geomatics

MDPI

Article

Collapsed Building Detection from Multi-Temporal Pléiades Imagery: Comparing CNN Architectures and Input Data Strategies

Tomáš Bouček [1,*], Ondřej Pešek [1], Karel Pavelka, Jr. [1], Eva Matoušková [1], Marzhan Yessenbekovna Rakhymberdina [2], Zhanna Aleksandrovna Assylkhanova [2] and Karel Pavelka [1]

[1] Department of Geomatics, Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7/2077, 166 29 Prague, Czech Republic; ondrej.pesek@fsv.cvut.cz (O.P.); karel.pavelka@cvut.cz (K.P.J.); eva.matouskova@fsv.cvut.cz (E.M.); karel.pavelka@fsv.cvut.cz (K.P.)

[2] School of Earth Sciences, D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk 070000, Kazakhstan; mrahymberdina@edu.ektu.kz (M.Y.R.); zhassylkhanova@edu.ektu.kz (Z.A.A.)

* Correspondence: tomas.boucek@fsv.cvut.cz

Abstract

A rapid and spatially explicit assessment of structural damage is essential following major disasters, particularly where access to the field is limited and reliable information on damage is scarce. This study evaluates how semantic segmentation architecture, temporal input design, and class formulation influence the detection of collapsed buildings in very high-resolution Pléiades imagery from Jindires, Syria, following the earthquakes of 6 February 2023. The U-Net, SegNet, DeepLabv3+, and fully convolutional network (FCN) models were compared together with a reference random forest model. Methodologically, this study employs a controlled factorial experiment designed to isolate the effects of using either post-event images alone or a combination of pre- and post-event imagery, across both class-based and binary label formulations. Performance was evaluated in a spatially separate validation region using Dice loss, overall accuracy, and pixel-level F1-score, supplemented by location-level evaluation for 102 visually interpreted sites of collapsed buildings. The FCN with multitemporal binary input achieved the highest F1-score at the pixel level (79.1%), followed by the U-Net network with multitemporal three-class input (76.0%) and post-event binary input (74.8%). At the site level, both the U-Net network with post-event binary input and the FCN with multitemporal binary input achieved an F1-score of 73.0%; the U-Net network prioritized recall, while the FCN prioritized precision. The results reveal that the optimal configuration depends heavily on the chosen architecture's inductive biases. Therefore, these findings highlight the importance of selecting not only the right architecture, but also the appropriate temporal input and class formulation. The results provide reproducible information for screening; although they do not constitute a definitive inventory of damage, they offer a truly independent perspective on the consequences of the earthquake in a region from which verified data cannot be obtained.

Keywords: collapsed buildings; earthquake damage assessment; semantic segmentation; convolutional neural networks; multi-temporal satellite imagery; Pléiades

Check for updates

Academic Editor: Alessandro Vitale

Received: 25 August 2026
Revised: 29 September 2026
Accepted: 1 October 2026
Published: 3 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Human settlements can change abruptly when buildings, transport links, utilities, and essential services are damaged or destroyed. Such disruption may be caused by human-induced events, including armed conflict [1], or by natural hazards such as earthquakes [2],

landslides [3], and floods [4]. When these processes affect densely inhabited areas, the physical transformation of the built environment can occur within minutes and immediately develop into a humanitarian, logistical, and information-management crisis.

Rapid identification of collapsed buildings is therefore required to estimate the magnitude and spatial distribution of damage, prioritize search-and-rescue operations, identify inaccessible neighborhoods, and support early recovery planning. Conventional field surveys may be delayed or impossible because roads are blocked, structures remain unstable, the affected area is remote, or access is constrained by security and political conditions. In some regions, active war, mine contamination, military or legally prohibited zones, border controls, or extreme remoteness can prevent entry altogether; direct comparison of satellite results with contemporaneous ground measurements is then not feasible. Decision-makers may consequently need a first building-level assessment before reliable ground reports become available. Very-high-resolution (VHR) satellite imagery can provide a rapid, repeatable, and spatially comprehensive observation of the affected settlement under these conditions.

Nevertheless, validation of satellite-based damage assessments against reliable ground truth is often problematic. Post-disaster inventories may be unavailable, incomplete, delayed, spatially imprecise, or internally inconsistent because access is restricted, damage evolves after the event, and institutions use different reporting criteria. Official records may aggregate damage by administrative unit, omit underreported or politically marginalized areas, or reflect limited reporting capacity and institutional bias. Satellite-derived results therefore cannot always be compared with an independent building-level reference. Where access is impossible, the absence of a field comparison must be reported as a validation limitation and must not be taken to mean that image interpretation is equivalent to ground truth. In such cases, systematically interpreted multi-temporal imagery does not replace field evidence, but it provides a transparent and spatially consistent proxy that can reveal damage where precise or unbiased ground information is absent.

The use of artificial intelligence and neural networks is currently state-of-the-art and, in many cases, yields high-quality and rapid results independent of manual, human data evaluation. The goal is to create a sufficiently large database for the system to learn from and then process the data automatically. In many cases, however, it is not possible to secure a statistically significant database, or the issue may involve a local event or a specific phenomenon being studied. The research in this paper focused precisely on a localized phenomenon and limited image data resources and sought to determine whether the current capabilities of neural networks allow for sufficiently high-quality automatic interpretation.

A prominent example is the earthquake doublet that struck Türkiye and Syria on 6 February 2023. Two major earthquakes, separated by approximately nine hours and with magnitudes of 7.8 and 7.6, caused extensive loss of life [5], widespread destruction of buildings and infrastructure [6,7], and very large volumes of debris and disaster waste [8,9]. The spatial extent of the affected region and the unequal availability of information across the international border made rapid Earth-observation analysis particularly relevant.

Research on the 2023 Türkiye-Syria earthquakes has addressed ground-motion characteristics [10–12], rupture processes [13–15], and surface displacement [16–18]. Other studies examined geotechnical and infrastructure effects, including liquefaction-related structural damage [19], remote-sensing detection of liquefaction manifestations [20], topographic controls on damage patterns [21], and earthquake impacts on dams [22]. These contributions provide essential hazard context, but they do not resolve the operational problem of mapping individual collapsed buildings where detailed field inventories are incomplete.

The present study focuses on the detection of collapsed buildings from VHR optical imagery. Remote sensing can support the mapping of large, inaccessible areas across

applications ranging from archeology to landscape-change detection [23]; however, the types of observable damage are constrained by spatial resolution and image geometry. Complete or substantial roof collapse, exposed interiors, debris fields, and major footprint changes may be visible, whereas wall cracks, internal damage, and many forms of moderate structural impairment generally remain below the detection capability of nadir or near-nadir satellite imagery [24]. Cotrufo et al. [25] therefore proposed a damage scale specifically adapted to VHR vertical imagery. Damage-mapping methods can use post-event imagery alone or compare pre- and post-event acquisitions [26]. Optical data offers comparatively intuitive spectral and spatial information but depend on cloud-free illumination conditions; synthetic aperture radar (SAR) operates independently of daylight and most weather yet introduces speckle and a more complex relationship between backscatter, geometry, and damage [27,28]. Multi-temporal stereoscopic surface-model differencing is another possibility, although suitably matched acquisitions are rarely available immediately after a disaster.

Methods applied to optical imagery can be broadly separated into conventional feature-based approaches and deep-learning approaches. Conventional techniques depend on manually designed spectral, textural, edge, or change features, whereas deep networks learn hierarchical spatial representations directly from labeled examples.

Within the first group, Foulser-Piggott et al. [29] investigated crowdsourced manual tagging of damaged buildings. Klonus et al. [30] combined adaptive filtering, edge extraction, texture measures, principal component analysis, and segmentation for change detection, while Zhang et al. [31] fused multiscale spectral and textural features. Romaniello et al. [32] classified image tiles rather than individual structures to estimate neighborhood-scale damage intensity. These methods can be interpretable and computationally efficient, but their transferability is often limited when acquisition geometry, illumination, urban morphology, and debris appearance differ between sites.

Recent building-damage research is dominated by convolutional neural networks (CNNs), and most studies use only post-event imagery [33]. The task may be formulated as image or patch classification, object detection, instance segmentation, or semantic segmentation. Alisjabhana et al. [34], for example, combined Mask R-CNN [35] and a ResNet-50 feature-pyramid network for damage mapping. Collapsed-building studies have frequently used object detectors, including a custom CNN [36], Faster R-CNN [37,38], and transfer learning with several pre-trained CNNs [39]. Semantic segmentation is less common, although it is well suited to spatially irregular debris fields because it assigns a class to every pixel rather than imposing a rectangular bounding box; Rashidian et al. [40] demonstrated this formulation with U-Net [41].

A growing body of work modifies established architectures or proposes specialized networks for damaged- and collapsed-building extraction. Examples include modified U-Net variants [42–44], YOLOv3 [45], YOLOv4 [46,47], RescueNet based on DeepLabv3 and DeepLabv3+ [48–50], the two-stage U-Net-derived BDANet [51], and the transformer-based HRTBDA model designed to capture long-range context [52]. Further purpose-built architectures include BDHE-Net [53], D2ANet [54], and PCDASNet [55]. Despite this methodological expansion, less attention has been given to controlled comparisons of widely used semantic-segmentation architectures under identical data and reference conditions. In particular, the interaction between architecture, single- versus multi-temporal input, and binary versus multiclass label formulation remains insufficiently examined.

Beyond these damage-specific adaptations, general-purpose semantic segmentation has changed substantially since the CNN architectures compared in this study were introduced. Vision transformers (ViT) [56] replace local convolutions with self-attention over image patches and can therefore model global context from the first layer. SETR [57] applied

a pure transformer encoder to dense prediction, SegFormer [58] combined a hierarchical transformer encoder with a lightweight all-MLP decoder, and the Swin Transformer [59] restricted self-attention to shifted local windows, yielding a hierarchical backbone whose computational cost grows linearly with image size. Mask2Former [60] subsequently unified semantic, instance, and panoptic segmentation through masked cross-attention. Hybrid designs such as UNetFormer [61], which couples a CNN encoder with a transformer decoder, were developed specifically for urban remote-sensing scenes, whereas ConvNeXt [62] showed that modernized, purely convolutional networks can compete with transformer backbones when trained with comparable recipes. The most recent developments include promptable foundation models such as the Segment Anything Model (SAM) [63], trained on more than one billion masks, remote-sensing foundation models pre-trained by self-supervision on large multispectral and multitemporal archives, such as SatMAE [64] and Prithvi [65], and selective state-space models (Mamba) [66] and their visual counterparts such as VMamba [67], which provide global receptive fields at a computational cost that grows linearly with input size.

This study does not try to find the best-performing model by implementing the bleeding-edge architectures. Instead, its aim is to explore whether the move towards context-knowledgeable CNNs helps the damaged building segmentation. To make the results as transferable and understandable as possible, it does so by implementing the still-most-common CNN architectures—U-Net [41], SegNet [68], DeepLabv3+ [50] with three ResNet backbones, and FCN [69]. The collapsed-building segmentation is performed over the city of Jindires, Syria, using VHR Pléiades imagery. The experiments isolate two design choices: whether the model receives post-earthquake imagery alone or a co-registered pre-/post-earthquake image stack, and whether it separates non-collapsed buildings from both background and collapse or detects collapse as a single target class. A random forest (RF) classifier [70] provides a non-contextual pixel-based baseline.

This study thus makes three main contributions to the field. First, it provides a systematic factorial evaluation that isolates the effects of temporal input strategies and label formulations. Second, it offers a performance-oriented evaluation aimed at determining not only which model achieves the highest numerical accuracy, but also which architecture provides a practical balance between missed collapses, false alarms, and computational complexity. Finally, by utilizing commonly used CNN architectures, it establishes a reference against which newer models can subsequently be evaluated under the same design.

2. Materials and Methods

This section describes the study area, the Pléiades acquisitions and preprocessing workflow, the tested classification architectures, the factorial design of the collapse-detection experiments, and the preparation of spatially separate training and validation reference data. The overall workflow is shown in Figure 1. The individual steps are described in detail in the following subsections.

