Vice-President of the IEEE Geoscience and Remote Sensing Society, in charge of meetings and symposia from 2017 to 2019. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote Sensing. He was the Chair, from 2009 to 2011, and the Co-Chair of GRS Data Fusion Technical Committee from 2005 to 2008. He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society from 2006 to 2008 and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing in 2009. He is an Associate Editor for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, and *Proceedings of the IEEE*. He was an Editor-in-Chief for the *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* from 2011 to 2015. In 2014, he was a Guest Editor for the *IEEE Signal Processing Magazine*. He was a member of the Institut Universitaire de France from 2012 to 2017 and a Highly Cited Researcher (Clarivate Analytics/Thomson Reuters).

**Danfeng Hong** (Senior Member, IEEE) received the Dr.-Ing degree (summa cum laude) in remote sensing from the Signal Processing in Earth Observation (SiPEO), Technical University of Munich (TUM), Munich, Germany, in 2019.

Since 2022, he has been a Full Professor with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His research interests include artificial intelligence, multimodal remote sensing, foundation models, hyperspectral imaging, and Earth observation.

Dr. Hong is an Associate Editor for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS) and IEEE TRANSACTIONS ON IMAGE PROCESSING, and the Editorial Board Member of *Information Fusion, ISPRS Journal of Photogrammetry and Remote Sensing*, and *International Journal of Applied Earth Observation and Geoinformation*. He was the recipient of the Jose Bioucas Dias Award for recognizing an outstanding paper at WHISPERS in 2021, the Remote Sensing Young Investigator Award in 2022, the IEEE GRSS Early Career Award in 2022, and a Highly Cited Researcher (Clarivate Analytics/Thomson Reuters) in 2022 and 2023.