2.1. Site Description

Northern Syria and southeastern Türkiye form part of a highly active tectonic region near the interaction of the Anatolian, Arabian, and African plates (Figure 2). On 6 February 2023, the region was struck by two major earthquakes with magnitudes Mw 7.8 and Mw 7.6. Although the first epicenter was in southern Türkiye, strong shaking affected a broad transboundary area that included settlements in the Syrian governorates of Aleppo and Idlib.

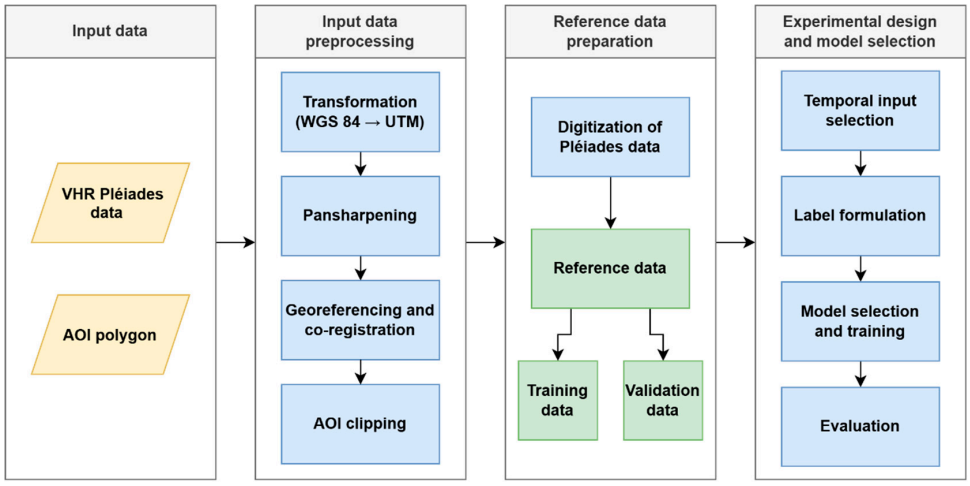


Figure 1. The overall workflow applied in this study. Orange—input data; blue—processing steps; green—an intermediate product created based on the input data.

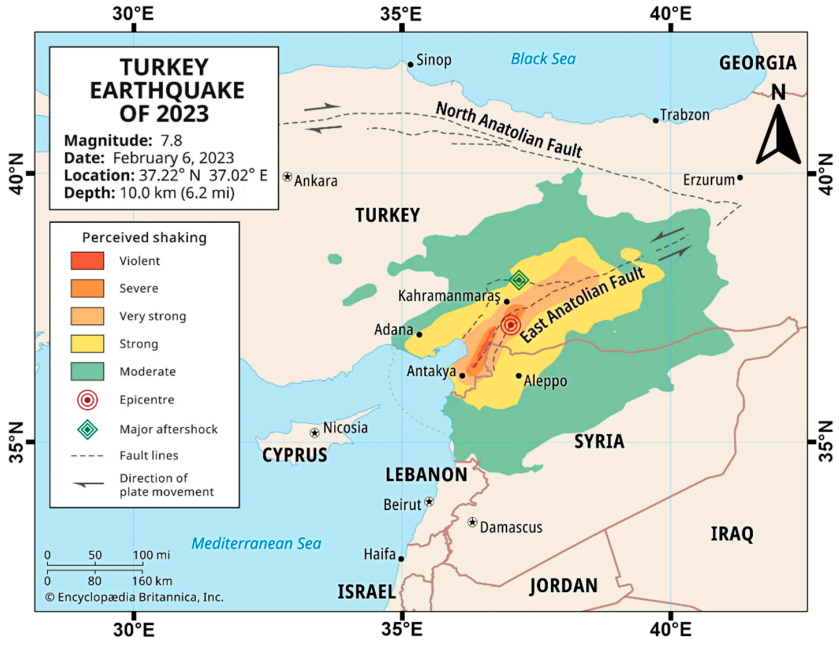


Figure 2. East Anatolian fault [71].

The Syrian part of the disaster received less systematic international reporting because earthquake damage overlapped with destruction and access restrictions associated with the long-running conflict. Available estimates were consequently heterogeneous in spatial extent and reporting criteria. Local sources reported that there were more than 800 demolished and nearly 6000 damaged buildings across the wider Idlib region [72], while satellite-based assessment in Idlib city identified at least 79 damaged or potentially damaged structures [73]. Jindires, in Afrin District of Aleppo Governorate, was among the settlements with the most concentrated building collapse.

This information environment illustrates the operational motivation of the study. Direct field verification was constrained, reports combined different definitions of damage, and pre-existing conflict-related destruction complicated attribution to the earthquake. A repeatable satellite-image analysis can therefore provide building-level spatial evidence while also making its observational limitations explicit.

Jindires is a compact historic town near the Turkish–Syrian border. According to OpenStreetMap, its mapped area is approximately 2.2 km^2 [74], and the 2004 census reported a

population of around 14,000 [75]. Sources from the period following the earthquake estimated 228 damaged and 318 potentially damaged buildings [76], approximately 300 destroyed houses [77], and, in a later retrospective report, 489 completely destroyed buildings [78]. The discrepancies between these figures confirm that there is no single authoritative inventory at the individual-building level, and no readily verifiable information exists from these areas; therefore, there is research potential here. Jindires was selected precisely because no verifiable independent information was available. Only a few pairs of satellite images were available; extensive building collapse had occurred in the area, and the compact urban structure allowed for a comparison of the entire area. Figure 3 illustrates the observed destruction, and Figure 4 shows the location and boundaries of the area of interest (AOI).



Figure 3. Jindires, photo by Alaa Ealyawi on 25 February 2023 [79].

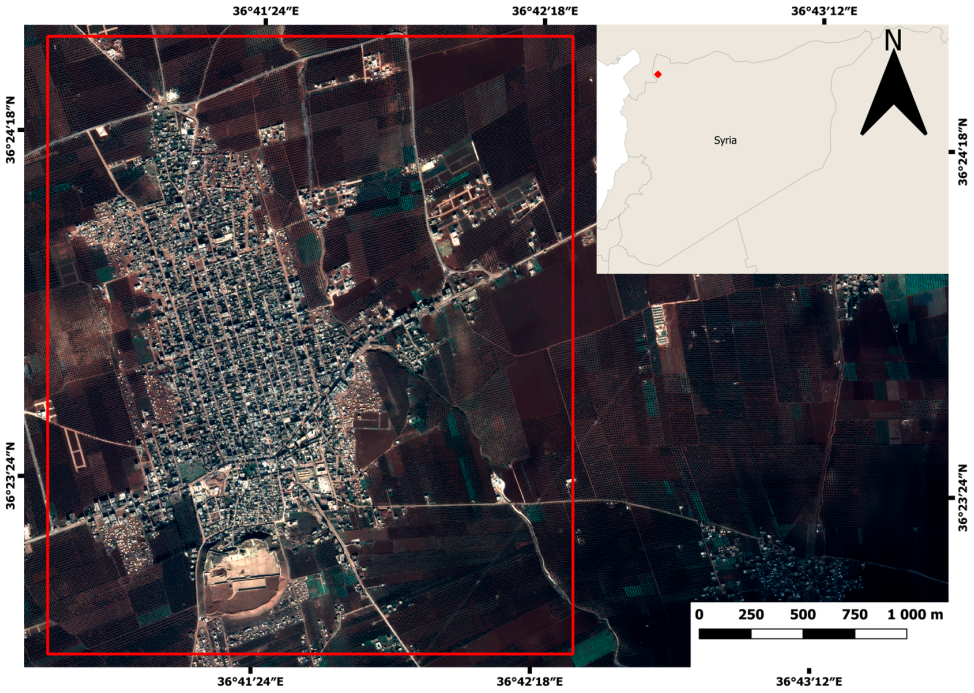


Figure 4. Location of the study area (Jindires) with the AOI marked by the red rectangle. The background is RGB imagery data from the Pléiades satellite. The overview in the upper right corner with a red dot representing the study location uses a vector map of the world available directly in QGIS 3.34.15 software [80].

2.2. Satellite Data

A great deal of data was available from the earthquake-affected area in Turkey, whereas very little data was available from the affected area in northern Syria. The search focused primarily on satellite data with the highest available resolution, taken near nadir, as well as data in the mid-infrared band. Stereoscopic pairs would also have been suitable. No such data existed; ultimately, the only suitable data were obtained from the Pléiades system. These consisted of two four-channel VHR Pléiades multispectral images, one taken before the earthquake and one after [81]. The High-Resolution Imager sensor captures blue, green, red, and near-infrared bands with a native ground sampling distance of approximately 2.8 m and a panchromatic band with a resolution of approximately 0.7 m [82]. Pléiades primary products preserve the sensor geometry and are not delivered as fully orthorectified map products. In contrast, the standard Ortho product is projected and resampled to 2.0 m for multispectral data and to 0.5 m for the panchromatic band [83,84]. The images available for this study were primary products; their acquisition dates and imaging parameters are listed in Table 1. Although there are sensors with finer spatial resolution or a broader spectral range, no alternative provided a suitable pair of cloud-free images of Jindires immediately before and after the earthquake. The pair of images from the Pléiades satellite therefore represented the only available coherent, high-spatial-resolution (VHR) multitemporal dataset for this study. Its resolution is sufficient to capture the distribution of debris and significant damage to roofs or building floor plans; however, it cannot reliably detect fine structural damage. Differences in viewing angle between individual images also cause geometric and visual discrepancies that must be considered when interpreting multitemporal classifications. This is precisely why it was a research challenge; suitable data are not always available, yet it is important to determine, at least approximately, the actual extent of the damage.

Table 1. Sensing parameters of the acquired Pléiades data for the center of the captured scenes.

Sensing Date	Azimuth	Viewing Angle [°]			Incidence Angle [°]		
		Across Track	Along Track	Overall	Across Track	Along Track	Overall
27 December 2022	180.11	−10.20	19.79	21.92	14.85	−20.09	24.31
10 February 2023	180.07	−3.22	17.10	17.36	6.91	−18.05	19.18

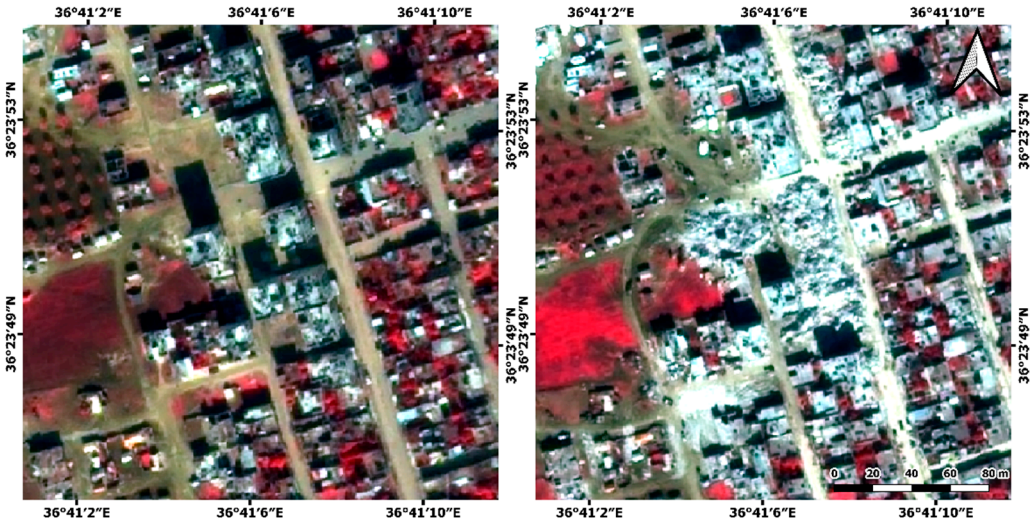
2.3. Satellite Data Preprocessing

Preprocessing consisted of four main stages. First, both primary scenes were transformed into WGS 84/UTM zone 37N (EPSG:32637). Second, the multispectral bands were pansharpened with the panchromatic image to retain four-band spectral information at sub-meter spatial detail. Third, the pre- and post-event products were georeferenced and co-registered. This step was critical for the multi-temporal experiments because even a small spatial offset can create artificial changes along roof edges and narrow streets. Co-registration also supported visual comparison of the acquisitions, as illustrated in Figure 5. Finally, both scenes were clipped to the common AOI. All preprocessing was performed in QGIS [80].

2.4. Classification Methods

The experiment compares semantic-segmentation architectures with different mechanisms for retaining spatial detail and incorporating context. U-Net [41] uses a symmetric encoder–decoder structure with skip connections that transfer high-resolution features from the contracting path to the expanding path. This design is well suited to small training datasets and precise boundary reconstruction and has performed strongly in

remote-sensing segmentation [85]. SegNet [68] also follows an encoder–decoder design, but its decoder reuses the max-pooling indices recorded by the encoder rather than concatenating full encoder feature maps. The result is a different trade-off between memory use, contextual abstraction, and boundary localization.



(**a**) Before earthquake (**b**) After earthquake

Figure 5. Example of destroyed buildings in the scenes before (**a**) and after (**b**) the earthquake. The imagery data are displayed in false colors to better distinguish green from built-up areas.

DeepLabv3+ [50] combines atrous convolution (dilated convolution), atrous spatial pyramid pooling (ASPP), and a lightweight decoder to capture context at multiple effective scales; ResNet-50, ResNet-101, and ResNet-152 backbones [86] were tested to examine whether increasing feature-extractor depth improved collapse mapping. The tested FCN [69] replaces fully connected classification layers with convolutions and restores dense spatial predictions through learned upsampling and skip-based feature fusion. It therefore differs from U-Net in the way multiscale features are combined, but it is not an encoder-only classifier. RF [70,87] was included as a classical baseline: unlike the CNNs, it classifies pixels from their feature vectors without learning neighborhood shape or debris texture. All models received the same imagery and reference data. Therefore, only all available Pléiades spectral bands (B-G-R-NIR) were used as input image data, which also applied to the RF. RF was implemented with scikit-learn [88], and the CNNs with TensorFlow [89].

Collapsed Building Detection

The experimental design separates the effects of temporal input and label formulation. Two input approaches were tested with every classifier:

(1) Detection of collapsed buildings using only post-earthquake imagery.
(2) Detection of collapsed buildings combining both pre- and post-earthquake imagery.

Each input approach was evaluated with two label variants:

(a) Classification of both collapsed and non-collapsed buildings together.
(b) Classification of collapsed buildings alone.

In approach (1), the four pansharpened post-event spectral bands formed the model input. In approach (2), the corresponding four-band pre- and post-event images were concatenated into an eight-channel multi-temporal stack. The same spatial training and validation areas were used for all experiments so that differences could be attributed to architecture, temporal input, and class formulation rather than to a different sample.

The CNN implementations followed the defining principles of their original architecture descriptions; the total and trainable parameter counts are reported in Table 2.

Table 2. Number of parameters per CNN architecture.

Architecture	Total Parameters	Trainable Parameters
U-Net	31,056,003	31,044,227
SegNet	62,502,530	62,481,538
FCN	467,513,236	467,485,332
DeepLabV3+_Resnet50	17,835,698	17,800,914
DeepLabV3+_Resnet101	36,906,162	36,819,154
DeepLabV3+_Resnet152	52,618,930	52,485,842

Model selection and performance reporting used data excluded from parameter estimation. Dice loss [90] was reported because overlap-based objectives are less dominated by the large background class than raw pixel accuracy. Overall accuracy was retained for comparability but was interpreted cautiously because collapsed pixels constitute only a small fraction of the AOI. The F1-score, defined as the harmonic mean of precision and recall, was therefore used as the principal pixel-level discrimination metric. Pixel-based evaluation was complemented by the site-level analysis described in Section 3.3, which assesses whether mapped collapse objects correspond to interpreted collapse locations. Source code is available under the MIT license at https://github.com/pesekon2/cnn-lib-syria-earthquake-paper (accessed on 10 August 2026). CNN training and inference were performed on the BDAP platform (formerly JEODPP) [91].

Network weights were initialized with the He normal initializer [92], and all CNNs were optimized with Adam (adaptive moment estimation) [93] at a learning rate of 0.001, using the minimal Dice loss as the training objective. Owing to the memory limitations of the computational server and the size of the image data, the batch size was set to 1. The maximum number of training epochs was set to 1000, but this limit was never reached because training was terminated by early stopping once the validation loss had not improved for 200 consecutive epochs; the weights from the epoch with the lowest validation loss were retained. No data augmentation was applied (during an early experiment, rotation data augmentation lowered the accuracy due to the direction-dependent features of the remote sensing data and area of interest), and the same training settings were used for all architectures and approach–variant combinations.

For RF, the number of trees (n_estimators), maximum tree depth (max_depth), and minimum number of samples required to split a node (min_samples_split) were optimized separately for each approach by grid search. Candidate combinations were ranked by F1-score using five-fold cross-validation within the training data. The selected configuration was then evaluated on an independent test subset reserved before model fitting. This procedure reduces tuning bias, although the RF remains fundamentally limited to the supplied pixel-level features and cannot learn the spatial morphology of collapsed structures.

2.5. Reference Data Preparation

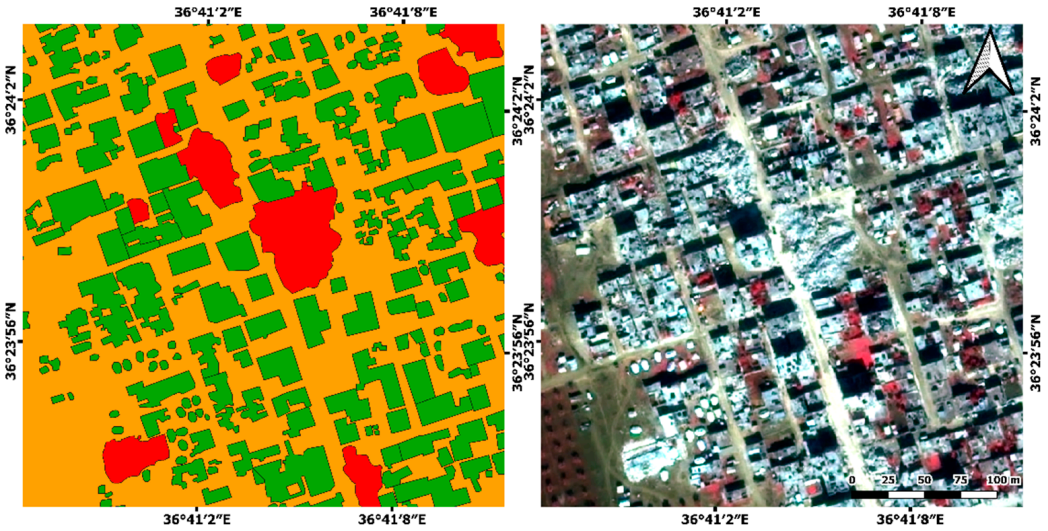
Supervised semantic segmentation requires spatially coherent labels that represent both target objects and their surroundings. The spatial representativeness, internal consistency, and harmonization of reference samples are known to influence the generalizability of satellite-image classifiers [94,95]. No field-verified building-level dataset with the required spatial detail was available for Jindires. A systematic ground survey could not be conducted under the prevailing conflict-related security and access restrictions, and existing vector data could not be used directly because the primary imagery had only approximate geolocation before co-registration. Reference masks were therefore created

manually. Two spatially separate areas were delineated: one for model training and one for validation. Spatial separation reduces direct leakage of neighboring pixels between the two subsets, although both areas still belong to the same urban and acquisition domain. The two datasets are spatially contiguous, meeting approximately in the center of the city. The training dataset thus covers the northern part of the city, and the validation dataset covers the southern part. Both parts of the city are visually similar. The exception is the southernmost part of the validation dataset, which likely contains the historic center with smaller buildings and narrower streets.

Buildings and visibly collapsed areas were digitized from the imagery. Many recent damage datasets use four levels (no damage, minor damage, major damage, and destroyed), particularly following the release of xBD, which contains more than 800,000 labeled buildings from disasters worldwide [96]. Such grading was not defensible here because the Pléiades viewing geometry and spatial resolution do not reveal many diagnostic signs of moderate or internal structural damage.

The annotation scheme therefore distinguished only non-collapsed buildings and collapsed areas. This binary physical distinction is deliberately narrower than a complete engineering damage assessment: a structure classified as non-collapsed may still have sustained severe but externally inconspicuous damage.

Digitization was performed primarily on the post-earthquake Pléiades scene, with Bing Maps [97] and OpenStreetMap [74] used to clarify ambiguous building locations. Approximately 800 buildings and 22 collapse sites were delineated in the training area, and approximately 1200 buildings and 10 collapse sites in the validation area. Only visible debris or rubble was labeled as collapsed; for partially collapsed buildings or blocks, only the visibly affected portion was included. Table 3 summarizes the class definitions and Figure 6 shows an example. Because the labels were derived from image interpretation rather than a complete post-event field survey, they should be understood as a consistent remote-sensing reference rather than absolute ground truth. The validation therefore quantifies agreement with independently delineated image evidence, not accuracy against on-site measurements. Ambiguous roof conditions, occlusion, shadows, residual conflict damage, and the sensor resolution remain sources of label uncertainty.



(**a**) Reference data (**b**) Pléiades data

Figure 6. Example of the reference dataset (**a**) and corresponding Pléiades imagery (**b**). Buildings are shown in green, collapsed buildings in red, and other areas in orange.

Table 3. Classification classes for each approach.

Label	Classes
Approach (1a)	Others Buildings Collapsed buildings
Approach (1b)	Others Collapsed buildings
Approach (2a)	Others Buildings Collapsed buildings
Approach (2b)	Others Collapsed buildings

3. Results

Results are reported first for the held-out validation area at pixel level and then for the complete AOI at collapse-site level. The two views are complementary: pixel metrics measure spatial agreement with the reference masks, whereas site-level metrics evaluate whether a model identifies the locations that would be relevant for rapid damage screening.

3.1. Approach (1)

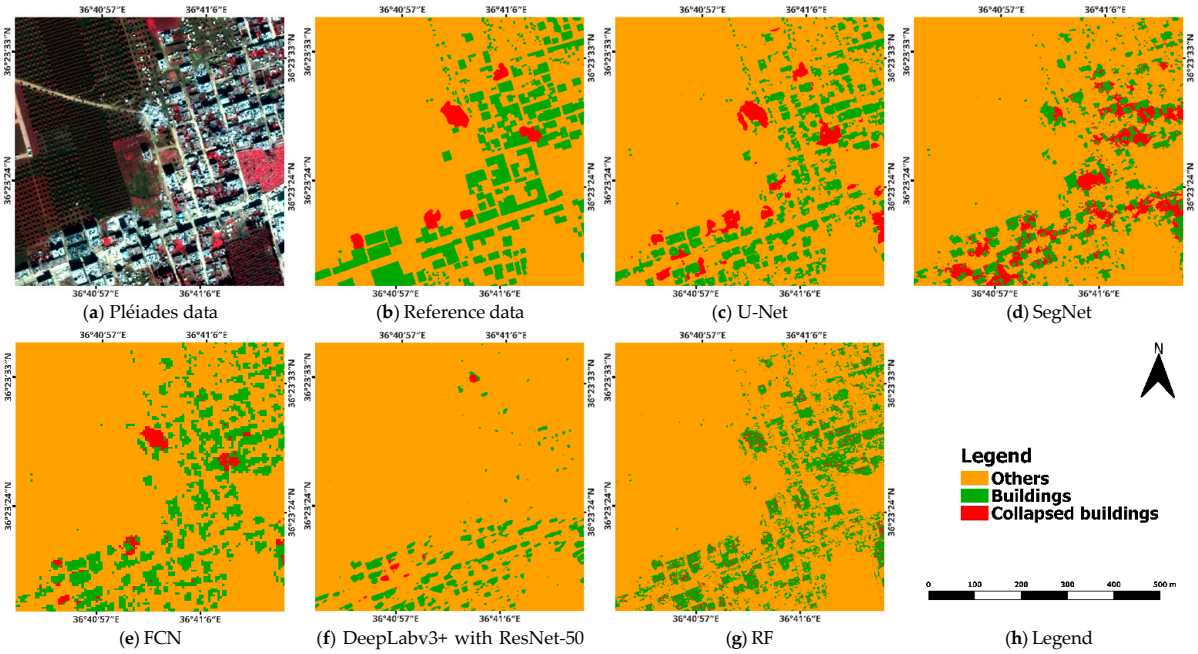
Approach (1) used only post-earthquake imagery and therefore required each model to infer collapse from post-event spectral, textural, and morphological evidence without an explicit pre-event reference. Variant (1a) separated background, non-collapsed buildings, and collapsed buildings, whereas variant (1b) treated collapse as the only foreground class.

Table 4 reports Dice loss, overall accuracy, and F1-score for both variants. Because the background class is dominant, the F1-score provides a more informative comparison of collapsed-building discrimination than overall accuracy alone.

Table 4. Dice loss, overall accuracy values, and F1-scores over the validation dataset for RF and different CNN architectures for approach (1). The lowest Dice loss and highest overall accuracy and F1-score values are shown in bold.

	Variant (1a)			Variant (1b)		
Architecture	**Dice Loss**	**Overall Accuracy [%]**	**F1-Score [%]**	**Dice Loss**	**Overall Accuracy [%]**	**F1-Score [%]**
U-Net	**0.49**	**93.8**	64.7	**0.17**	**99.4**	**74.8**
SegNet	0.57	90.0	50.6	0.34	98.4	53.9
FCN	0.51	**93.8**	**67.9**	0.21	99.0	66.6
DeepLabv3+_ResNet50	0.71	91.4	46.1	0.44	98.2	54.9
DeepLabv3+_ResNet101	0.77	86.0	31.3	0.50	97.8	52.7
DeepLabv3+_ResNet152	0.73	0.8	1.3	0.47	76.4	44.2
RF	—	93.0	57.0	—	99.3	54.1

For the three-class variant (1a), FCN achieved the highest F1-score (67.9%), followed by U-Net (64.7%). SegNet reached 50.6%, while all DeepLabv3+ variants remained below 50%; the ResNet152 configuration failed to produce a useful segmentation in this experiment. These findings show that a deeper backbone does not automatically produce better results. In this case, the deeper backbone was probably too large relative to the available labeled samples. The qualitative comparison in Figure 7 is consistent with these values: FCN and U-Net most closely reproduced the interpreted collapse pattern, whereas the other models showed greater omission or confusion.



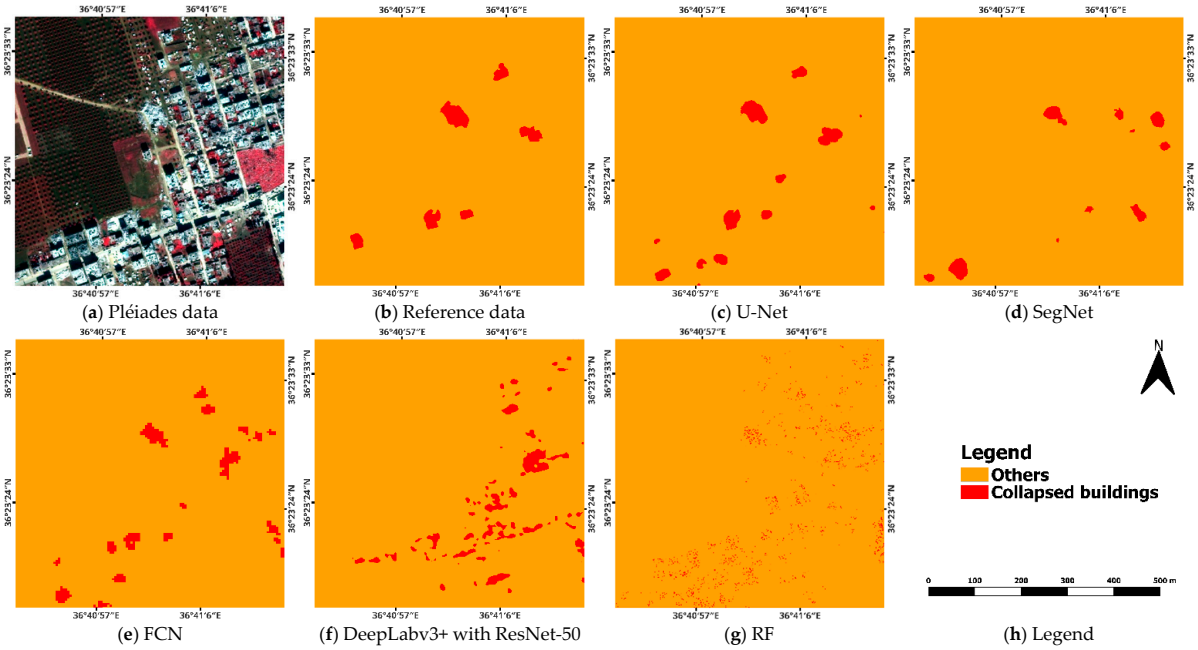
(**a**) Pléiades data (**b**) Reference data (**c**) U-Net (**d**) SegNet

(**e**) FCN (**f**) DeepLabv3+ with ResNet-50 (**g**) RF (**h**) Legend

Figure 7. Results of the classification of individual architectures on a subset of validation data using variant (1a).

For the binary variant (1b), U-Net achieved the highest F1-score (74.8%), 8.2 percentage points above FCN (66.6%). SegNet and the DeepLabv3+ variants produced lower and more variable scores, and RF reached 54.1%.

Dice loss and overall accuracy improved for all architectures when the non-collapsed-building class was removed. This increase should not be interpreted automatically as better operational performance, because binary classification changes both the class balance and the decision problem. The F1-score improved for most, but not all, models; FCN and RF performed better under the three-class formulation. Figure 8 further shows that the binary U-Net result retained collapse regions more consistently than competing outputs.



(**a**) Pléiades data (**b**) Reference data (**c**) U-Net (**d**) SegNet

(**e**) FCN (**f**) DeepLabv3+ with ResNet-50 (**g**) RF (**h**) Legend

Figure 8. Results of the classification of individual architectures on a subset of validation data using variant (1b).

3.2. Approach (2)

Approach (2) supplied the co-registered pre- and post-earthquake images jointly, allowing the networks to learn both post-event appearance and temporal change. Variant (2a) used three classes, while variant (2b) used collapse as the sole foreground class.

Table 5 reports the validation metrics for both multi-temporal variants. As for approach (1), F1-score is emphasized over overall accuracy because non-collapse pixels dominate the scene.

Table 5. Dice loss, overall accuracy values, and F1-scores over the validation dataset for RF and different CNN architectures for approach (2). The lowest Dice loss and highest overall accuracy and F1-score values are shown in bold.

	Variant (2a)			Variant (2b)		
Architecture	**Dice Loss**	**Overall Accuracy [%]**	**F1-Score [%]**	**Dice Loss**	**Overall Accuracy [%]**	**F1-Score [%]**
U-Net	**0.46**	**95.1**	**76.0**	0.15	98.9	67.1
SegNet	0.60	89.1	55.2	0.30	98.8	63.7
FCN	0.49	94.1	70.3	**0.14**	**99.7**	**79.1**
DeepLabv3+_ResNet50	0.61	92.9	55.7	0.29	99.4	64.9
DeepLabv3+_ResNet101	0.67	89.8	51.7	0.45	98.4	49.9
DeepLabv3+_ResNet152	0.73	89.3	31.5	0.50	99.6	49.9
RF	—	93.9	63.0	—	99.5	60.4

The effect of multi-temporal input depended strongly on label formulation and architecture. In variant (2a), U-Net achieved the highest F1-score (76.0%), followed by FCN (70.3%). In variant (2b), FCN was clearly best at 79.1%, while U-Net reached 67.1%. Thus, the architecture ranking reversed between the two variants rather than improving uniformly when pre-event information was added.

U-Net and FCN were the only architectures to produce consistently coherent three-class maps (Figure 9). SegNet achieved a substantially lower pixel-level F1-score, but it still localized several collapse sites, suggesting that object presence and pixel-boundary agreement should be evaluated separately.

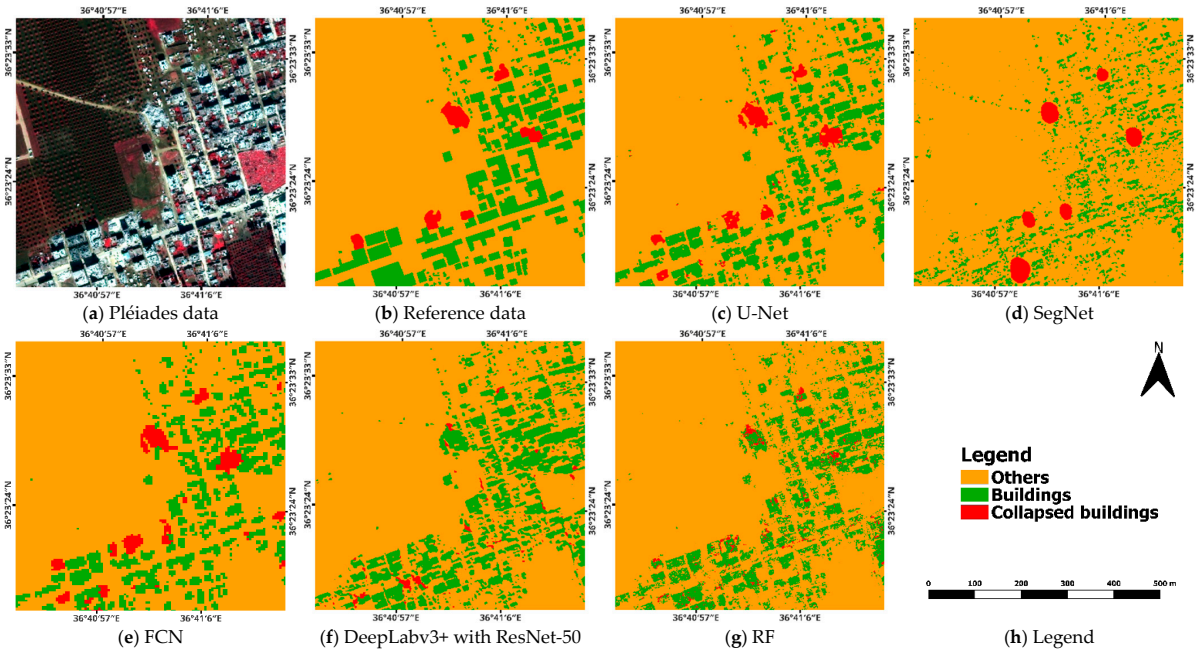


Figure 9. Results of the classification of individual architectures on a subset of validation data using variant (2a).

In variant (2b), FCN exceeded U-Net by 12.0 percentage points in F1-score even though their reported Dice losses were similar (0.14 and 0.15, respectively). The divergence shows that the optimization loss and the F1-score of thresholded output should not be treated as interchangeable summaries. SegNet and DeepLabv3+_ResNet50 reached 63.7% and 64.9%, respectively, whereas deeper DeepLabv3+ backbones did not improve performance.

The qualitative results in Figure 10 explain part of this difference. U-Net tended to over-map collapse, increasing false-positive area, whereas SegNet and DeepLabv3+ tended to omit portions of the interpreted collapse regions. FCN produced a more selective spatial response that yielded the strongest balance between precision and recall for this variant.

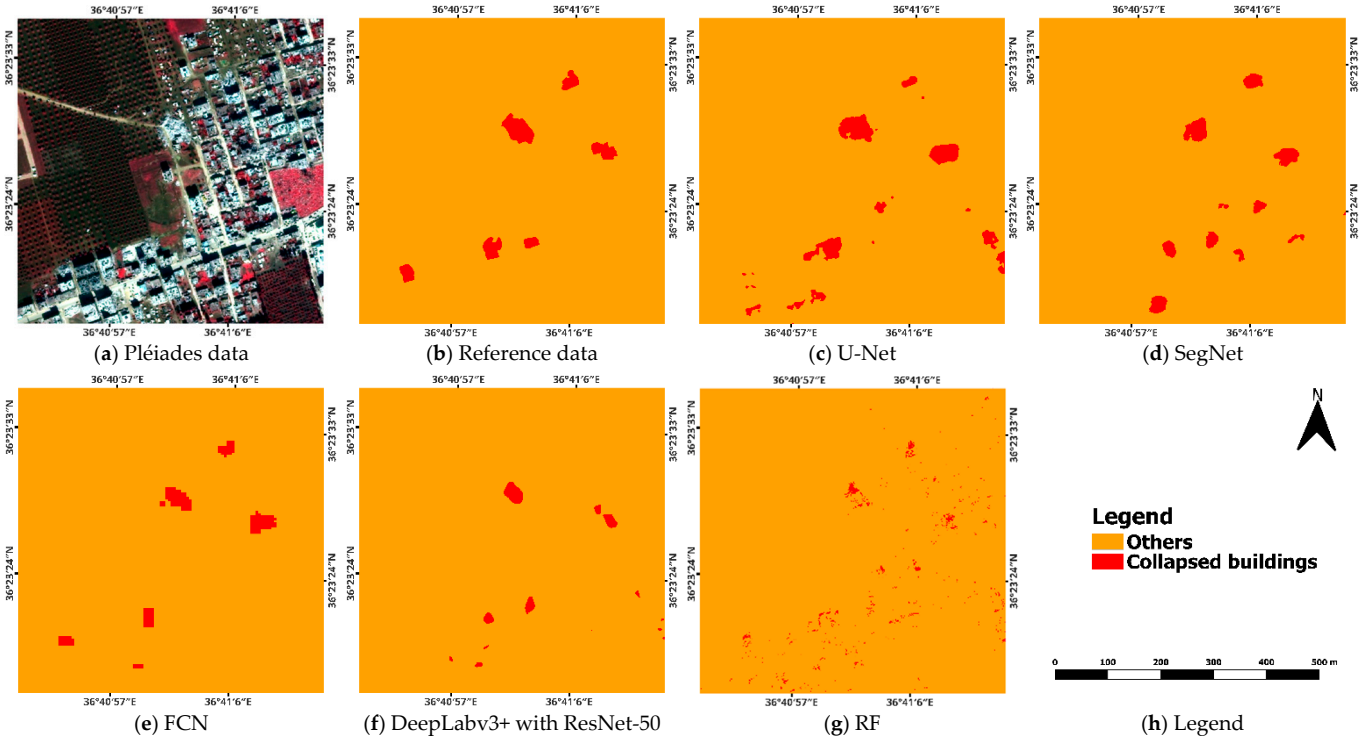


Figure 10. Results of the classification of individual architectures on a subset of validation data using variant (2b).

3.3. Overall Assessment of Tested Approaches

Pixel-level agreement does not fully describe operational utility. A prediction can have imperfect boundaries yet correctly identify a damaged site, or it can achieve high overall accuracy while missing rare collapse objects. The best-performing model-variant combinations were therefore evaluated over the complete AOI against collapse locations identified by systematic visual interpretation. For approach (1), U-Net and FCN were evaluated in both variants and SegNet in variant (1b). For approach (2), U-Net, FCN, and SegNet were evaluated in both variants.

Predicted collapse masks were vectorized, and polygons smaller than 50 m^2 were removed to suppress isolated pixel noise and objects unlikely to represent a collapsed building. This 50 m^2 filtering threshold was chosen based on two main factors. The first is the quality of the satellite data itself, which often made it impossible to determine with certainty what type of object was present at a given location. In some cases, certain objects might have appeared to be rubble, even though they were not. The second factor relates to the size of the buildings themselves; based on a visual inspection, it was determined that the area between a collapsed building and other debris roughly corresponds to 50 m^2. The resulting site-level metrics are, however, influenced by this threshold. A lower or no

threshold would result in a higher number of false-positive detections, which would reduce precision. Conversely, a higher threshold would eliminate correctly detected collapses and reduce the resulting recall. The chosen threshold is thus a compromise between these two extremes.

The remaining polygons were spatially matched to 102 visually interpreted collapse locations across the AOI. The criterion for spatial matching was the intersection of the detected collapsed building with the approximate centroid of the visually marked collapsed building. Correct detections and unmatched predicted sites were counted to derive site-level precision, recall, and F1-score (Table 6). These metrics measure localization at the site level and are therefore not numerically equivalent to the pixel metrics in Tables 4 and 5.

Table 6. Comparison of visually detected collapsed buildings with automatic detection using selected architectures for individual approaches, together with accuracy metrics.

Approach	Number of Correctly Detected Sites by Chosen Approaches	Number of Extra Sites Detected by Chosen Approaches	Precision [%]	Recall [%]	F1-Score [%]
U-Net (1a)	94	246	27.6	92.2	42.5
FCN (1a)	86	164	34.4	84.3	48.9
U-Net (1b)	88	51	63.3	86.3	73.0
FCN (1b)	95	168	36.1	93.1	52.1
SegNet (1b)	40	62	39.2	39.2	39.2
U-Net (2a)	93	79	54.1	91.2	67.9
FCN (2a)	96	200	32.4	94.1	48.2
SegNet (2a)	64	30	68.1	62.7	65.3
U-Net (2b)	82	114	41.8	80.4	55.0
FCN (2b)	69	18	79.3	67.6	73.0
SegNet (2b)	70	49	58.8	68.6	63.3

Most selected configurations achieved high recall but substantially lower precision, indicating systematic over-detection. FCN reached recall above 93% in variants (1b) and (2a), and U-Net exceeded 91% in variants (1a) and (2a). Precision reached at least 60% in only three cases. The highest site-level F1-score was 73.0%, obtained by both U-Net (1b) and FCN (2b), but through different error profiles: U-Net (1b) combined 63.3% precision with 86.3% recall, whereas FCN (2b) combined 79.3% precision with 67.6% recall. U-Net (2a) ranked next with an F1-score of 67.9%. These differences are operationally important because a rapid-screening workflow may prioritize recall, while a workflow with limited capacity for manual verification may favor precision.

The number of false detections must also be interpreted in relation to the data. At approximately 0.7 m panchromatic resolution, roof texture, shadows, narrow streets, demolition debris, and pre-existing damage can resemble earthquake collapse. Conversely, partial collapse may occupy too few pixels to form a stable segment. The site-level comparison therefore evaluates consistency with observable image evidence, not the complete engineering condition of every building.

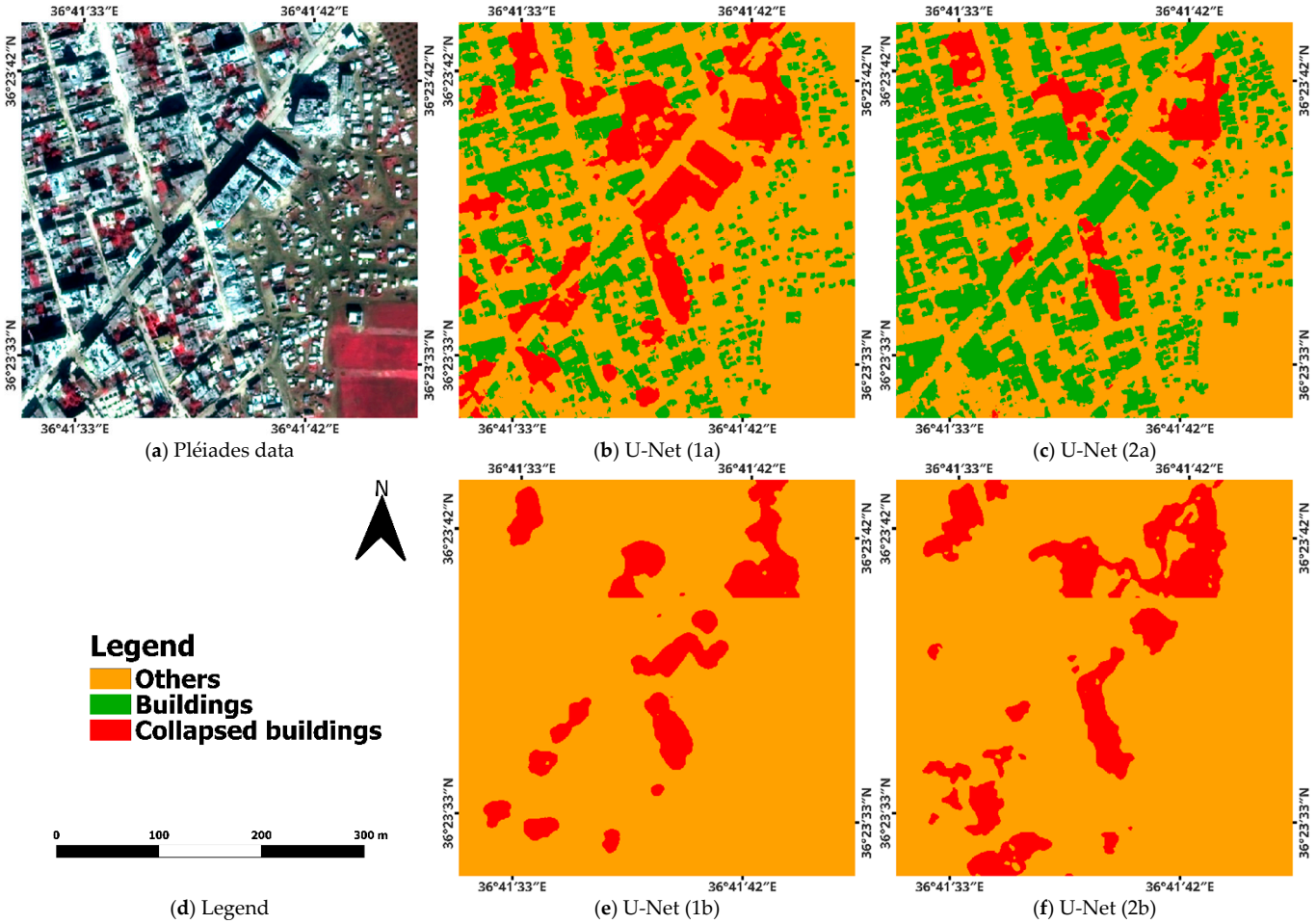
4. Discussion

The experiments show that multi-temporal input is not intrinsically superior to post-event imagery. A pre-event image can supply evidence of prior building form and make abrupt roof or footprint change easier to recognize. At the same time, the additional channels increase model dimensionality and introduce differences caused by viewing angle, illumination, season, and residual co-registration error. Whether the extra image improves

segmentation therefore depends on the architecture's ability to separate true structural change from acquisition-related variation and on how the output classes are defined.

U-Net and FCN were the only architectures that performed consistently across both temporal strategies, but they responded differently to the label variants. This behavior reflects differences in their inductive biases rather than a simple ordering by network size. U-Net explicitly fuses high-resolution encoder features with decoder features, which supports boundary recovery and the use of local spatial context. The tested FCN has much greater capacity and reconstructs dense output through learned upsampling and feature fusion, producing a different balance between spatial smoothness and class selectivity.

With post-event input, U-Net performed best when collapse was the only foreground class: the pixel-level F1-score was 10.1 percentage points higher than in the three-class variant, and the number of site-level false positives decreased from 246 to 51. When pre-event imagery was added, the pattern reversed: U-Net performed best with explicit separation of non-collapsed buildings, reaching 76.0% pixel-level F1-score and 67.9% site-level F1-score. The pre-event channels appear to have helped U-Net establish an intact building form when that class was modeled directly. Figure 11 illustrates this interaction; it should be viewed as architecture- and formulation-specific rather than as proof that multi-temporal input always improves U-Net.



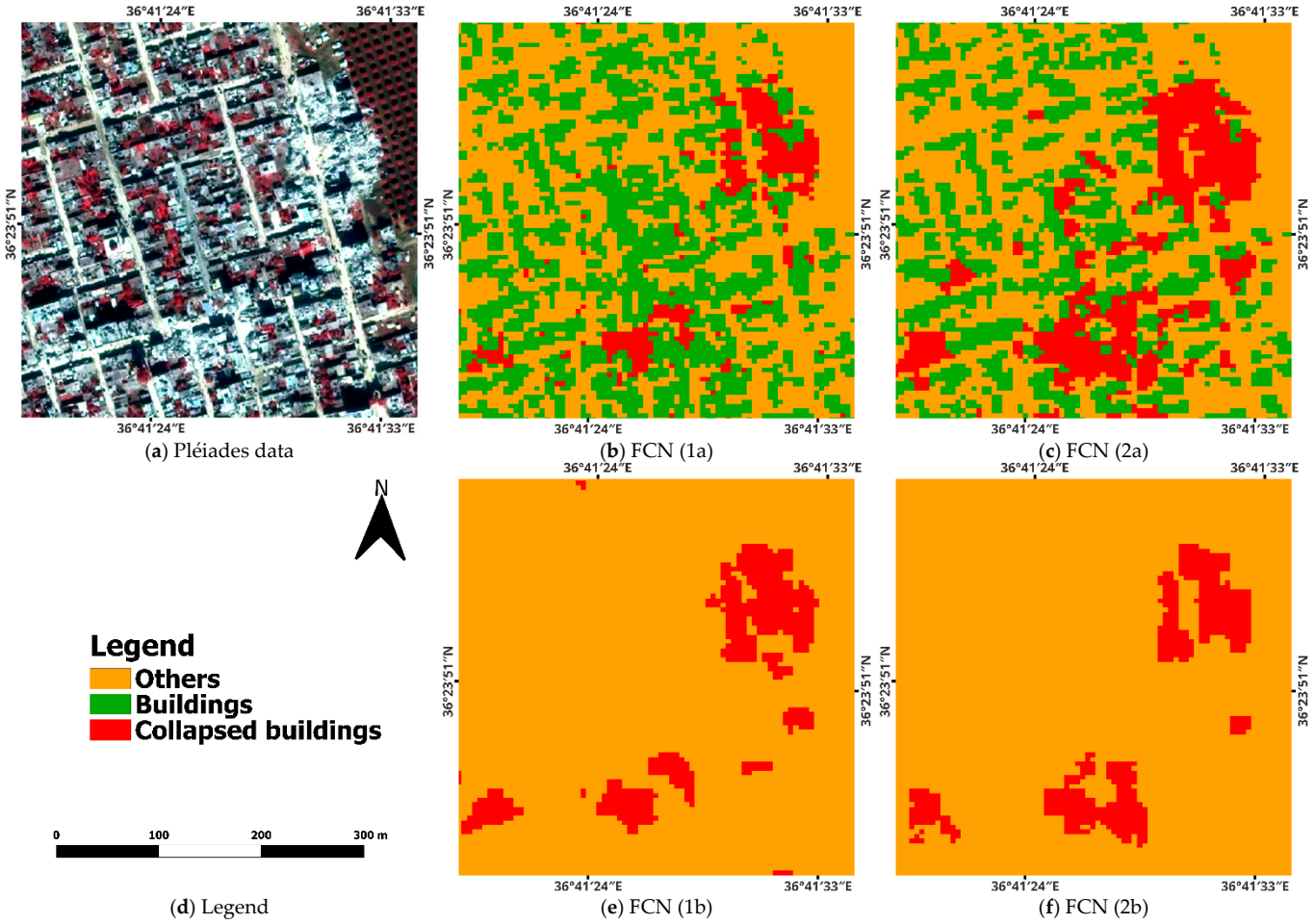
(**a**) Pléiades data (**b**) U-Net (1a) (**c**) U-Net (2a)

(**d**) Legend (**e**) U-Net (1b) (**f**) U-Net (2b)

Figure 11. Comparison of U-Net classification results across different approaches outside the validation data.

FCN showed the opposite interaction. With post-event imagery, its three-class variant outperformed the binary variant at pixel level, whereas the multi-temporal binary configuration (2b) produced the best pixel-level F1-score in the study and the highest site-level

precision. FCN produces dense predictions through convolutional score maps, skipping fusion, and learned upsampling, so its behavior cannot be attributed simply to the absence of a decoder. A more plausible interpretation is that its high-capacity feature representation and comparatively coarse-to-fine reconstruction favored a selective collapse response in variant (2b), reducing small false-positive regions at the expense of recall. Figure 12 supports this error profile, but the experiment does not isolate which architectural component caused it. However, it is important to note that these explanations represent plausible interpretations based more on architectural considerations than on confirmed findings, and verifying them would require a targeted ablation study.



(**a**) Pléiades data (**b**) FCN (1a) (**c**) FCN (2a)

(**d**) Legend (**e**) FCN (1b) (**f**) FCN (2b)

Figure 12. Comparison of FCN classification results across different approaches outside the validation data.

The experimental results reveal that the optimal label formulation is not universal, but rather highly dependent on the chosen network architecture and temporal input strategy. For instance, U-Net demonstrated robust performance when utilizing the three-class formulation with multitemporal input. This result is likely due to U-Net's symmetric encoder–decoder structure and skip connections, which are well-equipped to preserve fine-grained spatial details and leverage the richer, nuanced gradient signals provided by a multi-class objective. Conversely, the FCN architecture achieved its highest pixel-level F1-score using the multi-temporal binary formulation. By streamlining the learning objective to a straightforward binary paradigm, the FCN can focus its representational capacity more efficiently on temporal feature differencing, avoiding the added optimization complexity of separating intermediate classes. Ultimately, these disparities underscore the critical role of

label design in experimental workflows; the formulation must be carefully aligned with a model's specific architectural mechanisms for spatial reconstruction and feature extraction.

SegNet and DeepLabv3+ illustrate why architectural sophistication alone does not guarantee better results with limited disaster-specific training data. SegNet's pooling-index decoder retained enough localization information to achieve a site-level F1-score of 65.3% in variant (2a), despite modest pixel agreement. DeepLabv3+'s atrous multiscale context did not consistently resolve small, irregular collapse regions, and increasing the ResNet backbone depth did not improve performance. The very deep variants likely offered more capacity than could be supported by the available labeled examples. RF obtained apparently high overall accuracies because background pixels dominate the AOI, but its lower F1-scores confirm that spectral values without learned spatial context are insufficient for reliable collapse discrimination.

The comparison also demonstrates the importance of reporting metrics at more than one spatial level. Overall accuracy is inflated by class imbalance, pixel-level F1-score penalizes boundary displacement, and site-level F1-score focuses on whether a collapse location is detected at all. These measures answer different questions and should be interpreted together. For emergency screening, a high-recall model such as U-Net (1b) may be appropriate because it minimizes missed sites, if analysts can review false alarms. When field verification capacity is scarce, the higher precision of FCN (2b) may be more useful even though more true collapse sites are omitted.

The large differences in model size reported in Table 2 also have practical consequences. During training with Adam, the gradients and the two moment estimates roughly quadruple this parameter-related memory before the memory needed for intermediate activations is considered; together with the size of the image data, these requirements contributed to the need to restrict the batch size to 1 on the available server (Section Collapsed Building Detection). Larger models require more GPU memory for both training and inference, more storage and transfer time for model checkpoints, and more capable hardware for deployment and more parameters in general mean slower training or inference and higher probability of overfitting; training FCN took more time than training both U-Net and SegNet together. Moreover, this study showed that raising the number of parameters does not equal raising the performance of the model and it is easy to overshoot it, especially with complex design choices (as is the case of DeepLabV3+ with the deepest backbone). For rapid-response mapping, in which a model may need to be retrained for a new event or run on limited hardware, U-Net therefore offers the most favorable balance between accuracy and computational cost, whereas FCN is justified mainly when its higher site-level precision outweighs its substantially greater hardware requirements. Direct measurements of training and inference time on standardized hardware would quantify these differences more precisely and should accompany future comparisons.

Several limitations constrain generalization. The training and validation areas are spatially separate but come from the same city and image pair; consequently, the experiment measures within-domain transfer rather than geographic robustness. The number of annotated collapse sites is small, the classes are strongly imbalanced, and the reference masks inherit uncertainty from image resolution, shadows, occlusion, pre-existing conflict damage, and the absence of complete field records. Results are also affected by co-registration quality and by removal of predicted polygons below 50 m^2.

Furthermore, due to the limited access to the computational server, each configuration was trained only once, so the values in Tables 4–6 come from single training runs without repeated trials or variance estimates. Because network training depends on random weight initialization and other stochastic elements, small differences, such as those between the site-level results of SegNet in variants (2a) and (2b) or the tie between U-Net (1b) and FCN

(2b) at the site level, should not be read as stable rankings, whereas larger gaps, such as the 8.2-percentage-point difference in pixel-level F1-score between U-Net and FCN in variant (1b), are less likely to be explained by run-to-run variation alone. Repeating at least the best-performing configurations with different random seeds and reporting the mean and standard deviation would be needed to confirm these rankings. In addition, all CNNs were trained with the Dice loss only, although compound objectives that combine cross-entropy with the Dice loss, such as BCE+Dice, have been recently reported to be more robust than single losses, mainly in medical image segmentation [98–100]. A joint evaluation of such compound losses under the same factorial design would therefore be a natural extension of this work.

Future studies should include multi-city and cross-sensor validation, controlled tests of registration error, uncertainty-aware prediction, data augmentation or transfer learning for rare collapse examples, and independent object-level verification where field information is available.

Independent ground validation was not feasible under the conflict-related security and access restrictions affecting the study area. The same constraint can arise in active war zones, mine-contaminated land, military or legally prohibited territories, and disaster areas that are physically unsafe or logistically unreachable. In such settings, an independently interpreted image reference enables reproducible within-domain evaluation, but it cannot establish absolute field accuracy or replace validation against on-site measurements when those measurements later become safely and legally obtainable.

5. Conclusions

This study compared four CNN semantic-segmentation architectures and an RF baseline for collapsed-building detection in Jindires using VHR Pléiades imagery. A controlled factorial design tested post-earthquake imagery alone versus a co-registered pre-/post-earthquake stack and compared a three-class formulation with a binary collapse-focused formulation. The same spatially separate training and validation areas were used throughout, and pixel-level metrics were complemented by site-level evaluation across 102 visually interpreted collapse locations.

The comparative evaluation demonstrated that no single architecture or input strategy is universally optimal. The results show that performance depends heavily on the operational priority. U-Net offers a strong accuracy-to-complexity compromise and is preferable when missed collapse sites are especially costly. FCN is preferable when reducing false alarms is the main priority, but it contains approximately fifteen times more trainable parameters than U-Net (Table 2), with correspondingly greater training and deployment requirements. The results also show that pre-event imagery should be treated as an experimental design choice rather than an automatic improvement.

The findings are limited to a single city, a single pair of pre- and post-event images, a small number of marked collapse sites, and reference data derived from image interpretation rather than from exhaustive field verification. Future work should test the geographical transferability across different types of cities and disaster mechanisms, quantify the sensitivity to co-registration and the 50 m^2 threshold in post-processing, evaluate the uncertainty calibration, and compare computationally fewer demanding models or ensemble-based models. External validation using independent field data—provided it can be collected safely and legally—is essential before this method is used for a definitive damage assessment. Higher-quality data in terms of spatial resolution and spectral range would certainly be helpful. Although current performance is insufficient for a definitive damage assessment, the results demonstrate the feasibility of rapid and independent screening of collapsed buildings in inaccessible areas using only satellite imagery. In such situations,

the results can serve as a basis for prioritizing follow-up verification, but not as a substitute for on-site assessment.

Author Contributions: Conceptualization, T.B., O.P. and K.P.J.; methodology, T.B. and O.P.; software, T.B. and O.P.; validation, T.B. and O.P.; formal analysis, T.B. and O.P.; investigation, K.P.J. and E.M.; resources, K.P.; data curation, T.B., O.P. and K.P.J.; writing—original draft preparation, T.B., M.Y.R., Z.A.A. and O.P.; writing—review and editing, O.P., K.P.J., E.M. and K.P.; visualization, T.B.; supervision, E.M. and K.P.; project administration, E.M. and K.P.; funding acquisition, K.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Grant Agency of the Czech Technical University in Prague, grant no. SGS26/053/OHK1/1T/11.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data derived are available from the corresponding author upon reasonable request. The Pléiades imagery cannot be redistributed because of licensing restrictions.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (GPT-5.6 Luna) for the purposes of correcting the grammar. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Aimaiti, Y.; Sanon, C.; Koch, M.; Baise, L.G.; Moaveni, B. War Related Building Damage Assessment in Kyiv, Ukraine, Using Sentinel-1 Radar and Sentinel-2 Optical Images. *Remote Sens.* **2022**, *14*, 6239. [CrossRef]
2. Liu, C.; Fang, D.; Zhao, L. Reflection on earthquake damage of buildings in 2015 Nepal earthquake and seismic measures for post-earthquake reconstruction. *Structures* **2021**, *30*, 647–658. [CrossRef]
3. Liu, W.; Yan, S.; He, S. Landslide damage incurred to buildings: A case study of Shenzhen landslide. *Eng. Geol.* **2018**, *247*, 69–83. [CrossRef]
4. Pistrika, A.K.; Jonkman, S.N. Damage to residential buildings due to flooding of New Orleans after hurricane Katrina. *Nat. Hazards* **2010**, *54*, 413–434. [CrossRef]
5. Mai, P.M.; Aspiotis, T.; Aquib, T.A.; Cano, E.V.; Castro-Cruz, D.; Espindola-Carmona, A.; Li, B.; Li, X.; Liu, J.; Matrau, R.; et al. The Destructive Earthquake Doublet of 6 February 2023 in South-Central Türkiye and Northwestern Syria: Initial Observations and Analyses. *Seism. Rec.* **2023**, *3*, 105–115. [CrossRef]
6. Kırtel, O.; Aydın, F.; Boru, E.; Toplu, E.; Aydın, E.; Sarıbıyık, A.; Dok, G.; Akkaya, A.; Vural, İ.; Öntürk, K.; et al. Seismic damage assessment of under-construction industrial buildings: Insights from the February 2023 Türkiye-Syria earthquakes. *Case Stud. Constr. Mater.* **2024**, *21*, e03507. [CrossRef]
7. Tobita, T.; Kiyota, T.; Torisu, S.; Cinicioglu, O.; Tonuk, G.; Milev, N.; Contreras, J.; Contreras, O.; Shiga, M. Geotechnical damage survey report on February 6, 2023 Turkey-Syria Earthquake, Turkey. *Soils Found.* **2024**, *64*, 101463. [CrossRef]
8. Mavroulis, S.; Mavrouli, M.; Vassilakis, E.; Argyropoulos, I.; Carydis, P.; Lekkas, E. Debris Management in Turkey Provinces Affected by the 6 February 2023 Earthquakes: Challenges during Recovery and Potential Health and Environmental Risks. *Appl. Sci.* **2023**, *13*, 8823. [CrossRef]
9. Bektaş, N.; Zhang, J. Sustainable waste management strategies for earthquake debris: Lessons from the 2008 China and 2023 Türkiye-Syria disasters. *Int. J. Disaster Risk Reduct.* **2025**, *116*, 105153. [CrossRef]
10. Hu, J.; Liu, M.; Taymaz, T.; Ding, L.; Irmak, T.S. Characteristics of strong ground motion from the 2023 *Mw* 7.8 and *Mw* 7.6 Kahramanmaraş earthquake sequence. *Bull. Earthq. Eng.* **2025**, *23*, 1225–1254. [CrossRef]
11. Buckreis, T.E.; Pretell, R.; Sandikkaya, M.A.; Kale, Ö.; Askan, A.; Brandenberg, S.J.; Stewart, J.P. Engineering attributes of ground motions from February 2023 Türkiye earthquake sequence. *Earthq. Spectra* **2024**, *40*, 2268–2284. [CrossRef]
12. Tang, Y.; Şeşetyan, K.; Mai, P.M. Comprehensive ground-motion characterization of the 6 February 2023 M_W 7.8 Pazarcık earthquake in Kahramanmaraş, Türkiye: Insights into attenuation effects, site responses and source properties. *Bull. Earthq. Eng.* **2024**, *22*, 6829–6857. [CrossRef]
13. Gabriel, A.; Ulrich, T.; Marchandon, M.; Biemiller, J.; Rekoske, J. 3D Dynamic Rupture Modeling of the 6 February 2023, Kahramanmaraş, Turkey M_W 7.8 and 7.7 Earthquake Doublet Using Early Observations. *Seism. Rec.* **2023**, *3*, 342–356. [CrossRef]

14. Okuwaki, R.; Yagi, Y.; Taymaz, T.; Hicks, S.P. Multi-Scale Rupture Growth with Alternating Directions in a Complex Fault Network During the 2023 South-Eastern Türkiye and Syria Earthquake Doublet. *Geophys. Res. Lett.* **2023**, *50*, e2023GL103480. [CrossRef]
15. Ren, C.; Wang, Z.; Taymaz, T.; Hu, N.; Luo, H.; Zhao, Z.; Yue, H.; Song, X.; Shen, Z.; Xu, H.; et al. Supershear triggering and cascading fault ruptures of the 2023 Kahramanmaraş, Türkiye, earthquake doublet. *Science* **2024**, *383*, 305–311. [CrossRef] [PubMed]
16. Özkan, A.; Solak, H.İ.; Tiryakioğlu, İ.; Şentürk, M.D.; Aktuğ, B.; Gezgin, C.; Poyraz, F.; Duman, H.; Masson, F.; Uslular, G.; et al. Characterization of the co-seismic pattern and slip distribution of the February 06, 2023, Kahramanmaraş (Turkey) earthquakes (Mw 7.7 and Mw 7.6) with a dense GNSS network. *Tectonophysics* **2023**, *866*, 230041. [CrossRef]
17. Wang, W.; Liu, Y.; Fan, X.; Ma, C.; Shan, X. Coseismic Deformation, Fault Slip Distribution, and Coulomb Stress Perturbation of the 2023 Türkiye-Syria Earthquake Doublet Based on SAR Offset Tracking. *Remote Sens.* **2023**, *15*, 5443. [CrossRef]
18. Tung, S.; Sippl, C.; Shirzaei, M.; Taymaz, T.; Masterlark, T.; Medved, I. Structural Controls on Fault Slip Models of the 6 February 2023 Kahramanmaraş, Türkiye Earthquake Doublet with Finite Element Analyses. *Geophys. Res. Lett.* **2024**, *51*, e2023GL107472. [CrossRef]
19. Öser, C.; Sarğın, S.; Yıldırım, A.K.; Korkmaz, G.; Altınok, E.; Keleşoğlu, M.K. Geotechnical aspects and site investigations on Kahramanmaras earthquakes, February 06, 2023. *Nat. Hazards* **2025**, *121*, 5637–5668. [CrossRef]
20. Taftsoglou, M.; Valkaniotis, S.; Papathanassiou, G.; Karantanellis, E. Satellite Imagery for Rapid Detection of Liquefaction Surface Manifestations: The Case Study of Türkiye–Syria 2023 Earthquakes. *Remote Sens.* **2023**, *15*, 4190. [CrossRef]
21. Karray, M.; Karakan, E.; Kincal, C.; Chiaradonna, A.; Gül, T.O.; Lanzo, G.; Monaco, P.; Sezer, A. Türkiye Mw 7.7 Pazarcık and Mw 7.6 Elbistan earthquakes of February 6th, 2023: Contribution of valley effects on damage pattern. *Soil Dyn. Earthq. Eng.* **2024**, *181*, 108634. [CrossRef]
22. Bayraktar, A.; Ventura, C.E.; Yang, T.Y.; Hökelekli, E.; Taş, Y. Observed Damage Behavior of Earth Dams During the 2023 Kahramanmaraş, Turkey, Earthquakes. *Geotech. Geol. Eng.* **2024**, *42*, 7025–7040. [CrossRef]
23. Richter, C.; Teichert, B.; Pavelka, K. Astronomical Investigation to Verify the Calendar Theory of the Nasca Lines. *Appl. Sci.* **2021**, *11*, 1637. [CrossRef]
24. Dell'Acqua, F.; Gamba, P. Remote Sensing and Earthquake Damage Assessment: Experiences, Limits, and Perspectives. *Proc. IEEE* **2012**, *100*, 2876–2890. [CrossRef]
25. Cotrufo, S.; Sandu, C.; Tonolo, F.G.; Boccardo, P. Building damage assessment scale tailored to remote sensing vertical imagery. *Eur. J. Remote Sens.* **2018**, *51*, 991–1005. [CrossRef]
26. Dong, L.; Shan, J. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS J. Photogramm. Remote Sens.* **2013**, *84*, 85–99. [CrossRef]
27. Brunner, D.; Lemoine, G.; Bruzzone, L. Earthquake Damage Assessment of Buildings Using VHR Optical and SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2403–2420. [CrossRef]
28. Ge, P.; Gokon, H.; Meguro, K. A review on synthetic aperture radar-based building damage assessment in disasters. *Remote Sens. Environ.* **2020**, *240*, 111693. [CrossRef]
29. Foulser-Piggott, R.; Spence, R.; Eguchi, R.; King, A. Using Remote Sensing for Building Damage Assessment: GEOCAN Study and Validation for 2011 Christchurch Earthquake. *Earthq. Spectra* **2016**, *32*, 611–631. [CrossRef]
30. Klonus, S.; Tomowski, D.; Ehlers, M.; Reinartz, P.; Michel, U. Combined Edge Segment Texture Analysis for the Detection of Damaged Buildings in Crisis Areas. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 1118–1128. [CrossRef]
31. Zhang, R.; Duan, K.; You, S.; Wang, F.; Tan, S. A novel remote sensing detection method for buildings damaged by earthquake based on multiscale adaptive multiple feature fusion. *Geomat. Nat. Hazards Risk* **2020**, *11*, 1912–1938. [CrossRef]
32. Romaniello, V.; Piscini, A.; Bignami, C.; Anniballe, R.; Stramondo, S. Earthquake damage mapping by using remotely sensed data: The Haiti case study. *J. Appl. Remote Sens.* **2017**, *11*, 016042. [CrossRef]
33. Wang, L.; Wu, J.; Yang, Y.; Tang, R.; Ya, R. Deep Learning Models for Hazard-Damaged Building Detection Using Remote Sensing Datasets: A Comprehensive Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 15301–15318. [CrossRef]
34. Alisjahbana, I.; Li, J.; Strong, B.; Zhang, Y. DeepDamageNet: A Two-Step Deep-Learning Model for Multi-Disaster Building Damage Segmentation and Classification Using Satellite Imagery. *arXiv* **2024**, arXiv:2405.04800. [CrossRef]
35. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
36. Miura, H.; Aridome, T.; Matsuoka, M. Deep Learning-Based Identification of Collapsed, Non-Collapsed and Blue Tarp-Covered Buildings from Post-Disaster Aerial Images. *Remote Sens.* **2020**, *12*, 1924. [CrossRef]
37. Ding, J.; Zhang, J.; Zhan, Z.; Tang, X.; Wang, X. A Precision Efficient Method for Collapsed Building Detection in Post-Earthquake UAV Images Based on the Improved NMS Algorithm and Faster R-CNN. *Remote. Sens.* **2022**, *14*, 663. [CrossRef]
38. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef] [PubMed]

39. Toby, T.; Gopalakrishnan, U.; Rao, S.N. A Deeper CNN Approach for Detection of Collapsed Buildings in Drone Images. In Proceedings of the 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), Sonepat, India, 8–9 July 2022; pp. 404–410. [CrossRef]
40. Rashidian, V.; Baise, L.G.; Koch, M. Detecting Collapsed Buildings After a Natural Hazard on VHR Optical Satellite Imagery Using U-Net Convolutional Neural Networks. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 9394–9397. [CrossRef]
41. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Springer: Cham, Switzerland, 2015; pp. 234–241. [CrossRef]
42. Wu, C.; Zhang, F.; Xia, J.; Xu, Y.; Li, G.; Xie, J.; Du, Z.; Liu, R. Building Damage Detection Using U-Net with Attention Mechanism from Pre- and Post-Disaster Remote Sensing Datasets. *Remote. Sens.* **2021**, *13*, 905. [CrossRef]
43. Deng, L.; Wang, Y. Post-disaster building damage assessment based on improved U-Net. *Sci. Rep.* **2022**, *12*, 15862. [CrossRef] [PubMed]
44. Ahmadi, S.A.; Mohammadzadeh, A.; Yokoya, N.; Ghorbanian, A. BD-SKUNet: Selective-Kernel UNets for Building Damage Assessment in High-Resolution Satellite Images. *Remote Sens.* **2024**, *16*, 182. [CrossRef]
45. Ma, H.; Liu, Y.; Ren, Y.; Yu, J. Detection of Collapsed Buildings in Post-Earthquake Remote Sensing Images Based on the Improved YOLOv3. *Remote Sens.* **2020**, *12*, 44. [CrossRef]
46. Shi, L.; Zhang, F.; Xia, J.; Xie, J.; Zhang, Z.; Du, Z.; Liu, R. Identifying Damaged Buildings in Aerial Images Using the Object Detection Method. *Remote Sens.* **2021**, *13*, 4213. [CrossRef]
47. Wang, Y.; Cui, L.; Zhang, C.; Chen, W.; Xu, Y.; Zhang, Q. A Two-Stage Seismic Damage Assessment Method for Small, Dense, and Imbalanced Buildings in Remote Sensing Images. *Remote Sens.* **2022**, *14*, 1012. [CrossRef]
48. Gupta, R.; Shah, M. RescueNet: Joint Building Segmentation and Damage Assessment from Satellite Imagery. In *Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021*; IEEE: Piscataway, NJ, USA, 2021; pp. 4405–4411. [CrossRef]
49. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv* **2017**, arXiv:1706.05587. [CrossRef]
50. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In *Computer Vision—ECCV 2018*; Springer International Publishing: Cham, Switzerland, 2018; pp. 833–851. [CrossRef]
51. Shen, Y.; Zhu, S.; Yang, T.; Chen, C.; Pan, D.; Chen, J.; Xiao, L.; Du, Q. BDANet: Multiscale Convolutional Neural Network with Cross-Directional Attention for Building Damage Assessment from Satellite Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5402114. [CrossRef]
52. Chen, F.; Sun, Y.; Wang, L.; Wang, N.; Zhao, H.; Yu, B. HRTBDA: A network for post-disaster building damage assessment based on remote sensing images. *Int. J. Digit. Earth* **2024**, *17*, 2418880. [CrossRef]
53. Liu, J.; Luo, Y.; Chen, S.; Wu, J.; Wang, Y. BDHE-Net: A Novel Building Damage Heterogeneity Enhancement Network for Accurate and Efficient Post-Earthquake Assessment Using Aerial and Remote Sensing Data. *Appl. Sci.* **2024**, *14*, 3964. [CrossRef]
54. Mei, J.; Zheng, Y.B.; Cheng, M.M. D2ANet: Difference-aware attention network for multi-level change detection from satellite imagery. *Comput. Vis. Media* **2023**, *9*, 563–579. [CrossRef]
55. Wang, J.; Guo, H.; Su, X.; Zheng, L.; Yuan, Q. PCDASNet: Position-Constrained Differential Attention Siamese Network for Building Damage Assessment. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5622318. [CrossRef]
56. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image Is Worth 16 × 16 Words: Transformers for Image Recognition at Scale. *arXiv* **2020**, arXiv:2010.11929. [CrossRef]
57. Zheng, S.; Lu, J.; Zhao, H.; Zhu, X.; Luo, Z.; Wang, Y.; Fu, Y.; Feng, J.; Xiang, T.; Torr, P.H.S.; et al. Rethinking Semantic Segmentation from a Sequence-to-Sequence Perspective with Transformers. In *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Piscataway, NJ, USA, 2021; pp. 6877–6886. [CrossRef]
58. Xie, E.; Wang, W.; Yu, Z.; Anandkumar, A.; Alvarez, J.M.; Luo, P. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. *arXiv* **2021**, arXiv:2105.15203. [CrossRef]
59. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 11–17 October 2021; pp. 9992–10002. [CrossRef]
60. Cheng, B.; Misra, I.; Schwing, A.G.; Kirillov, A.; Girdhar, R. Masked-attention Mask Transformer for Universal Image Segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 1280–1289. [CrossRef]
61. Wang, L.; Li, R.; Zhang, C.; Fang, S.; Duan, C.; Meng, X.; Atkinson, P.M. UNetFormer: A UNet-like Transformer for Efficient Semantic Segmentation of Remote Sensing Urban Scene Imagery. *ISPRS J. Photogramm. Remote Sens.* **2022**, *190*, 196–214. [CrossRef]

62. Liu, Z.; Mao, H.; Wu, C.-Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A ConvNet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 11966–11976. [CrossRef]
63. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. Segment Anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2–6 October 2023; pp. 3992–4003. [CrossRef]
64. Cong, Y.; Khanna, S.; Meng, C.; Liu, P.; Rozi, E.; He, Y.; Burke, M.; Lobell, D.B.; Ermon, S. SatMAE: Pre-training Transformers for Temporal and Multi-Spectral Satellite Imagery. In Proceedings of the Advances in Neural Information Processing Systems 35 (NeurIPS 2022), New Orleans, LA, USA, 28 November–9 December 2022; pp. 197–211.
65. Jakubik, J.; Roy, S.; Phillips, C.E.; Fraccaro, P.; Godwin, D.; Zadrozny, B.; Szwarcman, D.; Gomes, C.; Nyirjesy, G.; Edwards, B.; et al. Foundation Models for Generalist Geospatial Artificial Intelligence. *arXiv* **2023**, arXiv:2310.18660. [CrossRef]
66. Gu, A.; Dao, T. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. *arXiv* **2023**, arXiv:2312.00752. [CrossRef]
67. Liu, Y.; Tian, Y.; Zhao, Y.; Yu, H.; Xie, L.; Wang, Y.; Ye, Q.; Jiao, J.; Liu, Y. VMamba: Visual State Space Model. In Proceedings of the Advances in Neural Information Processing Systems 37 (NeurIPS 2024), Vancouver, BC, Canada, 10–15 December 2024.
68. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
69. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.
70. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
71. Encyclopaedia Britannica. Map of Earthquakes and Aftershocks in Turkey and Syria 2023. Available online: https://cdn.britannica.com/68/240568-050-3956F362/map-earthquakes-aftershocks-Turkey-Syria-2023.jpg (accessed on 10 July 2025).
72. Suleiman, A.H.; Hezaber, H. Engineers Count Cost of Rebuilding in Quake-Hit, War-Weary Syria. Available online: https://www.aljazeera.com/news/2023/2/22/syria-building-damage (accessed on 10 July 2025).
73. UNOSAT. Damage Assessment in Idlib City, Syria as of 9 February 2023. Available online: https://unosat.org/products/3487 (accessed on 10 July 2025).
74. Haklay, M.; Weber, P. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Comput.* **2008**, *7*, 12–18. [CrossRef]
75. Results of the 2004 Population and Housing Census at the City/Village Level. Available online: https://archive.ph/20121220144705/http://www.cbssyr.org/new%20web%20site/General_census/census_2004/NH/TAB02-17-2004.htm (accessed on 9 June 2025).
76. Damage Assessment in Jindires, Afrin District, Aleppo Governorate, Syria as of 11 February 2023. Available online: https://reliefweb.int/map/syrian-arab-republic/damage-assessment-jindires-afrin-district-aleppo-governorate-syria-11-february-2023-imagery-analysis-11022023-published-11022023-v1-enar (accessed on 9 June 2025).
77. Sallon, H. In Syria, Villagers Despair After Earthquake: 'We Are Treated like Dogs'. Available online: https://www.lemonde.fr/en/international/article/2023/02/25/in-syria-villagers-despair-after-earthquake-we-are-treated-like-dogs_6017288_4.html (accessed on 9 June 2025).
78. Baladi, E. Jindires Residents Recall Memories of February 6 Earthquake. Available online: https://english.enabbaladi.net/archives/2024/02/jindires-residents-recall-memories-of-february-6-earthquake/ (accessed on 9 June 2025).
79. Ealyawi, A. The Town of Jindires in Syria After the 2023 Turkey–Syria Earthquakes. Available online: https://commons.wikimedia.org/wiki/File:%D8%B2%D9%84%D8%B2%D8%A7%D9%84_%D8%B3%D9%88%D8%B1%D9%8A%D8%A7_%D9%88%D8%AA%D8%B1%D9%83%D9%8A%D8%A7_722023%D8%AC%D9%86%D8%AF%D9%8A%D8%B1%D8%B3_02.jpg (accessed on 10 July 2025).
80. QGIS Development Team. QGIS Geographic Information System. Available online: https://www.qgis.org (accessed on 9 June 2025).
81. European Space Agency. Pléiades–Earth Online. Available online: https://earth.esa.int/eogateway/missions/pleiades (accessed on 9 June 2025).
82. Mukhopadhyay, A.; Pal, I.; Hati, J.P.; Pramanick, N.; Acharyya, R.; Kumar, A.; Jana, S.K.; Mitra, D. High-resolution Pléiades data: An in-depth analysis of applications and future prospects. *Spat. Inf. Res.* **2024**, *32*, 739–755. [CrossRef]
83. DINAMIS/Data Terra. Pléiades Products. Available online: https://dinamis.data-terra.org/en/pleiades-products/ (accessed on 9 June 2025).
84. Perko, R.; Raggam, H.; Roth, P.M. Mapping with Pléiades—End-to-End Workflow. *Remote. Sens.* **2019**, *11*, 2052. [CrossRef]
85. Pešek, O.; Brodský, L.; Halounová, L.; Landa, M.; Bouček, T. Convolutional neural networks for urban green areas semantic segmentation on Sentinel-2 data. *Remote Sens. Appl. Soc. Environ.* **2024**, *36*, 101238. [CrossRef]
86. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.

87. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
88. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
89. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Available online: https://www.tensorflow.org/ (accessed on 9 June 2025).
90. Dice, L.R. Measures of the amount of ecologic association between species. *Ecology* **1945**, *26*, 297–302. [CrossRef]
91. Soille, P.; Burger, A.; De Marchi, D.; Kempeneers, P.; Rodriguez, D.; Syrris, V.; Vasilev, V. A versatile data-intensive computing platform for information retrieval from big geospatial data. *Future Gener. Comput. Syst.* **2018**, *81*, 30–40. [CrossRef]
92. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In Proceedings of the International Conference on Computer Vision, Las Condes, Chile, 11–18 December 2015; pp. 1026–1034. [CrossRef]
93. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
94. Brodský, L.; Landa, M.; Bouček, T.; Pešek, O.; Halounová, L. The LUCAS dataset revisited: Enhancing spatial representativeness for machine learning land cover mapping. *Int. J. Digit. Earth* **2026**, *19*, 2644671. [CrossRef]
95. Syrris, V.; Pešek, O.; Soille, P. SatImNet: Structured and harmonised training data for enhanced satellite imagery classification. *Remote Sens.* **2020**, *12*, 3358. [CrossRef]
96. Gupta, R.; Hosfelt, R.; Sajeev, S.; Patel, N.; Goodman, B.; Doshi, J.; Heim, E.; Choset, H.; Gaston, M. xBD: A Dataset for Assessing Building Damage from Satellite Imagery. *arXiv* **2019**, arXiv:1911.09296. [CrossRef]
97. Microsoft. Bing Maps Aerial Imagery. Available online: https://www.bing.com/maps (accessed on 9 June 2025).
98. Taghanaki, S.A.; Zheng, Y.; Zhou, S.K.; Georgescu, B.; Sharma, P.; Xu, D.; Comaniciu, D.; Hamarneh, G. Combo loss: Handling input and output imbalance in multi-organ segmentation. *Comput. Med. Imaging Graph.* **2019**, *75*, 24–33. [CrossRef] [PubMed]
99. Isensee, F.; Jaeger, P.F.; Kohl, S.A.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2021**, *18*, 203–211. [CrossRef] [PubMed]
100. Ma, J.; Chen, J.; Ng, M.; Huang, R.; Li, Y.; Li, C.; Yang, X.; Martel, A.L. Loss odyssey in medical image segmentation. *Med. Image Anal.* **2021**, *71*, 102035. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.